Running head: INPUT DRIVEN DIFFERENCES IN ACQUISITION

Input driven differences in toddler's perception of a disappearing phonological contrast

Suzanne van der Feest and Elizabeth K. Johnson

- - - - Draft Only - - - -

Abstract

How does phonological development differ in children exposed to one versus two variants of a single language? If children receive mixed evidence for a phonological contrast (i.e. one language variant in the environment maintains a contrast while another neutralizes it) will they treat this contrast as non-contrastive (i.e. as allophonic)? Or will they learn that only some speakers maintain the contrast and use this information to strategically optimize online word recognition? We examine these issues in two groups of Dutch 24-month-olds. One group had exposure to a single variant of Dutch that devoices all fricatives; the other had exposure to two Dutch variants, only one of which devoices all fricatives. We find that children who receive mixed evidence for a phonological contrast rapidly adapt their signal processing strategies to suit different speakers. Moreover, children routinely exposed to only one language variant demonstrate similar capabilities if given time to adapt to an unfamiliar variant.

1. Introduction

Comprehending spoken language requires attention to language-specific acoustic-phonetic detail, as well as flexibility in signal-to-word mapping strategies. Attention to detail allows listeners to perceive crucially important information for word recognition, such as the presence or absence of a pre-voicing cue that distinguishes the Dutch word *baard* ('beard') from the Dutch word *paard* ('horse'; Cho & Ladefoged, 1999; Iverson & Salmons, 1995; Van Alphen & McQueen, 2006; Van Alphen & Smits, 2004). Children are thought to learn which details are important to attend to in their native language by tracking the distribution of speech sounds in their linguistic environment (e.g., Maye, Werker, & Gerken, 2002). However, substantial variation can exist in the pronunciation of words by different members of a community (e.g., Labov, Ash, & Boberg, 2006). For example, a person from California typically pronounces the words 'cot' and 'caught' with the same vowel whereas a speaker from New York does not. And the words 'pander' and 'panda' sound far more similar in /r/-dropping variants of English than in American English. Thus, for children growing up in communities where they are exposed to multiple variants of their native language, the challenge of learning the relevant phonological contrasts of their native language may not be as straightforward as tracking the overall distribution of sounds in the input: young listeners need to be able to pay attention to the characteristics of specific speakers' (or groups of speakers') pronunciation of words. In the current study, we begin to explore these issues by examining online word recognition abilities in two-year-olds exposed to one versus two variants of their native language.

Adult language users skillfully attend to acoustic-phonetic detail (e.g., McMurray, Tanenhaus, Aslin, 2009; Salverda, Dahan, & McQueen, 2003; Shatzman & McQueen, 2006) while at the same time maintaining flexibility in their signal-to-word mapping strategies (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004; Kraljic & Samuel, 2006; Maye, Aslin & Tanenhaus, 2008; Norris, McQueen, & Cutler, 2003). Children, in contrast, may take some time to master this balancing act. Are toddlers and young children capable of coping with the fact that speakers of different variants of their native language may maintain different phonological systems? And if so, can they use this information strategically to adjust their expectations for how a particular interlocutor will pronounce particular words? Studies examining speech perception in various age groups and using different methodologies have reported mixed findings, with some studies focusing on how much children initially struggle with accented speakers (e.g., Best, Tyler, Orlando, Gooding & Quann, 2009; Floccia, Delle Luche, Durrant, Bulter, & Goslin, 2012; Nathan, Wells, & Donlan, 1998), and others focusing on how well children cope with accented speakers (e.g., Schmale, Cristia, & Seidl, 2012; Van Heugten & Johnson, 2014; Van Heugten, Krieger & Johnson, 2014; White & Aslin, 2011). Here, we contribute to this debate by asking whether toddlers from different linguistic backgrounds can efficiently draw on their personal real life experiences with accent variation to optimize their word recognition strategies. We also ask whether brief exposure in the lab is sufficient to enable children to adjust their signal-to-word mapping strategies to a speaker of an unfamiliar variant of Dutch.

Contemporary models of developmental speech perception suggest that young children possess very detailed representations of words (e.g., Ballem & Plunkett, 2005; Johnson, 2005; Swingley & Aslin, 2000; White & Morgan, 2008), although these representations are not necessarily completely adult-like (Van der Feest, 2007; Altvater-Mackensen, Van der Feest & Fikkert, 2014). Indeed, many studies have suggested that early on, young children are overly attentive to certain types of acoustic detail (see Werker & Curtin, 2005, for review). For example, segmentation studies using the Headturn Preference Procedure have shown that infants have difficulty recognizing the commonality between words spoken in different emotional affects, dialects, and genders. Infants who are familiarized to an isolated word in a happy register subsequently recognize that word when it is spoken again in a happy register, but not when it is spoken in a neutral register (Singh, Morgan, & White, 2004). Similarly, infants tested in the same paradigm recognize words when they are first spoken by one female and then spoken by another female, but not when the words are first spoken by a female and then spoken by a male (Houston and Jusczyk, 2000; see, however, Van Heugten and Johnson, 2012). Finally, infants under one year of age are also challenged by dialectal- and accent- related variation in the pronunciation of word forms (e.g., Schmale, Cristià, Seidl & Johnson, 2010).

The picture that emerges from the studies summarized above is that infants and toddlers may initially form over-specified representations of words, in the sense that they are storing information in their lexical representations that is not directly part of the native language's phonological system. In other words, infants appear to attend so closely to acoustic-phonetic detail in the speech signal that they lack flexibility in the mapping of this information to their developing lexicon. A possible explanation for this observation is that infants do not yet know exactly which dimensions of the speech signal carry linguistically meaningful information in their native language; so all information is attended to. That is, children may initially form exemplar-based representations of words, where all details of heard word tokens are stored (e.g., Jusczyk, 1997; Werker & Curtin, 2005). Early in development, when children have only had limited exposure to different realizations of a particular word, their representations are not robust enough to recognize words that vary greatly in their surface realization from earlier heard tokens. Support for this view comes from studies showing that children's ability to recognize new tokens of a familiarized word improves if children are initially familiarized to highly variable tokens of the word (e.g., Singh, 2008; see also Nazzi, Mersad, Sundara, Iakimova, & Polka, 2014; Singh, Nestor & Bortfeld, 2008; for related findings). Another way to characterize children's initial difficulty in coping with variability in spoken language is to focus on children's lack of abstract phonological representations (e.g., Best et al., 2009; Mulak, Best, Tyler, Kitamura & Irwin, 2013), which have been argued to be a prerequisite for efficiently coping with speaker-related variation (e.g., Cutler, Eisner, McQueen & Norris, 2010). Regardless of the theoretical framework chosen to explain the findings, it appears that infants take some time to reach an adult-like balance between attention to phonetic detail and flexibility in the mapping of speech sounds to speech categories.

This brings us to the question we address in the current study. How do children who hear more than one variant of their native language in their everyday environment cope when one of those variants neutralizes a particular phonological contrast while the other does not? As mentioned above, most contemporary models of language acquisition assume that children could learn the phoneme inventory of their native language by tracking the distributional statistics of segmental features in the input (e.g., Werker & Curtin, 2005). For instance, segmental features that appear to have a frequency distribution in a unimodal or overlapping fashion are most likely allophonic whereas those with a bimodal or non-overlapping frequency distribution are likely to be phonemically contrastive (e.g., Maye et al., 2002; White, Peperkamp, Kirk, & Morgan, 2008; Seidl & Cristía, 2012). The situation where children are exposed to two regional variants of the native language, with mismatching phoneme distributions, could be viewed as analogous to a situation of learning two languages. In that case children exposed to two regional variants may track two sets of statistics, one for each regional variant as bilingual children have been argued to do for different languages (e.g., Sundara & Scutellaro, 2011). Alternatively, when dealing with two variants of the same language, children may collapse the distributional statistics of sounds across the two variants. If children were to do this, then they might infer that contrasts that are distinctive in only one of two variants of the native language are in fact not distinctive in either variant. That is, they will classify the contrast as allophonic. At this point in time, little evidence exists to adjudicate between these two alternative possibilities.

To date, the growing numbers of studies that have examined children's perception of accents and dialects have typically used an approach where children are brought to the lab and exposed to a highly familiar and/or a relatively unfamiliar variant of their native language. This approach has taught us a great deal about infants' perception of native language variation. For example, we know that by 5 to 7 months of age children readily distinguish between familiar and unfamiliar varieties of their own native language, and that this ability is likely experience-dependent (Butler, Floccia, Goslin & Panneton, 2011; Nazzi, Jusczyk & Johnson, 2000). At the same time, as summarized above, we also know that infants initially struggle to recognize the commonality between the same word pronounced in two different accents (e.g., Best et al., 2009; Schmale et al., 2010; Van Heugten & Johnson, 2014; Van Heugten et al., 2014). There is also evidence that early word segmentation abilities and the ability to segment words in an unfamiliar accent may develop differently across infants acquiring different dialects of the same language (Nazzi et al., 2014). Importantly for the current study, studies exposing children to unfamiliar accents have shown that under some circumstances, children nearing their second birthday begin to show some ability to adapt to novel pronunciations of familiar words (e.g., Schmale et al., 2012; Van Heugten & Johnson, 2014; White & Aslin, 2011).

The first study demonstrating that children can adapt to a particular person's speaking style presented 19-month-olds with a training phase during which familiar objects were labeled with a standard vowel or shifted vowel (White & Aslin, 2011). Only those children who had been assigned to the atypical pronunciation training condition recognized familiar words when they were pronounced with the shifted vowel. That is, only those children who had heard words like ‘dog’ pronounced like ‘dag’ looked longer towards a picture of a block than a distractor item when they heard ‘black’. Follow-up studies allowed the authors to argue that the children were not simply broadening their categories for what was and was not an acceptable pronunciation. Rather, a potential explanation offered by the authors was that children might be using top-down lexical knowledge to learn the specific phonological vowel shift they were being presented with (e.g., if this speaker pronounces dog as ‘dag’ then they will probably pronounce block as ‘black’). In a more recent study, English-learning 24-month-olds have been reported to cope with Spanish-accented speech better after experiencing some exposure to a Spanish-accented speaker (Schmale et al., 2012).

To date, there is only one published study that has compared word recognition in toddlers who are naturally exposed to one versus two local variants of a single language (Floccia et al., 2012). Perhaps part of the reason for the scarcity of perceptual work on monolingual language-learning populations naturally exposed to one versus two varieties of the same language is due to the challenge of finding large numbers of toddlers to test who live in the same location but clearly fall into distinct uni- and a bi-accent exposure groups. Unless sizeable populations of uni- and bi-accent groups of toddlers are readily available for testing, it can be difficult to have enough statistical power to draw any firm conclusions regarding the role of accent exposure on early perceptual development. Floccia et al. succeeded in this task, testing two groups of 20-month-olds living in a community where a rhotic (r-preserving) variety of British English is spoken. Although in this particular community the rhotic accent was dominant, many parents who moved into this region from other areas spoke /r/-dropping variants of British English to their children at home. In the Floccia et al. study, one group of children had exposure to only the locally dominant rhotic variety, whereas the other group lived in the same community but had at least one parent who spoke an /r/-dropping variety of English. In the lab, children were presented with pictures of familiar objects and asked to look at one or the other (e.g., 'arm'). On different trials, children heard different voices. Two voices spoke in the rhotic variant of English and two spoke in the /r/-dropping variant. Interestingly, both groups of children recognized words with /r/'s in them produced in the local rhotic variety of English, but neither group recognized these words produced in the /r/-dropping variety. Thus, Floccia et al. concluded that children who hear two variants of their native language lack the flexibility to recognize words in both variants, and preferentially acquire the socially dominant variant, defined as the language spoken in their community. The notion that children with regular exposure to two varieties of their native language only recognize words produced in the socially dominant variant of a language in a particular region is fascinating, and is in line with recent studies on children's social preference for different accents and languages: Children show social preference for people speaking their same language variant and accent and privilege accent over other factors such as race (Kinzler, Shutts, DeJesus & Spelke, 2009), and they show a preference for speakers of a language that has higher status in society even over speakers of their own language (Kinzler, Shutts & Spelke, 2012). However, this finding does not fit well with studies suggesting that children can adapt to novel accents following as little as two minutes of exposure (e.g., Schmale et al., 2012; Van Heugten & Johnson, 2014; White & Aslin, 2011). If children can adapt to novel accents so quickly, why can children with routine exposure to two language variants not recognize words spoken in both variants?

In the current study, we re-visit the question of how children exposed to two variants of their native language acquire the phoneme inventory of their language and learn to efficiently recognize words produced in each language variant. Our primary question is whether two-year-olds (who are exposed to one or two variants of their native language) are capable of adopting distinct signal-to-word mapping strategies for different speakers in a word recognition task. Like Floccia et al. (and unlike White & Aslin, 2011, and Schmale et al., 2012), we test children naturally exposed to either one or two local varieties of a single language in their everyday lives.

Although our study is similar to Floccia et al. (2012) in its testing of two naturally occurring populations of language learners, our study is also different in many key respects. Crucially, like White and Aslin (2012), and unlike Floccia et al. (2012), we use a mispronunciation detection paradigm (Swingley & Aslin, 2000). In mispronunciation detection paradigms, children are presented with both canonical and mispronounced variants of familiar words while they view a picture of the named target and a distractor. Typically, the mispronunciations vary minimally from canonical (i.e., correct) pronunciations. In the mispronunciation detection paradigm, if children show that their word recognition is hindered when they hear mispronounced variants of familiar words (indicated by slower and shorter looking times to the target picture), then this is taken as an indication that their phonological representation for the target word includes the feature that distinguishes the properly pronounced word from the mispronounced word. For example, the label ‘pook’ might be given for the word ‘book’. (White & Morgan, 2008). Since the label ‘pook’ only differs from the canonical pronunciation of ‘book’ by a single voicing feature in the word initial segment, presenting children with this mispronunciation allows testing for their sensitivity to voicing mispronunciations.

In both experiments here we focus on monolingual Dutch-learning children's sensitivity to voicing mispronunciations on word initial fricatives. Importantly for the current study, fricative voicing contrasts are disappearing in most regions of the Netherlands remaining contrastive in only a few variants of Dutch (Van de Velde, Gerritsen & Van Hout, 1996; Van de Velde & Van Hout, 2001). The local variant of Dutch spoken in the Nijmegen region devoices all word initial fricatives. That is, it fails to maintain a voicing contrast on word initial fricatives, which are all realized as voiceless (Van de Velde et al., 1996). The Dutch children we test all lived in the Nijmegen region, but were classified as belonging to one of two distinct groups: Uniform Input or Mixed Input. The Uniform Input children had parents from the same Nijmegen region, meaning that they receive nearly uniform evidence that their native language maintains no voicing contrasts on word-initial fricatives. In contrast, the Mixed Input group had parents who had relocated to the Nijmegen area late in life and spoke a different variant of Dutch. Specifically, the Mixed Input children's parents were from the Southern Limburg region, where speakers systematically differentiate between voiced and voiceless fricatives. The Mixed Input children therefore received regular exposure to the disappearing Dutch fricative voicing contrast through their parents' speech. But since the Mixed Input children were being raised in Nijmegen, they also had substantial exposure to the speakers in their community who failed to maintain a voicing contrast on word initial fricatives. Thus, the Mixed Input children received mixed evidence as to whether fricative voicing is contrastive. Other differences between the Limburg accent and Standard Dutch lie in the realization of the velar /χ/ as well as differences in the realizations of vowels, pitch and intonation: A Standard Dutch speaker would thus be easily identified as coming from a different region than a speaker from Limburg by any native speaker of Dutch after hearing just a few words or sentences, in the same manner as an American English speaker can rapidly distinguish a Southern US accent from a New York accent (for more detailed information on the characteristics of the Southern Limburg accent, see Grondelaers, Van Hout & Steegs, 2010; Van Hout, Adank & Van Heuven, 2000).

In the first experiment reported in this paper, we investigated the online word recognition skills in Dutch-learning children exposed to one versus two variants of their native language (the Mixed Input and Uniform Input groups) on speech materials produced by a speaker who maintains a fricative voicing contrast (like the variant of Dutch spoken in Limburg) or a speaker who devoices all fricatives (like the variant of Dutch spoken in Nijmegen). In the second experiment, we further examine how quickly children routinely exposed to only one variant of their native language can adapt to an unfamiliar variant. The results of the two experiments reported demonstrate that even two-year-olds are well-equipped to respond adaptively to linguistic environments involving exposure to multiple variants of their native language.

2. Experiment 1

In Experiment 1, we ask whether children in the Mixed Input group differ from children in the Uniform Input group in their expectations for how speakers from different regions pronounce word-initial fricatives. We focus on children's perception of fricative voicing, because the voicing contrast in word initial fricatives is disappearing in most variants of Dutch spoken in the Netherlands (including the variant spoken in the region where both the Mixed and Uniform input children were being raised). One set of Uniform and Mixed Input children were tested on speech materials produced by a speaker from Limburg, who naturally maintains fricative voicing contrasts (henceforth the Fricative Contrast condition). Another set of Uniform and Mixed Input children were tested on speech materials produced by a Standard Dutch speaker who naturally devoices all fricatives (henceforth the Fricative Devoicing condition). All children in all conditions were tested on both correctly pronounced (e.g. 'sok' *, sock*) and mispronounced versions (e.g. 'zok') of familiar fricative initial words. In addition, all children were presented with correct (e.g. 'teen' *toe*) and mispronounced versions (e.g. 'deen') of familiar stop-initial words. The addition of stop-initial targets provided an interesting comparison for the fricative-initial targets since the two variants of Dutch used in this experiment both maintain a contrast on word-initial stops but only one variant maintains a voicing contrast on word-initial fricatives. Note that we included only voiced mispronunciations of voiceless segments, because previous studies have found possible perceptual asymmetries of voiced and voiceless Dutch stops in both toddlers (Van der Feest & Fikkert, accepted pending minor revisions) and adults (Van Alphen & Smits, 2004). To avoid the issue of perceptual asymmetries and maximize the potential mispronunciation effects in this study, we included only voiceless-initial target words.

Since fricative voicing is not contrastive in the local Nijmegen region of Dutch where the Uniform Input children and their parents have been born and raised, we predict that the Uniform Input children may simply ignore fricative voicing mispronunciations (i.e., they will recognize fricative-initial target words equally well regardless of whether the word-initial fricative is voiced or not). This prediction follows from what we know about listeners' tendency to map sounds that do not occur in the native language (e.g., voiced fricatives) onto speech sounds that do occur in the native language (e.g., voiceless fricatives; cf. Best, McRoberts & Goodell, 2001). Predicting how the Mixed Input children will perform with fricative voicing mispronunciations, however, is more complicated. We can see three likely outcomes, all of which provide us with novel insights into the phonological development of young children with exposure to more than one variant of their native language. If receiving mixed evidence for a voicing contrast leads the Mixed input toddlers to perceive the fricative voicing contrast as allophonic, then the children in this group should treat fricative voicing mispronunciations in the same manner as we expect the children in the Uniform Input Group to treat them (i.e., they should ignore them), regardless of the speaker they are presented with. Another possibility is that although the Mixed Input children receive mixed evidence for a fricative voice contrast in Dutch, they receive enough positive evidence that they will always notice fricative voicing mispronunciations (even if they are produced by a speaker who does not maintain a fricative voicing contrast). Finally, the third and in our mind most interesting possibility, is that the children in the Mixed Input Group may understand that some speakers in their daily environment maintain a fricative voicing contrast whereas others do not. If this is the case, children in the Mixed Input Group may ignore fricative voicing contrasts when they are produced by speakers from the Nijmegen region, but attend to fricative voicing contrasts (i.e. demonstrate a mispronunciation effect for incorrectly voiced fricatives) when they are produced by speakers from the Limburg region. This last outcome would demonstrate that children who receive mixed evidence for a phonological contrast due to exposure to more than one variant of a single language can actually flexibly adapt their expectations about how speakers will pronounce words depending on what variant of the native language that speaker is producing. Such an outcome would be the first evidence that children exposed to more than one variety of their native language in their everyday life can develop different phonological expectations for these different varieties, and use this information strategically to optimize their word recognition capabilities.

Although the main goal of this experiment is to examine how Mixed and Uniform Input children respond to fricative mispronunciations (as reflected by our labeling of the speaker conditions as Fricative Contrast versus Fricative Devoicing), we also have predictions concerning stop consonant voicing mispronunciations. Unlike the fricative voicing contrast that is maintained in the Limburg but not Nijmegen variant of Dutch, a stop voicing contrast is maintained in all variants of Dutch. Therefore, our predictions for children's sensitivity to stop consonant voicing mispronunciations are different from our predictions for fricative voicing mispronunciations. Based on past research demonstrating that Dutch-learning toddlers' develop sensitivity to the Dutch stop consonant voicing contrast in their native variant of Dutch some time between 20 and 24 months of age (Van der Feest, 2007; Van der Feest & Fikkert, under review), we predict that both Uniform and Mixed Input children tested in the current study will detect voicing mispronunciations of stop consonants when produced by the speaker of a familiar variant of Dutch (i.e., the speaker in the Fricative Contrast condition and the speaker in the Fricative Devoicing condition for the Mixed Input children, and only the speaker in the Fricative Devoicing condition for the Uniform Input children). However, stop consonant voicing in Dutch is realized in a relatively acoustically subtle fashion (unlike English, voiceless stops are not marked by aspiration; rather, the distinction is between prevoiced stops and unaspirated voiceless stops similar to e.g. the contrast in Spanish; Cho & Ladefoged, 1999). This is likely why Dutch-learning children fail to detect stop consonant mispronunciations under the age of 24 months of age (Van der Feest, 2007). Moreover, there has been very little work to date examining children's sensitivity to subtle mispronunciations in an unfamiliar variant of the native language. Thus, it is not as easy to predict how sensitive the Uniform Input children will be to stop consonant voicing mispronunciations in the unfamiliar variant of Dutch that they heard in the Fricative Contrast condition.

2.1. Method

2.1.1. Participants.

Sixty-four 24-month-old monolingual Dutch children participated (34 male; mean age 24 months, 13 days or 744 days; range: 730 days to 769 days). Half of the children (32) participated in the Fricative Contrast condition, and half (32) participated in the Fricative Devoicing condition (more details are in the Stimuli and Design sections below). Within each of these two conditions, half of the children (16) were in the Mixed Input group and half of the children (16) were in the Uniform Input group. All children were born and lived in the city of Nijmegen, a medium-sized city in the Southeastern central region of the Netherlands. Twenty-four additional participants were excluded from the study (15 in the Uniform Input group - 8 in the Fricative Contrast condition, 7 in the Fricative Devoicing condition – and 9 in the Mixed Input group - 4 in the Fricative Contrast condition, 5 in the Fricative Devoicing condition). The reasons for exclusion of these twenty-four children were: failure to complete at least 10 test trials (out of which the participant had to attend to at least one correctly pronounced and one mispronounced test trial of each of the 4 test words word), and / or general fussiness or crying leading to unwillingness to participate in the task (13), experimenter error (7), recording equipment failure (3), hearing issues not mentioned at the time the appointment was made (1), and parental interference (1).

Child participants for this study were chosen very carefully. To meet our criteria to be included in the Mixed Input group, children's parents could not have lived in a region of the Netherlands other than Nijmegen or Limburg for more than 3 years. The parents of the children in the Mixed Input group, who had all relocated to Nijmegen from Limburg, spoke a variant of standard Dutch different from the regionally dominant variant, providing their children with input containing fricative voicing contrasts in the home. At the same time, however, the children in the Mixed Input group also receive substantial evidence for the absence of the contrast, for instance through interactions with the local community and television exposure (e.g., characters on the Dutch version of Sesame Street). Additionally, children in the Mixed Input group could *not* have a secondary caregiver who was from a different region than the parents (including from Nijmegen). In this way, we insured that the children in the Mixed Input group received substantial mixed evidence for a fricative voicing contrast. To meet our criteria for inclusion in the Uniform Input group, children in that group could not have any secondary caregiver (grandparents, nanny, etcetera) who was from Limburg or southern Brabant (the region adjacent to Limburg where a large part of the population may also devoice fricatives, see Van de Velde & Van Hout, 2001). This was to insure that the children in the Uniform Input group were rarely if ever exposed to fricative voicing.

2.1.2. Stimuli.

The visual stimuli for the Fricative Contrast and Fricative Devoicing conditions were the same. They consisted of photographs of familiar objects, which were all edited to be similar in size and brightness. The objects measured about 23 cm on the screen, and were presented on a TV screen horizontally separated by about 20 cm. We chose target words that we expected 24-month-olds to know, and confirmed this by asking the parents after the experiment if their children knew the target words. All participants reportedly knew all target words.

The auditory stimuli for both the Fricative Contrast condition and the Fricative Devoicing condition were recorded from the same script, but were produced by different speakers. In the

Fricative Contrast condition, children were presented with auditory stimuli produced by a female native-Dutch speaker from the same region (Limburg) as the parents of children from the Mixed Input group. In the Fricative Devoicing condition, children heard the same auditory stimuli produced by a female native-Dutch speaker from the central-north of the Netherlands. In this area, the most standard or neutral variety of Dutch is spoken. Word-initial fricative voicing contrasts in this region are not maintained: all word-initial fricatives are produced with no voicing, as they are in Nijmegen (Van de Velde & Van Hout, 2001). See Figure 1 for a map of the Netherlands.

Figure 1. The Netherlands: the large light-blue dot indicates the Nijmegen region, the hatch-marked area the region of Limburg. The solid red dot indicates the central-north area where the Standard Dutch speaker who does not naturally maintain fricative voicing contrasts was from.

All auditory stimuli were recorded in a quiet room, with a sampling rate of 44.100 Hz. The sentences were all spoken in a moderately infant-directed voice, and were of the structure *"Kijk eens naar de [target]!"* (Look at the [target]!). On each trial, the sentence was followed by a 750ms pause and a second phrase *"Leuk hè?"* (Nice, right?) *"Lekker hè?"* (Good, right?), *"Vind je hem mooi?"* (Do you like it?) or *"Kun je hem vinden?"* (Can you find it?). The children heard correctly or incorrectly pronounced versions of these familiar words, on 16 critical test trials (see Appendix A for more details). On 4 trials one of two stop-initial targets ('tand' *tooth* and 'teen' *toe*) was correctly pronounced, on 4 trials these targets were pronounced with a voicing mispronunciation ('dant' and 'deen'). On another 4 trials one of two fricative-initial targets ('soep' *soup* and 'sok' *sock*) was correctly pronounced, on 4 trials these were mispronounced as 'zoep' and 'zok'. Importantly, a Dutch speaker who typically devoices word-initial fricatives (such as the speaker in the Fricative Devoicing condition) can readily produce voiced fricatives when prompted to do so. Before we recorded her, the Standard Dutch speaker listened to the stimuli produced by the speaker from Limburg and was instructed to mimic the affect and speaking rate as much as possible in her own (naturally produced) recordings of the stimuli. The stimuli recorded for the Fricative Devoicing condition were all checked by two phonetically trained linguists (one of them the first author) to ensure that target words were voiced on the voicing-mispronunciation trials (and we checked whether the voicing was comparable to the voicing of the speaker in the Fricative Contrast condition). We also confirmed with an initial group of 5 additional native Dutch listeners that the appropriate stimuli sounded voiced.

As in other past mispronunciation detection studies (e.g. Altvater-Mackensen et al., 2014; Quam & Swingley, 2010), test items were paired with a distractor item that started with the same initial consonant. Such pairings ensured that a mispronunciation of that initial segment did not make it in any way closer to the initial segment of the distractor item than to the initial segment of the test item. In addition to the 16 test trials, there were 4 filler trials on which a car (‘auto’), a baby (‘baby’), cat (‘poes’) and a cow (‘koe’) appeared. The filler target words were always canonically pronounced. These filler trials not only helped to maintain children’s attention throughout the experiment, they also kept the number of correctly pronounced words high relative to the number of mispronounced words so that children did not adapt to the mispronunciations.

2.1.3. Design.

Two between subject factors with two levels each were manipulated to create four distinct test conditions. The factors were test group (Mixed Input, Uniform Input) and speaker region (Fricative Contrast with speaker from Limburg, Fricative Devoicing with speaker from the north-east). Thus, the four conditions were (1) Mixed Input children tested on stimuli produced by a speaker from Limburg, (2) Uniform Input children tested on stimuli produced by that same speaker from Limburg, (3) Mixed Input children tested on stimuli produced by a speaker from the north-east, and (4) Uniform Input children tested on stimuli produced by that same speaker from the north-east. Crucially, the Mixed Input children had substantial real life daily exposure to speakers from both the ‘fricative voicing contrast’ Limburg region as well as from the ‘fricative devoicing’ Nijmegen region, whereas the Uniform Input children only had substantial exposure to speakers from the Nijmegen region. All children in all four conditions were presented with both fricative and stop correct pronunciation (CP) and mispronunciation (MP) trials. Note that the fricative voicing contrast is maintained in the Limburg (but not the

Nijmegen) variant of Dutch whereas the stop voicing contrast is maintained in both the Limburg and Nijmegen variants of Dutch.

2.1.4. Apparatus and procedure.

All participants were tested in the same location. During the 3-minute test phase of the experiment, participants sat on their parents' lap in a quiet darkened room in front of a 192 cm diagonal Sony LCD Projection Data Monitor. Throughout the experiment, the caregivers listened to music mixed with speech over Sennheiser Noisegard headphones so that they were unable to hear the audio of the experiment and were unable to provide (involuntary) cues to their child regarding the target locations. Caregivers were instructed not to speak to their child or point at the screen. The pre-recorded audio was played over external speakers, and a concealed video camera (Sony CVX-V18NSP) located 30cm below the screen recorded the participants' face close-up onto a DV cassette (with a Sony DV cassette recorder SR-40P). Participants heard *familiar* stop-initial and fricative-initial words labeled by the speaker from Limburg while watching two pictures side by side on a screen. The side on which the labeled target object appeared was counter-balanced across trials such that each target appeared 4 times on the left – the same two objects always appeared together, alternating which object was the named target on a particular trial. On each test trial the two objects appeared immediately, and audio started exactly 2.5 seconds later. Each test trial lasted exactly 6.5 seconds. After every 5 trials, short animations of a bouncing duck or a fish swimming across the screen were played to maintain children's interest in the movie. The duck and fish were not named.

2.1.5. Coding and analyses.

Children's eye movements were coded off-line by trained coders who analyzed the silenced video using the SuperCoder program (Hollich, 2005). Eye-movements were coded for each 40ms frame: the coders indicated whether the child was looking at the left picture, the right picture or was looking away from the screen. The beginning and end of each test-trial was clearly indicated on the video by a change in background light. The coder was blind to target side and test order. Coder reliability was determined by comparing codings by two different coders of ten percent of the data (from each of the reported experiments). The mean agreement between coders was 97%.

For each trial we calculated the ratio between the fixations to the target object and the sum of all fixations (to the target or distractor object) to the screen. Following previous studies, we calculated fixations over a two second time window starting 365ms after target word onset (e.g., Johnson, McQueen, & Huettig, 2011; Swingley, 2009). A difference score was computed for each trial. To calculate the difference scores, we calculated the ratio (of the sum of the total time subjects looked at the screen) of fixations to the target object during this two-second window *after* target word onset. We then compared this with the ratio of fixations to the target during a window of 2 seconds *before* target word onset (when the pictures were shown in silence), during the same trial. This way we could evaluate the effect of the auditory stimulus for each individual trial. All Figures in the Results sections depict these difference scores (see e.g., White & Aslin, 2011; Quam & Swingley, 2010 for a similar use of difference scores). By directly comparing the fixations in the two different windows, we maximize the power of our analyses by taking the absolute baseline visual preferences and possible changes in those preferences over the course of the experiment into account for each individual child and trial. Proportion of time spent looking at the familiar object during the window before target word onset did not differ significantly from change across the test groups (Mixed Input Contrast Condition group mean = .47; Mixed Input Devoicing Condition group mean = .47; Uniform Input Contrast Condition group mean = .48; Uniform Input Devoicing Condition group mean = .49). Trials on which the subject did not look at the screen for at least 400 ms in each of the two windows of analyses were excluded from the analyses (excluding a total of about 8% of trials across all subjects). All children included in the analyses (in Experiment 1 as well as in Experiment 2 reported below) ended up being attentive to at least 13 of the 16 test trials (not including fillers).

2.2. Results and Discussion

The difference scores calculated for each of the 64 participants in Experiment 1 were entered into a four-way mixed design ANOVA, with speaker region (Fricative Contrast, Fricative Devoicing) and test group (Mixed Input, Uniform Input) as between-subjects factors, and pronunciation (Correct, Mispronounced) and word type (Stop-initial, Fricative-initial) as within-subjects factors. The 2 x 2 x 2 x 2 ANOVA showed a significant main effect of pronunciation ($F(1,60 = 10.1$, $p = .001$), as well as a significant three-way interaction between pronunciation, test group and speaker region ($F(1,60 = 4.36$, $p = .04$). No other main effects or interactions were significant. Although we failed to find a significant main effect of word type, or a significant four-way interaction between pronunciation, test group, speaker region and word type we proceeded to investigate our a priori predictions by analyzing the Mixed and Uniform Input groups separately (Note that a significant 4-way interaction is exceedingly rare in perception studies of this type with young age groups, especially when power is limited due to limitations on the sample size with very carefully selected participant groups. Moreover, as the detailed analyses below reveal, the lack of a main effect of Word Type could be explained by the fact that fricative and stop MPs were treated differently in some but not all conditions.).

2.2.1. Mixed Input Group.

For the Mixed Input group, a 2 x 2 x 2 ANOVA with speaker region (Fricative Contrast, Fricative Devoicing) as between-subjects factor, and pronunciation (Correct, Mispronounced) and word type (Stops-initial, Fricative-initial) as within-subjects factors showed a significant main effect of pronunciation ($F(1,30) = 6.04$, $p = .01$), and a significant interaction between speaker region and pronunciation ($F(1,30) = 5.48$, $p = .02$). There were no other significant main effects or interactions. This indicates that the Mixed Input children reacted differently to the mispronunciations depending on the accent of the speaker they were presented with (in the Fricative Contrast condition versus the Fricative Devoicing condition). Figure 2 illustrates the difference scores (target fixations before versus after target word onset) for the Mixed Input groups.

Given our a priori predictions, the design of the study, and the significant interaction between speaker region and pronunciation, we next looked at the effects of pronunciation for the Mixed Input Group with each of the speaker conditions separately. First, we looked at the Mixed Input children in the Fricative Contrast condition, where children heard the speaker who (like the Mixed Input children's parents) naturally maintains a voicing contrast on both stops and fricatives. A 2 x 2 planned comparisons ANOVA with pronunciation (Correct, Mispronounced) and word type (Stop-initial, Fricative-initial) as within-subjects factors showed a significant main effect of pronunciation (across both stop-initial and fricative-initial word types) ($F(1,15) = 4.16$, $p = .02$). There was no significant main effect of word type, and no significant interaction. The lack of an interaction between word type and pronunciation shows that the Mixed Input children detected voicing mispronunciations of both stops and fricatives when presented with utterances produced by the speaker from Limburg. This is precisely the result one would expect if the Mixed Input children realized that speakers from the Limburg region maintain a voicing contrast on word-initial fricatives. In a parallel analysis of the Mixed Input children in the Fricative Devoicing condition, where they heard the speaker who naturally maintained a voicing contrast on stops but not fricatives, a 2 x 2 planned comparisons ANOVA with pronunciation (Correct, Mispronounced) and word type (Stop-initial, Fricative-initial) as within-subjects factors revealed no significant main effect of pronunciation ($F(1,15) = .01$ $p = .9$), and no other main effects or interactions. Note that the lack of a significant main effect of pronunciation with this speaker (as opposed to an overall effect for the speaker from the Limburg region) could be because the children ignored fricative voicing mispronunciations in a speaker that does not produce this contrast, and detection of stop voicing mispronunciations alone was not enough to drive a significant main effect (or significant interactions). We analyze this in further planned comparisons reported below.

Planned one-tailed *t*-tests were first conducted for the Fricative Contrast condition, comparing the difference scores of the Mixed Input children (depicted in Figure 2) to chance, or 0. Note that we report one-tailed tests here since in most studies, single-feature mispronunciations of the type we report do not result in significant looks to the distractor (we know of only one study showing increased looks to the distractor, White & Morgan (2008), where the distractor picture was an unfamiliar object and the mispronunciations involved multiple features). Word recognition is indicated by an increase in target fixations (significantly different from 0) after target word onset, compared to when the pictures were watched in silence.

For the Fricative Contrast condition, we found a significant increase in target fixations only on CP trials (stop-initial: $t(15) = 2.3$, $p = .02$, and fricative-initial: $t(15) = 4.5$, $p < .001$), but not on MP trials (stop-initial: $t(15) = .42$, $p = .34$, and fricative-initial: $t(15) = -.8$, $p = .22$). The results of the t-tests indicates that the Mixed Input children in the Fricative Contrast condition treated changes in word-initial voicing of both stop- and fricative-initial words as mispronunciations.

Next, planned one-tailed *t*-tests were conducted for the Fricative Devoicing condition, again comparing the difference scores of the Mixed Input children to chance. There were significant increases in target fixations on CP trials (stop-initial: $t(15) = 2.14$, $p = .02$, and fricative-initial: $t(15) = 3.2$, $p = .003$). On MP trials, the increase in target fixations was again not significant on stop-initial words but was significant on the fricative-initial words (stop-initial: $t(15) = 1.02$, $p = .16$, and fricative-initial: $t(15) = 2.91$, $p = .005$). This indicates that these Mixed Input children again treated changes in voicing on stop-initial words as mispronunciations, but did not treat changes in fricative voicing by the speaker in the Fricative Devoicing condition as mispronunciations (unlike the changes in fricative voicing produced by the speaker in the Fricative Contrast condition).

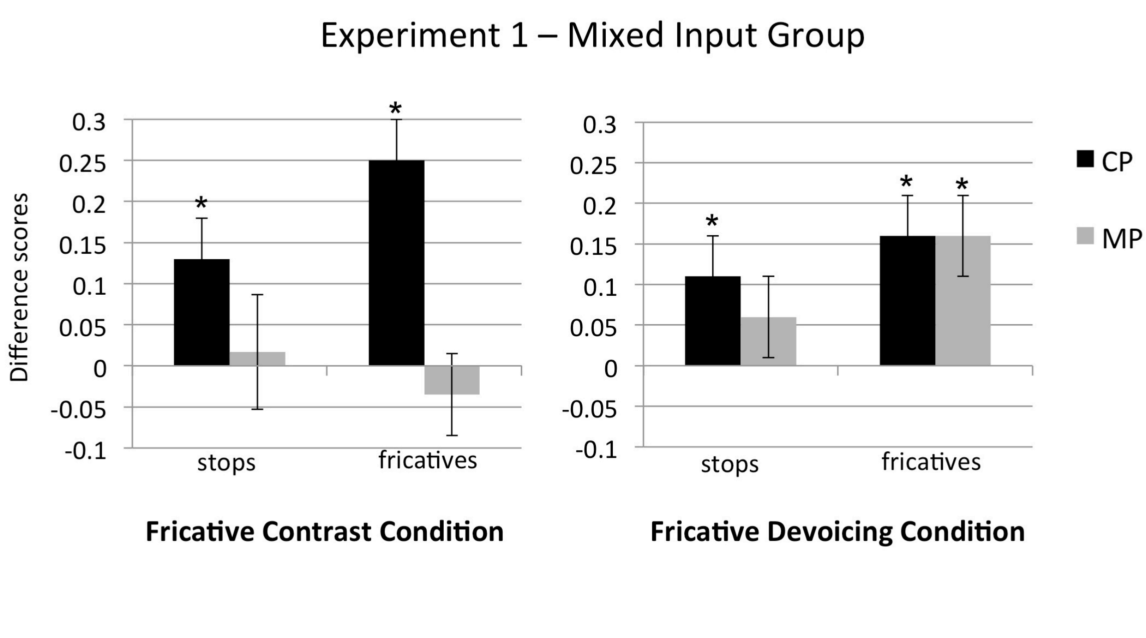

Figure 2. Experiment 1, Mixed Input children: Mean difference scores (target fixations, before versus after target word onset) and standard errors. Data is broken down by pronunciation (correct pronunciations (CP), voicing mispronunciations (MP), word type (Stop-initial, Fricative-initial), and speaker dialect (Fricative Contrast condition, Fricative Devoicing condition). Fricative Contrast condition = speaker from Limburg, Fricative Devoicing condition = speaker from the central-north. Stars indicate significant differences from chance.

2.2.2. Uniform Input Group.

We now turn our attention to the Uniform Input Group, and perform the same 2 x 2 x 2 ANOVA as reported above for the Mixed Input children, with speaker region (Fricative Contrast, Fricative Devoicing) as between-subjects factor and pronunciation (Correct, Mispronounced) and word type (Stop-initial, Fricative-initial) as within-subjects factors. For the Uniform Input children these analyses showed a main effect of pronunciation ($F(1,30) = 4.09$, $p = .04$) and a trend towards a main effect of word type ($F(1,30) = 3.7$, $p = .06$), but no interactions. Importantly, note that for these Uniform Input children, there was no effect of speaker region: this indicates that unlike the Mixed Input children, the Uniform Input children behave similarly, regardless of the accent of the speaker they were presented with. This is precisely the result one would expect if the Uniform Input children were ignoring fricative voicing MPs in all speech samples they were presented with, regardless of whether the speaker was from a region where a fricative voicing contrast is maintained or not. Figure 3 shows the difference scores (target fixations before versus after target word onset) for the Uniform Input groups.

Planned one-tailed *t*-tests were first conducted for the Fricative Contrast condition, comparing the difference scores of the Uniform Input children (depicted in Figure 3) to chance, or 0. For the Fricative Contrast condition, we found significant increases in target fixations on CP trials (stop-initial: $t(15) = 2.19$, $p = .02$, and fricative-initial: $t(15) = 3.56$, $p = .001$). Increase in target fixations was also significant on MP trials with fricative-initial words, but not on MP trials with stop-initial words (stop-initial: $t(15) = .79$, $p = .2$, and fricative-initial: $t(15) = 2.19$, $p = .02$). Again, this can be seen as an indication that for the Uniform Input children, word recognition was inhibited when stop- but not fricative-initial words were pronounced with a voicing mispronunciation.

Next, planned one-tailed *t*-tests were conducted for the Fricative Devoicing condition, again comparing the difference scores of the Uniform Input children to chance. There were significant increases in target fixations on CP trials (stop-initial: $t(15) = 2.4$, $p = .01$, and fricative-initial: $t(15) = 2.95$, $p = .004$). On MP trials, the increase in target fixations was again not significant on stop-initial words but was significant on the fricative-initial words (stop-initial words: $t(15) = -0.92$, $p = .18$, and fricative-initial words: $t(15) = 2.06$, $p = .03$).

These overall patterns confirm our predictions. In addition, the results indicate that the Uniform Input children did recognize the target words even in the Fricative Contrast condition: this is illustrated by the significant increases in target fixations on the CP trials, which shows that they recognized the test words even when they were listening to the speaker from Limburg, who had an accent the Uniform Input children are not regularly exposed to. Keep in mind that in the Fricative Contrast condition the Uniform Input children were presented with a variant of Dutch that differs from the variant of Dutch that is dominant in the region where they live, and that they have had very little if any exposure to (unlike the Mixed Input children whose parents speak this variant). This finding fits well with studies demonstrating comprehension of novel accents by two-year-olds (Best et al., 2009; Schmale et al., 2012; Van Heugten & Johnson, 2014; White & Aslin, 2011), but could be seen as contrasting with the findings reported by Floccia et al. (2012). But, as we saw above, there were no clear general significant effects of pronunciation or interactions between pronunciation and word type in the ANOVA, which may be interpreted to illustrate a more limited sensitivity to voicing mispronunciations by the Uniform Input children when they are presented with an unfamiliar accent. We will return to a discussion of this issue in the General Discussion.

In the Fricative Devoicing condition, the children receiving uniform input (as we also saw for the children receiving mixed input) treated changes in voicing on stop-initial words but not fricative-initial words as mispronunciations when they were presented with a standard Dutch speaker who naturally devoices all fricatives. Figure 3 illustrates that the children in the Uniform Input group in the Fricative Devoicing condition showed robust sensitivity to voicing mispronunciations on stop-initial words.

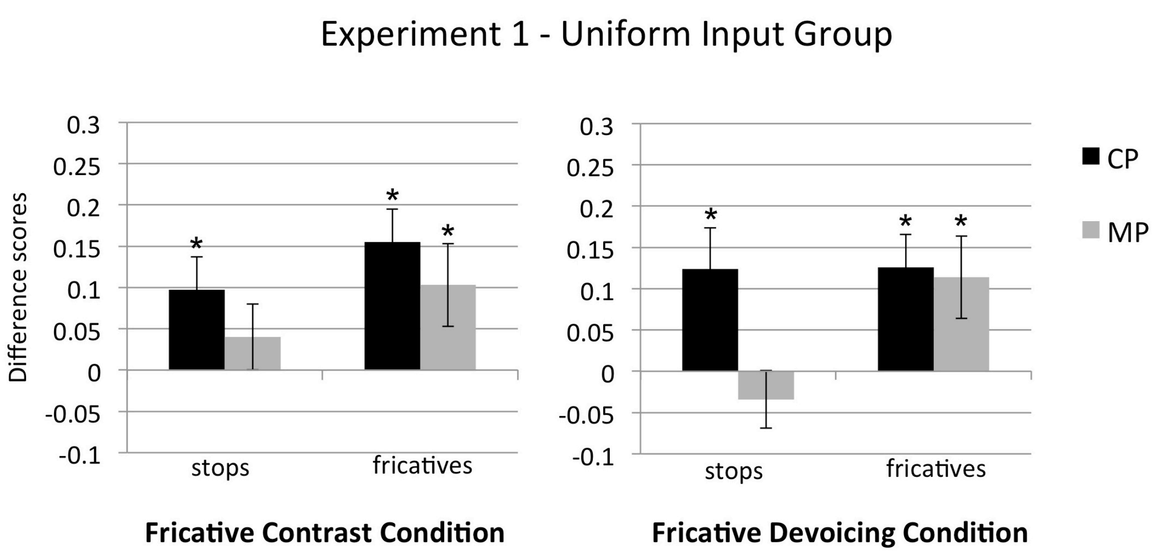

Figure 3. Experiment 1, Uniform Input children: Mean difference scores (target fixations, before

versus after target word onset) and standard errors. Data is broken down by pronunciation (correct pronunciations (CP), voicing mispronunciations (MP), speaker dialect (Fricative Contrast condition, Fricative Devoicing condition), word type (Stop-initial, Fricative-initial), and speaker dialect (Fricative Contrast condition, Fricative Devoicing condition). Fricative Contrast condition = speaker from Limburg, Fricative Devoicing condition = speaker from the central-north. Stars indicate significant differences from chance.

When we consider the results from Experiment 1 for each of the groups separately, the data indicates that children who are exposed to one versus two variants of their native language (the Mixed Input versus the Uniform Input children) differ in their expectations for how Dutch speakers from different regions will pronounce word-initial fricatives, and that the Mixed Input children can adjust their speech processing strategies depending on the speaker they are presented with. In the Fricative Contrast condition, when we look only at the children receiving mixed input we see that they treated changes in voicing of both stops and fricatives as mispronunciations. At the same time, these Mixed Input children ignored fricative voicing mispronunciations produced by our standard Dutch speaker who normally devoices all her fricatives. Thus, the mixed input children adapted their speech processing strategies to the speaker at hand, only noticing fricative voicing MPs when they were produced by the speaker who maintains fricative voicing contrasts. In contrast, when we only consider the data from the children in the Uniform Input group, we see that they ignored all fricative MPs regardless of who produced them.

Overall, our findings suggest that the type of natural daily input that the children in our study are exposed to changed the way they responded to changes in fricative voicing. The children who receive mixed input showed flexibility and speaker adaptation, the children who receive uniform input did not. But why did the children who receive uniform input ignore fricative voicing contrasts, regardless of the speaker's accent? One could take this result to suggest that these children have learned to always ignore fricative voicing because they have mainly been exposed to speakers who produce only voiceless fricatives. However, the fact that the Uniform Input children ignored fricative voicing mispronunciations from the speaker from Limburg may have been a result of their unfamiliarity with that speaker's accent. Their lack of everyday experience with the Limburg variant may have hindered the Uniform Input children to adapt to the speaker's use of fricative voicing when they had no pre-test exposure to that speaker / accent. Results from recent studies indicate that toddlers and young children can adapt to a new accent, if given some exposure to an accent. White & Aslin (2011) showed that toddlers were able to adapt to an accent with an (artificial) vowel shift after a brief pre-test exposure phase, and Schmale et al. (2012) showed that 24-month-olds were able to generalize across different accents as long as they were presented with the more unfamiliar (Spanish-accented) language variant first.

We predicted that our 24-month-olds in the Uniform Input group might be able to adapt to a speaker's fricative voicing, if they had some exposure to the variant of Dutch where the contrast is maintained prior to the test phase of the experiment. In Experiment 2, we set out to test this prediction by adding a pre-test exposure phase, to give the children more information about the speaker's accent and thus even more chance to adapt. If we find that the children in the Uniform Input group are able to adapt to a speaker's fricative voicing after a brief pre-test speaker exposure, this would provide evidence that young toddlers already have the ability to adapt to different speakers and use speaker-specific information to optimize word recognition.

3. Experiment 2

Experiment 2 was essentially identical to the Fricative Contrast condition of Experiment 1, apart from the addition of a pre-test phase where participants were exposed to a 2-minute story read by the same speaker from Limburg who recorded the stimuli for the Fricative Contrast condition. The goal of this second experiment was to examine whether children who are less familiar with the Limburg variety of Dutch *can* be prompted to pay attention to fricative voicing if they get more evidence that a speaker uses such a contrast. Since the Mixed Input children are already familiar with Limburg-accented Dutch, we predict that children from the Mixed Input group will behave exactly as they did in the Fricative Contrast condition of Experiment 1. For the Uniform Input children, we predict that if a few minutes of exposure to the unfamiliar Limburg variant of Dutch is sufficient to drive adaptation, the Uniform group in Experiment 2 should behave differently from the Uniform group in the Fricative Contrast condition of Experiment 1. That is, in contrast to Experiment 1, they will now notice the fricative voicing mispronunciations.

3.1. Method

The apparatus and procedure were identical to Experiment 1, except for the addition of the 2-minute pre-test story phase preceding the test phase, during which the participants watched a short animated movie on the same screen where the test was played. The test phase immediately followed the story phase, without any break or further instructions: the whole experiment was presented as a single movie.

3.1.1. Participants.

Thirty-two additional 24-month-old monolingual Dutch children participated in the study (15 male; mean age 24 months, 11 days or 743 days; range: 730 days to 757 days). As in Experiment 1, all children were from Nijmegen: half (16, 7 male) of the participants had parents who were both from the same Nijmegen region (the Uniform group) and the other half (16, 8 male) had parents who were both from Limburg (the Mixed Input group). Seven additional children were excluded from the study (4 in the Uniform Input group, 3 in the Mixed Input group) due to fussiness / unwillingness to participate (6), and experimenter error (1).The same inclusion criteria for all participants were used as in Experiment 1.

3.1.2. Stimuli.

The pre-test story phase directly preceded and continued into the test-phase without any breaks. For the pre-test story, the Limburg speaker from the Fricative Contrast condition in Experiment 1 was recorded with the same equipment, the same style of infant-directed speech, and the same speaking rate as in Experiment 1. The story (see Appendix B), about a boy and his sister who had a picnic on the beach, consisted of 19 short sentences, containing a total number of 13 word-initial voiceless fricative /s/ and 25 word-initial voiced fricative /z/. There were 10 word-initial voiceless stops /t/ and 34 word-initial voiced stops /d/ (including 11 instances of the prosodically weak definite determiner ‘de’). It is important to note that although the story was designed to contain a high number of (voiced) fricative-initial words, the story and the rate of fricatives in it were designed to sound natural – the purpose of the story was to familiarize the child with the speaker in a natural setting (telling a story). While listening to the story during the pre-test phase, the children watched a slow animated picture of the brother and sister on the beach, which filled the entire screen; the main objects and characters in the 'movie' were similar in size and design to the visual stimuli in the test phase, but did not contain any of the test items. The auditory and visual stimuli used during the test phase of Experiment 2 were the exact same stimuli that were used in the Fricative Contrast condition of Experiment 1.

3.2. Results and Discussion

The difference scores calculated for each of the 32 participants in Experiment 2 were entered into a three-way ANOVA, with test group (Mixed Input, Uniform Input) as between-subjects factor, and pronunciation (Correct, Mispronounced) and word type (Stop-initial, Fricative-initial) as within-subjects factors showed a significant main effect of pronunciation ($F(1,30) = 18.1$, $p < .0001$). The 2 x 2 x 2 ANOVA showed no main effect of group or word type and no significant interactions. The lack of a main effect and interactions involving test group indicates that in marked contrast to Experiment 1, both groups of children tested in Experiment 2 responded similarly to both the stop and fricative mispronunciations. Figure 4 illustrates the difference scores (target fixations before versus after target word onset) on the different trial types. Proportion of time spent looking at the familiar object during the window before target word onset again did not differ significantly from chance across the test groups (Mixed Input group mean = .47; Uniform Input group mean = .48) and the mean baseline looking times also did not differ from those in Experiment 1 (Experiment 1 mean = .48, Experiment 2 mean = .47).

Planned one-tailed *t*-tests were first conducted for the Mixed Input group, comparing the difference scores to chance, or 0. There were significant increases in target fixations on CP trials, i.e. the subjects recognized the target words (stop-initial: $t(15) = 2.3$, $p = .02$, and fricative-initial: $t(15) = 2.5$, $p = .01$). Increase in target fixations were not significant on MP trials (stop-initial:

t(15) = .32, *p* = .37, and fricative-initial: *t*(15) = -.97, *p* = .17). Thus, as we saw in the Fricative Contrast condition (but not the Fricative Devoicing condition) of Experiment 1, the results indicate that the children in the Mixed Input group treated changes in word-initial fricative voicing as mispronunciations, which is in line with our predictions.

Next, planned *t*-tests were conducted for the Uniform Input group. There were significant increases in target fixations on CP trials (stop-initial: *t*(15) = 2.0, *p* = .03, and fricative-initial: *t*(15) = 4.7, *p* <.001). There was no significant increase in target fixations on MP trials (stop-initial: *t*(15) = .47, *p* = .32, and fricative-initial: *t*(15) = -1.0, *p* = .16). Crucially, the Uniform Input children in Experiment 2 also treated changes in fricative voicing as mispronunciations: this is in contrast with the Uniform Input children in the Fricative Contrast condition of Experiment 1, who did not treat changes in fricative voicing as mispronunciations.

To confirm that the Uniform Input children in Experiment 2 indeed reacted differently than the Uniform Input children in Experiment 1 (in the Fricative Contrast condition) to fricative voicing mispronunciations by the speaker who naturally maintains fricative voicing contrasts, we ran analyses directly comparing the results from Experiment 2 with the results from the Fricative Contrast condition of Experiment 1. Recall that participants in both experiments were presented with the exact same experimental stimuli in the test phase of the experiments: the only difference was the presence or absence of the pre-test exposure phase, which was recorded by the same talker from Limburg that participants heard in the test phase. We ran a four-way planned comparison ANOVA with experiment (Experiment 1, Experiment 2) and test group (Mixed Input, Uniform Input) as between-subjects factors and pronunciation (Correct, Mispronounced) and word type (Stops-initial, Fricative-initial) as within-subjects factors. There was a significant main effect of pronunciation (*F*(1,60 = 28.08, *p* < .0001), a significant interaction between pronunciation and word type (*F*(1,60) = 4.35, *p* = .04), and a significant three-way interaction between pronunciation, test group and experiment (*F*(1,60 = 4.17, *p* = .04). The four-way interaction was not significant, possibly due to a lack of power.

Further direct comparisons between Experiment 1 and 2 were performed to look at the significant effects of pronunciation and word type for the two test groups. For the Mixed Input group, a 2 x 2 x 2 ANOVA with experiment (Experiment 1, Experiment 2) as between-subjects factor and pronunciation (Correct, Mispronounced) and word type (Stop-initial, Fricative-initial) as within-subjects factors, showed a significant main effect of pronunciation (*F*(1,30) = 18.25, *p* < .0001), and no other significant main effects or interactions. For the Uniform Input group, the same three-way ANOVA with experiment (Experiment 1, Experiment 2) as between-subjects factor and pronunciation (Correct, Mispronounced) and word type (Stops-initial, Fricative-initial) as within-subjects factors showed a significant main effect of pronunciation (*F*(1,30) = 10.75, *p* = .001), and an interaction between experiment and pronunciation (*F*(1,30) = 3.94, *p* = .047). These analyses confirm our observations that only the Uniform Input children reacted differently to fricative voicing mispronunciations by the speaker from Limburg in Experiment 1 versus Experiment 2, where the presence or absence of the short exposure phase was the only difference between the two experiments. Just two minutes of pre-test exposure to the speaker who naturally produces fricative voicing put the two groups on equal footing in Experiment 2: the Uniform Input children who are less familiar with this speaker's accent treated changes in fricative voicing as mispronunciations and show that they are capable of rapidly adapting their signal-to-word mapping strategies depending on the accent of the speaker that they hear.

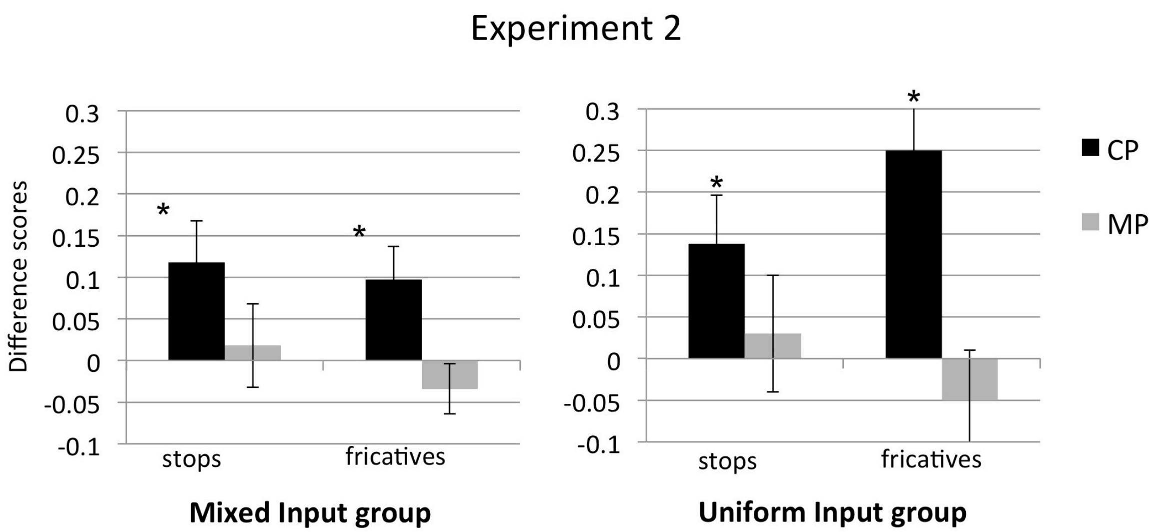

Figure 4. Experiment 2 (Fricative Contrast condition – speaker from Limburg - with 2-minute story training phase preceding the test phase): Mean difference scores (target fixations, before versus after target word onset) and standard errors. Data is broken down by pronunciation (correct pronunciations (CP), voicing mispronunciations (MP), participant group (Mixed Input, Uniform Input), and word type (Stop-initial, Fricative-initial). Stars indicate significant difference from chance.

4. General Discussion

There is a sizeable body of literature suggesting that children initially struggle with unfamiliar accents confronted for the first time in the lab (e.g., Best et al., 2009; Nathan et al., 1998; Schmale et al., 2010; 2012). Moreover, there is evidence that toddlers who are exposed to more than one variant of their native language in their daily environment only recognize words produced in the dominant language variant (Floccia et al., 2012). Taken together, these studies suggest a relatively large gap between adults' and toddlers' capabilities of handling speakers with different accents. Other work, however, has suggested that toddlers possess highly adaptive speech processing capabilities (e.g., Schmale et al., 2012; White & Aslin, 2011; Van Heugten & Johnson, 2014). The current study sheds new light on the question of how well young children can cope with variation in their language input. We find that two-year-olds routinely exposed to two variants of their native language demonstrate no difficulty in adapting their signal-to-word mapping strategies to suit the speaker at hand. Moreover, two-year-olds exposed to only the dominant variant of the native language at home and in the environment are able to adapt to the non-dominant variant of their native language, after only two minutes of exposure to that variant.

In Experiment 1, we saw that regardless of which variant(s) of Dutch they were exposed to in their everyday environment, all children recognized familiar words produced by a speaker of a less dominant variant of Dutch (the speaker from Limburg in the Fricative Contrast condition). However, when we consider the data of the different groups of children separately, we see that only the children who routinely hear both the dominant and the non-dominant variant of Dutch detected fricative voicing mispronunciations by this speaker, who naturally produces fricative voicing in her variant of Dutch. Results from Experiment 1 also showed that all children recognized familiar words produced by a speaker of the regionally dominant variant of Dutch (i.e., a variant that devoices all fricatives, in the Fricative Devoicing condition). When we considered only the children who are exposed to different variants of Dutch in their daily input, we saw that they rapidly adapted their lexical processing strategies to suit the variant of Dutch they were presented with, and they ignored fricative voicing ‘errors’ in a speaker who does not naturally maintain fricative voicing contrasts (note that we use the term ‘adapt’ loosely here, as our study was not designed to distinguish between the possibility that children either adapted on-line in response to the accent or triggered some stored knowledge about how speakers of different variants of Dutch produce fricatives). The overall results from Experiment 1 suggested that children who are exposed to multiple variants of their native language can use their knowledge of how sound contrasts are used in different variants of Dutch to adapt to different speaker's accents and efficiently recognize words. In Experiment 2, we found further evidence that even those children routinely exposed to only the standard variant of Dutch are capable of adapting and paying attention to fricative voicing if provided with a short exposure to the non-dominant variant of Dutch that maintains fricative voicing contrasts. In short, our work indicates that toddlers are able to adapt to speakers of different varieties of their language, and can be flexible in their treatment of a phonological contrast in a word recognition task, regardless of the amount of mixed input they hear in their environment.

Although our general findings fit well with laboratory studies showing children can adapt to a novel accent in the lab (Best et al., 2009; Schmale et al., 2012; Van Heugten & Johnson, 2014; White & Aslin, 2011), our results showed that all toddlers recognized words in both familiar and unfamiliar accents on the correct pronunciation trials. This last finding could be seen as conflicting with recent reports that bidialectal toddlers only recognize rhotic-containing words spoken in the socially dominant variant of their native language (Floccia et al., 2012). How can we reconcile our finding that even the children receiving uniform input recognized correctly pronounced words when listening to an unfamiliar accent from Limburg in Experiment 1 with the results of Floccia et al., the only other published study examining word comprehension in toddlers naturally exposed to one versus two variants of their native language? There are several potentially important differences between our study and the Floccia et al. (2012) study. First, one potential explanation for why our results differ from Floccia et al.'s could be related to the age of the children we tested. We tested 24-month-olds whereas Floccia et al. tested 20-month-olds. Support for an age-related explanation comes from work showing a dramatic improvement of toddlers' ability to cope with accent variation near their second birthday (Best et al., 2009; Mulak et al., 2013; Van Heugten & Johnson, 2014; Van Heugten, Krieger & Johnson, 2014). Another major difference between our study and the Floccia et al. study is the type of contrast we studied. Floccia et al. tested children on specific segments that were either present or absent in the language variants under investigation; a /r/-deleting versus /r/-preserving variant of English, and the children crucially did not recognize /r/-deleting words. In contrast, in this study we investigated children's sensitivity to a phonological featural contrast (voicing) present in both variants of Dutch under study, albeit only on the stop consonants for one variant. Acoustic input distributions that provided evidence for phonemic use of voicing in Dutch at least on a subset of consonants – namely, at least on stop consonants - may have enabled the children in the Uniform Input group to quickly adapt to a less familiar language variant where voicing is phonemic on a larger set of phonemes – namely, on both stops and fricatives. Other evidence in support of the notion that contrast type impacts listeners' performance on a perceptual learning task is provided by adult voice adaptation experiments, which have shown that perceptual adjustment in adults vary depending on the type of contrast they are tested on. For example, it has been argued that perceptual learning for speech sounds is speaker-specific for fricatives (Eisner & McQueen, 2005; Kraljic & Samuel, 2005; 2006; McQueen, Tyler & Cutler, 2012), but not for stops (Kraljic & Samuel, 2006). Finally, another potentially crucial difference between our study and Floccia et al. 2012, is that in our design each child was presented with only one voice. Our rationale for making accent familiarity a between- rather than within-subjects manipulation was that we wanted to make it easier for children to demonstrate their accent adaptation capabilities (e.g., see Fennell & Waxman, 2010, for evidence that higher task demands can influence young children's ability to demonstrate their speech processing capabilities). The study by Floccia et al. was more focused on how differential exposure to accents influences word recognition in children with different language inputs at home, rather than on the general accent adaptation abilities of children. White and Aslin (2011) also used one voice, and also found evidence for rapid adaptation to novel pronunciations of familiar words. An interesting question for future research might be to experimentally test whether children adapt faster to novel accents if they are exposed to only one versus many voices.

We saw in Experiment 1 of our study that children routinely exposed to two variants of their native language were able to rapidly adapt to different speakers' accents when presented with a single speaker in each experimental condition. A possible interpretation of this finding is that children exposed to two regional variants are able to 'tag' the different language variants in their environment, in a similar way bilingual children have been suggested to do (Sundara & Scutellaro, 2011), and track two sets of phoneme distributions. These children may then have a schema stored for how the different variants of Dutch in their environment use fricative voicing contrasts, and have in general formed more flexible representations due to their daily exposure to multiple accents. They are thus able to adapt very quickly as soon as they identify where a speaker is from. Such an interpretation could suggest that exposure to multiple accents might have cognitive benefits akin to the benefits associated with learning more than one language early in life (Bialystok, 2005; Kovacs & Mehler, 2009). This is an interesting issue for future studies to explore.

The Uniform Input children tested in Experiment 1, who were routinely exposed to only one variant of Dutch in their daily life, failed to demonstrate any flexibility in their processing of fricative voicing MPs. That is, without prior exposure to the unfamiliar variant of Dutch, they failed to adapt to the fricative voicing contrast produced by the speaker of a Limburg variant of Dutch. Nonetheless, after a mere two-minute exposure period in Experiment 2, these children were able to adapt to this unfamiliar fricative voicing variant of Dutch under study. White and Aslin (2011) argued that children might use top-down lexical knowledge when they are presented with specific phonological shifts, and this is what enables them to adapt (see also Van Heugten & Johnson, 2014; Cristia, Seidl, Vaugn, Schmale, Bradlow & Floccia, 2012). Lexical information has been shown to be helpful for adults and older children as well. (e.g., Eisner & McQueen, 2005; McQueen et al., 2012). Under such a view, the top-down lexical evidence provided in the animated story phase in Experiment 2 may have been a crucial factor enabling the children who receive uniform input to adapt.

We have argued that the Uniform Input children tested in the current study only adapted to the unfamiliar Limburg variant of Dutch when first exposed to a two-minute story read by a speaker from Limburg. But past studies have suggested that adults adapt to unfamiliar accents much faster, sometimes after hearing just a few words (e.g., Clarke & Garrett, 2004). Do the results of the current study demonstrate that children are far slower than adults when it comes to accent adaptation? This may be the case, but it is also possible that we would have observed adaptation over the course of Experiment 1 had we had enough statistical power (i.e., more trials and more participants) to analyze our results by trial block. That is, we may have observed better performance by the Uniform Input kids on the unfamiliar Limburg variant of Dutch towards the end of the experiment when compared to the beginning of the experiment. Future studies could be designed to have enough statistical power to directly examine this question (power in our study, for example could have been increased by testing children only on fricative voicing mispronunciations – without including stop-initial trials – or by increasing the number of participants who could be tested by setting up experiments in multiple labs).

The two populations tested in the current study existed for us to test thanks to a rapid historical change that the Dutch language is currently undergoing, with fricative voicing being described as a disappearing phonological contrast (Van de Velde et al., 1996; Van de Velde & Van Hout, 2001). An interesting thought to ponder is whether the current study is documenting language change in progress. Will a similar study be possible in 50 years time, or will the disappearing phonological contrast under investigation then no longer exist? Although it is not possible to conclude from the current study whether or how quickly the fricative voicing contrast will completely disappear in variants of standard Dutch spoken in the Netherlands, the results of the current study allow us to speculate on the topic because input received by the children in the Uniform Input group in our experiments represents the kind of non-consistent fricative voicing input that a large number of children in the Netherlands experience. Future studies could investigate the perceptual adaptation abilities of children who do not live in 'border' regions such as Nijmegen, but in regions where there may already be even less evidence that some Dutch speakers maintain fricative voicing. Finally, although we found that all children in the current study were able to *perceive* both the voiceless and voiced variety, it is likely that both groups of children in our study will ultimately learn to *produce* the dominant voiceless-fricative variety of Dutch, assuming they both continue growing up in the voiceless-dominant Nijmegen region (see Kinzler et al., 2009, for evidence that children show a social preference for speakers of their same language variant). This could ultimately lead to the complete disappearance of this phonological contrast in a future generation of listeners, who may no longer be exposed to any evidence for fricative voicing contrasts in their input.

5. Conclusions

To become successful listeners of their native language, young children need to balance the need to attend to acoustic-phonetic detail with the need to remain flexible in their signal-to-word mapping strategies. We investigated whether children who receive mixed evidence for a phonological contrast by being routinely exposed to multiple variants of their language learn that only some speakers maintain this contrast, and if they can use this information to optimize online word recognition. Our study shows that toddlers who receive mixed distributional evidence for a phonological contrast due to variation in accents in their input do not simply treat such the contrast in question as allophonic. Nor do they ignore the contrast all together. Instead, they are capable of adjusting their signal-to-word mapping strategies depending on the speaker at hand (Experiment 1). Furthermore, we have shown that not only those children who routinely hear multiple accents, but even young toddlers who grow up in a uniform language environment are able to adapt to a speaker's accent after brief exposure to that speaker, which is a crucial skill for successful communication and word learning for so many young speakers around the world (Experiment 2). We conclude that even two-year-olds are well-equipped to handle linguistic environments involving multiple variants of their native language.

References

Altvater-Mackensen, N., Van der Feest, S.V.H. & Fikkert, P. 2014. Asymmetries in early word recognition: The case of stops and fricatives. *Language Learning and Development, 10*(2), 149-178.

Ballem, K.D., & Plunkett, K. 2005. Phonological specificity in children at 1;2. *Journal of Child Language, 32,* 159-173.

Best, C.T., McRoberts, G.W. & Goodell, E. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109,* 775-794.

Best, C., 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.

Bialystok, 2005. Consequences of bilingualism for cognitive development. In J. R. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 417–432). Oxford, UK: Oxford University Press.

Bradlow, A.R., & Bent, T. 2008. Perceptual adaptation to non-native speech. *Cognition 106,* 707-729.

Butler, J., Floccia, C., Goslin, J., & Panneton, R. 2011 Infants' discrimination of familiar and unfamiliar Accents in Speech. *Infancy, 16*(4), 392-417

Cho, T., & Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27*, 207–229.

Clarke, C.M., & Garrett, M. 2004. Rapid adaptation to foreign accented speech. *Journal of the Acoustical Society of America, 116,* 3647–3658.

Cristia, A., Seidl, A., Vaugn, C., Schmale, R., Bradlow, A., & Floccia, C. 2012. Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology, 3,* 479.

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* pp. 91-111. Berlin: Mouton de Gruyter.

Eisner, F., & McQueen, J.M. 2005. The specificity of perceptual learning in speech processing. *Perception & Psychophysics, 67*, 224–238.

Fennell, C.T. & Waxman, S.R. 2010. What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development, 815*, 1376-1383.

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.

Grondelaers, S., van Hout, R., & Steegs, M. (2010). Evaluating regional accent variation in Standard Dutch. *Journal of Language and Social Psychology, 29*, 101–116.

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

Houston, D.M., & Juszcyk, P.W. 2000. The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1570–1582.

Iverson, G.K., & Salmons, J.C. 1995. Aspiration and laryngeal representations in Germanic. *Phonology, 12*(3), 369-396.

Johnson, E.K. 2005. English-learning infants' representations of word-forms with iambic stress. *Infancy*, *7*, 95-105.

Johnson, E.K., McQueen, J., & Huettig, F. 2011. Toddlers' language-mediated visual search: They need not have the words for it. *Quarterly Journal of Experimental Psychology, 64,* 1672-1682.

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

Kinzler, K.D., Shutts, K., DeJesus, J., & Spelke, E.S. 2009. Accent trumps race in guiding children's social preferences. *Social Cognition, 27*(4), 623-634.

Kinzler, K.D., Shutts, K., & Spelke, E.S. (2012). Language-based social preferences among multilingual children in South Africa. *Language Learning and Development*, *8*, 215-232.

Kovacs, A.M., & Mehler, J. 2009. Flexible learning of multiple speech structures in bilingual infants. *Science, 325,* 611-612.

Kraljic, T., & Samuel, A.G. 2005 Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51*(2), 141-78

Kraljic, T., & Samuel, A.G. 2006 Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review, 13*(2), 262-268

Labov, W., Ash, S., & Boberg, C. 2006. *Atlas of North American English: Phonology and phonetics.* Berlin: Mouton de Gruyter.

Maye, J., Aslin, R.N., & Tanenhaus, M.K. 2008. The weckud wetch of the wast: lexical adaptation to a novel accent. *Cognitive Science, 32*, 543–562.

Maye, J., Werker, J.F., & Gerken, L., 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101-B111.

McMurray, B., Tanenhaus, M.K., & Aslin, R.N. 2009. Within-category VOT affects recovery from "lexical" garden-paths; Evidence against phoneme-level inhibition. *Journal of Memory and Language, 60*(1), 65-91.

McQueen, J.M., Tyler, M.D., & Cutler, A. 2012 Lexical retuning of children's speech perception: Evidence for knowledge about words' component sounds. *Language Learning and Development, 8*, 317-339.

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

Nathan, L., Wells, B., & Donlan, C. 1998. Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of Child Language, 25*, 343-365.

Nazzi, T., Jusczyk, P.W., & Johnson, E.K. 2000. Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language, 43*, 1-19.

Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., & Polka, L. 2014. Early word segmentation in infants acquiring Parisian French: task-dependent and dialect-specific aspects. *Journal of Child Language*, 41*(3),* 600-633.

Norris, D., McQueen, J.M., & Cutler, A. 2003. Perceptual learning in speech. *Cognitive Psychology, 47*, 204–238.

Quam, C. & Swingley, D. 2010. Phonological knowledge guides two-year-olds' and adults' interpretation of salient pitch contours in word learning. *Journal of Memory and Language, 62*(2), 135-150.

Salverda, A.P., Dahan, D., & McQueen, J.M. 2003. The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, *90*, 51-89.

Schmale, R., Cristià, A., & Seidl, A. 2012. Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science, 6*, 732-738.

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*(6), 650-662.

Seidl, A., & Cristia, A. 2012. Infants' learning of phonological status. *Frontiers in Psychology, 3*, 448.

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

Singh, L. 2008. Influences of high and low variability on infant word recognition. *Cognition*, *106*, 833-870.

Singh, L., Morgan, J.L., & White, K.S. 2004. Preference and processing: the role of speech affect in early spoken word recognition. *Journal of Memory and Language*, *51*, 173-189.

Singh, L., Nestor, S.S., & Bortfeld, H. 2008. Overcoming the effects of variation in infant speech segmentation: Influences of word familiarity. *Infancy, 13*(1), 57-74.

Sundara, M., & Scutellaro, A. 2011. Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics, 39*(4), 505-513.

Swingley, D. 2009. Onsets and codas in 1.5-year-olds' word recognition. *Journal of*

Memory and Language, 60, 252-269.

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

Van Alphen, P., & Smits, R. 2004. Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics*, *32,* 455–491.

Van Alphen, P. M., & McQueen, J. M. 2006. The effect of voice onset time differences on lexical access in Dutch. *Journal of Experimental Psychology: Human Perception and Performance, 32*(1), 178-196.

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.

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*, 430-450..

Van Heugten, M., Krieger, D., & Johnson. E.K. 2014. The developmental trajectory of toddlers' comprehension of unfamiliar regional accents. *Language Learning and Development.*

Van Hout, R., Adank, P., Van Heuven, V.J. (2000). Akoestische meting van Nederlandse klinkers in algemeen Nederlands en in Zuid-Limburg. *Taal en Tongval, 52*, 150-162.

Van der Feest, S.V.H. 2007. *Building a phonological lexicon. The acquisition of the Dutch voicing contrast in perception and production.* Doctoral dissertation, Radboud University Nijmegen. Utrecht: Prince Productions BV.

Van der Feest, S.V.H. & Fikkert, P. Accepted pending minor revisions. Building phonological lexical representations. *Phonology.*

Van de Velde, H., Gerritsen, M., & Van Hout, R. 1996. Watching Dutch change: a real time study of variation and change in standard Dutch pronunciation. *Journal of Sociolinguistics, 1*(3), 361-391.

Van de Velde, H., & Van Hout, R. 2001. The devoicing of fricatives in a reading task. *Linguistics in the Netherlands, 21*(1), 47-57.

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

White, K.S., & Aslin, R.N. 2011. Adaptation to novel accents by toddlers. *Developmental Science, 14*(2), 372-384.

White, K.S., & Morgan, J.L. 2008. Sub-segmental detail in early lexical representations. *Journal of Memory and Language, 591,* 114-132.

White, K.S., Peperkamp, S., Kirk, C., & Morgan, J.L. 2008. Rapid acquisition of phonological alternations by infants. *Cognition, 107*(1), 238-265.