Shhh… I need quiet! Children’s understanding of American, British, and Japanese speakers

Tessa Bent

Department of Speech and Hearing Sciences

Indiana University

Rachael Frush Holt

Department of Speech and Hearing Science

Ohio State University

Corresponding author: Tessa Bent, Department of Speech and Hearing Sciences, Indiana University, 200 S. Jordan Ave., Bloomington, IN, 47405, USA; Email: tbent@indiana.edu; Phone: 812-856-3279

Abstract

Children's ability to understand speakers with a wide range of dialects and accents is essential for efficient language development and communication in a global society. Here, the impact of regional dialect and foreign-accent variability on children's speech understanding was evaluated in both quiet and noisy conditions. Five- to 7-year-old children (n = 90) and adults (n = 96) repeated sentences produced by three talkers with different accents – American English, British English, and Japanese-accented English – in quiet or noise. Adults had no difficulty understanding any talker in quiet. Their performance declined for the nonnative talker with a moderate amount of noise; their performance only substantially declined for the British talker (i.e., below 93% correct) when their understanding of the American talker was also impeded. In contrast, although children showed accurate word recognition for the American and British talkers in quiet, they had difficulty understanding the nonnative talker, even under ideal listening conditions. With a moderate amount of noise, their perception of British talker declined substantially and their ability to understand the nonnative talker was particularly poor. These results suggest that although school-aged children can understand unfamiliar native dialects under ideal listening conditions, their abilities to recognize words in these dialects may be highly susceptible to the influence of environmental degradation. Fully adult-like word identification for talkers with unfamiliar accents and dialects may exhibit a protracted developmental trajectory.

Keywords

Foreign accent, regional dialect, word recognition, development

Introduction

Very early in life children have difficulty recognizing a familiar word when the word's pronunciation differs from the initial encounter due to a change in the talker's gender, dialect, or accent or the affect of the production (Houston & Jusczyk, 2000, 2003; Schmale, Cristia, Seidl, & Johnson, 2010; Schmale & Seidl, 2009; Singh, Morgan, & White, 2004). However, later in the first year of life and into the second year, children's abilities to recognize known words with these variations improves (Best, Tyler, Gooding, Orlando, & Quann, 2009; Houston & Jusczyk, 2000; Mulak, Best, Tyler, Kitamura, & Irwin, 2013; Schmale et al., 2010; Schmale & Seidl, 2009; Singh et al., 2004; van Heugten, Krieger, & Johnson, 2015). The essential skills needed to recognize words produced in unfamiliar regional dialects and foreign accents may be in place by 2.5 years of age (Schmale, Hollich, & Seidl, 2011; van Heugten & Johnson, 2016) and can be evidenced even earlier if children are provided with a short period of exposure (van Heugten & Johnson, 2014; White & Aslin, 2011). Although foundational skills may emerge within the first few years of life, research with older children suggests that children continue to have immature abilities to contend with dialect and accent variation well into the school-aged years (Bent, 2014; Bent & Atagi, 2015, 2017; Nathan, Wells, & Donlan, 1998; Newton & Ridgway, 2016; O'Connor & Gibbon, 2011).

The differences in methods employed across varying age groups (i.e., toddlers versus school-age children) may account for the conflicting claims in the literature regarding the age at which children can successfully recognize words with unfamiliar accents or dialects. Specifically, tasks used with infants and toddlers tend to be closed-set or require recognition of items as "familiar" as measured through looking times or head turns. These types of tasks may not fully capture children's difficulties with unfamiliar speech varieties compared to open-set tasks. For example, preschoolers show very little difference in behavior between familiar and unfamiliar accents with a four-alternative forced-choice task, but a substantial difference in open-set word repetition tasks (Creel, Rojo, & Paullada, 2016). In addition to task differences, all experiments with infants and toddlers have been conducted in ideal, quiet listening conditions whereas many of the experiments with older children have used noise-added conditions (Bent, 2014; Bent & Atagi, 2015, 2017; Holt & Bent, 2017; Newton & Ridgway, 2016). The lack of environmental degradation may overestimate children's abilities to contend with accent and dialect variation under conditions that more closely simulate real-world listening conditions. The environments where language learning primarily occurs—classrooms and the home—tend to be noisy (Bradley & Sato, 2008; Evans & Lepore, 1993). For example, measurements of the average signal-to-noise ratio (SNR) in active classrooms has varied from -7 to +10 dB SNR across studies (Crandell & Smaldino, 2000; Sato & Bradley, 2008). Within these noisy environments, children are likely to encounter both peers and adults who speak with a dialect or accent different than their own. For example, within the United States, nearly 10% of public school children are learning English as a second language (National Center for Education Statistics, 2016). In the grades that the children in the current study would most likely be enrolled (i.e., kindergarten through second grade), the percentage is even higher: approximately 16%. Furthermore, 20% of the United States population speaks a language other than English at home (Ryan, 2013). Of course, not all of these children or adults would speak with a detectable foreign accent, but these statistics suggest that children are very likely to encounter at least some speakers in their everyday lives with accents or dialects that are different than their own.

Although children clearly show improvements in contending with speech variability within the first few years of life, little work indicates when children's performance matches that of adults. Only three studies to date have compared children's understanding of foreign-accented speech to adults' (Bent, 2014; Bent & Atagi, 2015, 2017) and there are no comparisons across adults and children with regional dialect variation. Although both adults and school-aged children show intelligibility decrements in noise for foreign-accented speech compared to native-accented speech, children show greater intelligibility declines than adults with sentences (Bent & Atagi, 2015, 2017). Furthermore, adults show near perfect word identification accuracy with foreign-accented sentences in quiet, whereas children demonstrate difficulty understanding foreign-accented speakers in the same listening conditions (Bent & Atagi, 2015). Therefore, children's abilities to contend with foreign-accented speech have not reached adult-like performance levels in the early school-aged years. However, these studies have employed foreign-accented speech and have not investigated how children's abilities to identify sentences produced in an unfamiliar regional dialect compare to adults' and how this perception may be influenced by noise. The two studies that investigated school-aged children's perception of regional dialect variation were conducted in quiet and did not have an adult comparison group (Nathan et al., 1998; O'Connor & Gibbon, 2011).

Very few studies have compared the impact of regional dialects versus foreign accents on speech perception or adaptation under the same experimental conditions for either adults or children. In the few extant investigations of this issue, conflicting results have been reported. Two adult studies reported that foreign-accented speech resulted in greater processing costs or intelligibility decrements than unfamiliar regional varieties (Adank, Evans, Stuart-Smith, & Scott, 2009; Bent, Baese-Berk, Borrie, & McKee, 2016). In contrast, Floccia, Butler, Goslin, and Ellis (2009) reported similar processing costs across regional and nonnative varieties for adults listeners. In addition to these discrepancies in behavioral results across studies, there is still debate over whether foreign accents represent extreme forms of variation along the same continuum as regional varieties or whether foreign accents and regional dialects are perceived and processed in distinct manners (see discussion in Cristia et al. (2012) for example). Several studies suggest that these two forms of speech variability may be processed or represented in distinct ways. For example, Goslin, Duffy, and Floccia (2012) indicate that there could be different normalization mechanisms for regional versus foreign-accent variation. In metalinguistic tasks, adults classify regional varieties and foreign-accented varieties in distinct clusters, even when a wide range of native varieties are included in the task (Bent, Atagi, Akbik, & Bonifield, 2016). Further, school-aged children are more sensitive to foreign accents than regional dialects (Girard, Floccia, & Goslin, 2008; Wagner, Clopper, & Pate, 2014), even when the varieties are equated for accent strength based on judgements made by adult listeners (Floccia, Butler, Girard, & Goslin, 2009).

The representational and processing differences reported across regional and nonnative varieties may be influenced by the ways in which the varieties tend to differ from the home dialect in the acoustic-phonetic domain. In particular, regional varieties primarily differ from one another in vowel realizations (Clopper, Pisoni, & de Jong, 2005), although there can be differences across dialects in consonants and suprasegmental aspects as well (Clopper & Smiljanic, 2015; Hickey, 2004; Hughes, Trudgill, & Watt, 2012). In contrast, nonnative varieties frequently differ from native varieties across all levels of phonetic and phonological structure including differences in phonotactics, consonants, vowels, and suprasegmentals (Carlisle, 1991; Sereno, Lammers, & Jongman, 2016). Furthermore, there is some evidence that nonnative varieties display more inter- and intra-talker variation than native varieties (Baese-Berk & Morrill, 2015; Hanulikova & Weber, 2012; Wade, Jongman, & Sereno, 2007). Some authors have suggested that the acoustic-phonetic differences between regional and nonnative varieties (i.e., primarily vowels for regional dialects vs. more consonantal differences for foreign accents) may underlie their findings of differential metalinguistic sensitivity for regional versus nonnative varieties in children (Floccia, Butler, Girard, et al., 2009).

The outcomes of these studies may crucially depend on the specific varieties included as well as the talkers' accent strengths. That is, native varieties will differ in their perceptual distances to the listener's home dialect; nonnative varieties will be strongly influenced by the relationship between the phonetic and phonological features of the first and second languages as well as the talker's proficiency in the second language. Much more work is needed before definitive claims can be made about how perception and processing of regional and nonnative varieties may diverge. Therefore, the current work represents a first step in comparing both children and adults' perception of one regional variety to one foreign accent. The current study expands on earlier work by comparing children's and adults' perception of three speech varieties—the home dialect, an unfamiliar regional dialect, and a foreign accent—in both quiet and noise-added listening conditions.

Method

Participants

The participants included monolingual American English speaking 5- to 7-year-old children with equal numbers in each year bracket (n = 90; 53 female) and 18- to 35-year-old adults (n = 96; average age = 24 years; 44 female). All participants had typical speech, language, and hearing based on parental- or self-report. An additional 34 children were tested, but their data were not included due to an inability to complete the experimental task (i.e., prematurely stopping the experiment of their own volition) (n = 11), experimenter error (n = 11), incomplete paperwork (n = 6), speech, language, hearing, or attention disorder (n = 3), bilingual language status (n = 2), or not meeting the age requirement (n = 1). An additional nine adults were tested, but their data were not included due to bilingual language status (n = 5), inadequate exposure to the Midland dialect (n = 3), or performance that was greater than 3 standard deviations below the mean (n = 1). Using a language background questionnaire, parents rated their child's exposure and adults rated their own exposure to various accents and dialects on a scale of 1 – 5, where 1 = no exposure and 5 = frequent daily exposure. The average exposure ratings for the three linguistic varieties used in this study were: 4.6 (children; range = 1 – 5) and 4.8 (adults; range = 2 – 5) for the Midland dialect; 1.4 (children; range = 1 – 4) and 1.9 (adults; range = 1 – 5) for British English; and 1.1 (children; range = 1 – 3) and 1.4 (adults; range = 1 – 4) for Japanese-accented English. Most participants spoke with a Midland dialect (n = 145). Other dialects included North Central (n = 14), Southwest (n = 7), Southern (n = 6), Appalachian (n = 5), Western Pennsylvanian (n = 4), Northwest (n = 2), Mid-Atlantic (n = 1), and a mixture of Midland and North Central (n = 1). One parent did not indicate their child's dialect on the language history form. None of the parents reported that their children spoke a language other than English. Most of the adults had studied a language other than English, but none reported extensive use of another language or fluency in a language other than English. Participants were not paid for their participation.

Stimuli

The stimuli included sixty sentences from the Hearing-in-Noise-Test for Children, HINT-C (Nilsson, Soli, & Gelnett, 1996). Using read speech ensures that lexical and syntactic content are consistent with only differences in pronunciation across speakers. Three female talkers were recorded reading the sentences including an American speaker with a Midland dialect, a British English speaker, and a native Japanese speaker who spoke English as a second language. The American speaker was 23 years of age, had lived exclusively in Indiana, and spoke only English. The British speaker was 20 years of age and had grown up primarily in London. She had been living in Bloomington, Indiana for the prior 8 months as part of a study abroad program. The Japanese speaker was 27 years of age, began studying English at the age of 13 years, and had lived in the U.S. for 10 months prior to the recording. Recordings of the American and Japanese-accented talkers were taken from the Hoosier Database of Native and Nonnative Speech for Children (Bent, 2014). The strengths of the American and Japanese talkers' accents were rated as 1.1 and 6.5, respectively, on a scale of 1 – 9, where 1 = no foreign accent and 9 = very strong foreign accent (Atagi & Bent, 2013). The British talker was recorded specifically for this study using the same procedures, recording location (a sound-attenuated booth), and equipment as were used to create the recordings in the Hoosier Database. Sentences were equated for root-mean-squared amplitude.

The three selected talkers were typical representatives of their respective speech varieties. The British speaker showed differences from Midland American English in productions of intervocalic /t/ (i.e., producing aspirated /t/ where the midland speaker flapped (e.g., [yɛsɾɚdeɪ] vs. [yɛstʰədeɪ]), derhotacized vowels (e.g., [pʰɑrk˺tʰ] vs. [pʰɑk˺tʰ]), and vowels participating in the bath-trap split (e.g., [læftʰ] vs. [lɑftʰ]) as well as some differences influencing syllable structure of the word (e.g., [strɔbɛriz] vs. [strɔbriz]). The Japanese speaker's productions included frequent substitutions across a variety of consonants (e.g., [klɔs] for *cloth*, [kwɪrɪŋ] for *clearing*, and [mɛdə] for *mother*) and vowels (e.g., [bɛs] for *bus*, [blɑk] for *black*, [dɑn] for *down*, and [dʌg] for *dog*). There were also instances of epenthesis (e.g., [ʧɑkəlɪt] for *chocolate*), phoneme deletion (e.g., [ʊmən] for *woman*), and differences in stress (e.g., [ˈlɛmənpaɪ] vs. [lɛmənˈpaɪ]). Transcriptions and waveforms of two example sentences, as produced by the three speakers, are shown in Appendix A.

Procedure

Participants were recruited from and tested at the Center for Science and Industry (COSI), a museum located in Columbus, Ohio. Participants were randomly assigned to either a quiet or an 8-talker babble condition (see Van Engen, Phelps, Smiljanic, & Chandrasekaran, 2014 for additional details about the babble). Children assigned to the babble condition received a +4 dB signal-to-noise ratio (SNR); adults received a +4 dB or a 0 dB SNR. The sentences were each individually embedded in a random selection from the babble file that was 1 second longer than the sentence.

To familiarize the listeners with the task, listeners were presented with nine practice trials in either quiet or babble (matching the listener's assigned experimental condition) including three sentences from each talker. Following the practice trials, listeners were presented with all sixty sentences blocked by talker. Twenty sentences were assigned to each talker. The sentences assigned to the talkers and talker presentation order were counter-balanced across listeners. Sentence order within a block was randomized for each participant. Each sentence was only presented once. After the presentation of the sentence, the participants verbally repeated back what they heard. Responses were scored by hand at the time of testing by a trained research assistant. Audio recordings of children's responses were not made because previous work with children in a similar age range (5 – 8 years) using a very similar method with both native and nonnative talkers revealed that the rate of discrepancies was only on 1% of keywords between initial transcription and a second transcription used to assess reliability (Bent & Atagi, 2017). Stimulus presentation was controlled by E-Prime v. 2.0 (Psychology Software Tools, 2007). Stimuli were presented binaurally via Audiotechnica headphones (model 8TH-770COM) on a Dell Optiplex 790 desktop computer in a quiet room at a comfortable listening level. Participants were not given feedback during practice or test trials as to the accuracy of their responses, but were given general encouragement.

Results

All words, including both content and function words, were scored for accuracy. Words with added or deleted morphemes were counted as incorrect. Percent correct scores were converted to rationalized arcsine units (RAU) to facilitate meaningful comparisons across the entire range of the scale and to stabilize the error variance (Studebaker, 1985). The RAU scores were analyzed in three ways with the first two analyses comparing adult and child performance and the final analysis focused on the children exclusively.

The first analysis included data from 64 adults and 90 children, who were tested under identical conditions. Half of the participants in each age group were tested in quiet and the other half in babble. Data were analyzed with a repeated-measures ANOVA with one within-subjects variable (accent: American, British, Japanese) and two between-subjects variables (listening environment: quiet, babble at +4 dB SNR; listener age: child, adult) (see Figure 1 for percent correct scores and Table 1 for RAU scores). All three main effects were significant. Accuracy was highest on the American talker followed by the British talker with lowest accuracy on the Japanese-accented talker, $F\ (2, 300) = 670.451$, $p \leq .0001$, $\eta_p^2 = .817$; higher scores were found in quiet compared to noise, $F\ (1, 150) = 190.205$, $p \leq .0001$, $\eta_p^2 = .559$; adults were significantly more accurate than the children, $F\ (1, 150) = 150.618$, $p \leq .0001$, $\eta_p^2 = .501$. An accent by listening environment interaction arose due to a greater influence of noise on the Japanese and British talkers compared to the American talker, $F\ (2, 300) = 11.290$, $p \leq .0001$, $\eta_p^2 = .070$. The accent by listener age interaction was significant, $F\ (2, 300) = 14.540$, $p \leq .0001$, $\eta_p^2 = .088$. The performance difference between the children and adults was relatively small for the American talker, whereas larger performance differences were observed between the children and adults for the British and Japanese talkers. Finally, the three-way interaction among accent, listening environment, and listener age was significant, F (2, 300) = 5.737, $p \underline{=} .004$, $\eta_p{}^2$ = .037. To aid in the interpretation of the three-way interaction, we ran three univariate ANOVAs with one for each accent. These tests allowed us to determine the effects of age and listening environment for each accent separately. For all three ANOVAs, the main effects of age and listening environment were significant ($p < .0001$). The interaction between listening environment and age was not significant for the Midland or Japanese talkers ($p > .3$). However, the interaction between listening environment and age was significant for the British talker ($p = .025$, $\eta_p{}^2 = .033$). These results suggest that noise resulted in similar performance decrements for adults and children when they were presented with the Midland or Japanese talkers, but children showed greater performance decrements than adults in the noise condition relative to the quiet condition for the British talker.

The second analysis included the same child (n = 90) and adult (n = 32) data for the quiet condition as in the first analysis, but the adult data for the noise condition was taken from the adults tested in a 0 dB SNR (n = 32). This analysis was conducted to determine effects of accent and noise when performance across the age groups was equated for the Japanese talker through the use of different SNRs. All three main effects were significant: age, F (1, 150) = 7.423, $p \underline{=} .007$, $\eta_p{}^2$ = .047, accent, F (2, 300) = 556.485, $p \leq .0001$, $\eta_p{}^2$ = .788, and listening environment, F (1, 150) = 425.163, $p \leq .0001$, $\eta_p{}^2$ = .739. The main effects were in the same direction as in the first analysis. All two-way interactions were significant. For the interaction between accent and listening environment, noise affected the American talker the least, the British talker an intermediate amount, and the Japanese talker the most, F (2, 300) = 22.481, $p \leq .0001$, $\eta_p{}^2$ = .130. The accent by listener age interaction was also significant, F (2, 300) = 11.146, $p \leq .0001$, $\eta_p{}^2 = .069$. Follow-up Independent Samples t-tests showed that the children and adults were significantly different on the British and Midland talkers (p-values of .013 and <.0001, respectively) with the children outperforming the adults, likely because of the more favorable SNR for the children. However, the two age groups were not significantly different for the Japanese talker (p = .265). The listener age by listening environment interaction was significant, F (1, 150) = 58.024, $p \leq .0001$, $\eta_p{}^2 = .279$. Lastly, the three-way interaction among listener age, listening environment, and accent was not significant. These results demonstrate that when adults and children are tested at different SNRs to equate performance in the most difficult condition (i.e., nonnative in noise), the main effects observed were consistent with the first set of analyses, in which the adults and children were tested with the same SNRs. However, in this analysis, noise had a greater overall effect on the adults' performance than that of the children likely due to both the adults' superior performance in quiet (average across accents = 112 RAU for adults vs. 95 RAU for children) as well as the more difficult SNR employed with the adults. When compared with the results from the first analysis, we see that adults' performance for the British talker only declined substantially when their performance on the American talker also declined, suggesting that adults have much more robust word recognition abilities for unfamiliar native dialects than children.

The final analysis investigated whether there were differences within the children across the three years included (5-, 6-, and 7-year-old children) (see Figure 2 for percent correct scores and Table 2 for RAU scores). These data were analyzed with a repeated-measures ANOVA with planned post-hoc tests using a conservative Bonferroni adjustment to examine group differences by age in years (5, 6, and 7 years). There were main effects of accent, F (2, 168) = 414.617, p < .0001, $\eta_p{}^2$ = .832, listening environment, F (1, 84) = 113.222, p < .0001, $\eta_p{}^2$ = .574, and age, F (2, 84) = 13.231, p < .0001, $\eta_p{}^2$ = .240. The older children outperformed the younger children with follow-up tests showing that the 7- and 6-year-olds were more accurate than the 5-year-olds (p < .0001 and p = .001, respectively); the 6- and 7-year-olds were not significantly different from one another. The other main effects were in the same directions as the previous analyses. There was a significant interaction between talker accent and listening environment, F (2, 168) = 10.012, p < .0001, $\eta_p{}^2$ = .106, with the largest effect of noise on the British talker (28 RAU difference between quiet and noise), followed by the Japanese talker (21 RAU difference) with the least effect for the Midland talker (16 RAU difference). The other two-way interactions and the three-way interaction were not significant. This analysis demonstrated that within the 5- to 7-year-old age range, children's performance on all speech varieties improved with increasing age. However, there were not significant differences in the extent of improvement across the 5-, 6-, and 7-year-olds for the accents tested (i.e., the age by accent interaction was not significant). However, the age range assessed in this study was quite narrow so it remains possible that an interaction between children's ages and talker accent could arise in future work with a wider age range.

Discussion

The current study examined adults' and school-aged children's abilities to understand speakers representing three speech varieties—the home dialect (Midland American English), an unfamiliar regional dialect (British English), and a nonnative accent (Japanese-accented English)—in quiet and noise-added conditions. This study is the first to directly compare children's perception of these three speech varieties under the same experimental conditions. Adults' word recognition in quiet was highly accurate, regardless of the accent of the talker. In moderate amounts of noise, they began to show difficulty understanding the nonnative talker (72% correct) whereas their performance for the Midland and British talkers was still robust (≥93% correct). Performance for the British talker only declined below these high levels of accuracy when performance on the American talker also was affected and performance on the nonnative talker was particularly poor. Children showed a quite different pattern of results. In quiet, children were highly accurate with both the American and British talkers, but had difficulty understanding the Japanese talker. With even a moderate amount of noise, performance on the British and Japanese talkers declined substantially (75% and 52% correct, respectively). Therefore, children showed significantly less resistance to noise when listening to an unfamiliar regional dialect than their home dialect and had difficulty with a nonnative accent even without the addition of noise, unlike adults who performed well with the nonnative in quiet.

Children did not have difficulty understanding the British talker in quiet listening conditions. This result contrasts with two previous findings in which children were found to have difficulty understanding talkers with unfamiliar regional dialects (Nathan et al., 1998; O'Connor & Gibbon, 2011). It is possible that the acoustic-phonetic or psycholinguistic distance between the two regional dialects tested here (American English vs. British English) was smaller than the distances between the dialects tested in the previous studies (i.e., London vs. Glasgow; two Irish varieties). Incorporating metrics that can capture these acoustic-phonetic (Huckvale, 2004; Wieling, Nerbonne, & Baayen, 2011) or psycholinguistic distances (e.g., ladder task - Bent, Atagi, et al., 2016) into work on children's perception of regional variation would help to elucidate how the relationship between the talker's and listener's dialects impacts speech understanding. Before definitive claims can be made about how regional-dialect or foreign-accent variation influences children's speech understanding, a larger range of varieties including both regional varieties from within and outside of their home country as well as a variety of foreign accents should be tested under the same experimental conditions. It could be informative to then identify the specific acoustic-phonetic properties of those dialects or accents that influence the perceptual effects.

Task differences across studies could also have resulted in conflicting results. Specifically, children in the current study had to repeat short meaningful sentences whereas Nathan et al. (1998) used isolated words, which children had to repeat and define. In this study, children may have capitalized on sentence context to support their top-down processing of the unfamiliar dialect. Children can benefit from sentence context during the perception of a nonnative accent (Creel et al., 2016; Holt & Bent, 2017). Although the previous comprehension task with Irish accents (O'Connor & Gibbon, 2011) also used sentences, children had to carry out commands based on the sentences rather repeating the stimuli. The increased memory demands and task complexity may have reduced the resources available for processing the unfamiliar dialect resulting in lower scores for the younger children. More work is needed to determine how both stimulus length and task complexity contribute to children's ability to understand unfamiliar regional varieties.

Although children in the current study did not have difficulty recognizing words in sentences spoken by the British talker in quiet, their performance declined significantly in noise and they showed a greater performance decrement for the British talker than adults when tested under the same listening conditions. Further, when noise was added, the children's performance declined more for the British talker than either the American or Japanese talker. These results suggest that children's abilities to recognize words in unfamiliar regional dialects are highly susceptible to the influence of environmental degradation. Children's schools and homes are noisy places (Bradley & Sato, 2008; Evans & Lepore, 1993) and the noise condition used here is more similar to what they experience in daily life compared to the ideal, quiet laboratory listening conditions employed in most studies of infant and children's accent and dialect perception. Therefore, even in cases in which an unfamiliar dialect does not appear to have an impact on children's speech recognition or understanding in quiet, there might be an impact of a mismatch in dialect between the talker and listener under real-world listening conditions. These findings also bear on the claim that children show relatively mature abilities for understanding talkers with unfamiliar nonnative accents and dialects by 2 to 2.5 years in novel word learning tasks (Schmale et al., 2011) and word recognition tasks in which looks to a target object versus a distractor were measured (van Heugten & Johnson, 2016; van Heugten et al., 2015). Adult-like performance may not emerge until much later in development, if perception is considered outside of ideal, laboratory conditions.

It is well established that children have more difficulty with speech in noise compared to adults, even when the talker's dialect is matched to their own (Fallon, Trehub, & Schneider, 2000; Leibold & Buss, 2013; Nittrouer & Boothroyd, 1990). The current prevailing view regarding these age effects is that young children do not attend to signals of interest during listening-in-noise tasks in the same way as adults. Both investigations employing tonal stimuli (Bargones & Werner, 1994; Greenberg, Bray, & Beasley, 1970;

Leibold & Neff, 2011; Werner & Bargones, 1991) and a recent study using native-accented speech (Youngdahl, Healy, Yoho, Apoux, & Holt, in press) suggest that prior to approximately 6 to 7 years of age, infants and children do not listen selectively in the spectral region containing the anticipated signal of interest. This finding is in contrast to adults, who listen selectively in the spectral region of interest (Dai, Scharf, & Buus, 1991; Scharf, Quigley, Aoki, Peachey, & Reeves, 1987; Schlauch & Hafter, 1991). Infants' and young children's approach – listening across the frequency spectrum while not ignoring regions that do not aid in the task at hand – interferes with their detection and recognition of speech in noise, resulting in poorer performance relative to adults.

Here, we extend the findings on children's speech-in-noise perception to an unfamiliar dialect and show that children had difficulty in noise conditions with an unfamiliar dialect where adults did not and in a case in which their perception of the home dialect was relatively unimpaired. Children’s more limited language experience, including less experience with a wide range of talkers from their home dialect and with talkers from other dialect regions, may leave them with a narrower pool of exemplars to draw from when attempting to map incoming signals onto words in their lexicons. When these less familiar productions are mixed with noise, children show deficits in mapping words that diverge from the majority of their stored exemplars. Furthermore, because the sentences are designed for use with children, the words are also highly frequent for adults, which would lead to robust performance for adults in noise-added conditions (Savin, 1963). In addition to children’s lower word identification possibly stemming from a lack of sufficient exemplars in their lexicon, another possibility is that children could experience difficulty retrieving accented phonological forms of items under the relatively high task demands of the current study, particularly in the noise-added condition. The current results cannot definitely adjudicate between these two possible underlying causes of word identification difficulty.

In comparison to the unfamiliar regional accent, children had difficulty understanding the nonnative talker, even in quiet conditions. In fact, the nonnative in quiet condition demonstrated one of the greatest differences between child and adult performance. This study is the first to directly compare children's perception of regional and nonnative varieties. Children and adults had greater difficulty with the nonnative accent than the unfamiliar regional dialect, in both quiet and noise-added conditions. These results mirror earlier studies with adults in which greater processing costs for a nonnative accent than an unfamiliar native dialect were observed (Adank et al., 2009; Bent, Baese-Berk, et al., 2016). There are several possible reasons for these results. First, regional dialects tend to primarily differ from one another in vowel realizations (Clopper et al., 2005), whereas nonnative accents frequently differ from native dialect norms on vowels, consonants, and suprasgmental dimensions (Sereno et al., 2016). Thus, the nonnative accent may have diverged from native dialect norms to a greater extent than the unfamiliar regional dialect. Second, there is some evidence that nonnative speakers exhibit more intra-talker variability (Baese-Berk & Morrill, 2015; Wade et al., 2007), which could inhibit listeners' abilities to tune in to the divergences from home dialect norms. Certainly, there were instances of some of these types of dialect and accent differences in our talkers' utterances (see Method and Appendix A). However, this study cannot establish whether the processing of nonnative accents and regional dialects are fundamentally different or whether processing nonnative accents is "an extreme form of the variation in native accents" (Adank et al., 2009, p. 527). Additionally, it is likely that the Japanese talker's accent was stronger than the British speakers', which may contribute to the performance differences across the two varieties. Future work should attempt to match the regional and nonnative varieties for accent strength (Floccia, Butler, Girard, et al., 2009) and also incorporate a larger range of both native and nonnative varieties to determine whether processing nonnative variation involves unique processes. Incorporating regional dialects that diverge further from the listener's dialect (e.g., Glaswegian English for American listeners) and nonnative talkers who diverge less (e.g., a Dutch speaker with a mild foreign accent) may suggest that it is not a difference between regional variation and nonnative variation, but rather the acoustic-phonetic/psycholinguistic distance between the talker's and listener's accents or dialects that determines children's speech understanding.

These results may bear on work showing that children demonstrate social preferences for people who share their language or accent (Kinzler, Corriveau, & Harris, 2011; Kinzler, Dupoux, & Spelke, 2007). There is some evidence that this preference is evidenced from the very beginning of life with newborns showing a preference for their native language (Byers-Heinlein, Burns, & Werker, 2010; Moon, Cooper, & Fifer, 1993), although other work shows that the preference for the native *accent* diminishes over the first year (Kitamura, Panneton, & Best, 2013). Later in development, children demonstrate explicit preference for friends who share their native language or accent (Kinzler & DeJesus, 2013; Kinzler et al., 2007; Kinzler, Shutts, Dejesus, & Spelke, 2009). Speech comprehension has been explored to some extent in these investigations (Kinzler et al., 2011; Kinzler et al., 2009). In Kinzler et al. (2009), who reported that children explicitly show a friend preference for speakers who share their native accent over speakers with a foreign accent, children were presented with samples from a foreign-accented speaker and a foreign-language speaker and asked which talker they understood. In this task, children reliably selected the speaker with the foreign accent over the talker speaking in a foreign language. This result suggests that children are able to understand foreign-accented speech, at least at levels higher than an unknown language. Furthermore, when exposed to speech from native- and foreign-accented speakers who later silently demonstrated functions of novel objects, children selectively learned from the native speaker over the foreign-accented speaker even when the initial exposure period included Jabberwocky (nonsense) speech, which did not provide any semantic content (Kinzler et al., 2011). Because the Jabberwocky condition does not provide meaningful linguistic content, the children's selective learning could not have been based on differences in understanding across the two accents, but must have been based on their preferences for the phonological properties of native-accented speech. Together, these findings were taken as evidence that a child's ability to understand the foreign-accented speech was not guiding their social preferences.

However, intelligibility or processing differences may still contribute to children's social preferences. Children may link unfamiliar dialects and accents to difficult communication encounters or may predict that the unfamiliar varieties would be challenging in naturalistic listening environments. These linkages could lead to their social preferences and selective learning orientations. Indeed, there is a link between adult listeners' negative evaluations of foreign-accented speakers and their indications of speech processing difficulty (Dragojevic & Giles, 2016).

Children’s social preferences and selective learning orientations for those with whom they share an accent may be guided by their real-world experiences interacting with speakers whose dialects and accents are different than their own. Even without experience outside of the laboratory setting, their difficulties understanding the speaker during the experiment or decreased certainty about their responses could be influencing their social preferences. As shown in this study, children have difficulties understanding speakers whose dialects and accents differ from their home dialect, particularly under conditions that more closely mimic what is experienced in real-world communication situations. Further work explicitly investigating the connection between social preferences and understanding of speech from talkers with unfamiliar accents and dialects is warranted.

Conclusion

Under ideal quiet listening conditions, school-aged American children were highly adept at understanding American and British talkers. However, their understanding of a nonnative talker in both quiet and noise as well as a British talker in noise were much lower compared to adults. Thus, children are still developing the ability to map pronunciations of words, which deviate from their home dialect, to items in their lexicons. Fully adult-like perceptual constancy may exhibit a protracted developmental trajectory, particularly when considered under adverse environmental listening conditions. The current study highlights the importance of testing children’s speech comprehension with a range of accents and dialects in conditions that more closely simulate real-world listening conditions.

Acknowledgments

We thank Akemi Jones, Sarah Mabie, Chelsea Mason, Katherine Miller, Leah Neczypor, and Tiarah Wilcox for their assistance with data collection, Dr. Michele Morrisette for her transcription assistance, the children and families who participated in the research, and the Language Sciences Research Lab at the Center for Science and Industry for supporting data collection efforts. This work was supported by the National Science Foundation [grant number 1461039].

References

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520-529. doi:10.1037/a0013552

Atagi, E., & Bent, T. (2013). Auditory free classification of nonnative speech. *Journal of Phonetics, 41*(6), 509-519. doi:doi:10.1016/j.wocn.2013.09.003

Baese-Berk, M. M., & Morrill, T. H. (2015). Speaking rate consistency in native and non-native speakers of English. *Journal of the Acoustical Society of America, 138*(3), EL223-EL228. doi:10.1121/1.4929622

Bargones, J. Y., & Werner, L. A. (1994). Adults listen selectively; infants do not. *Psychological Science, 5*(3), 170-174.

Bent, T. (2014). Children's perception of foreign-accented words. *Journal of Child Language, 41*(6), 1334-1355. doi:10.1017/S0305000913000457

Bent, T., & Atagi, E. (2015). Children's perception of nonnative-accented sentences in noise and quiet. *Journal of the Acoustical Society of America, 138*(6), 3985-3993. doi:10.1121/1.4938228

Bent, T., & Atagi, E. (2017). Perception of nonnative-accented sentences by 5- to 8-year-olds and adults: The role of phonological processing. *Language and Speech, 60*(1), 110-122. doi:10.1177/0023830916645374

Bent, T., Atagi, E., Akbik, A., & Bonifield, E. (2016). Classification of regional dialects, international dialects, and nonnative accents. *Journal of Phonetics, 58*, 104-117.

Bent, T., Baese-Berk, M., Borrie, S. A., & McKee, M. (2016). Individual differences in the perception of regional, nonnative, and disordered speech varieties. *The Journal of the Acoustical Society of America, 140*(5), 3775-3786.

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

Bradley, J. S., & Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *Journal of the Acoustical Society of America, 123*(4), 2078-2086. doi:10.1121/1.2839285

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science, 21*(3), 343-348. doi:10.1177/0956797609360758

Carlisle, R. S. (1991). The influence of environment on vowel epenthesis in Spanish English interphonology. *Applied Linguistics, 12*(1), 76-95. doi:Doi 10.1093/Applin/12.1.76

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

Clopper, C. G., & Smiljanic, R. (2015). Regional variation in temporal organization in American English. *Journal of Phonetics, 49*, 1-15.

Crandell, C. C., & Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Language, speech, and hearing services in schools, 31*(4), 362-370.

Creel, S. C., Rojo, D. P., & Paullada, A. N. (2016). Effects of contextual support on preschoolers' accented speech comprehension. *Journal of Experimental Child Psychology, 146*, 156-180. doi:10.1016/j.jecp.2016.01.018

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

Dai, H. P., Scharf, B., & Buus, S. (1991). Effective attenuation of signals in noise under focused attention. *J Acoust Soc Am, 89*(6), 2837-2842.

Dragojevic, M., & Giles, H. (2016). I don't like you because you're hard to understand: The role of processing fluency in the language attitudes process. *Human Communication Research, 42*, 396-420.

Evans, G. W., & Lepore, S. J. (1993). Nonauditory effects of noise on children: A critical review. *Children's environments, 10*(1), 31-51.

Floccia, C., Butler, J., Girard, F., & Goslin, J. (2009). Categorization of regional and foreign accent in 5-to 7-year-old British children. *International Journal of Behavioral Development, 33*(4), 366-375. doi:10.1177/0165025409103871

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research, 38*(4), 379-412. doi:10.1007/s10936-008-9097-8

Girard, F., Floccia, C., & Goslin, J. (2008). Perception and awareness of accents in young children. *British Journal of Developmental Psychology, 26*, 409-433. doi:10.1348/026151007x251712

Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain and Language, 122*(2), 92-102. doi:10.1016/j.bandl.2012.04.017

Greenberg, G. Z., Bray, N. W., & Beasley, D. S. (1970). Children's frequency-selective detection of signals in noise. *Perception & Psychophysics, 8*(3), 173-175.

Hanulikova, A., & Weber, A. (2012). Sink positive: Linguistic experience with th substitutions influences nonnative word recognition. *Atten Percept Psychophys, 74*(3), 613-629. doi:10.3758/s13414-011-0259-7

Hickey, R. (2004). A Sound Atlas of Irish English (pp. 43 - 50). Berlin, Germany: Mouton de Gruyter.

Holt, R. F., & Bent, T. (2017). Children's use of semantic context in perception of foreign-accented speech. *Journal of Speech, Language, and Hearing Research, 60*, 223-230.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology-Human Perception and Performance, 26*(5), 1570-1582.

Houston, D. M., & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology-Human Perception and Performance, 29*(6), 1143-1154. doi:Doi 10.1037/0096-1523.29.6.1143

Huckvale, M. (2004). *ACCDIST: a metric for comparing speakers' accents.* Paper presented at the ICSLP, Jeju, Korea.

Hughes, A., Trudgill, P., & Watt, D. (2012). English accents and dialects: An introduction to social and regional varieties of English in the British Isles (Fifth edition. ed., pp. 141 - 142). London: Routledge.

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science, 14*(1), 106-111. doi:10.1111/j.1467-7687.2010.00965.x

Kinzler, K. D., & DeJesus, J. M. (2013). Children's sociolinguistic evaluations of nice foreigners and mean Americans. *Dev Psychol, 49*(4), 655-664. doi:10.1037/a0028740

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America, 104*(30), 12577-12580. doi:10.1073/pnas.0705345104

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. doi:10.1521/soco.2009.27.4.623

Kitamura, C., Panneton, R., & Best, C. T. (2013). The Development of Language Constancy: Attention to Native Versus Nonnative Accents. *Child Development, 84*(5), 1686-1700. doi:Doi 10.1111/Cdev.12068

Leibold, L. J., & Neff, D. L. (2011). Masking by a remote-frequency noise band in children and adults. *Ear Hear, 32*(5), 663-666. doi:10.1097/AUD.0b013e31820e5074

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). 2-day-olds prefer their native language. *Infant Behavior & Development, 16*(4), 495-500. doi:Doi 10.1016/0163-6383(93)80007-U

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

Nathan, L., Wells, B., & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of Child Language, 25*(2), 343-365.

National Center for Education Statistics. (2016). English language learners in public schools. Retrieved from http://nces.ed.gov/programs/coe/indicator_cgf.asp

Newton, C., & Ridgway, S. (2016). Novel accent perception in typically-developing school-aged children. *Child Language Teaching and Therapy, 32*(1), 111-123. doi:10.1177/0265659015578464

Nilsson, M., Soli, S. D., & Gelnett, D. J. (1996). *Development of the Hearing in Noise Test for Children (HINT-C)*. Los Angeles, CA: House Ear Institute.

O'Connor, C., & Gibbon, F. E. (2011). Familiarity of speaker accent on Irish children's performance on a sentence comprehension task. *Journal of Clinical Speech and Language Studies, 18*, 1-17.

Psychology Software Tools. (2007). E-Prime 2.0. Pittsburgh, PA.: Psychology Software Tools.

Ryan, C. (2013). *Language Use in the United States: 2011*.

Sato, H., & Bradley, J. S. (2008). Evaluation of acoustical conditions for speech communication in working elementary school classrooms. *The Journal of the Acoustical Society of America, 123*(4), 2064-2077.

Savin, H. B. (1963). Word-frequency effect and errors in perception of speech. *Journal of the Acoustical Society of America, 35*(2), 200-206.

Scharf, B., Quigley, S., Aoki, C., Peachey, N., & Reeves, A. (1987). Focused auditory attention and frequency selectivity. *Percept Psychophys, 42*(3), 215-223.

Schlauch, R. S., & Hafter, E. R. (1991). Listening bandwidths and frequency uncertainty in pure-tone signal detection. *J Acoust Soc Am, 90*(3), 1332-1339.

Schmale, R., Cristia, A., Seidl, A., & Johnson, E. K. (2010). Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy, 15*(6), 650-662. doi:10.1111/J.1532-7078.2010.00032.X

Schmale, R., Hollich, G., & Seidl, A. (2011). Contending with foreign accent in early word learning. *Journal of Child Language, 38*(5), 1096-1108. doi:10.1017/S0305000910000619

Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: flexibility of early word representations. *Developmental Science, 12*(4), 583-601. doi:10.1111/j.1467-7687.2009.00809.x

Sereno, J., Lammers, L., & Jongman, A. (2016). The relative contribution of segments and intonation to the perception of foreign-accented speech. *Applied Psycholinguistics, 37*(2), 303-322. doi:10.1017/S0142716414000575

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*(2), 173-189. doi:Doi 10.1016/J.Jml.2004.04.004

Studebaker, G. A. (1985). A rational arcsine transform. *Journal of Speech and Hearing Research, 28*(3), 455-462.

Van Engen, K. J., Phelps, J. E. B., Smiljanic, R., & Chandrasekaran, B. (2014). Enhancing Speech Intelligibility: Interactions Among Context, Modality, Speech Style, and Masker. *Journal of Speech Language and Hearing Research, 57*(5), 1908-1918. doi:10.1044/Jslhr-H-13-0076 van Heugten, M., & Johnson, E. K. (2014). Learning to contend with accents in infancy: benefits of brief speaker exposure. *Journal of Experimental Psychology: General, 143*(1), 340-350. doi:10.1037/a0032192

van Heugten, M., & Johnson, E. K. (2016). Toddlers' word recognition in an unfamiliar regional accent: The role of local sentence context and prior accent exposure. *Language and Speech, 59*(3), 353-363. doi:10.1177/0023830915600471

van Heugten, M., Krieger, D. R., & Johnson, E. K. (2015). The developmental trajectory of toddlers' comprehension of unfamiliar regional accents. *Language Learning and Development, 11*(1), 41-65.

Wade, T., Jongman, A., & Sereno, J. (2007). Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. *Phonetica, 64*(2-3), 122-144. doi:10.1159/000107913

Wagner, L., Clopper, C. G., & Pate, J. K. (2014). Children's perception of dialect variation. *Journal of Child Language, 41*(5), 1062-1084. doi:10.1017/S0305000913000330

Werner, L. A., & Bargones, J. Y. (1991). Sources of auditory masking in infants: distraction effects. *Percept Psychophys, 50*(5), 405-412.

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

Wieling, M., Nerbonne, J., & Baayen, R. H. (2011). Quantitative social dialectology: explaining linguistic variation geographically and socially. *PLoS One, 6*(9), e23613. doi:10.1371/journal.pone.0023613

Youngdahl, C. L., Healy, E. W., Yoho, S. E., Apoux, F., & Holt, R. F. (in press). The effect of remote masking on the reception of speech by young school-aged children. *Journal of Speech, Language, and Hearing Research*.

Table 1: Word identification accuracy for adults and children in RAU

		Midland	British	Japanese
Adults	Quiet	120.8	119.3	96.8
	Noise (+4 dB SNR)	106.3	100.1	72.1
	Noise (0 dB SNR)	82.7	65.4	47.7
Children	Quiet	109.1	103.6	72.5
	Noise (+4 dB SNR)	93.2	75.3	51.8

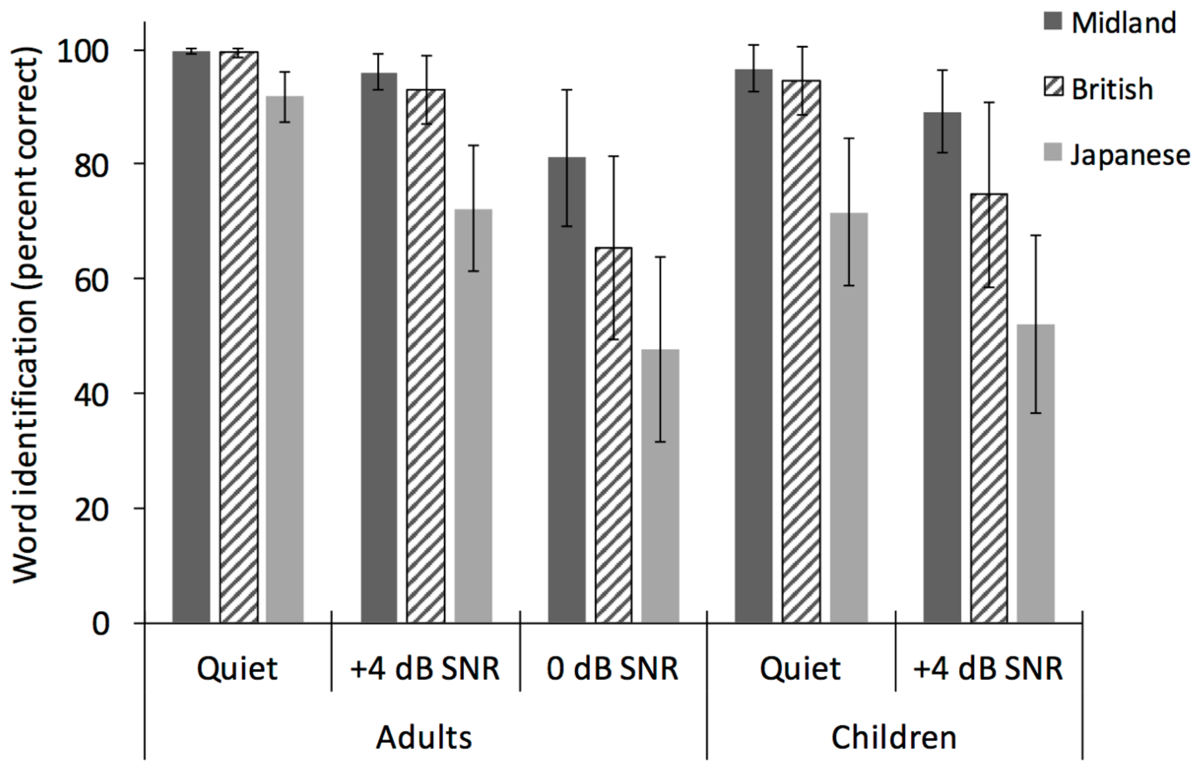
Table 2: Word identification accuracy scores for 5-, 6-, and 7-year-old children in RAU

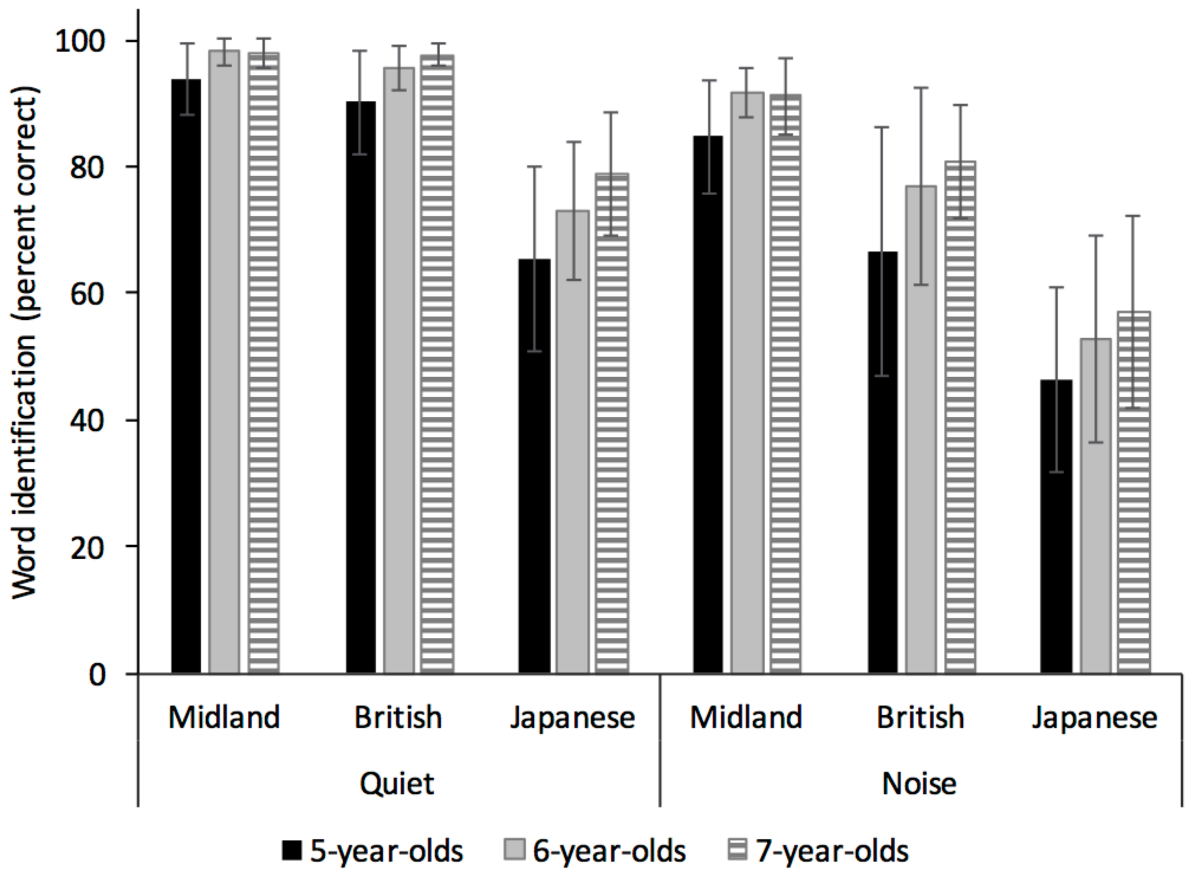
		Midland	British	Japanese
5-year-olds	Quiet	101.9	95.6	65.2
	Noise	87.0	66.3	46.1
6-year-olds	Quiet	114.1	105.0	73.0
	Noise	96.3	77.9	52.6
7-year-olds	Quiet	111.4	110.1	79.4
	Noise	96.2	81.7	56.7

Figure captions

Figure 1: Word identification accuracy (in percent correct) for the adults (left) and children (right) for the American talker (dark gray), British talker (striped), and Japanese-accented talker (light gray).

Figure 2: Word identification accuracy (in percent correct) for the three talkers (American, British, and Japanese-accented) in quiet (left) and in noise (right) for the 5-year-olds (black), 6-year-olds (gray), and 7-year-olds (stripes).


Midland
British
Japanese
Word identification (percent correct)
100
80
60
40
20
0
Quiet
+4 dB SNR
0 dB SNR
Quiet
+4 dB SNR
Adults
Children



Word identification (percent correct)
100
80
60
40
20
0
Midland
British
Japanese
Midland
British
Japanese
Quiet
Noise
5-year-olds
6-year-olds
7-year-olds