--This paper has been accepted for publication in *Language Learning and Development*, please do not distribute without permission--

The influence of accent distance on perceptual adaptation in toddlers and adults

Angela Cooper

Melissa Paquette-Smith

Caterina Bordignon

Elizabeth K. Johnson*

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

angela.cooper@utoronto.ca; paquettesmith@psych.ucla.edu; caterina.bordignon@mail.utoronto.ca; elizabeth.johnson@utoronto.ca

Word Count: 10,694

*Corresponding author.

Keywords: accent adaptation; speech perception; perceptual learning; language development

Abstract

Foreign accents can vary considerably in the degree to which they deviate from the listener's native accent, but little is known about how the relationship between a speaker's accent and a listener's native language phonology mediates adaptation. Using an artificial accent methodology, we addressed this issue by constructing a set of three artificial accents (Near, Far, and Farther), varying in the number and magnitude of pronunciation deviations from standard Canadian English. These accents were presented to toddlers and adults in an eye-tracking task. Regardless of accent type, adults readily adapted to the exposed pronunciation change. Adults exposed to the Farther accent were also more willing to accept novel pronunciation changes. Young toddlers exposed to Far or Farther accents showed no evidence of acquiring the exposed pronunciation change and demonstrated worse word recognition than for standard Canadian-accented words. These findings suggest that when a speaker's accent deviates substantially from a young toddler's native accent, this may lead to a significant decrement in their ability to recognize not only an unfamiliar accent but also native-accented speech. Overall, these findings provide a well-controlled test of competing models of accent adaptation and generate multiple hypotheses to be examined in the future using more ecologically valid stimuli.

Keywords: speech perception, word recognition, perceptual learning, accent, language development

1. Introduction

Successful communicative interactions are contingent on our ability to map a fluent, highly variable acoustic speech signal onto discrete linguistic representations stored in memory. Spoken language is rife with inherent variability, stemming for instance from differences in the physiological characteristics of the speaker, their emotional affect, speaking rate, to even the phonological or prosodic context in which a segment is produced. One pervasive source of pronunciation variation arises from a speaker's accent, whether they produce a different dialect of the native language or whether they are a non-native speaker of the language. There are a wide range of different accent types, varying in the degree to which they deviate from listeners' native-accented norms. While previous research has demonstrated that both young children and adults are capable of rapidly adapting to unfamiliar accents (Bradlow & Bent, 2008; Cooper & Bradlow, 2016; Johnson et al., 2021; van Heugten et al., 2018; van Heugten & Johnson, 2014), little is known about how the similarity between a specific accent and the listener's native language phonology may modulate these adaptive processes. In the present work, we examine how accent distance affects perceptual adaptation in young children and adults using an eye-tracking paradigm.

Accent variation introduces not only sub-phonemic variability, where the realization of specific speech sounds varies within a single phonemic category, but can also result in phonemic variability. That is, a change from one accent to another can yield categorical deviations. For example, a native Dutch speaker's difficulty in producing the English interdental fricative may lead to a categorical substitution of /t/ for /θ/, resulting in *think* and *thought* being pronounced as [tɪŋk] and [tɑt] (Hanulíková & Weber, 2012). Accent differences resulting in phonemic changes introduce significant challenges to comprehension, with generally slower and poorer recognition of accented words for adults and children (e.g., Bent, 2014; van Wijngaarden, 2001; see Cristia et al., 2012 for a review)[1]. A considerable body of research has investigated listeners' ability to contend with this acoustic-phonetic variation through perceptual adaptation (e.g., Adank et al., 2010; Baese-Berk et al., 2013; Bradlow & Bent, 2008; Cooper & Bradlow, 2016, 2018; Kraljic et al., 2008; Mitterer & McQueen, 2009; van der Feest & Johnson, 2016; van Heugten & Johnson, 2014; White & Aslin, 2011). Infants as young as 15 months of age show evidence of adaptive processes enhancing subsequent word recognition of accented speech (Paquette-Smith et al., 2020; van Heugten & Johnson, 2014). However, children's ability to cope with accent variation does not achieve an adult-like level until after adolescence, with 8-year-old children still displaying performance decrements with accented word recognition relative to adult listeners (Bent & Atagi, 2017; Bent & Holt, 2018). By adulthood, listeners are capable of adapting to talker-specific characteristics within a matter of minutes (Clarke & Garrett, 2004) and these adaptive adjustments have been found to subsist for days without any intervening exposure (Eisner & McQueen, 2006; Kraljic & Samuel, 2005).

Given that children and adults are capable of adapting to variation in their linguistic environment, what is the nature of these adaptive mechanisms that enable us to so efficiently improve our understanding of other speakers? Numerous studies have suggested that adaptation entails making targeted adjustments to specific phonemic categories (or sets of categories) in response to shifts perceived in the input (Kraljic & Samuel, 2011; Maye et al., 2008; McQueen et al., 2006; Norris et al., 2003; Reinisch et al., 2014). For example, Maye et al. (2008) exposed

[1] Although we acknowledge that mispronunciations of known words can be similarly detrimental to speech processing (see Altvater-Mackensen & Mani, 2013; Mani & Plunkett, 2008; Paquette-Smith et al., 2016; Swingley & Aslin, 2002), here we distinguish the work on mispronunciations (which is concerned with online sensitivity to phonemic or sub phonemic deviations) from work on accent adaptation (which involves learning or adapting to the specific phonetic deviations made by a particular speaker).

listeners to a novel, vowel chain-shifted accent of English where front vowels were systematically lowered (e.g., *keep* /kip/ ➔ “kip” /kɪp/; *witch* /wɪtʃ/ ➔ “wetch” /wɛtʃ/). Listeners first heard a story passage in a standard American English accent followed by a lexical decision task. In a second session, listeners heard the same story passage, but in the vowel-lowered accent, before completing the same lexical decision task. Results revealed that listeners were more likely to consider nonwords such as “wetch” to be real words following exposure to the novel accent, suggesting that they had adjusted their vowel category space to accommodate these shifted exemplars. However, they were not more willing to endorse items with front raised vowels, which were shifts not present in the exposure phase. The authors interpreted these findings as indicating that perceptual learning does not involve a general relaxing of criteria for what is an acceptable exemplar of a vowel category but rather constitutes targeted category shifts. This *targeted linguistically-guided strategy* has been posited to be predominantly driven by top-down linguistic knowledge, including lexical (e.g., Kraljic & Samuel, 2007; Norris et al., 2003) or phonotactic information (Cutler et al., 2008). For example, upon encountering an item such as “wetch”, listeners’ knowledge of the fact that *witch* but not “wetch” is a lexical item will guide their phonetic adjustments, so that future examples, such as “lev”, will be more efficiently and accurately categorized (i.e., *live*).

It is important to consider, however, that there are contexts in which a *targeted linguistically-guided strategy* might not be utilized, such as in situations where top-down information either is not informative (e.g., the pronunciation shift results in a different real word rather than a nonword) or is unavailable. Young children, for example, have incomplete linguistic knowledge and thus will likely encounter a host of lexical items that they do not yet have in their lexicon. As such, these lexical items would not always be able to provide top-down guidance for making adaptive adjustments. Additionally, it could also be the case that the *targeted linguistically-guided strategy* is effective for handling a single or constrained set of pronunciation shifts. However, an accent can have a considerable number of deviation patterns that could make tracking all of them and making all of the necessary adjustments initially quite challenging. Moreover, a non-native speaker could also be inconsistent in their substitutions (e.g., Hanulíková & Weber, 2012), making it difficult for the perceptual system to determine the optimal adjustment to make. Thus, an alternative strategy has been proposed, whereby listeners relax what they consider to be a permissible match between input and representation, expanding their categories to accommodate a certain amount of deviation from their native-accented norms (see Schmale et al., 2015 or Zheng & Samuel, 2020). Recent work has suggested that a *general expansion strategy* may be utilized by toddlers in certain contexts. Schmale et al. (2015) tested 24-month-olds on their word learning abilities, training them with a native English talker and testing them on a Spanish-accented talker. However, prior to training, groups underwent a high-variability exposure phase meant to induce a *general expansion strategy*, either indexical exposure (listening to four native English talkers varying in age and sex) or social exposure (viewing four people gesturing silently). They found that toddlers were capable of recognizing the newly-trained words when produced by a novel, accented speaker, without any prior exposure to that particular accent. These findings are in contrast to previous work using the same stimuli, where no evidence for adaptation was observed without a varied pre-exposure (Schmale et al., 2012). The authors argue that exposure to diversity, either indexical or social variability, led them to be more accepting of pronunciations that deviated from their stored representations. These findings are comparable to subsequent studies showing that exposure to an other-race speaker can facilitate accent adaptation in toddlers (Weatherhead & White, 2018).

Given the wealth of behavioural evidence of the flexibility inherent in successful speech perception, numerous models have been proposed to account for perceptual adaptation to speech variation (e.g., Guediche et al., 2014; Kleinschmidt & Jaeger, 2015; Mirman et al., 2006; Norris et al., 2003; Sohoglu & Davis, 2016). For instance, recent work has proposed a formally explicit Bayesian model of perceptual adaptation (the ideal adapter framework, Kleinschmidt & Jaeger, 2015), whereby speech perception involves a combination of "prediction and inference under uncertainty" (p. 76). The authors posit that for speech perception to occur, listeners build generative linguistic models, which can be defined as knowledge about the distribution of acoustic cues associated with each linguistic unit (e.g., a phonemic category). Listeners utilize knowledge of higher-level linguistic units, comparing them to determine how well each one predicts the incoming signal. Because of the variability inherent in speech perception (e.g., talker or accent-related differences), accurate perception also relies on utilizing the contextually-appropriate generative model. However, because listeners cannot ever truly know the exact nature of the generative model for any given talker or situation, they maintain uncertain beliefs about it. Thus, adaptation is a process where listeners update their beliefs about the cue distributions of a talker- or situation-specific generative model. This can entail shifting the mean of a particular category in the direction of the observed values or increasing the variance of that category (both routes are predicted to be possible by this model). That is, the model predicts that both *linguistically-guided* and *general expansion strategies* can be utilized depending on the particular context. Indeed, recent work has reported evidence for both types of strategies within the same set of listeners (Cooper & Bradlow, 2018).

Considerably less is known about the circumstances that might induce these different adaptation mechanisms. It is also less clear how and when these strategies develop and are utilized across the lifespan. What has not been considered in much detail in prior studies is how the nature of the specific accent and how it relates to the listeners' own native accent might influence the use of these adaptation strategies. Prior work on adaptation has typically provided a foreign or regional accent that listeners may or may not be familiar with (Bradlow & Bent, 2008; Cooper & Bradlow, 2016; Schmale et al., 2012; van Heugten & Johnson, 2014) or an artificial accent (White & Aslin, 2011). But how might the characteristics of the specific accent influence adaptation? There are certain accents that deviate relatively little from one's native accented norms. For instance, one of the primary features that differs across certain dialects of Canadian and American English is the presence of Canadian raising, where /aɪ/ is pronounced as [ʌɪ] in certain phonological contexts (Chambers, 1973). Idiolectal variation can also potentially introduce only a small number of differences from one's own pronunciation patterns; for instance, a speaker may have an unusual pronunciation of one particular segment (e.g., pronouncing [ʒ] instead /z/ in words like *vision*). Other accents, on the other hand, can deviate substantially from one's native accent, containing multiple segmental substitutions that may vary in their phonetic distance to segments present in the native accent. For example, Witteman et al. (2013) reported that accent strength can impact the speed of adaptation with the presence of strongly-accented items (that is, those that phonetically deviate further from native-accented productions) slowing adaptation to a greater extent relative to weakly-accented items. This effect was only present for participants who had minimal prior experience with the accent. For those with extensive experience, adaptation occurred for both strong and weakly-accented items (Witteman et al., 2013).

1.1 The current study

The present work sought to address the question of how the specific characteristics of an accent influence adaptation. To do this we constructed a set of artificial accents and manipulated their distance, defined for the purposes of this study as the number and magnitude of pronunciation deviations, from the listeners' native accent. Additionally, we investigated the influence of linguistic knowledge and how it might interact with accent distance by testing both toddlers and adults. The experimental paradigm consisted of two phases: exposure and test. The exposure phase presented images on a screen that were named with one of the three different artificial accent types (i.e., a Near, Far or Farther accent). In the test phase, participants were presented with pairs of familiar and novel objects accompanied by audio of the speaker producing the items. The test phase items included both exposed and unexposed pronunciation changes in order to gain insight into which strategy listeners might be employing. Because young children are not always able to reliably make selections or explicitly identify items, an eye-tracking paradigm was employed, which has been used extensively in previous research with infants and toddlers as an implicit measure of word recognition (e.g., Altvater-Mackensen & Mani, 2013; Delle Luche et al., 2015; Mani & Plunkett, 2007; Schmale et al., 2012; White & Aslin, 2011; White & Morgan, 2008). Due to low task demands, eye movements are generally considered a more sensitive measure of children's language competency than methodologies involving explicit choice or verbal responses. Furthermore, this ensured that the same task was utilized for both adults (Experiment 1) and young children (Experiment 2). Eye-movements to object images have been found to be time-locked to the spoken words referring to these objects (e.g., Eberhard et al., 1995). As such, word recognition can be inferred from fixations to the named image (target).

If listeners use a *targeted linguistically-guided strategy,* they would be predicted to fixate the target on test trials if labeled with an exposed pronunciation shift but not an unexposed shift, as a consequence of their making a targeted shift to a specific phonemic category. On the other hand, if using a *general expansion strategy*, listeners would be expected to fixate the target on trials if labeled with either exposed or unexposed pronunciation changes. If this *expansion* strategy yields a system-wide loosening of categories, then listeners might be more accepting of novel pronunciations of known words (e.g., “boat” → “sote”) to which they had not been previously exposed.

We hypothesized that the use of these strategies would differ as a function of accent distance. A Near accent, one that deviates relatively little from the native accent, might promote the use of a *targeted linguistically-guided strategy*, as the limited deviations from the listener’s native accent should be readily learnable and not trigger the need to generally expand the range of permissible pronunciations. In contrast, an accent that is farther from one’s own, with multiple pronunciation deviations, might induce the use of a *general expansion strategy*, increasing listeners’ tolerance for mismatches between input and representation and relaxing category boundaries across the phonological system. Moreover, listeners’ linguistic knowledge may interact with the type of accent presented and impact how and when these strategies are utilized. Children, with their relatively less developed linguistic knowledge (less robust representations, greater uncertainty about the phonology and lexicon of their language), might be quicker to utilize a *general expansion strategy* than adult listeners, as it does not rely on the use of top-down linguistic information. Such a finding would fit with the existing literature on real accent adaptation, suggesting children can learn specific re-mappings for some accents (van Heugten et al., 2018), but also readily use the general expansion strategy (Schmale et al., 2015).

To examine these possibilities, this study tested adult listeners (Experiment 1) and 27-month-old toddlers (Experiment 2) utilizing an eye-tracking paradigm. We manipulated the number and magnitude of pronunciation changes present in the accent, ranging from Near, Far to Farther, and participants were randomly assigned to receive exposure to one of those three accent types. During the test phase, listeners heard standard Canadian-accented real words, items with a pronunciation change heard during exposure (Exposed items), items with an unexposed pronunciation change (Generalization items) and nonwords. Listeners' performance with each of these item types will provide insight into which adaptive processes they may be employing.

2. Experiment 1

2.1 Methods

2.1.1 Participants

Seventy-two Canadian-English speaking adult listeners were randomly assigned to one of three accent exposure groups: 1) Near (n = 24, M_{age} = 20 years, 5 males), 2) Far (n = 24, M_{age} = 20 years, 7 males), and 3) Farther (n = 24, M_{age} = 19 years, 6 males). All listeners learned English before the age of 6 in North America, but did not necessarily learn English as their first language. The majority of adult participants were fluent in at least one other language (e.g., Mandarin, Cantonese, Tagalog, Gujarati, Somali, Urdu, Korean, Punjabi, etc.) but currently used English at least 70% of the time. They all self-reported no hearing or vision impairments at the time of testing. One additional participant was tested and excluded because they did not meet the language criterion (i.e., they did not learn English before the age of 6).

2.1.2 Stimuli

In order to control the number and magnitude of pronunciation deviations (and other difficult to control differences between natural accents, such as intonation differences), artificial accents were constructed. Materials for the exposure phase consisted of 16 words (i.e., 4 target items and 12 non-target items; see Appendix A for the complete list of exposure stimuli). The four target items all contained the same vowel. In order to ensure that adaptation effects were not restricted to a particular accent pattern, two sets of artificial accents were constructed: one set contained target items sharing the vowel /æ/ (e.g., “hand”) and the other set contained target items containing the vowel /ʌ/ (e.g., “truck”). In the exposure phase, the 12 non-target words were produced in either a Near, Far or Farther Accent. In the Near Accent, only the target items underwent a pronunciation change (either /æ/ “hand” ➔ [ɑ] “hond” or /ʌ/ “truck” ➔ [oʊ] “troke”). The remaining 12 non-target words were produced with a standard Canadian English accent. In the Far Accent, the same target pronunciation change was provided as in the Near accent (e.g., /æ/ ➔ [ɑ]), but now all 12 non-target words contained a range of vowel and consonantal pronunciation changes. Six different changes were included, two vowel (/ʊ/ ➔ [oʊ], (/ɪ/ ➔ [eɪ]) and four consonant (/s/ ➔ [ʃ], /z/ ➔ [ʒ], /θ/ ➔ [f], /l/ ➔ [w]). Finally, the Farther accent contained the same number of pronunciation changes as the Far Accent but also included larger phonetic deviations in the non-target items. For example, in the Far Accent, the consonantal deviations were a 1-feature change in place of articulation (e.g., /θ/ ➔ [f], changing place of articulation). In the Farther Accent, many of these became 3- feature changes (e.g., /θ/ ➔ [d], changing both place, manner and voicing). It is important to note that in the Near, Far and Farther Accents, the target pronunciation change (either /æ/ ➔ [ɑ] or /ʌ/ ➔ [oʊ]) remained the same. Accent distance was manipulated in the size and number of pronunciation changes present in the non-target items. Cartoon images were selected for each of the 16 exposure items (see Figure 1). Providing images to accompany the shifted pronunciations was intended to provide some context for the labels, helping listeners realize these were altered pronunciations of real words rather than novel words.

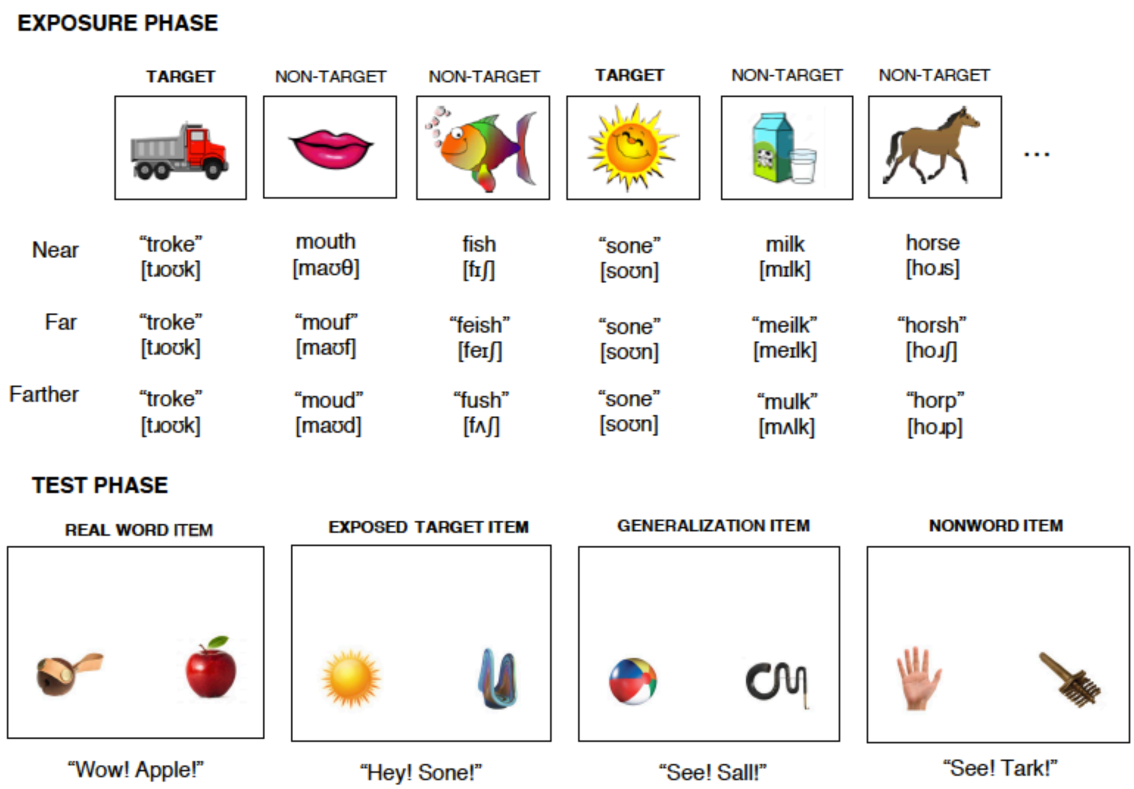

Figure 1. The top panel depicts a sample sequence of trials (left to right) in the exposure phase. Individual images were presented on the screen with concurrent auditory labeling from one of three accent exposure conditions (either Near, Far or Farther). Target item labeling was identical across all 3 conditions (e.g., /ʌ/ "truck" ➔ [oʊ] "troke"). Pronunciation changes in non-target items varied by accent type (Near: no changes to the non-target items; Far: 6 types of changes; Farther: 6 types of changes phonetically larger than in the Far accent). Note: each item was repeated aurally twice per trial. The bottom panel depicts a set of sample test phase trials for each trial type.

The test phase stimuli consisted of 16 items, divided into four types: 1) Canadian-accented real words, 2) Exposed target items, 3) Generalization items containing an unexposed pronunciation change and 4) Nonword items (see Figure 1 for examples; see Appendix B for the complete list of test phase items). Since the artificial accents were constructed (either with target changes /æ/ ➔ [ɑ] or /ʌ/ ➔ [oʊ]), the Exposed target items from one set served as the Canadian-accented real words for the other set. For instance, one set of listeners were exposed to the speaker pronouncing *hand* /hænd/ as [hɑnd] and *apple* /æpl̩/ as [ɑpl̩] in the exposure phase. The other set of listeners heard the Canadian-accented productions of those items (e.g., [hænd] and [æpl̩]) in the test phase. This design allows us to examine whether there are differences in listeners' ability to recognize the /æ/ stimuli set compared to the /ʌ/ stimuli set. Because all the test items were produced by the same speaker, it could be confusing for listeners to hear the same speaker produce both the standard and non-standard pronunciations, which is why we counterbalanced which items served as real words and which items served as Exposed target words across subjects. Since participants were exposed to altered pronunciations of familiar words in all three conditions (i.e., near, far, and farther), we anticipate the possibility that recognition of even Canadian-accented real words in the test phase may be impacted by our exposure manipulation. The generalization items contained an unexposed, 3-feature pronunciation change (/b/ "boat" ➔ [s] "sote") and the nonword items were not real words in English (e.g., mawg, ʃump). By design, the generalization items and non-word items did not contain any of the segments affected by the pronunciation changes in any of the accents.

A female, native speaker of Canadian English (specifically from Southern Ontario) was recorded naturally-producing the auditory stimuli in a child-directed manner. The speaker produced three tokens of each item in isolation, along with attention-getting tokens of "wow", "see", and "hey". Importantly these attention getting tokens did not contain the exposed vowel shifts of /æ/ ➔ [ɑ] or /ʌ/ ➔ [oʊ]. The use of child-directed speech was necessary to enable direct comparisons with the data collected from children in Experiment 2. We have no reason to believe that the use of child-directed speech would negatively impact adult listeners' ability to adapt to the accent. Indeed, infant and child-directed speech has been argued to be clearer than adult-directed speech (e.g., Uther et al., 2007) and has even been argued to facilitate adult word recognition in noise (van der Feest et al., 2019) and word learning in an unfamiliar language (Golinkoff & Alioto, 1995). Stimuli were recorded at a sampling rate of 48,000 Hz and edited in Praat (Boersma & Weenink, 2016).

Thirty-two photo-real images were selected (16 familiar images and 16 novel images, see Figure 1 for an example) as visual stimuli for the test phase. The test phase images were different from the cartoon images used in the exposure phase, which eliminates the possibility that participants could simply match the exposed target words to the images they saw in the exposure phase. Each known object (e.g., apple, boat) was paired with an image of a novel object because many of the auditory labels could be considered new words; thus, the novel object in the image set provided participants with a plausible referent to which the label could refer.

2.1.3 Procedure

Participants were seated in a sound-attenuated booth for the duration of the study and were instructed to watch the screen. The exposure phase began with two warm up trials in which two practice items (i.e., train and dog) were labeled, followed by 24 exposure trials. In each 6000 ms exposure trial, participants were presented with a cartoon image of the item accompanied by audio of the speaker labeling the object two times with either the Near, Far or Farther accent (depending on which exposure group they were assigned to). Each of the four exposed targets were labeled six times over the course of the exposure phase and each of the non-target words were labeled twice. In order to make the task more interesting, and maintain participant's attention to the screen, the images gradually increased and decreased in size throughout the trial before exiting the screen accompanied by a non-speech sound (e.g., a bell ringing).

Following this passive accent exposure, participants completed 24 test trials which utilized an eye-tracking (Intermodal Preferential Looking) task. In each 6000 ms test trial, participants were presented with a pair of objects containing a known object (e.g., apple) and a novel object. The visual stimuli were accompanied by an aurally-presented phrase labelling one of the objects ("Wow! [Target word]!"). The onset of the target word was always aligned to 3000ms after the onset of the trial.

Participants were randomly assigned to one of four versions of the /æ/ ➔ [ɑ] test phase (for participants exposed to the /æ/ ➔ [ɑ] vowel shift during the exposure phase) or the /ʌ/ ➔ [oʊ] test phase (for participants exposed to the /ʌ/ ➔ [oʊ] vowel shift during the exposure phase). Each test order consisted of 24 trials: 4 Canadian-accent real word trials, 8 Exposed target item trials, 8 Generalization trials and 4 Nonword trials (see Appendix B for the complete list of test phase stimuli). The Exposed target trials tested participants on the 4 exposed target items that they previously heard during the exposure phase (that contained either the /æ/ ➔ [ɑ] or the /ʌ/ ➔ [oʊ] shift). If participants had learned the shift from the exposure phase then they should look more to the target object (over the novel object) in the test phase. The Generalization trials tested whether the exposure might lead participants to be more willing to accept a novel 3-feature pronunciation change (e.g., /b/ ➔ [s]) that they were never exposed to. Finally, in the nonword trials, participants were presented with a novel word. Here if participants considered the nonword label to be a new word, we expect listeners to fixate more on the novel item compared to the familiar item. The side of the screen (left or right) in which each object appeared on was counterbalanced across the 4 versions of each test phase.

Each session was videotaped with a remote-controlled camera to allow for frame-by-frame off-line coding using SuperCoder (Hollich, 2005). Each 30-ms frame was coded for whether the participant 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 (i.e., the mean Pearson's r correlation between coder 1 and coder 2's look duration = 0.97 for 4 subjects).

2.1.4 Analysis

For each trial, the proportion of fixation time on the target picture was calculated [target fixation time/(total fixation time to target + distracter)], for the 3000ms window before target word onset (henceforth referred to as the "pre-naming window") and the 2750ms window after target word onset (henceforth referred to as the "post-naming window"). As in previous work, the post-naming window began 250 ms after word onset to allow time for participants to plan and execute an eye movement. To provide an estimate of listeners' word recognition, we subtracted the proportion of fixation time on the target picture in the pre-naming window from the proportion of fixation time on the target picture in the post-naming window to create a difference score for each trial. This enabled us to control for any inherent image preferences, particularly with respect to the unfamiliar items (e.g., White & Aslin, 2011). Owing to the need to establish a baseline looking behaviour for each participant, trials where the participant did not look at both images in the pair during the pre-naming period were excluded from the analysis.

Given the hypotheses of the present work, a series of linear mixed-effects models were conducted to predict difference scores for each of the four test item types: 1) Canadian-accented real words, 2) Exposed target items, 3) Generalization items and 4) Nonword items. All models included the Helmert-coded fixed effect of Accent Distance (A: Near vs. Far + Farther; B: Far vs. Farther) and the simple-coded fixed effect of Exposed Vowel Shift /æ/ ➔ [ɑ] vs /ʌ/ ➔ [oʊ]. Vowel shift was included in the models to examine whether there are differences in test trial performance for participants exposed to the /æ/ versus the /ʌ/ shift. The maximal random effects structure that would converge was implemented (Barr et al., 2013), including random intercepts for participants and items[2]. For the fixed effects, we report *b*, standard error, *t*-values and *P*-values calculated using Satterthwaite approximations to degrees of freedom and implemented using lme4 (Bates et al., 2018) the lmerTest (Kuznetsova et al., 2015) packages in R. In places where the model indicated that there were significant differences between the three Accent exposure groups ($p < .05$), a series of Bonferroni corrected one-sample t-tests were conducted to compare difference scores to chance (0).

2.2 Results and Discussion

As stated above, we conducted a series of LME models predicting the change in the proportion of looks to the target from the pre-naming phase to post-naming phase (difference scores) for each of the four trial types. First, we examined whether adult's adaptation to the exposed target items varied depending on the Accent Distance or the Exposed Vowel Shift. Here we find no differences depending on Accent Distance, (Near vs. Far + Farther; $b = 0.03$, $SE = 0.05$, $t = 0.52$, $p = .608$; Far vs. Farther, $b = 0.02$, $SE = 0.06$, $t = 0.35$, $p = .728$) or Vowel Shift (/æ/ ➔ [ɑ] vs. /ʌ/ ➔ [oʊ], $b = -0.05$, $SE = 0.05$, $t = -0.92$, $p = .360$). This suggests that the findings were not specific to one version of accents or to a specific exposed vowel shift. That

[2] Model structure: lmer(Difference Score ~ Accent Distance + Vowel Shift + (1 | Participant) + (1 | Item))

being said, we do find that overall participants looked significantly more towards the exposed target object in the post-naming period compared to the pre-naming period, indicating that after the exposure phase, participants were adapting to the accent and have learned that "troke", for example, refers to the truck and not the novel object (Mean difference = 0.19), $t(71) = 7.49$, $p < .001$, 95% CI [.14, .24], $d = 0.89$.[3]

As expected, participants were also well above chance (0) in shifting their gaze towards the target objects in the Canadian-accented Real Word trials (Mean difference = 0.19), $t(71) = 8.76$ $p < .001$, 95%CI [.15, .23], $d = 1.04$, indicating that they still recognized familiar words produced in a Canadian accent after exposure. Similar to the Exposed Target trials, recognition of the Canadian accent words was not significantly impacted by Accent Distance, (Near vs. Far + Farther; $b = 0.03$, $SE = 0.05$, $t = 0.64$, $p = .524$; Far vs. Farther, $b = -0.00$, $SE = 0.05$, $t = -0.07$, $p = .948$) or Exposed Vowel Shift, $b = -0.03$, $SE = 0.06$, $t = -0.57$, $p = .586$.

For the Generalization items, we predicted that participant's willingness to accept a novel 3-feature shift (i.e., /b/ ➔ [s]), might depend on whether they were exposed to the Near, Far or Farther accent in the exposure phase. Although differences scores did not seem to be impacted by the type of vowel shift participants were exposed to (/æ/ ➔ [ɑ] vs. /ʌ/ ➔ [oʊ], $b = -0.03$, $SE = 0.04$, $t = -0.83$, $p = .412$), as predicted we do observe a main effect of Accent Distance. Participants exposed to the Near accent were less accepting of the novel shift compared to those exposed to the "Far + Farther" accents, $b = -0.08$, $SE = 0.04$, $t = -2.19$, $p = .032$. There was also a marginally significant difference between the Far vs. Farther accents, $b = -0.08$, $SE = 0.04$, $t = -1.79$, $p = .077$. Follow-up t-tests comparing difference scores in the three

[3] Given that there was no main effect of Accent Distance in the model, here we report a single t-test to chance (0). If differences scores for Exposed Target items were computed separately for participants exposed to the Near, Far and Farther accents, participant's scores would be greater than 0 in all three groups (all *t*s> 3.80, all *p*s< .001)

exposure groups to chance (0)[4], indicated that performance on the generalization items did not differ significantly from zero for both the Near, $t(23) = -0.72$, $p = .482$, 95%CI [-.09, .04], $d = 0.15$, and Far, $t(23) = 0.84$, $p = .409$, 95%CI [-.04, .10], $d = .18$ accent exposure groups. That is, they demonstrated no increase in fixation to either the familiar or unfamiliar object post-naming. However, a significant increase was found in the Farther accent exposure group $t(23) = 3.82$, $p < .001$, 95%CI [.05, .16], $d = 0.80$, indicating that they were more likely to fixate the familiar object upon hearing items with the unexposed pronunciation change.

Finally, in the non-word trials there was no impact of Accent Distance (Near vs. Far + Farther; $b = -0.01$, $SE = 0.04$, $t = -0.26$, $p = .800$; Far vs. Farther, $b = -0.06$, $SE = 0.05$, $t = -1.11$, $p = .270$) or Vowel Shift, $b = 0.03$, $SE = 0.04$, $t = 0.74$, $p = .463$. However, overall there was a significant decrease from pre-naming window in the proportion fixation to the familiar object, suggesting that participants considered the nonword to be a new word referring to the novel object, $t(71) = -5.71$, $p < .001$, 95%CI [-.16, -.08], $d = 0.68$.

[4] To reduce the likelihood of a type I error, the alpha level was Bonferroni adjusted to ($\alpha/3 = .017$). Comparisons where the p <.017 were considered statistically significant.

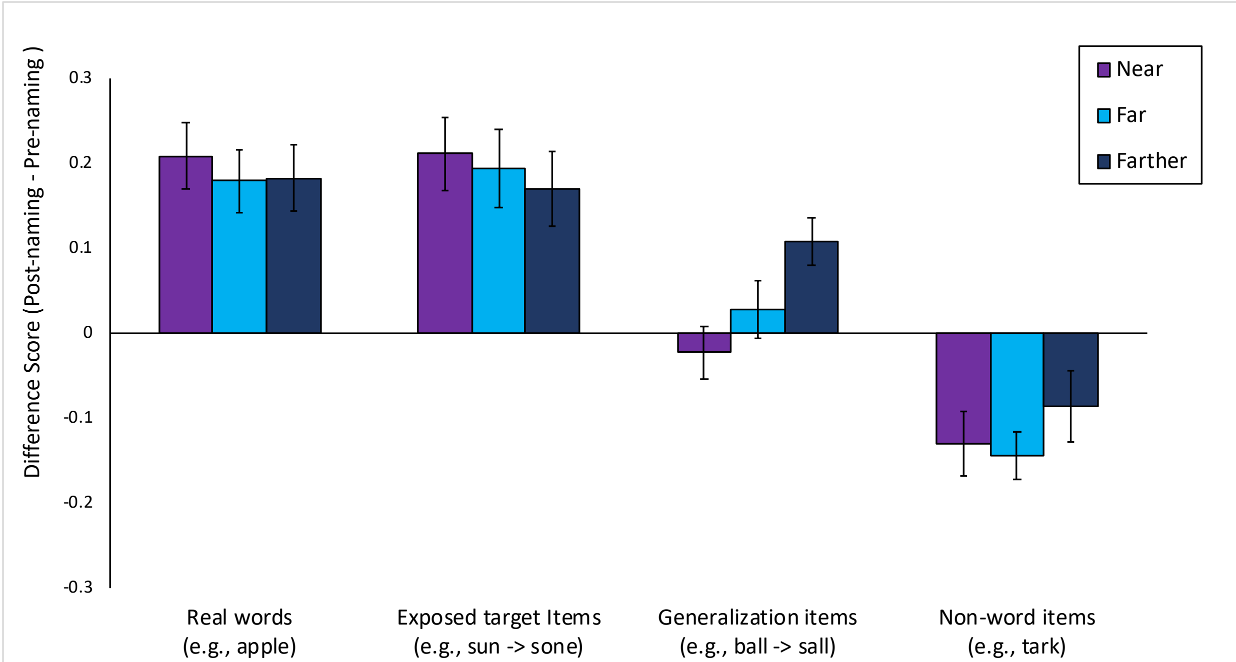

Figure 2. Mean difference scores for adult participants in Experiment 1 for each Accent Distance group (Near, Far, Farther) by item type. Positive values indicate greater proportion fixation to the familiar object; negative values indicate greater proportion fixation to the unfamiliar object. Error bars represent the standard error of the mean.

The results of Experiment 1 indicate that adult listeners adjusted to the exposed pronunciation change, which is consistent with prior research showing adults' rapid adaptive capabilities (e.g., Clarke & Garrett, 2004). This was acquired regardless of whether this pronunciation change was embedded in a Near, Far or Farther accent, suggesting that the presence of multiple pronunciation changes (in the Far and Farther accents) did not inhibit the learning of one specific pronunciation change. Critically, an effect of Accent Distance was found in the perception of items that contained an unexposed (/b/ ➔ [s]) pronunciation change. Participants who were first exposed to either the Near or Far accent were not willing to accept, for instance, [soʊt] "soat" as a possible label for *boat* during the test phase. Those who received Farther accent exposure, on the other hand, were more willing to consider such items as referring to the target image, despite never having been exposed to that specific pronunciation change.

This suggests that exposure to an accent that is relatively distant from one's native accent, containing multiple pronunciation deviations from native accented norms, can induce a *general expansion strategy*, whereby listeners loosen their categories and increase their tolerance for mismatches between input and representation.

It is worth noting that the majority of adult listeners in this experiment, as a product of recruiting from an undergraduate population in a diverse city like Toronto, were bilinguals, who have had exposure to greater linguistic variation in their environment than monolingual listeners. One could imagine that, as a result of this linguistic variation, they might (as a group) be more willing to apply a *general expansion strategy*. However, given that our listeners in the Near and Far accent conditions did not show evidence of utilizing a *general expansion strategy*, one could argue that the type of accent exposure may (at least in a sample of mostly bilinguals) play a greater role in driving the difference in adaptive strategies than listeners' bilingualism. It will be important for future research to compare monolingual and bilingual listeners and their willingness to employ *target linguistically-guided* versus *general expansion* strategies.

Taken together, these findings support our hypothesis that accent distance affects how adults adapt to an unfamiliar accent, utilizing a *targeted linguistically-guided strategy* to make adjustments for which they have evidence, but also resorting to a *general expansion strategy* when facing an accent containing numerous, sizeable pronunciation changes.

3. Experiment 2

While Experiment 1 found evidence for a differential usage of adaptive strategies as a function of accent distance, the question remains as to how these strategies develop and whether factors such as the listeners' linguistic knowledge may play a role in the use of these strategies. Children possess a less-developed linguistic system; they have smaller, newer vocabularies and a heightened level of uncertainty about the linguistic units of their language. Thus, it is conceivable that young children's adaptive processes may differ from that of adults. Relative to adult listeners, young children may be more willing to utilize a *general expansion strategy*, which does not require top-down linguistic information, such as lexical knowledge, but involves a general loosening of categories and what is considered a permissible match between input and representation (Schmale et al., 2015). This might predict that young children would be more willing than adults to accept deviant pronunciations, even for patterns to which they had not yet heard (e.g., unexposed pronunciation changes), if exposed to sufficient variation in the input.

On the other hand, children's heightened linguistic uncertainty could inhibit adaptation, as they may require more evidence than adult listeners before making the appropriate adaptive adjustments. Prior work with adult second language listeners (Cooper & Bradlow, 2018), who in some respects are similar to children in that they hold more linguistic uncertainty in their second language relative to native listeners, showed less successful adaptation in certain contexts. Moreover, this linguistic uncertainty might make children more careful in their adaptive adjustments, making only targeted adjustments to specific categories for which they have evidence in the input. Indeed, White and Aslin (2011) argued for a specificity of adaptation in infants, whereby 19-month-old infants were only willing to accept items that contained the exposed pronunciation change (in this case, /a/ *dog* pronounced as [æ] "dag"), but were not willing to accept the same items produced with a different pronunciation change by the same talker (e.g., /a/ *dog* ➔ [ɛ] "deg"). The authors suggest that young children are not more willing to accept just any deviant pronunciation, but rather, acquire a specific, targeted shift in response to a shift heard in the input. However, it could be the case that the infants in White and Aslin (2011) were not willing to accept novel pronunciations because the talker appeared to be inconsistent in their pattern of productions. In the exposure phase, they heard the talker produce one variant of a set of words during exposure (/a/ ➔ [æ]), but in the test phase, the same talker produced a different variant of the same set of words (/a/ ➔ [ɛ]). The talker's apparent inconsistency in how they produced the same set of words may have prevented children from accepting this new pronunciation change. In the present work, we circumvent this potential issue by providing a completely different set of words containing an unexposed pronunciation change.

Relatively little work has been conducted on young children's accent adaptation abilities and even less on their willingness to generalize their exposure to novel pronunciation changes. To investigate these issues, Experiment 2 tested 27-month-old toddlers using the same paradigm as Experiment 1. The current developmental literature on the perception of accented speech suggests that the ability to utilize accent exposure for adaptation is in place at some point between 19-24 months of age (e.g., Best et al., 2009; Schmale et al., 2011; van der Feest & Johnson, 2016; van Heugten et al., 2015). 27-month-olds were selected for the present study, as toddlers' perceptual and linguistic abilities at this age were considered to have matured sufficiently for adaptation to reliably take place (van Heugten & Johnson, 2016). Indeed, this age group was tested because the focus of the current study was not on *whether* children are able to adapt to accented speech but *how* this adaptation takes place.

3.1 Method

3.1.1 Participants

Seventy-two 27-month-old toddlers were tested, divided between one of three accent exposure groups: 1) Near (n = 24, M_{age} = 26.82 months, range = 26.13-27.99 months, 10 males), 2) Far (n = 24, M_{age} = 26.80 months, range = 26-27.95 months, 11 males), 3) Farther (n = 24, M_{age} = 26.94 months, range = 26.20-28.59, 9 males). All were Canadian-English learning toddlers raised in households where English was spoken at least 80% of the time (i.e., in contrast to our adults, all the toddler participants were learning English from birth). Parents reported no hearing impairments or recent ear infections at the time of testing. Caregivers completed the MacArthur-Bates Words and Sentences Communicative Development Inventory (CDI) to provide an approximate index of the child's expressive vocabulary size. A one-way ANOVA found no significant difference in vocabulary size between the three exposure groups, $F(2,65) = 1.79$, $p = .175$, $\eta^2 = .05$.[5] An additional 22 toddlers were tested but excluded due to fussiness (i.e., 6 children from Near Accent group, 6 from the Far Accent group, 10 from the Farther Accent group).

3.1.2 Stimuli, Procedure & Analysis

The materials and procedure were identical to Experiment 1. Children completed the same task as adults, except that they were seated on their caregiver's lap. Caregivers wore over-ear headphones playing masking music (i.e., a combination of music and speech stimuli) to prevent them from potentially biasing the child's looking behaviour. Children were instructed to watch the screen and the images being named.

As in Experiment 1, difference scores were calculated for each participant for each of the four test trial types and were used as the dependent variable in a series of linear mixed effects regression models (Baayen et al., 2008). The model included Accent Distance (Helmert coded; A: Near vs. Far + Farther; B: Far vs. Farther) and Exposed Vowel Shift (contrast-coded; /æ/ ➔ [ɑ] vs. /ʌ/ ➔ [oʊ]) as fixed effects and random intercepts for participants and items[6].

[5] N = 68; The caregivers of four children did not submit their vocabulary forms and were excluded from this ANOVA.

[6] Model structure: lmer(Difference Score ~ Accent Distance + Vowel Shift + (1 | Participant) + (1 | Item))

3.2 Results and Discussion

In the model predicting difference scores for the Canadian-accented real words, there was no difference between children trained on the /æ/ ➔ [ɑ] vs. the /ʌ/ ➔ [oʊ] vowel shift, b = -0.04, SE = 0.04, t = -0.97, p = .368. There was, however, a significant main effect of Accent Distance (Near vs. Far + Farther), b = 0.08, SE = 0.04, t = 2.03, p = .046, indicating that children exposed to the Near accent were more accurate at recognizing familiar, Canadian-accented words relative to those exposed to the Far and Farther accents. No differences were observed between the Far accent and Farther accent, b = -0.00, SE = 0.05, t = -0.01, p = .995. A series of Bonferroni corrected t-tests[7] were conducted to compare children's difference scores to chance (0) for each of the three exposure accents. Children were above chance in their recognition of the Canadian-accented real words after exposure to the Near accent, $t(23)$ = 5.32, p <.001, 95%CI [.10, .23], d = 1.11. Children in the Far, $t(23)$ = 2.31, p = .030, 95% CI [.01, .16], d = 0.48, and Father conditions, $t(23)$ = 2.20, p = .038, 95%CI [.004, .14], d = 0.46, *tended* to look towards the familiar object, but their difference scores were not statistically different from 0 after correcting for multiple comparisons. This suggests that although toddlers should be able to recognize familiar words produced in their native accent, exposure to productions that deviated substantially from their Canadian-accented norms seemed to disrupt their recognition of the known Canadian-accented real words.

For the Exposed Target words, there were no differences in performance as a result of Accent Distance (Near vs Far+Farther, b = 0.04, SE = 0.02, t = 1.44, p = .154; Far vs. Farther, b = -0.02, SE = 0.03, t = -0.68, p = .497). However, there was some indication that performance

[7] To reduce the likelihood of a type I error, the alpha level was Bonferroni adjusted ($\alpha/3$ = .017). T-tests where p < .017 were considered statistically significant.

varied as a result of the Exposed Vowel Shift, $b = 0.06$, $SE = 0.03$, $t = 2.34$, $p = .061$. Toddlers who were exposed to the /æ/ ➔ [ɑ] pattern tended to show stronger recognition of the exposed target words relative to children exposed to the /ʌ/ ➔ [oʊ] pattern. Although, this difference was not statistically significant, it could be an indication that certain vowel shifts might be easier to learn. For instance, here the /æ/ ➔ [ɑ] may be more in line with recent shifts in the production of Canadian English, and thus may be acquired more easily. Although there was no main effect of Accent Distance, we conducted a series of exploratory Bonferroni corrected one sample t-tests to compare the difference scores to 0 in each of the three exposure conditions. This suggested that children recognized the targets that they were exposed to in the Near accent, $t(23) = 2.79$, $p = .011$, 95%CI [.02, .10], $d = 0.58$. The Far and Farther groups, on the other hand, did not show evidence of a significant change from the pre-naming window (Far $t(23) = 0.42$, $p = .675$, 95% CI [-.04, .06], d =0 .09; Farther: t(23) = 1.00, p = .327, 95%CI [-.03, .09], $d = 0.21$). This could hint that accents that contain multiple feature changes might be more difficult for children to acquire.

For the generalization items (i.e., that contained an unexposed shift) there were no main effects of Accent Distance (Near vs. Far+Farther, $b = -0.00$, $SE = 0.03$, $t = -0.13$, $p = .898$; Far + Farther $b = -0.02$, $SE = 0.03$, $t = -0.69$, $p = .490$) or Exposed Vowel shift, $b = -0.00$, $SE = 0.03$, $t = -0.15$, $p = .883$. Overall, performance on the generalization items did not differ significantly from zero, indicating that children, regardless of which exposure phase they heard, were not willing to accept a novel pronunciation change (/b/ ➔ [s]) in the test phase, $t(71) = -0.82$, $p = .415$, 95%CI [-.04, .02], $d = 0.10$.

Similar to the generalization trials, for the nonword trials, we see no main effect of Vowel shift, $b = -0.01$, $SE = 0.04$, $t = -0.22$, p =.831, and no difference between Near vs Far +

Farther accents, *b* = 0.01, *SE* = 0.05, *t* = 0.20, *p* = .843. Although unexpected, there was a trend towards a difference in the Far versus Farther exposed children, *b* = 0.10, *SE* = 0.05, *t* = 1.81, *p* = .075. This association is relatively weak and could simply reflect differences in children's willingness to fixate the novel object upon hearing a novel word.

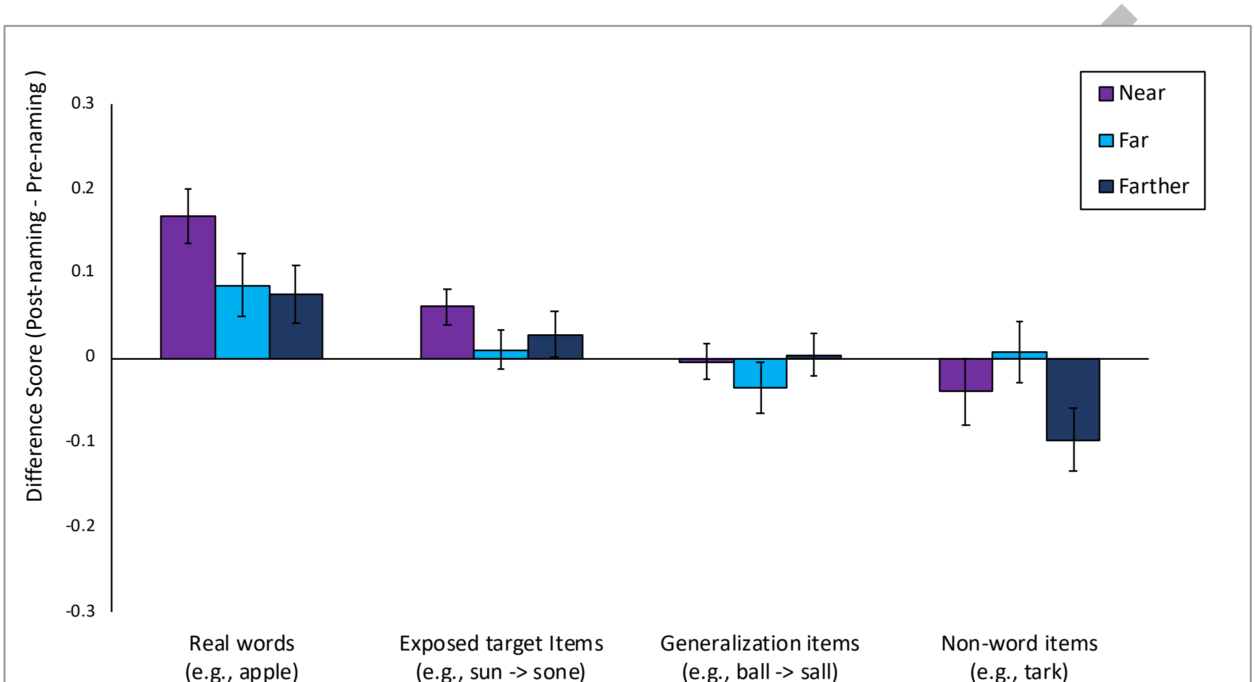

Figure 3. Mean difference scores for child participants in Experiment 2 for each accent distance group (Near, Far, Farther) by item type. Positive values indicate greater proportion fixation to the familiar object; negative values indicate greater proportion fixation to the unfamiliar object. Error bars represent the standard error of the mean.

Finally, to compare performance between the adult participants in Experiment 1 and the child participants in Experiment 2, LME models (Baayen et al., 2008) were conducted for each item type, with difference scores as the dependent measure, a contrast-coded fixed effect of Age (Adult vs Child), and Helmert simple-coded fixed effects of Accent Distance (A: Near vs. Far +

Farther; B: Far vs. Farther)[8]. There was a significant main effect of Age for the Real word trials, $b = 0.08$, $SE = 0.03$, $t = 2.81$, $p = .006$, Exposed Target items, $b = 0.16$, $SE = 0.03$, $t = 5.44$, $p < .001$ and Generalization items, $b = 0.05$, $SE = 0.02$, $t = 2.12$, $p = .036$, with adults showing a greater increase in looks to the target from pre-naming window compared to children. For the nonwords there was also a main effect of Age, $b = -0.07$, $SE = 0.03$, $t = -2.44$, $p = .016$, however here adults looked less to the known object (and more to the novel objects) than children. The main effects of Accent Distance are not particularly informative because they combine both the children and adult's data and have been discussed separately in the models above (all $|t| < 1.92$, $p > .057$). In order to determine whether children and adults reacted differently to the trial types based on their exposure we would have to examine the interactions between Age and Accent Distance. Here we observe a significant Age x Accent Distance (Far vs. Farther) interaction for the nonword items, $b = -0.15$, $SE = 0.07$, $t = -2.07$, $p = .040$, with children exposed to the Farther accent considering nonwords as referring to the novel object significantly more than those exposed to the Far accent, which was not the case for the adult listeners (who more uniformly thought the nonwords referred to the novel objects). Moreover, for the generalization items (that had the unexposed pronunciation change), there was a trend towards an interaction between Age and Accent Distance (Near vs. Far + Farther), $b = -0.08$, $SE = 0.05$, $t = -1.70$, $p = .091$. That is, accent distance seemed to influence the extent to which adult listeners considered items with the unexposed pronunciation change to refer to real words; however, this did not seem to be the case for young children. No other item types yielded significant Age x Accent Distance interactions (all $|t| < 1.00$, $p > .320$). The findings of Experiment 2 revealed that accent distance can have a significant influence on the perception not only of accented pronunciations but also of canonical,

[8] Model structure: lmer(Difference Score ~ Accent Distance * Age + (1 | Participant) + (1 | Item))

native-accented pronunciations by young children. While children demonstrated the ability to recognize familiar, Canadian-accented real words in the Near accent, toddlers exposed to Far or Farther accents were less accurate at recognizing these real words. In other words, exposure to an accent that deviated considerably from the listeners' native accent, containing numerous vowel and consonant pronunciation changes, led toddlers to be less certain about the identity of native-accented words.

Moreover, it seems that exposure to the Far or Farther accent may have also influenced children's learning of the exposed target shifts. Children who received Near accent exposure, with only a single pronunciation change, were above chance in recognizing those items in the test phase. This is consistent with prior work demonstrating that young children are capable of acquiring a specific shift in the exposure accent (White & Aslin, 2011). Those exposed to a Far or Farther accent, however, did not show evidence of having acquired the target pronunciation change. Unlike the adult listeners in Experiment 1, exposure to the Far or Farther accent may have impaired young children's ability to adapt to the artificial accent, perhaps because they were attempting to track multiple pronunciation changes during exposure. Taken together, these findings highlight that the type of accent to which young children are exposed, and its relationship to their native accent, can impact young children's adaptive processes.

4. General Discussion

The current study examined the influence of accent distance on adults' and toddlers' adaptation to an unfamiliar accent. Manipulating the number and size of pronunciation changes present in a set of artificial accents (ranging from Near, Far to Farther from the native accent) provided unique insight into how adaptation might change as a function of accent distance and the listener's age/linguistic experience. Consistent with prior work demonstrating their robust capacity for contending with pronunciation variation (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004; Samuel & Kraljic, 2009), adult listeners in Experiment 1 recognized items containing the exposed target pronunciation change, regardless of the type of accent to which they were initially exposed. Interestingly, adults exposed to an accent that contained numerous, sizeable pronunciation changes (Farther accent) also demonstrated use of a *general expansion strategy*, whereby they were more willing to accept items containing an unexposed pronunciation pattern that they had not heard previously (e.g., *boat* ➔ "sote"). This was in contrast to listeners exposed to the Near or Far accents, who did not show an increase from the pre-naming window in their proportion fixation to the familiar object for this trial type (Figure 2).

Adult listeners have been posited to track the distributional statistics for a given talker's productions, maintaining uncertain beliefs about the linguistic generative model, including beliefs about the amount of potential category variation that that talker produces (Kleinschmidt & Jaeger, 2015). If their exposure to that talker contains minimal shifts (Near) from the cue distributions in the model, then only targeted adjustments appear to be made. Therefore, if they encounter new productions that do not align with those specifically-adjusted distributions, those items will not be recognized. The findings of the present work suggest that exposure to considerable linguistic variation can lead to a *general expansion strategy*, whereby adults update their beliefs about permissible category variance in the model. Though this expansion appears to not simply permit just any kind of pronunciation variation, as evidenced by the Far accent group. It could be the case that if adult listeners have prior evidence that the talker tends to produce a certain size of shift, this strategy will expand categories throughout the linguistic system to accommodate similarly-sized shifts. This would account for the discrepancy between the Far and

Farther accent groups. Despite the fact that the two accent types had the same number of pronunciation changes, the deviations were smaller (1- and 2-feature changes for instance) in the Far relative to the Farther accent (containing a number of 3-feature changes). It is important to note that the unexposed pronunciation change presented in the generalization trials was a 3-feature change. Exposure to the Farther accent may have induced an expansion strategy that expanded categories to accommodate more sizeable pronunciation shifts, allowing them to accept the unexposed pronunciation change in this case. The Far accent group may have also utilized a *general expansion strategy*; however, if the magnitude of category expansion was congruent with prior evidence (in this case, smaller shifts), then they would not accept this particular unexposed pronunciation change. Based on this hypothesis, if the unexposed change was a smaller shift (e.g., /b/ *boat* ➔ [p] "pote"), the Far accent group would be predicted to accept such items. It remains for future work to investigate this issue. We also acknowledge that the positioning of the unexposed pronunciation change (at word onset), might have influenced participant's willingness to accept the pronunciation.

Additionally, the present research examined how listeners' age and linguistic knowledge interacted with the type of accent to influence perceptual adaptation. To examine this, we tested toddlers on an identical paradigm in Experiment 2. While there was no significant main effect of Accent Distance on children's looks to the exposed target, only children who were exposed to an accent that deviated relatively little from their own native accent (Near) were capable of recognizing exposed target items during the test phase. This is consistent with prior research demonstrating that young children are capable of making specific, targeted adjustments in response to a shift in the input (e.g., White & Aslin, 2011). In contrast, children who were exposed to an accent that substantially differed from their native accent (Far or Farther) showed no evidence of recognizing the exposed target items. It is possible that their inability to successfully acquire the exposed pronunciation change could be the result of a higher processing load during the exposure phase, as they may have attempted to track the multiple pronunciation changes present in the Far/Farther accents. Moreover, the Far and Farther accents contained changes to consonants and vowels, whereas the Near accent only shifted the vowels of the target words. Given the greater reliance on consonants to encode words (Nazzi & Cutler, 2019), children's word recognition could have been more disrupted in listening to the Far and Farther accents compared to the Near accent. The current results hint that the specific characteristics of the accent (e.g., the size and number of pronunciation changes) may have an impact on young children's ability to adapt to the accent and could actually inhibit word recognition if the accent is sufficiently deviant from the listeners' native-accented norms.

Moreover, we were surprised to see that these same children (Far or Farther accent exposure) were also less accurate in recognizing Canadian-accented real words relative to children exposed to the Near Accent. Although the differences between groups are admittedly small (thus necessitating replication in future work), one potential explanation for this pattern of results could be that the Far and Farther accent groups actually adopted a general expansion strategy, and that the general loosening of permissible variation resulted in heightened uncertainty about native-accented words. However, if they had adopted such a strategy, we would have expected to see them recognize, at the very least, the exposed pronunciation changes along with the untrained pronunciation changes. Instead, it is conceivable that just the presence of all the pronunciation deviations during the exposure phase heightened their uncertainty about the speaker's overall pronunciation accuracy. Thus, even when encountering real words that were unaffected by the speaker's accent, this uncertainty might have resulted in their being more cautious and considering both options in the image pairs at test to a greater extent. These findings are in line with recent research reporting slower and less accurate word recognition of familiar, Canadian English-accented words by 24-month-old infants who received regular exposure to both Canadian and non-native variants of English relative to children exposed to only Canadian-accented productions (Buckler et al., 2017; see Durrant et al., 2015 and Floccia et al., 2012 for related findings). Buckler et al., (2017) noted two possible explanations for this finding: 1) a difference in robustness of representation, or 2) a difference in word recognition strategy. In the first case, multi-accent children necessarily receive less exposure to Canadian-accented tokens and so may have weaker representations, leading to slower response times. Alternatively, multi-accent children, as a product of regularly experiencing highly variable pronunciations, may generally be "more conservative in accessing lexical representations" (p. 97, Buckler et al., 2017). The results of the current study provide evidence in support of the latter explanation. As children in the present work were predominantly exposed to Canadian-accented English, we cannot attribute differences between the Near and Far/Farther accent groups in recognizing Canadian-accented real words to differences in the robustness of their lexical representations. Rather, it appears that highly variable accent exposure (even relatively short-term exposure in a laboratory context) can have an impact on the recognition of native-accented words.

Finally, children showed no differences as a function of accent distance for the generalization items with the unexposed pronunciation change. This may be surprising, particularly in light of the findings from Schmale et al. (2015), who reported that young children exposed to indexical or social variation were more willing to accept non-standard pronunciations to which they had never been exposed. We predicted that linguistic variation, in the form of accent distance, would have a similar effect, enabling children to be more accepting of non-standard forms. One critical difference, however, between Schmale et al. (2015) and the current research is that the former study was a word learning task rather than recognition of familiar words. It is conceivable that young children resort to a *general expansion strategy* when faced with newly-formed lexical representations in conjunction with considerable variation. Indeed, even adult listeners have been found to be more accepting of small relative to large mispronunciations for newly-learned words (White et al., 2013). When presented in the context of substantial variation, young children might be less certain about new, relatively weaker lexical representations and so utilize an *expansion strategy,* which will be more accommodating of variation along with any potential encoding errors during word learning. The results of the current study might have been different if our artificial accents included other clues that they were actually different accents, like prosodic differences (natural accents often differ along prosodic as well as segmental dimensions, and children are particularly attuned to prosodic info).

Why then would adult listeners utilize a *general expansion strategy* when exposed to the Farther accent, considering their lexical representations should be even more robust than young children's? One speculation is that children may utilize a general expansion strategy less readily because it may be too computationally expensive. This kind of expansion would increase the number of possible representations that the input could be matched with, which would make lexical retrieval more processing intensive. While there has been some evidence for the use of a *general expansion strategy* (Schmale et al., 2015), its use may be constrained to exposure contexts that provide no relevant linguistic information (e.g., watching the hand gestures of a diverse set of people) or contexts where the child is uncertain about the accuracy of their encoding (e.g., word learning). Moreover, adults have considerable experience with cross-talker and cross-accent variation and are thus implicitly aware that highly-accented speakers will likely produce new pronunciation changes. It is conceivable that as an individual gains more experience with linguistic variation over their lifespan, they are more willing to utilize a *general expansion strategy* that would accommodate these novel pronunciation variants, as the possible benefits outweigh the potential processing costs. This is the first study to compare accent adaptation, specifically the influence of accent distance on those adaptive mechanisms, with adults and toddlers using the same task. Future work is needed to fully understand the differences in how adaptation takes place across the lifespan.

After exposure to the Far or Farther accents, children had difficulty recognizing both real words and the exposed target words. However, after exposure to the same accents adults did not show the same degree of impairments in their word recognition. There are a number of reasons why might this be the case. Adult listeners have more robust linguistic knowledge and greater experience with accented speakers. Furthermore, young adults also benefit from having superior executive function (e.g., Thomason et al., 2008), including working memory and inhibition, which has been implicated in the adaptation to accented speech (e.g., Banks et al., 2015; Janse & Adank, 2012). Such linguistic and cognitive advantages may have enabled adult listeners to better contend with an unfamiliar accent that is highly deviant from their own native accent, enabling them to better track the multiple pronunciation changes presented. It is possible that exposure to highly-accented speech could impact adaptation (Witteman et al., 2013) and even the recognition of standard forms for adult listeners, but the test task, which was designed to be suitable for young children, was not sensitive enough to detect it. Future work may consider employing a more challenging task targeted specifically for adults to further investigate this question.

The findings of the present work raise a number of additional questions for future research. What is the impact of exposure to a “far” accent for young children on their processing of familiar, native-accented words? In Experiment 2, we saw a performance decrement for native-accented real words for children exposed to the Far and Farther accents. The question remains as to whether or not this is a temporary effect following immediate exposure to a more distant accent that would rebound relatively quickly, perhaps with intervening exposure to a native-accented speaker. Moreover, is it talker-specific? Children exposed to a Far or Farther accent heard that same speaker in the test phase producing words in their Canadian-accented form (as those items were unaffected by the accent). However, children knew that speaker possessed a non-standard accent and might have been uncertain as to whether what they were hearing was actually an appropriate match to their stored representations. Would we see this same performance decrement if the items were all produced by a different speaker? The majority of previous research on accent adaptation with young children has involved exposing them to a particular speaker and then testing them on the same speaker (e.g., van Heugten & Johnson, 2014; White & Aslin, 2011). There have been a handful of studies involving exposing young children to several speakers before testing them on a different speaker (Potter & Saffran, 2017; Schmale et al., 2012), but more research is needed to tease apart whether their adaptation is talker-specific or generalizes to new talkers sharing the same accent.

To conclude, the current study provides insight into how young children and adults contend with a highly variable speech signal. Findings demonstrated that how deviant a speaker’s accent is from the listeners’ native accent can have a significant impact on the recognition of not only non-standard accented speech but also native-accented speech. Moreover, evidence was found for the use of different adaptive strategies (*targeted linguistically-guided* *strategy* vs. *general expansion strategy*), dependent on the magnitude of variation present during accent exposure as well as listeners' linguistic knowledge.

Acknowledgments

We would like to thank Momina Raja, Yazad Bhathena, and Lisa Hotson, as well as the other members of the Child Language and Speech Studies Lab for their assistance in completing this study. We also would like to extend our thanks to the families who participated in this study. 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 awarded to EKJ. Portions of this work were presented at Laboratory Phonology 16.

References

Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science*, *21*(12), 1903–1909. https://doi.org/10.1177/0956797610389192

Altvater-Mackensen, N., & Mani, N. (2013). The impact of mispronunciations on toddler word recognition: Evidence for cascaded activation of semantically related words from mispronunciations of familiar words. *Infancy*, *18*(6), 1030–1052. https://doi.org/10.1111/infa.12022

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

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America*, *137*(4), 2015–2024. https://doi.org/10.1121/1.4916265

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

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., & Green, P. (2018). *Linear Mixed-Effects Models using "Eigen" and S4* (1.1-17). http://lme4.r-forge.r-project.org/

Bent, T. (2014). Children's perception of foreign-accented words. *Journal of Child Language*, *41*, 1334–1335. https://doi.org/10.1017/S0305000913000457

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., & 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., Gooding, T., Orlando, C., & Quann, C. (2009). Development of phonological constancy. *Psychological Science*, *20*(5), 539–542. https://doi.org/10.1111/j.1467-9280.2009.02327.x

Boersma, P., & Weenink, D. (2016). *Praat: doing phonetics by computer* (6.0.05). http://www.praat.org

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

Chambers, J. K. (1973). Canadian raising. *Canadian Journal of Linguistics/Revue Canadienne De Linguistique*, *18*(2), 113–135. https://doi.org/10.1017/S0008413100007350

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

Cooper, A., & Bradlow, A. R. (2016). Linguistically guided adaptation to foreign-accented speech. *The Journal of the Acoustical Society of America*, *140*(5), EL378–EL384. https://doi.org/10.1121/1.4966585

Cooper, A., & Bradlow, A. R. (2018). Training-induced pattern-specific phonetic adjustments by first and second language listeners. *Journal of Phonetics*, *68*, 32–49. https://doi.org/10.1016/j.wocn.2018.02.002

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*, 1–15. https://doi.org/10.3389/fpsyg.2012.00479

Cutler, A., McQueen, J. M., Butterfield, S., Norris, D., & Planck, M. (2008). Prelexically-driven perceptual retuning of phoneme boundaries. In J. Fletcher, D. Loakes, M. Wagner, & R. Goecke (Eds.), *Proceedings of Interspeech 2008*.

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

Durrant, S., Delle Luche, C., Cattani, A., & Floccia, C. (2015). Monodialectal and multidialectal infants' representation of familiar words. *Journal of Child Language*, *42*(2), 447–465. https://doi.org/10.1017/S0305000914000063

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye Movements as a Window into Real-Time Spoken Language Comprehension in Natural Contexts. *Journal of Psycholinguistic Research*, *24*(6), 409–436.

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *The Journal of the Acoustical Society of America*, *119*(4), 1950–1953.

https://doi.org/10.1121/1.2178721

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

Golinkoff, R. M., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: implications for language acquisition. *Journal of Child Language*, *22*, 703–726.

Guediche, S., Blumstein, S. E., Fiez, J. A., & Holt, L. L. (2014). Speech perception under adverse conditions: insights from behavioral, computational, and neuroscience research. *Frontiers in Systems Neuroscience*, *7*(January), 126. https://doi.org/10.3389/fnsys.2013.00126

Hanulíková, A., & Weber, A. (2012). Sink positive: linguistic experience with th substitutions influences nonnative word recognition. *Attention, Perception & Psychophysics*, *74*(3), 613–629. https://doi.org/10.3758/s13414-011-0259-7

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

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology*, *65*(8), 1563–1585. https://doi.org/10.1080/17470218.2012.658822

Johnson, E. K., van Heugten, M., & Buckler, H. (2021). Navigating accent variation: a developmental perspective. *Annual Review of Linguistics*.

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

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition*, *107*(1), 54–81. https://doi.org/10.1016/j.cognition.2007.07.013

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*(2), 141–178. https://doi.org/10.1016/j.cogpsych.2005.05.001

Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, *56*, 1–15. https://doi.org/10.1016/j.jml.2006.07.010

Kraljic, T., & Samuel, A. G. (2011). Perceptual learning evidence for contextually-specific representations. *Cognition*, *121*(3), 459–465. https://doi.org/10.1016/j.cognition.2011.08.015

Kuznetsova, A., Brockhoff, B., & Christensen, R. H. B. (2015). *Tests in Linear Mixed Effects Models* (2.0-29). https://cran.r-project.org/web/packages/lmerTest/index.html

Mani, N, & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252–272. https://doi.org/10.1016/j.jml.2007.03.005

Mani, Nivedita, & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*(1), 53–59. https://doi.org/10.1111/j.1467-7687.2007.00645.x

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*(3), 543–562. https://doi.org/10.1080/03640210802035357

McQueen, J. M., Norris, D., & Cutler, A. (2006). The Dynamic Nature of Speech Perception.

Language and Speech, *49*(1), 101–112.

Mirman, D., McClelland, J. L., & Holt, L. L. (2006). An interactive Hebbian account of lexically guided tuning of speech perception. *Psychonomic Bulletin & Review*, *13*(6), 958–965. https://doi.org/10.3758/BF03213909

Mitterer, H., & McQueen, J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS ONE*, *4*(11), e7785. https://doi.org/10.1371/journal.pone.0007785

Nazzi, T., & Cutler, A. (2019). How Consonants and Vowels Shape Spoken-Language Recognition. *Annual Review of Linguistics*, *5*, 25–47. https://doi.org/10.1146/annurev-linguistics-011718-011919

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238. https://doi.org/10.1016/S0010-0285(03)00006-9

Paquette-Smith, M., Cooper, A., & Johnson, E. K. (2020). Targeted adaptation in infants following live exposure to an accented talker. *Journal of Child Language*, 1–25. https://doi.org/10.1017/S030500092000029X

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

Reinisch, E., Wozny, D. R., Mitterer, H., & Holt, L. L. (2014). Phonetic category recalibration: What are the categories? *Journal of Phonetics*, *45*, 91–105.

https://doi.org/10.1016/j.wocn.2014.04.002

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception & Psychophysics*, *71*(6), 1207–1218. https://doi.org/10.3758/APP

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., 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

Sohoglu, E., & Davis, M. H. (2016). Perceptual learning of degraded speech by minimizing prediction error. *Proceedings of the National Academy of Sciences* , 1–10. https://doi.org/10.1073/pnas.1523266113

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

Thomason, M. E., Race, E., Burrows, B., Whitfield-Gabrieli, S., Glover, G. H., & Gabrieli, J. D. E. (2008). Development of Spatial and Verbal Working Memory Capacity in the Human Brain. *Journal of Cognitive Neuroscience*, *21*(2), 316–332.

Uther, M., Knoll, M. a., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, *49*(1), 2–7.

https://doi.org/10.1016/j.specom.2006.10.003

van der Feest, S., Blanco, C. P., & Smiljanic, R. (2019). Influence of speaking style adaptations and semantic context on the time course of word recognition in quiet and in noise. *Journal of Phonetics*, *73*, 158–177. https://doi.org/10.1016/j.wocn.2019.01.003

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., & Johnson, E. K. (2016). Toddlers' word recognition in an unfamiliar regional accent: The role of local sentence context and prior accent exposure. *Language and Speech*, *59*, 353–363. https://doi.org/10.1177/0023830915600471

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

van Wijngaarden, S. J. (2001). Intelligibility of native and non-native Dutch speech. *Speech Communication*, *35*(1–2), 103–113. https://doi.org/10.1016/S0167-6393(00)00098-4

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

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114–132. https://doi.org/10.1016/j.jml.2008.03.001

White, K. S., Yee, E., Blumstein, S. E., & Morgan, J. L. (2013). Adults show less sensitivity to phonetic detail in unfamiliar words, too. *Journal of Memory and Language*, *68*(4), 362–378. https://doi.org/10.1016/j.jml.2013.01.003

Witteman, M. J., Weber, A., & McQueen, J. M. (2013). Foreign accent strength and listener familiarity with an accent codetermine speed of perceptual adaptation. *Attention, Perception & Psychophysics*, *75*(3), 537–556. https://doi.org/10.3758/s13414-012-0404-y

Zheng, Y., & Samuel, A. G. (2020). The relationship between phonemic category boundary changes and perceptual adjustments to natural accents. *Journal of Experimental Psychology: Learning Memory and Cognition*, *46*(7), 1270–1292. https://doi.org/10.1037/xlm0000788

Appendix A. Stimuli list. Note participants were randomly assigned to either hear the 4 target items containing /ʌ/ or the four items containing /æ/. The non-target words were produced in either a Near, Far or Farther accent.

Exposure Trial Type	Word	Near	Far	Farther
Practice trial	dog	dog	dog	dog
Practice trial	train	train	train	train
Non-target word	fish	fɪʃ	feɪʃ (ɪ → eɪ)	fʌʃ (ɪ → ʌ)
Non-target word	milk	mɪlk	meɪlk (ɪ → eɪ)	mʌlk (ɪ → ʌ)
Non-target word	mouth	maʊθ	maʊf (θ → f)	maʊd (θ → d)
Non-target word	pillow	pʰɪloʊ	pʰeɪwoʊ (l → w); (ɪ → eI)	pʰʌwoʊ (l → w); (ɪ → ʌ)
Non-target word	lion	laɪən	waɪən (l → w)	waɪən (l → w)
Non-target word	horse	hoɹs	hoɹʃ (s → ʃ)	hoɹp (s → p)
Non-target word	foot	fʊt	foʊt (ʊ → oʊ)	foʊt (ʊ → oʊ)
Non-target word	nose	noʊz	noʊʒ (z → ʒ)	noʊb (z → b)
Non-target word	goose	gus	guʃ (s → ʃ)	gup (s → p)
Non-target word	pig	pʰɪg	pʰeɪg (ɪ → eɪ)	pʰʌg (ɪ → ʌ)
Non-target word	teeth	tʰiθ	tʰif (θ → f)	tʰid (θ → d)
Non-target word	chicken	tʃɪkn̩	tʃeɪkn̩ (ɪ → eɪ)	tʃʌkn̩ (ɪ → ʌ)
Target /ʌ/	truck	tʰɹoʊk	tʰɹoʊk	tʰɹoʊk
Target /ʌ/	duck	doʊk	doʊk	doʊk
Target /ʌ/	sun	soʊn	soʊn	soʊn
Target /ʌ/	tongue	tʰoʊŋ	tʰoʊŋ	tʰoʊŋ
Target /æ/	apple	apl̩	apl̩	apl̩
Target /æ/	candy	kʰandi	kʰandi	kʰandi
Target /æ/	pants	pʰans	pʰans	pʰans
Target /æ/	hand	hand	hand	hand

Appendix B. Participants heard 24 test trials that consisted of: 4 Canadian-accent real word trials, 8 Exposed target item trials, 8 Generalization trials and 4 Nonword trials. The order of the trials and the size of the screen that objects were positioned on was counterbalanced across participants. Note that the exposed target trials and generalization trials occurred twice.

Participants Exposed to the /ʌ/→ [oʊ] Vowel Shift		
Trial Type	**Word**	**Audio**
Canadian-accented real word	apple	æpl̩
Canadian-accented real word	candy	kʰændi
Canadian-accented real word	pants	pʰæns
Canadian-accented real word	hand	hænd
Exposed target item	truck	tʰɹoʊk
Exposed target item	duck	doʊk
Exposed target item	sun	soʊn
Exposed target item	tongue	tʰoʊŋ
Generalization item	boat	soʊt
Generalization item	block	slɑk
Generalization item	ball	sɑl
Generalization item	bird	sɚd
Non-word item	mawg	mawg
Non-word item	shoomp	ʃump
Non-word item	tark	tʰaɹk
Non-word item	chone	tʃoʊn

Participants Exposed to the /æ/ → [ɑ] Vowel Shift		
Trial Type	**Word**	**Audio**
Canadian-accented real word	truck	tʰɹʌk
Canadian-accented real word	duck	dʌk
Canadian-accented real word	sun	sʌn
Canadian-accented real word	tongue	tʰʌŋ
Exposed target item	apple	ɑpl̩
Exposed target item	candy	kʰɑndi
Exposed target item	pants	pʰɑns
Exposed target item	hand	hɑnd
Generalization item	boat	soʊt
Generalization item	block	slɑk
Generalization item	ball	sɑl
Generalization item	bird	sɚd
Non-word item	mawg	mawg
Non-word item	shoomp	ʃump
Non-word item	tark	tʰaɹk
Non-word item	chone	tʃoʊn