

Developmental Changes in Infants' Ability to Cope with Dialect Variation in Word Recognition

Rachel Schmale*

*Department of Psychology
North Park University*

Alejandrina Cristià*

*Département d'Etudes Cognitives
Laboratoire de Sciences Cognitives et Psycholinguistique*

Amanda Seidl

*Department of Speech, Language, and Hearing Sciences
Purdue University*

Elizabeth K. Johnson

*Department of Psychology
University of Toronto*

Toward the end of their first year of life, infants' overly specified word representations are thought to give way to more abstract ones, which helps them to better cope with variation not relevant to word identity (e.g., voice and affect). This developmental change may help infants process the ambient language more efficiently, thus enabling rapid gains in vocabulary growth. One particular kind of variability that infants must accommodate is that of dialectal accent, because most children will encounter speakers from different regions and backgrounds. In this study, we explored developmental changes in infants'

*Both authors contributed equally to this work.

Correspondence should be sent to Rachel Schmale, North Park University, 3225 W Foster Avenue, Box 16, Chicago, IL 60625. E-mail: rschmale@northpark.edu

ability to recognize words in continuous speech by familiarizing them with words spoken by a speaker of their own region (North Midland-American English) or a different region (Southern Ontario Canadian English), and testing them with passages spoken by a speaker of the opposite dialectal accent. Our results demonstrate that 12- but not 9-month-olds readily recognize words in the face of dialectal variation.

Regionally driven dialectal differences produce phonetic variation that straddles the boundary between linguistically relevant and linguistically irrelevant variation. Even for mutually comprehensible dialectal accents, such as North Midland-American and Southern Ontario Canadian English, phonetic differences affect the realization of contrasts, which may complicate word recognition. As a result of the Canadian shift, both /æ/ and /I/ are lowered and more backed in Southern Ontario Canadian English, compared with North Midland-American English (Labov, Ash, & Boberg, 2006). For example, [ma:p] may be perceived as “map” in this Canadian dialect, but as “mop” in this American dialect. This may fetter perception for American listeners unfamiliar with the variation introduced by this dialect (e.g., Kraljic, Samuel, & Brennan, 2008). In fact, phonetic variation as a result of dialectal accent affects adults’ perception in very complex ways (see Sumner & Samuel, 2009 for a recent review), and it interferes with children’s word recognition at age 4 (Nathan, Wells, & Donlan, 1998) and sentence discrimination at age 5 (Floccia, Butler, Girard, & Goslin, 2009). Dialect variation may also be problematic for infant learners, who have less language experience. However, less is known about how such phonetic variation may impact infant speech perception, particularly word recognition (although, see Best, Tyler, Gooding, Orlando, & Quann, 2009 for its impact on budding semantic representations).

As infants gain experience with their ambient language, they attune to phonetic information that is linguistically relevant. Language experience may also help infants ignore information irrelevant to word identity, such as variation attributable to gender, affect, and accent (foreign and dialectal). From an early age, infants exhibit some ability to deal with irrelevant speaker variability. Two-month-olds detect a syllable change when produced by multiple speakers (Jusczyk, Pisoni, & Mullenix, 1992) and 6-month-olds discriminate a phonetic contrast between vowels, despite variability across speaker age and gender (Kuhl, 1979, 1983). Although infants can cope with linguistically irrelevant variability in sound discrimination, this ability does not translate to word recognition. Indeed, 7.5-month-olds fail to recognize a word when spoken by two speakers with dissimilar voices (e.g., male versus female; Houston & Jusczyk, 2000) and the same word spoken in different affective states (e.g., happy versus neutral; Singh, Morgan, & White, 2004).

It is not until 10.5 months that infants ignore irrelevant gender and affect variability in word recognition (Houston & Jusczyk, 2000; Singh et al., 2004).

Surprisingly little is known, however, about whether infants can accommodate the linguistically irrelevant variation introduced by dialectal accent when recognizing words in fluent speech. Although infants as young as 5–7 months of age can discriminate different dialectal accents (Kitamura, Panneton, Deihl, & Notley, 2006; Nazzi, Jusczyk, & Johnson, 2000), it is unknown how the aspects that differ across accents impact word recognition. One exception is Schmale and Seidl (2009), where 9- and 13-month-olds were tested on their ability to generalize words from a native speaker of infants' ambient dialectal accent (North Midland-American English) to a foreign-accented speaker (Spanish-accented English). Results showed that, although the 13-month-olds recognized words across these accents, 9-month-olds failed. The authors suggest that one explanation for this developmental pattern may relate to an increase in the flexibility of infants' word representations, with older infants being better able to ignore linguistically irrelevant variation introduced by different accents. For example, Spanish-accented speakers tend to use different cues to signal stress contrasts, show more overlap in vocalic categories, and demonstrate shorter voice onset time (VOT) in stop consonants (Jongman & Wade, 2007; Shah, 2004; Steinlen, 2005). In fact, there were many linguistically irrelevant subphonemic and suprasegmental differences between the Spanish-accented and American speakers (Schmale & Seidl, 2009). Thus, it is possible that 9-month-olds failed because the differences between the accents were substantial. This is plausible, given that younger infants are worse at "harder" word recognition tasks, as it has been shown for vowel-initial words (Mattys & Jusczyk, 2001; Seidl & Johnson, 2008), iambic words (Jusczyk, Houston, & Newsome, 1999; Nazzi, Dille, Jusczyk, Shattuck-Hufnagel, & Jusczyk, 2005), and words in nonsalient prosodic positions (Seidl & Johnson, 2006).

Thus, it was unclear which differences were responsible for the 9-month-olds' difficulty. For example, Spanish-accented English deviates from North Midland-American English by way of subphonemic and suprasegmental (sentence and word) differences. Here, instead, we examine developmental changes in infants' word recognition abilities across two regional accents that differ minimally: North Midland-American English (infants' ambient dialect) and Southern Ontario Canadian English (Labov et al., 2006). These dialectal accents should differ only in vowel implementation, as no reports have been made of differences at the consonantal or suprasegmental level (Clarke, Elms, & Youssef, 1995; Labov et al., 2006; Wells, 1982). Investigating the impact of vowel variation on word recognition provides insight into the relative specificity of early word representations in responding to irrelevant phonetic information.

EXPERIMENT

Both 9- and 12-month-olds were familiarized with words spoken in isolation, and subsequently tested with passages that either contained the familiar words or not, as spoken by a speaker of a different dialectal accent. If infants recognized the familiar words in the passages during test, despite the speaker (and dialectal accent) change, they should exhibit a preference for passages containing the familiar words (e.g., Jusczyk & Aslin, 1995).

Methods

Participants

A total of twenty-four 9-month-olds (M age = 9.01 months; range = 8.52–9.44 months; 11 females) and twenty-four 12-month-olds (M age = 12.14 months; range = 11.58–12.76 months; 13 females) raised in the Midwest participated. Fifteen additional infants were excluded (11 owing to fussing, of which 2 were 12-month-olds; 1 as a result of parental interference; 1 because of prematurity; and 2 owing to foreign language exposure). After data were collected, parents of participants were invited to report both spouses' dialect, and 33 responded. No parent had a Canadian accent, and all but one (English) had an American accent; there was only one case in which a child had both parents from non-Midwestern origins.

Procedure

In this version of the Headturn Preference Procedure (Jusczyk & Aslin, 1995), the infant is seated on a caregiver's lap in the middle of a room, which has a green light in front and red lights on the sides. Each trial begins with the green light flashing. Once the infant orients to it, it extinguishes and one of the sidelights begins to flash. When the infant orients toward the sidelight, speech plays from the speakers hidden behind it, and continues playing until the infant orients away for more than 2 sec. When this happens, the sidelight extinguishes and the front light begins flashing, in preparation for the next trial. If the infant reorients in less than 2 sec the trial continues, but time spent looking away is not counted. A computer program randomly specifies the activation of the sidelights and the stimuli presentation. Both the caregiver and experimenter (who monitor the headturns through an opening in the front) are blind to the stimuli the infant hears.

Design

Following Jusczyk et al. (1999) and Schmale and Seidl (2009), infants were familiarized with 14 different repetitions of each of two target words (either *kingdom* and *hamlet*, for half the infants; or *candle* and *raptor*, for the other half) until they accumulated 30 sec of looking time to each word, and were then tested with three blocks of four trials. During test trials, a six-sentence passage was presented, for a total of six repetitions of each target word. To control for a possible speaker or dialect preference, half of the infants were familiarized by the American speaker and tested by the Canadian speaker. The other half heard the speakers in the opposite order. Infants were randomly, equally assigned to one of two conditions (familiarized with *kingdom/hamlet* or *candle/raptor*) and one of two familiarization orders (familiarized by American or Canadian speaker). All infants were tested on the same passages.

Stimuli

Two speakers were selected from a sample of five North Midland-American speakers and five Southern Ontario Canadian speakers (all women) because they had the greatest voice similarity of all pairs, established by listener ratings following Houston (2000) and Schmale and Seidl (2009). The American speaker was also used in Schmale and Seidl (Experiments 1–3). Further, the speakers' voices used in this work differed much less than the two same-dialect speakers used in Experiment 1 of Schmale and Seidl.¹ Because 9-month-olds successfully recognized words in their work, voice dissimilarity is unlikely to prevent recognition here.

Recordings of American speakers were conducted in a double-walled sound-attenuated booth with an Audio-Technica 100HE Hypercardioid dynamic microphone (Stow, OH). Recordings of Canadian speakers were conducted in a double-walled Industrial Acoustics Company booth (Bronx,

¹Adult listener similarity ratings were carried out using both natural and sinewave speech. Sinewave speech annuls voice differences leaving only accent characteristics (Krentz & Corina, 2008; Remez, Fellowes, & Rubin, 1997; Remez, Van Dyk, Fellowes, & Rubin, 1998). Thus, by subtracting the sinewave (accent-based) similarity ratings from the natural similarity ratings, speakers with the most similar voices (beyond accent) were chosen through multidimensional scaling (Houston, 2000; Houston & Jusczyk, 2000; Wannemacher, 1995; see also Sheffert, Pisoni, Fellowes, & Remez, 2002). In Experiment 1 of Schmale and Seidl (2009), 9-month-olds successfully recognized words across female talkers with highly dissimilar voices (i.e., large change in the mean-squared distances in sinewave and natural speech, 11.03 in natural speech, and 6.08 in natural speech). The talkers used in this work had more similar voices, as indicated by a much smaller change in the mean-squared distances in sinewave and natural speech (1.39 in natural speech and 1.26 in sinewave speech).

NY) with an Edirol wave recorder (Bellingham, WA). Stimuli were digitized at 44.1 kHz, normalized to ~70 dB, and all target words and passages were equated in duration. The average duration of the American speaker's stimuli was 17.94 sec for target word lists (range = 16.53–19.41 sec) and 19.81 sec for passages (range = 18.31–20.59 sec). The average duration of the Canadian speaker's stimuli was 17.33 sec for target word lists (range = 16.85–17.80 sec) and 20.24 sec for passages (range = 18.77–21.53 sec).

An important consideration is how the speakers used in this work compare with those in the cross-accent experiments of Schmale and Seidl (2009). As noted earlier, the 9-month-olds' failure to recognize words across a native and a Spanish-accented speaker in Schmale and Seidl may have been owing to the accents varying on several suprasegmental and subphonemic dimensions. In contrast, the speakers used here were predicted to deviate primarily on vowel implementation. Thus, an examination of acoustic and perceptual differences between these speakers increases our understanding of the type of variation present in these stimuli, and may shed light on the causes of the 9-month-olds' failure in previous work.

Acoustic measurements and analyses of variance (ANOVAs) with F1 and F2 in /ae/ and /ɪ/ as dependent measures and talker (North Midland-American speaker ["MidW"], and either Spanish-accented speaker ("Span") or Southern Ontario Canadian speaker ["Can"]) support the prediction that talkers would differ on vowel implementation, see Figure 1, particularly with respect to the backing of /ae/ by the Canadian speaker.²

These dialectal accents were chosen because they should diverge minimally, unlike in nonnative speech, which should diverge at other levels (including general characteristics, such as fluency, and subphonemic characteristics, such as coarticulation). This claim is supported by an investigation of the rate of speech, voicing, and coarticulation of the three speakers, which show that the MidW and Can speakers differ less than the MidW and Span speakers, as evident in Figure 2. First, nonnative speakers lack the fluency

²First and second formant frequencies (F1, F2) were measured at the vowel midpoint of the strong syllable in all target vowels using a Praat script (Boersma & Weenik, 2005; four Can tokens had to be excluded for nonmodal phonation). Can and MidW differ in F1 and F2 for /ae/ (F1: $F[1, 114] = 5.1, p < .03$; F2: $F[1, 106] = 7.1, p < .01$), but only in F2 for /ɪ/ (F1: $F[1, 38] = 2.4, p > .12$; F2: $F[1, 34] = 7.88, p < .01$). Span and MidW differ in both correlates for /ae/ (F1: $F[1, 118] = 5.66, p < .02$; F2: $F[1, 118] = 10.18, p < .002$); but only in F1 for /ɪ/ (F1: $F[1, 38] = 12.96, p < .001$; F2: $F[1, 38] = .07$). For the speech rate calculations, the duration of each word or sentence was divided by its number of syllables, including both passages (P) and words in isolation (W). For VOT, the onset of the release of the /k/ and the onset of vibration were hand-tagged in all tokens of *kingdom* and *candle*. Finally, coarticulation was assessed through F1 measured at .75 of the duration of /ae/ in *hamlet* and *candle*, the two targets that differ in a single feature.

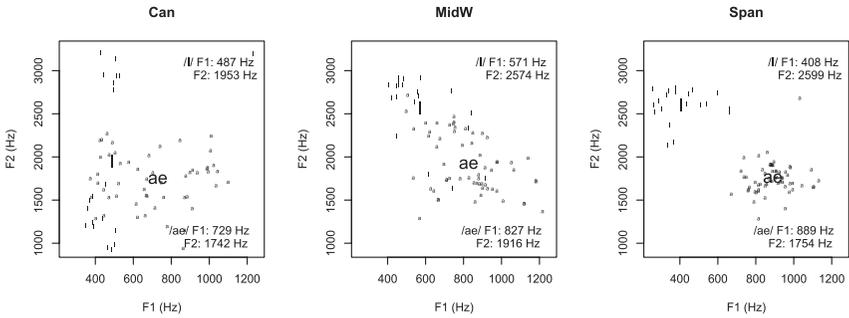


Figure 1 Vowel differences in /ae/ and /I/, across three speakers (North Midland-American speaker [MidW]) in the center, for ease of comparison; Southern Ontario Canadian speaker [Can] on the left, and Spanish-accented speaker [Span] on the right). In each panel, the vowels uttered by the relevant speaker are plotted with /ae/ denoted by “a” and /I/ by “I”. Average F1 and F2 values for each of these two categories are plotted in a larger font, as well as reported within each panel.

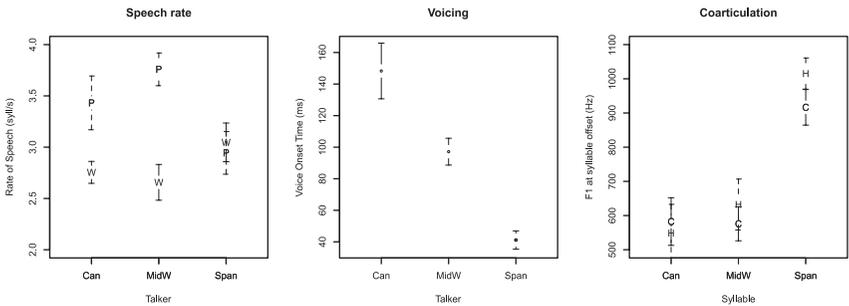


Figure 2 Speech rate (left pane, measured in syllables per second), voicing (middle pane, measured in voice onset time), and coarticulation (right pane) across three speakers (North Midland-American [MidW], Southern Ontario Canadian [Can], and Spanish-accented [Span]).

that characterizes native speakers, which affects global characteristics, including speech rate (although individual variation exists; naturally, a comparison with someone who stutters would not reveal this native advantage). For example, Span exhibited a relatively constant speech rate, whereas the native speakers differ less from each other by talking much slower when uttering words in isolation (I) than within passages (P); ANOVAs with rate as outcome and talker (Midwestern and either Canadian or Spanish) and type (passage, isolation) as factors confirm that the interaction talker \times type is much larger in the MidW-Span comparison, $F(1, 156) = 32.01$ for MidW-Span; 5.34 for MidW-Can. As for consonants, the Spanish-accented

speaker produces the /k/ in *candle* and *kingdom* with a much shorter VOT than either of the English-speaking speakers, and the VOT differs more, $F(1, 78) = 120.72$, than in the comparison among the native talkers, $F(1, 78) = 27.87$. Finally, nonnative speakers differ in subphonemic, coarticulatory, and allophonic patterns (Speeter Beddor, Harnsberger, & Lindemann, 2002) in manners that affect adult listeners' word segmentation (Mondini & Miller, 2004). This is evident in all vowels; our example compares the first formant (F1) location at .75 of the duration of the vowel /ae/ in *hamlet* and *candle*. Although there is no effect of word (*hamlet*, *candle*) on F1 for the native speakers (both $p > .19$), the Spanish-accented speaker produces different F1s depending on the word, $F(1, 38) = 8.9$, $p < .005$, either because these sounds are coarticulated more in Spanish or because the slower movements involved in the production of nonnative sounds affects coarticulation. In addition, findings from a listening experiment provided perceptual evidence that stimuli produced by Can and MidW are more similar as compared with MidW-Span.³

Results and discussion

A repeated-measures ANOVA with average looking time as dependent measure, age group (younger, older), condition (kingdom/hamlet, candle/raptor), and order (American test, Canadian test) as factors, and familiarity (familiar, unfamiliar) as repeated measures revealed a main effect of familiarity, $F(1, 44) = 10.88$, $p = .002$, main effect of order, $F(1, 44) = 8.41$, $p = .005$, significant interaction between age group and familiarity, $F(1, 44) = 4.55$, $p = .04$, and no other significant interactions, $F(1, 44) < .18$. Follow-up paired, two-tailed comparisons of looking time averaged across blocks revealed that familiar and unfamiliar trials differed significantly in the older age group, $t(1, 23) = 3.77$, $p = .001$, but not in the younger group, $t(1, 23) = 0.88$, $p = .39$, as shown in Figure 3. The main effect of order emerges because both groups showed higher looking times when tested with the American speaker. As evidenced by the lack of interac-

³A listening experiment was conducted to determine adults' ratings of the speaker(s) perceived as most different. In each trial, participants heard three tokens of the same target word (interstimulus interval = 500 msec), each spoken by a different talker (the two speakers used in this work and the Spanish-accented speaker used in the cross-accent generalization experiments in Schmale & Seidl, 2009). Participants then responded to which of the three sounded the most different. Listeners were not explicitly instructed to detect accents, or to rate their strength. Fourteen adults responded to a total of 24 different stimuli (four target words *hamlet*, *candle*, *kingdom*, *raptor*, in six different orders, to counterbalance talker order), each of which was presented three times. Participants rated the target words produced by Span as the most different of the three words 85% of the time, compared with 9% for MidW and 6% for Can.

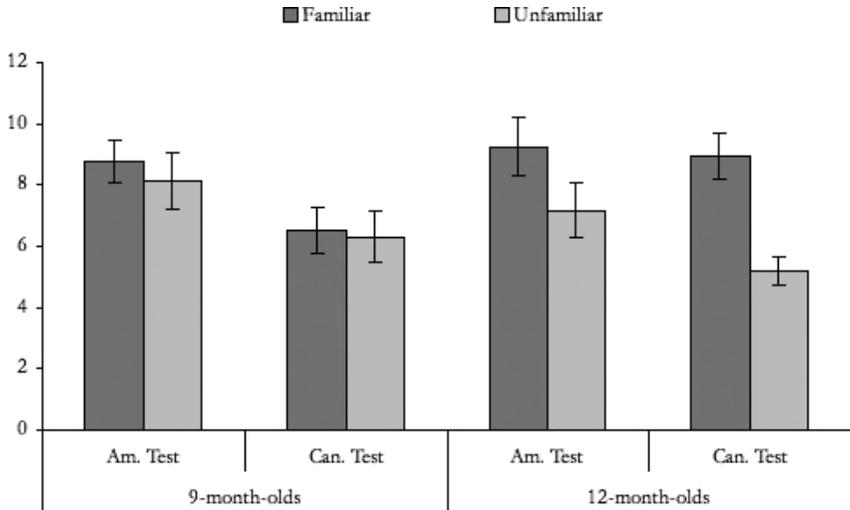


Figure 3 Average looking times in seconds by age group, order (speaker in test), and familiarity (error bars represent standard error).

tion with order and familiarity, the pattern of looking remained the same in both the novel and familiar test trials, and only 12-month-olds showed a significant difference in looking time between passages containing familiar and novel words.

These findings suggest that 12-month-olds successfully recognized words in the face of variation in dialectal accent, as evidenced by the significant preference for test passages containing familiar words. In contrast, 9-month-olds showed no preference, suggesting that dialectal differences were large enough to impede word recognition.

GENERAL DISCUSSION

This work extends the finding that infants are sensitive to dialect differences by showing the functional relevance of this sensitivity for word recognition in 9-month-olds. The 9-month-olds' poor performance could be attributed to their lack of familiarity with dialectal accents, perhaps complicating the representation of words in unfamiliar speech streams. However, this hypothesis could not account for results showing that English-learning 9-month-olds are successful in recognizing words in foreign-accented speech (produced by a single Spanish-accented speaker; Schmale & Seidl, 2009, Experiment 4), a foreign language (Houston, Jusczyk, Kuijpers, Coolen, & Cutler, 2000), and

even synthetic speech (Johnson & Tyler, 2009; Thiessen & Saffran, 2003). Instead, it is more likely that 9-month-olds can perform pattern-matching, but fail because they lack more abstract representations that encompass irrelevant phonetic variability.

In interpreting these findings, an important consideration is the particular type of variation responsible for the 9-month-olds' failure. Based on acoustic and perceptual evidence, the American and Canadian speakers only appear to deviate markedly on vowel implementation (and not on fluency, subphonemic, or consonantal dimensions). It is reasonable to conclude that 9-month-olds' failure is because of attention to linguistically irrelevant vowel variation across dialectal accents. Moreover, this attention to irrelevant vowel variation may have played an important role in 9-month-olds' inability to recognize words across accents in Schmale and Seidl (2009). Therefore, this work provides further evidence for the relative rigidity of infants' early word representations: words varying slightly in vowel implementation may escape 9-month-olds' recognition.

The developmental change documented for word recognition in the face of gender and affect variation (Houston & Jusczyk, 2000; Singh et al., 2004) could be explained through *semantic constancy*, as older infants are more likely to have accumulated experience hearing an object talked about by male and female speakers, in different affects. Additionally, *exposure to specific dialectal accents* influences infants' listening preference. After exposure to American accents, Australian 6-month-olds do not show a preference for Australian English, whereas American infants *do* show a preference for their native dialect (Kitamura et al., 2006). In contrast, neither semantic constancy nor exposure to Canadian dialectal accents provides a compelling explanation for these results. Taken together with the findings of Schmale and Seidl (2009), an alternative account is that increased language exposure in general leads to more robust representations, through which infants may accommodate irrelevant variation.

One possibility is that infants' representations become generally laxer over time, such that even an inexact match activates word representations. Alternatively, infants do not simply come to accept variation along *any* dimension, but rather disregard variation along specific dimensions they have identified as highly variable across speakers. Training studies with adults (e.g., Lively, Logan, & Pisoni, 1993) and infants (e.g., Rost & McMurray, 2009) provide indirect evidence for the latter possibility, as learners come to identify linguistically relevant dimensions through exposure to more speakers. For example, slight vowel variation could be liable to being ignored, as vowels are inherently more variable than consonants across speakers, even within a homogeneous linguistic community. This possibility is bolstered by research on reduced attention to vowels in word

recognition (Cutler, Sebastián-Gallés, Soler-Vilageliu, & van Ooijen, 2000; Havy & Nazzi, 2009). Future work investigating the impact of variation in consonantal implementation would shed light on this matter.

Overall, these results suggest that, by 12 months, children can segment words from continuous speech across minimally different dialectal accents. Nonetheless, the learning task is not over, as toddlers may still have difficulty with this type of variation when recognizing or learning lexical items. Indeed, a recent article by Best et al. (2009) reports that toddlers do not show a preference for high-frequency words spoken in an unfamiliar dialect until 19 months, and cross-accent word learning may not be possible until 30 months (Schmale, Hollich, & Seidl, 2009). Importantly, these findings underline the importance of piecing together infants' representations along different stages of language development (e.g., Werker & Curtin, 2005).

In sum, this work is the first to demonstrate that in word segmentation from continuous speech, even minimal, regionally driven vowel variation can only be processed by older, more experienced infants. Although future research should explore the relative sensitivity of these processing abilities, these findings make an important contribution to our understanding of how infants learn to equate dissimilar instances of the same word, and approximate the abilities of adults in weighting irrelevant phonetic variation. Thus, this investigation affords an invaluable opportunity to approach the complex question of how infants' early word forms are represented.

REFERENCES

- 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.
- Boersma, P., & Weenik, D. (2005). *Praat: Doing phonetics by computer (version 5.0.09)* [computer program]. Retrieved February 17, 2008, from <http://www.praat.org/>
- Clarke, S., Elms, F., & Youssef, A. (1995). The third dialect of English: Some Canadian evidence. *Language Variation and Change, 7*, 209–228.
- Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & van Ooijen, B. (2000). Constraints of vowels and consonants in lexical selection: Cross-linguistic comparisons. *Memory & Cognition, 28*, 746–755.
- 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.
- Havy, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy, 14*, 439–456.
- Houston, D. M. (2000). The role of talker variability in infant word representations (Doctoral Dissertation, Johns Hopkins University, 1999). *Dissertation Abstracts International, 60*, 5802.

- 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*, 1570–1582.
- Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Both Dutch- and English-learning 9-month-olds segment Dutch words from fluent speech. *Psychonomic Bulletin and Review*, *7*, 504–509.
- Johnson, E. K., & Tyler, M.. (2009). Testing the limits of statistical learning for word segmentation. *Developmental Science*, *13*, 339–345.
- Jongman, A., & Wade, T. (2007). Acoustic variability and perceptual learning. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 135–150). Amsterdam: John Benjamins Publishing.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*, 159–207.
- Jusczyk, P. W., Pisoni, D. B., & Mullenix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, *43*, 253–291.
- Kitamura, C., Panneton, R., Deihl, M., & Notley, A. (2006). Attuning to the native dialect: When more means less. In *Proceedings of the Eleventh Australian International Conference on Speech Science & Technology* (pp. 124–129). Auckland, New Zealand: Australasian Speech.
- Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008). First impressions and last resorts: How listeners adjust to speaker variability. *Psychological Science*, *19*, 332–338.
- Krentz, U. C., & Corina, D. P. (2008). Preference for language in early infancy: The human language bias is not speech specific. *Developmental Science*, *11*, 1–9.
- Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, *66*, 1668–1679.
- Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, *6*, 263–285.
- Labov, W., Ash, S., & Boberg, C. (2006). *Atlas of North American English*. Berlin & New York: Mouton de Gruyter.
- Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, *94*, 1242–1255.
- Mattys, S. L., & Jusczyk, P. W. (2001). Do infants segment words or continuous recurring patterns? *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 644–655.
- Mondini, M., & Miller, J. L. (2004). *Perceiving non-native speech: Word segmentation*. Poster presented at the 147th meeting of the Acoustical Society of America, New York, New York.
- 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., Dille, L., Jusczyk, A. M., Shattuck-Hufnagel, S., & Jusczyk, P. W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*, *48*, 279–298.
- 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.
- Remez, R. E., Fellowes, J. M., & Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 651–666.
- Remez, R. E., Van Dyk, J. L., Fellowes, J. M., & Rubin, P. E. (1998). On the perception of qualitative and phonetic similarities of voice. In P. K. Kuhl & L. A. Crum (Eds.), *Proceedings*

- of the 16th International Congress on Acoustics and the 135th Meeting of the Acoustical Society of America (pp. 2063–2064). New York: Acoustical Society of America.
- Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science, 12*, 339–349.
- Schmale, R., Hollich, G. J., & Seidl, A. (2009). *Contending with foreign accent variability in early lexical acquisition*. Manuscript submitted for publication.
- Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science, 12*, 583–601.
- Seidl, A., & Johnson, E. K. (2006). Infant word segmentation revisited: Edge alignment target extraction. *Developmental Science, 9*, 566–574.
- Seidl, A., & Johnson, E. K. (2008). Boundary alignment enables 11-month-olds to segment vowel initial words from speech. *Journal of Child Language, 35*, 1–24.
- Shah, A. P. (2004). Production and perceptual correlates of Spanish-accented English. *Proceedings of the MIT conference: From sound to sense: 50+ years of discoveries in speech communication*. Cambridge, MA: MIT.
- Sheffert, S. M., Pisoni, D. B., Fellowes, J. M., & Remez, R. E. (2002). Learning to recognize talkers from natural, sinewave, and reversed speech samples. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 1447–1469.
- 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.
- Speeter Beddor, P., Harnsberger, J. D., & Lindemann, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates. *Journal of Phonetics, 30*, 591–627.
- Steinlen, A. (2005). *The influence of consonants on non-native vowel production: A cross-linguistic study*. G. Narr: Tübingen.
- Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*, 487–501.
- Thiessen, E., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word segmentation by 7- to 9-month-old infants. *Developmental Psychology, 39*, 706–716.
- Wannemacher, C. A. (1995). The relationship between processing of talker voice and phonetic information in speech perception (Doctoral Dissertation, University of New York at Buffalo, 1995). *Dissertation Abstracts International, 56*, 2488.
- Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.
- Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development, 1*, 197–234.