

DEVELOPMENT OF PERCEPTUAL FLEXIBILITY

Tessa Bent

Department of Speech and Hearing Sciences, Indiana University

E-mail: tbent@indiana.edu

ABSTRACT

In multicultural and multilingual societies, perceptual flexibility – the ability to recognize words with novel pronunciations, such as those encountered in unfamiliar dialects and accents – is an essential skill for children’s receptive language development. Without this ability frequent errors in mapping input to words in the lexicon would occur. To assess how unfamiliar accents impact word recognition during development, the word recognition abilities of 5- to 12-year-old children and young adults were tested with native- and non-native-accented speech in quiet and noise. Results showed that perception of non-native speech develops slowly with adult-like abilities emerging only in adolescence. Perceptual flexibility may emerge late in development because extensive linguistic input and knowledge are required to accurately map novel acoustic-phonetic patterns onto known words. Compared to familiar native accents, the accurate perception of non-native accents may also more heavily recruit cognitive skills that continue to develop into adolescence (e.g., working memory, executive functions).

Keywords: non-native speech, first language acquisition, speech perception, word recognition

1. INTRODUCTION

During language learning, children must develop the perceptual flexibility to recognize words that differ from initially encountered acoustic-phonetic forms. A particular challenge is in the recognition of words produced in unfamiliar regional dialects or non-native accents, which can introduce non-canonical pronunciations ranging from sub-phonemic mismatches to substantial segmental and suprasegmental differences (e.g., segmental substitutions, deletions, and additions). Non-native speech presents the added perceptual complexity of higher rates of nonsystematic variation both within and across talkers [14]. Adult listeners can have perceptual difficulty with unfamiliar accents, but with exposure show adaptation [7].

Research into how children develop the vital ability to understand speech produced with a variety of accents is just beginning [5]. Dialect and accent

variability can cause word recognition and word learning difficulties for infants and toddlers [6, 19, 26, 30], but improvements are seen with continued development [6, 19, 25, 27] and exposure [24, 30, 32]. Some mechanisms necessary for processing novel accents appear to develop by the second year of life, including the use of top-down lexical feedback to adapt to an unfamiliar regional accent [19, 30]. However, the influence of novel accents on open-set word recognition cannot be tested explicitly with infants and young toddlers. Research on children beyond toddlerhood has focused on their metalinguistic skills or word recognition abilities. By age 5, children can explicitly distinguish native from non-native speakers, but the same metalinguistic ability develops more slowly with native varieties [8, 12, 13, 18, 31]. The few studies on word recognition with children beyond toddlerhood suggest that word recognition abilities for both non-native accents and unfamiliar regional dialects continue to develop throughout childhood [4, 5, 16, 20]. Because these studies focus on children in relatively narrow age ranges and have not always included comparisons with adults, our knowledge of the developmental trajectory for regional dialect and non-native accent perception – including the age at which mature abilities emerge – is severely limited.

Word recognition under other types of adverse listening conditions (e.g., noise, reverberation, or spectral degradation) does not reach adult-like levels until 13 or 15 years of age [9, 17, 21] and even later with two simultaneous sources of environmental degradation (i.e., noise and reverberation) [17]. Similarly, older children (i.e., 9 – 12 years of age) continue to have decrements relative to adults in their abilities to perceive vowels under high variability conditions, such as multiple talkers varying in age, gender, and dialect [16]. This study extends our knowledge of the developmental time course for word recognition with an unfamiliar non-native accent. Specifically, 5- to 12-year-old children’s recognition of sentences produced by one non-native and one native speaker of English in both quiet and noise-added conditions was tested. Their performance in these conditions was compared with a group of young adults.

2. METHOD

2.1. Stimuli

Eighty sentences from the Hearing in Noise Test [22] were selected from the Hoosier Database of Native and Non-native Speech for Children [2, 5]. Each sentence contained three to four keywords. Two adult male talkers produced the sentences: a native speaker of American English and a non-native speaker of English with a first language of Japanese. The Japanese-accented talker's foreign accent strength was rated in an earlier study as 5.5 on a scale from 1 – 9, where 1 represents no foreign accent and 9 represents a strong foreign accent [2].

2.2. Participants

Ninety-nine monolingual American English listeners participated. The listeners represented four age groups: 5- and 6-year-olds ($n=50$), 8- and 9-year-olds ($n=15$), 11- and 12-year-olds ($n=10$), and young adults between the ages of 18 and 24 years ($n=24$). All participants passed a hearing screening and all children demonstrated developmentally appropriate articulation and language skills. None of the participants had extensive experience with Japanese-accented English as determined by parental report (child participants) or self-report (adult participants).

2.3. Procedure

Listeners were presented with all 80 sentences in four blocked conditions: native in quiet, native in noise, non-native in quiet, and non-native in noise. For the noise-added conditions, sentences were embedded in speech-shaped noise at 0 dB SNR with 500ms of noise proceeding and following the sentence. The assignment of sentences to conditions and the order of presentation for the native and non-native talkers were counterbalanced across participants. Within a condition, sentences were randomized. Before the start of the experimental trials, listeners were presented with four practice trials including one from each condition type.

Sentences were presented over a loudspeaker (Yamaha MSP7 Studio) at approximately 68 dB. Listeners repeated the sentence they heard and were encouraged to guess if they were unsure. The experimenter typed in the participant's response, which was also audio recorded for later accuracy re-checking. No feedback was provided. Stimulus presentation and response entry were controlled by custom software written in Python.

2.4. Analysis

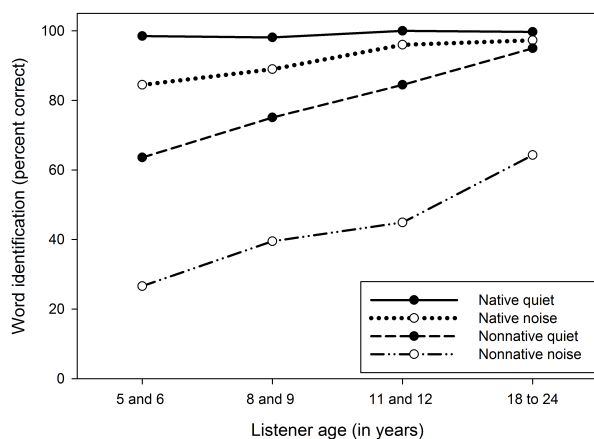
Participants' responses were scored for keyword accuracy in each of the four conditions. Words with added or deleted morphemes were counted as incorrect. Per cent correct scores were converted to rationalized arcsine transform units (RAU) [28] to facilitate meaningful statistical comparisons across the entire range of the scale.

3. RESULTS

Keyword accuracy scores (in RAU) were first analysed with a repeated-measures ANOVA including two within-subjects variables – accent (native, non-native) and listening environment (quiet, noise) – and one between-subjects variable – listener age (5- and 6-year-olds, 8- and 9-year-olds, 11- and 12-year-olds, and young adults). Results showed main effects of accent [$F(1, 95) = 1829.79$, $p < .001$], listener age [$F(3, 95) = 94.23$, $p < .001$], and listening environment [$F(1, 95) = 907.46$, $p < .001$]. These main effects resulted from less accurate word recognition performance for the non-native talker, for the younger listeners, and in noisy conditions (Figure 1). Further, all two-way interactions were significant. The interaction between listener age and listening environment [$F(3, 95) = 6.25$, $p = .001$] arose due to younger listeners showing greater decrements in noise-added conditions compared to older listeners. Similarly, the interaction between listener age and accent [$F(3, 95) = 28.94$, $p < .001$] resulted from younger listeners showing greater decrements for the non-native talker relative to the native talker compared to older listeners. Finally, the interaction between listening environment and accent [$F(1, 95) = 57.62$, $p < .001$] was due to word recognition accuracy on the non-native talker declining more steeply in the presence of noise than scores for the native talker. The three-way interaction was also significant [$F(3, 95) = 3.29$, $p = .024$]. To facilitate interpretation of the three-way interaction, separate ANOVAs were conducted on the data from the native talker and the non-native talker. For both the native and non-native talkers, there were main effects of listening environment [native: $F(1, 95) = 260.33$, $p < .001$; non-native: $F(1, 95) = 536.19$, $p < .001$] and listener age [native: $F(3, 95) = 33.56$, $p < .001$; non-native: $F(3, 95) = 88.49$, $p < .001$]. Additionally, for the native talker, there was a significant interaction between listening environment and listener age [$F(3, 95) = 10.46$, $p < .001$]. In contrast, for the non-native talker, the two-way interaction between listening environment and listener age was not significant ($p = .47$). Therefore,

the younger listeners had more difficulty overcoming the presence of noise than the older listeners when presented with a native talker. However, all age groups showed similar decrements from the quiet condition to the noise-added condition for the non-native talker.

Figure 1: Word identification scores for the native talker in quiet (solid line), native talker in noise (dotted line), non-native talker in quiet (dashed line), and non-native talker in noise (dashed/dotted line) for the four listener age groups.



An additional set of analyses was conducted to determine when performance reached adult-like levels for each condition. This analysis was not conducted for the native in quiet condition because all age groups demonstrated near ceiling performance ($\geq 98\%$ correct on average). For the other three conditions (native in noise, non-native in quiet, and non-native in noise), comparisons were made between each of the child age groups and the adult group using independent sample *t*-tests. Because there were a number of comparisons involved, the conservative Bonferroni correction was applied, which indicated that *p*-values must be .006 or less to be considered significant. For the native talker in noise condition, performance by the 11- and 12-year-old listeners was adult-like [$t(32) = .05$, *ns*] whereas the younger groups were significantly less accurate than the adults [5- and 6-year olds: $t(72) = 9.13$, $p < .001$; 8- and 9-year olds: $t(37) = 4.56$, $p < .001$]. Similarly, the two younger age groups were significantly less accurate on the non-native talker in quiet than adults [5- and 6-year olds: $t(66.23) = 12.43$, $p < .001$; 8- and 9-year olds: $t(37) = 6.21$, $p < .001$] and there was a trend for the 11- and 12-year-olds to be less accurate than the adults [$t(32) = 2.89$, $p = .007$]. Lastly, for the non-native talker in noise, children in all three age groups were significantly less accurate than the

adults [5- and 6-year olds: $t(72) = 14.69$, $p < .001$; 8- and 9-year olds: $t(37) = 7.27$, $p < .001$; 11- and 12-year olds: $t(32) = 6.65$, $p < .001$].

4. DISCUSSION

The current study investigated how listeners between the early school age years and young adulthood perceive native and non-native speech in quiet and noise-added conditions. The results demonstrated that there is a protracted developmental trajectory for the perception of non-native speech. Although the oldest children in the study (11- and 12-year-olds) showed adult-like performance for the native talker in noise, they still had not reached adult-like performance levels for the non-native talker in quiet or noise. Similar to the ability to perceive native-accented speech in very difficult environmental conditions (e.g., noise and reverberation), the ability to accurately recognize speech from non-native talkers appears to continue maturing into adolescence.

The factors that underlie age-related improvements in listeners' abilities to overcome challenging listening conditions may stem from sensory or cognitive development. For speech in noise, previous work suggests that the decrements seen in younger children can largely be accounted for by sensory factors [10], specifically an increased effect of masking [23]. Thus, noise-induced decrements to word recognition may be more related to peripheral factors and not closely tied to linguistic experience. In contrast, sensory development cannot fully account for the current findings because the children had significantly more difficulty than adults understanding the non-native talker, even in quiet conditions. Therefore, the children's difficulty perceiving words produced by non-native talkers may be accounted for by cognitive-linguistic developmental factors. Recognition of words that deviate from native talker norms likely requires additional skills that are not heavily recruited during the perception of native talkers whose pronunciation patterns are closely matched to the listener's own dialect. These skills may include both linguistic and general cognitive abilities.

To accurately recognize non-native accented words, listeners can use top-down knowledge to resolve mismatches between the input and the representations of words in the lexicon. In this study, the use of top-down knowledge from the sentential context could have been beneficial for many of the stimulus items (e.g., "The house has nine bedrooms"). However, the sentence context would have provided less beneficial information for some

others (e.g., “The boy did a handstand”). Young children can use sentence context to facilitate word recognition for native-accented speech in noise [11]. However, it is possible that children are less able to utilize sentential context in adult-like ways when the listening difficulty stems from talker-related factors, such as non-native speech. To test this hypothesis, children and adults’ recognition of non-native accented words in high and low predictability sentences could be evaluated.

Growth in other linguistic domains – such as expanding vocabulary – could account for some of the age-related improvements for the non-native talker seen here. There is a correlation between vocabulary size and young children’s (4- to 7-years of age) recognition of isolated non-native accented words in noise [5]. Although the children in the current study represented a wider age range than in the previous work, vocabulary knowledge is continuing to rapidly develop in the age range tested. Lexical development could at least partially account for performance improvements seen across the age groups as well as the performance gap between the 11- and 12-year-old children and the young adults. The mechanism that underlies the relationship between vocabulary size and speech perception in adverse conditions has not been fully specified. Greater lexical connectivity has been proposed as the mechanism that promotes better speech-in-noise perception for adult listeners with larger vocabularies [29]. Greater lexical connectivity could also assist listeners to access words whose pronunciations deviate from native language norms. Another account for the vocabulary size effect, particularly for children, may be the relationship between vocabulary size and language input [15]. Children with greater language input may have experienced a wider range of talkers and exemplars for each word. Therefore, in contrast to the age-related gains for speech-in-noise perception that appear to be primarily derived from sensory development, the age-related gains seen here for perception of non-native speech, including in quiet conditions, may be related to language experience. The amount of experience with the specific accent tested (i.e., Japanese-accented English) was controlled, such that none of the listeners included in the study had extensive experience with the accent. However, it is possible that greater linguistic experience more broadly – including experiences with other dialects and accents – promotes word recognition with an unfamiliar non-native accent [3].

Perception of speech that deviates from native language norms may also rely on cognitive abilities that are continuing to develop in adolescence.

Mapping words produced in an unfamiliar accent onto items in the lexicon requires flexibility in what listeners will accept for a pronunciation of a specific word. The cognitive skills that allow a listener to adapt to regularities in the talker’s speech while simultaneously ignoring cues that are uninformative may depend on cognitive skills, such as cognitive flexibility, executive function, selective attention, and working memory [1].

The current study was limited in several ways. First, the number of children in the middle age groups was smaller than in the youngest child and adult groups. More listeners at these ages should be tested. Second, the oldest listeners in the current study did not demonstrate fully adult-like abilities for the non-native talker. Therefore, future work should include older children so that full developmental trajectories for non-native accent perception can be mapped. Third, only one non-native accent and one type of noise were employed. Future research should investigate additional non-native accents and extend to regional dialect variation. Other types of noise could be included, particularly those more closely simulating real-world listening conditions, such as multi-talker babble or reverberant listening spaces. These additions would provide a more comprehensive view of how talker- and environment-related adverse listening conditions impact children’s word recognition.

5. CONCLUSION

This study demonstrated that adult-like perception of non-native accented speech does not emerge until adolescence. This protracted developmental trajectory suggests that the ability to perceive non-native accents, even in quiet listening conditions, relies on cognitive and linguistic skills that are continuing to develop throughout childhood.

6. ACKNOWLEDGEMENTS

This research would not have been possible without the technical support of C. Brandt and data collection assistance from E. Atagi, E. Bonifield, T. Burris, E. Byers, K.-Y. Chao, H. Craig, N. Eastman, S. Elmlinger, K. Fishman, V. Filimonova, J. Frye, A. Helms, M. Loughnane, M. McKee, D. Phillips, K. Quinones, R. Shepherd, L. Shidlofsky, A. Simeur, K. Taelman, and Z. Smith.

7. REFERENCES

- [1] Adank, P., Janse, E. 2010. Comprehension of a novel accent by young and older listeners. *Psychology and Aging* 25, 736-740.

- [2] Atagi, E., Bent, T. 2013. Auditory free classification of nonnative speech. *Journal of Phonetics* 41, 509-519.
- [3] Baese-Berk, M.M., Bradlow, A.R., Wright, B.A. 2013. Accent-independent adaptation to foreign accented speech. *The Journal of the Acoustical Society of America* 133, EL174-EL180.
- [4] Barker, B.A., Turner, L.M. in press. Influences of foreign accent on preschoolers' word recognition and story comprehension. *Applied Psycholinguistics*.
- [5] Bent, T. 2014. Children's perception of foreign-accented words. *Journal of Child Language* 41, 1334-1355.
- [6] Best, C.T., Tyler, M.D., Gooding, T.N., Orlando, C.B., Quann, C.A. 2009. Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science* 20, 539-542.
- [7] Bradlow, A.R., Bent, T. 2008. Perceptual adaptation to non-native speech. *Cognition* 106, 707-729.
- [8] Edwards, J., Gross, M., Chen, J., MacDonald, M.C., Kaplan, D., Brown, M., Seidenberg, M.S. 2014. Dialect awareness and lexical comprehension of mainstream american english in african american english-speaking children. *J Speech Lang Hear Res* 57, 1883-1895.
- [9] Elliott, L.L. 1979. Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *J Acoust Soc Am* 66, 651-653.
- [10] Fallon, M., Trehub, S.E., Schneider, B.A. 2000. Children's perception of speech in multitalker babble. *J Acoust Soc Am* 108, 3023-3029.
- [11] Fallon, M., Trehub, S.E., Schneider, B.A. 2002. Children's use of semantic cues in degraded listening environments. *J Acoust Soc Am* 111, 2242-2249.
- [12] 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, 366-375.
- [13] Girard, F., Floccia, C., Goslin, J. 2008. Perception and awareness of accents in young children. *British Journal of Developmental Psychology* 26, 409-433.
- [14] Hanulíková, A., Weber, A. 2012. Sink positive: Linguistic experience with th substitutions influences nonnative word recognition. *Attention Perception & Psychophysics* 74, 613-629.
- [15] Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., Lyons, T. 1991. Early Vocabulary Growth - Relation to Language Input and Gender. *Developmental Psychology* 27, 236-248.
- [16] Jacewicz, E., Fox, R.A. 2014. The effects of indexical and phonetic variation on vowel perception in typically developing 9- to 12-year-old children. *J Speech Lang Hear Res* 57, 389-405.
- [17] Johnson, C.E. 2000. Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research* 43, 144-157.
- [18] 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, 12577-12580.
- [19] 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, 2064-2078.
- [20] Nathan, L., Wells, B., Donlan, C. 1998. Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of Child Language* 25, 343-365.
- [21] Neuman, A.C., Hochberg, I. 1983. Children's perception of speech in reverberation. *J Acoust Soc Am* 73, 2145-2149.
- [22] Nilsson, M., Soli, S.D., Sullivan, J.A. 1994. Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America* 95, 1085-1099.
- [23] Nitttrouer, S., Boothroyd, A. 1990. Context effects in phoneme and word recognition by young children and older adults. *J Acoust Soc Am* 87, 2705-2715.
- [24] Schmale, R., Cristia, A., Seidl, A. 2012. Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science* 15, 732-738.
- [25] 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, 650-662.
- [26] Schmale, R., Hollich, G., Seidl, A. 2011. Contending with foreign accent in early word learning. *Journal of Child Language* 38, 1096-1108.
- [27] Schmale, R., Seidl, A. 2009. Accommodating variability in voice and foreign accent: flexibility of early word representations. *Developmental Science* 12, 583-601.
- [28] Studebaker, G.A. 1985. A rational arcsine transform. *Journal of Speech and Hearing Research* 28, 455-462.
- [29] Tamati, T.N., Gilbert, J.L., Pisoni, D.B. 2013. Some factors underlying individual differences in speech recognition on PRESTO: a first report. *Journal of the American Academy of Audiology* 24, 616-34.
- [30] 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, 340-50.
- [31] Wagner, L., Clopper, C.G., Pate, J.K. 2013. Children's perception of dialect variation. *Journal of Child Language* 40, 1062-1084.
- [32] White, K.S., Aslin, R.N. 2011. Adaptation to novel accents by toddlers. *Developmental Science* 14, 372-384.