

BRIEF RESEARCH REPORT

Contending with foreign accent in early word learning*

RACHEL SCHMALE

North Park University, USA

GEORGE HOLLICH AND AMANDA SEIDL

Purdue University, USA

*(Received 11 February 2010 – Revised 13 July 2010 – Accepted 4 October 2010 –
First published online 11 February 2011)*

ABSTRACT

By their second birthday, children are beginning to map meaning to form with relative ease. One challenge for these developing abilities is separating information relevant to word identity (i.e. phonemic information) from irrelevant information (e.g. voice and foreign accent). Nevertheless, little is known about toddlers' abilities to ignore irrelevant phonetic detail when faced with the demanding task of word learning. In an experiment with English-learning toddlers, we examined the impact of foreign accent on word learning. Findings revealed that while toddlers aged 2;6 successfully generalized newly learned words spoken by a Spanish-accented speaker and a native English speaker, success of those aged 2;0 was restricted. Specifically, toddlers aged 2;0 failed to generalize words when trained by the native English speaker and tested by the Spanish-accented speaker. Data suggest that exposure to foreign accent in training may promote generalization of newly learned forms. These findings are considered in the context of developmental changes in early word representations.

INTRODUCTION

In a linguistically diverse society it is likely that young children will encounter foreign-accented speech. This speech typically deviates from the native dialect in several ways (e.g. modifications to subphonemic

[*] Address for correspondence: Rachael Schmale, North Park University – Psychology, 3225 W. Foster Ave. Box 16, Chicago, Illinois 60625, United States. e-mail: rschmale@northpark.edu

and suprasegmental properties such as voice onset time duration, vowel formants and syllable duration; Shah, 2004), and it is uncertain how these departures impact early word learning. For example, a non-native English speaker whose native language does not have both /ɔ/ and /o/ may produce a word with these target sounds halfway between these two categories (e.g. producing the target *ball* somewhere between *ball* and *bowl*). Thus, this word may be ambiguous to a young child just beginning to map meaning to form. Unfortunately, little research has examined the impact of foreign accent on early word-learning abilities.

Nonetheless, it is clear that young children are particularly sensitive to the relevant/phonemic contrasts that are present in their native language, well before they begin segmenting and learning words (e.g. Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992). One might predict that this language-specific phonemic sensitivity would serve an important function in word learning, so as to highlight the phonological distinctions of the target language. However, while young children can discriminate words that vary in only one place feature very early on, interpreting those changes as relevant to word identity entails additional difficulties (Stager & Werker, 1997; Werker, Fennell, Corcoran & Stager, 2002). For example, even though toddlers aged 1;6 can detect phonetic mispronunciations (e.g. *car* vs. *gar*) when assigning meaning to words, they fail to appropriately interpret those mispronunciations as referents for novel objects (Mani & Plunkett, 2007; Swingley & Aslin, 2007; White & Morgan, 2008; see also Swingley & Aslin, 2000, 2002, for other work on how mispronunciation affects familiar word recognition). Toddlers succeed in this task, however, when the words are not phonological neighbors, when the differences between them are more salient, or when task demands are reduced (Ballem & Plunkett, 2005; Nazzi, 2005; Swingley & Aslin, 2007; Thiessen, 2007). In short, this work suggests that much of early word learning depends on appropriately interpreting fine phonetic detail, and processing load seems to modulate this ability.

Given that accents (dialectal and non-native) are characterized by deviation from native pronunciation norms, they may be comparable to mispronunciations. Thus, it is possible that toddlers will exhibit similar difficulties recognizing words produced in unfamiliar accents. Indeed, recent work demonstrates that toddlers prefer high- to low-frequency words spoken in an unfamiliar dialect at age 1;7, but not at age 1;3 (Best, Tyler, Gooding, Orlando & Quann, 2009), whereas a preference for high frequency words is evident at age 0;11 when tested with a familiar dialect (Hallé & Boysson-Bardies, 1994). Nevertheless, it remains unclear how unfamiliar accents might impact the learning of NOVEL words, as it is likely that learning words is more demanding than recognizing high-frequency familiar ones. Thus, in order to succeed, toddlers must first recognize

phonological structure in non-standard phonetic instantiations and then relate this novel form to a novel referent.

We explored this question by testing the impact of foreign-accented speech on early word-learning abilities. Specifically, we tested toddlers abilities at 2;0 and 2;6 to generalize words learned in a training period to a test period when the speakers had different accents. We predicted that younger toddlers might experience more difficulty appropriately interpreting relevant and irrelevant phonetic information than older toddlers. For example, toddlers aged 2;0 may not be able to rapidly process non-standard phonetic instantiations, which would lead to an inability to generalize the structure for the same novel word produced by talkers who pronounce them differently. In contrast, it is plausible that older toddlers might be more successful because they have more experience encoding a variety of word forms and relating sound to meaning. For example, Quam and Swingley (2010) found that toddlers aged 2;6 can successfully learn a novel word by disregarding irrelevant pitch variation. This ability to appropriately interpret relevant and irrelevant variation may promote better learning of dissimilar-sounding words.

EXPERIMENT

English-learning toddlers aged 2;0 and 2;6 were tested on their abilities to learn two novel words when trained by a native English speaker and tested by a speaker of Spanish-accented English (and vice versa). If toddlers can successfully map novel meanings to novel words, despite a change in speaker and foreign accent, this may demonstrate that they can successfully extract and encode the relevant, identifying features of words in the face of phonetic deviation not relevant to word identity.

METHOD

Participants

Thirty-two English-learning toddlers aged 2;0 (M age = 23.97 months; range = 23.63–24.67 months; 9 males) and twenty-four aged 2;6 (M age = 29.87 months; range = 29.47–30.67 months; 12 males) raised in the Midwest participated. Ten additional toddlers aged 2;0 were excluded (5 due to fussing, 3 owing to experimenter error, 1 because of parental interference and 1 due to more than 30% exposure to another language). Eight additional toddlers aged 2;6 were excluded (4 due to fussing, 3 owing to experimenter error and 1 because of prematurity). All included toddlers' parents reported normal hearing and full-term status. Further, as measured by the short form of the MacArthur-Bates Communicative Development Inventory (CDI): Words and Sentences (Fenson, Dale, Reznick, Bates,

Thal & Pethick, 1994), toddlers aged 2;0 averaged 53 words in their productive vocabulary (range=2–99 words) and those aged 2;6 averaged 76 words (range=6–100 words).

Auditory stimuli

The auditory stimuli consisted of four novel words (*neech*, *moof*, *feem*, *choon*), presented within the carrier phrase format in all training and test trials: “Do you see a ____? Look, it’s a ____! A ____!”, in an exaggerated, infant-directed register. The words consisted of vowels and consonants in both English and Spanish phonological inventories. To avoid the possibility that differences in voice onset time (VOT) between native and foreign-accented speech might affect toddlers’ learning, the novel words did not contain any stop consonants. The duration of three carrier phrases was 5 s for each speaker. To allow toddlers to look at both objects on the screen before hearing the object labels, 1 s of silence was added to the beginning of all sound files, minimizing possible object preference at visual onset.

Employing the same method as previous work on cross-talker word recognition (e.g. Houston & Juszyk, 2000; Schmale & Seidl, 2009), the two speakers used in this work were selected from a sample of ten different female speakers (5 native speakers of North Midland American English, 5 native speakers of Spanish who spoke English with an intelligible Spanish accent). Since there is little consensus on which acoustic dimensions are most important in voice and accent perception (e.g. Gelfer, 1993; Houston, 2000), and acoustic measurements do not always accurately represent speaker similarity, adult listener ratings were used as the basis for speaker selection.

Because voice and accent characteristics cannot be judged independently in natural speech (Remez, Fellowes & Rubin, 1997; Remez, Van Dyk, Fellowes & Rubin, 1998), adult listeners rated the similarity of all speaker pairs in natural speech and sinewave speech (which eliminates voice characteristics, while retaining only accentual information; Krentz & Corina, 2008). By subtracting sinewave (accent-based) similarity ratings from natural ratings, speakers with the most similar voices were selected through multidimensional scaling analyses (MDS; Houston & Juszyk, 2000; Schmale & Seidl, 2009; see also Sheffert, Pisoni, Fellowes & Remez, 2002). MDS yields speaker similarity by configuring average dissimilarity between each speaker pair. Thus, two speakers are determined to have SIMILAR voices if a small change exists between the average dissimilarities in natural and sinewave speech (e.g. 0.53 and 0.54, respectively) and to have DISSIMILAR voices if a large change exists (e.g. 11.03 and 6.08, respectively).

Using this method, two female speakers were selected to produce the auditory stimuli. The native English speaker was from Indianapolis,

Indiana and spoke the toddlers' ambient dialect (North Midland American English). The foreign-accented speaker was a university-educated native speaker of Spanish from the Dominican Republic who spoke English with an intelligible Spanish accent (as judged by adult listeners). These speakers had an average judged dissimilarity rating of 0.53 in natural speech and 0.54 in sinewave speech, indicating that their voices were highly similar as there was almost no change between the ratings in natural and sinewave speech. These speakers were also used in the cross-accent word recognition studies in Schmale and Seidl (2009). Speaker recordings were conducted in a double-walled sound-proofed booth with an Audio-Technica 100HE Hypercardioid dynamic microphone. Stimuli were digitized at 44 100 Hz, normalized to an approximate amplitude of 70 dB, and matched for average and maximum pitch.

Visual stimuli

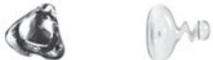
The visual stimuli consisted of pictures of four novel objects, which were constructed of different colors of glass (see Table 1 for pictures). Objects were paired with auditory stimuli and assembled into movies.

Design

The experiment consisted of four blocks, each six trials in length (to obtain a more reliable measure of word learning, Blocks 1 and 2 were presented twice in sequential order). To bolster attention, an attention-getting stimulus played between each trial. All blocks followed the same format: one Salience trial, three Training trials and two Test trials. In the Salience trial, two test objects were presented on right and left sides of the video display, in silence for the duration of a 6-second trial (e.g. orange object on left, green object on right). The purpose of this trial was to familiarize participants to the novel objects, so as to prevent a novelty preference from emerging to non-trained objects. In the Training trials, one novel object was presented on the center of the video display, which was paired with the carrier phrase format for that label-object pairing (e.g. *feem* + green object). There were two types of test trials: Trained Test and Novel Test. In these trials, two objects were presented on right and left sides of the video display. In Trained Test trials, toddlers were presented with the previously trained label-object pairing (e.g. *feem* + green object). In Novel Test trials, toddlers were presented with a label-object pairing that had NOT been presented previously (e.g. *choon* + orange object), which functioned as a control for possible familiarity preference to the trained object.

Toddlers in each age group were tested on the same novel words and objects, but were randomly assigned to four Conditions that were

TABLE 1. *Example of label-object pairings and trial order* in first half of experiment*

| Block | Trial | Label | Video |
|-------|---------------|--------------|---|
| 1 | Saliency | |  |
| | Training (×3) | <i>Neech</i> |  |
| | Trained Test | <i>Neech</i> |  |
| | Novel Test | <i>Moof</i> |  |
| 2 | Saliency | |  |
| | Training (×3) | <i>Feem</i> |  |
| | Trained Test | <i>Feem</i> |  |
| | Novel Test | <i>Choon</i> |  |

* Order of test trials and label-object pairings were counterbalanced across experimental orders.

counterbalanced for test trial order and label-object pairings (see Table 2 for details). Toddlers were also randomly assigned to two Generalization Orders (Native-to-Accented; Accented-to-Native) that differed according to which speaker produced the stimuli in Training and Test. In the Native-to-Accented Generalization Order, the native English speaker produced the stimuli in Training, and the Spanish-accented speaker produced the stimuli in Test. Alternatively, in the Accented-to-Native Generalization Order, the Spanish-accented speaker produced the stimuli in Training and the native English speaker produced the stimuli in Test (see Table 2). Equal numbers of participants within each age group were assigned to each Condition and Generalization Order.

Apparatus

Toddlers were tested in a three-sided booth constructed out of three wooden panels, approximately 6 feet high. A camcorder was mounted to the back of the front panel and two speakers were mounted on top of the booth,

TABLE 2. *Experimental design conditions*

| Condition | Trained words | Test trial order | Generalization order |
|-----------|--------------------|------------------|----------------------|
| 1 | <i>Neech, Feem</i> | Trained Test | Native-to-Accented |
| 2 | <i>Neech, Feem</i> | Trained Test | Accented-to-Native |
| 3 | <i>Moof, Choon</i> | Novel Test | Native-to-Accented |
| 4 | <i>Moof, Choon</i> | Novel Test | Accented-to-Native |

102 cm apart. A hole was cut in the front panel to allow the experimenter to videotape the toddlers' eye movements while they watched the experiment on the video display (102 cm \times 137 cm) on the front panel of the booth. The video display was projected by an InFocus X3 LCD projector.

Procedure

In this version of the Preferential Looking Procedure (PLP; Fagan, 1971; Spelke, 1979), the toddlers sat on the lap of a caregiver in the middle of the testing booth facing the video display. An experimenter conducted the experiment on a computer hidden behind the front panel and recorded the toddlers' looking patterns via camcorder. In order to prevent caregivers from inadvertently influencing toddlers' looking, they wore opaque-coated sunglasses or closed their eyes for the duration of the experiment.

Coding

The participants were videotaped for the duration of the experiment and videos were digitized for coding. The durations of toddlers' eye movements to the center, left or right of the video display were then coded off-line, frame-by-frame by highly trained coders. In order to prevent the coders from inadvertently influencing the results, they were blind to the location of the target object. Toddlers' looking times to the target objects were subsequently used as a measure of their success at learning the labels. If the toddlers looked longer at the labeled object when that label was requested, this pattern of results indicated that they learned the new word. One coder coded all of the data, while another coder coded 25%. The intercoder agreement was 99%.

RESULTS

Following Swingley and Aslin (2000; 2002; 2007), toddlers' mean looking time (LT) to the target and non-target objects in each test trial were calculated over a period that began approximately 367 ms after the onset of the first target word and ended 2 s later. In order to achieve a measure of overall

learning across Trained and Novel Test trials, raw LT were converted to difference scores within Trained and Novel Test trials. Thus, in Trained Test trials, LT to the non-target object (non-trained label-object pairing) were subtracted from LT to the target object (trained label-object pairing). Similarly, in Novel test trials, LT to the non-target object (trained label-object pairing) were subtracted from LT to the target object (non-trained label-object pairing). Because the target object differs in Trained and Novel test trials, difference scores were subsequently calculated across the above LT differences. This calculation therefore gives a single measure of the change in LT to target and non-target objects across test trials. Thus, it represents the degree to which children look longer at the target objects when they are labeled versus when they are not, which provides a measure of overall learning.

To explore differences in performance between toddlers aged 2;0 and 2;6, LT difference scores for each group were analyzed independently. For those aged 2;0, a Repeated Measures ANOVA with Word Label (labeled, not labeled) as repeated measure and Condition (1, 2) and Generalization Order [Native-to-Accented (1), Accented-to-Native (2)] as factors was conducted. This analysis revealed no main effects of Condition ($F(1, 28) = 0.08$, $\eta^2 = 0.05$) or Generalization Order ($F(1, 28) = 2.08$, $p = 0.16$, $\eta^2 = 0.26$), but a main effect of Word Label ($F(1, 28) = 5.23$, $p = 0.03$, $\eta^2 = 0.40$). There was no significant 2-way interaction of Word Label and Condition ($F(1, 28) = 0.01$, $\eta^2 = 0.02$) or 3-way interaction ($F(1, 28) = 1.23$, $p = 0.28$, $\eta^2 = 0.21$). However, there was a significant interaction of Generalization Order and Word Label ($F(1, 28) = 5.94$, $p = 0.02$, $\eta^2 = 0.42$).

To examine the interaction of Generalization Order and Word Label, post-hoc *t*-tests comparing LT difference scores in both Generalization Orders were conducted. The analyses revealed a significant difference in LTs in Generalization Order 2 (Accented-to-Native; $t(15) = -3.17$, $p = 0.006$, $\eta^2 = 0.54$), but not in Generalization Order 1 (Native-to-Accented; $t(15) = 0.12$, $p = 0.91$, $\eta^2 = 0.02$). This indicates that the main effect of Word Label resulted from toddlers' success in Generalization Order 2. These results are illustrated in Figure 1 and raw mean LTs are presented in Table 3. These findings demonstrate that the toddlers aged 2;0 could generalize novel words from training to test when trained by the Spanish-accented speaker, but not when trained by the native speaker.

For toddlers aged 2;6, the same Repeated Measures ANOVA was conducted. This analysis revealed no significant main effects of Condition ($F(1, 20) = 0.17$, $\eta^2 = 0.09$) or Generalization Order ($F(1, 20) = 0.23$, $\eta^2 = 0.11$), but a significant main effect of Word Label ($F(1, 20) = 11.51$, $p = 0.003$, $\eta^2 = 0.60$). There were no significant 2-way interactions ($F_s(1, 20) < 2.76$, $p_s > 0.11$, $\eta^2 < 0.35$) and no significant 3-way interaction ($F(1, 20) = 1.48$, $p = 0.24$, $\eta^2 = 0.26$). The results are once again summarized in Figure 1 and

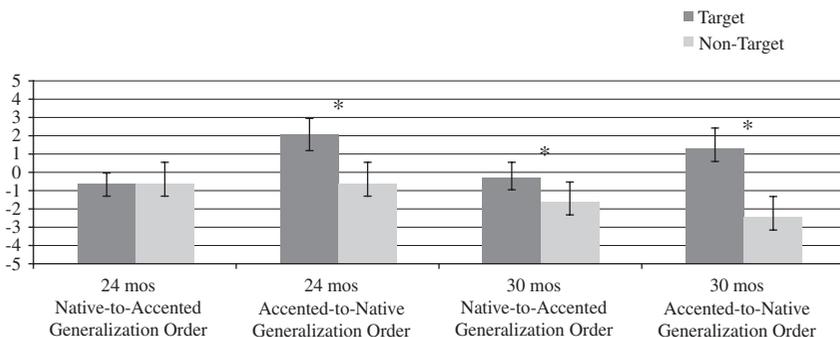


Fig. 1. Mean looking time difference scores (in seconds) to target and non-target objects with error bars showing standard error; * = $p < 0.05$.

TABLE 3. Raw mean looking times (in seconds) to target and non-target objects in all Trained and Novel Test trials for toddlers aged 2;0 and 2;6

| 2;0 | | |
|--------------|---|------------|
| | Native-to-Accented Generalization Order | |
| | Target | Non-Target |
| Trained Test | 0.74 | 0.89 |
| Novel Test | 0.88 | 0.75 |
| | Accented-to-Native Generalization Order | |
| | Target | Non-Target |
| Trained Test | 1.12 | 0.59 |
| Novel Test | 0.90 | 0.77 |
| 2;6 | | |
| | Target | Non-Target |
| Trained Test | 0.93 | 0.80 |
| Novel Test | 1.05 | 0.55 |

raw mean LTs are shown in Table 3. These results suggest that toddlers aged 2;6 could successfully generalize novel words from training to test when produced by speakers with different accents, regardless of the accent of the speaker in training.

To directly compare the abilities of toddlers aged 2;0 and 2;6 to generalize novel words from training to test when trained by a native speaker, a repeated-measures ANOVA with Word Label (labeled, not labeled) as repeated measure and Generalization Order and Age (younger, older) as factors was conducted. This analysis revealed a significant main effect of Word Label ($F(1, 52) = 17.59$, $p = 0.0001$, $\eta^2 = 0.51$) and a significant interaction between Word Label and Generalization Order

($F(1, 52) = 8.20$, $p = 0.006$, $\eta^2 = 0.37$) but no other main effects ($F_s(1, 52) < 1.79$, $p_s > 0.19$, $\eta^2 < 0.18$) or interactions ($F_s(1, 52) < 1.92$, $p_s > 0.17$, $\eta^2 = 0.19$). The lack of a significant main effect of Age is likely due to the success of the toddlers aged 2;0 in the Accented-to-Native Generalization Order. Simple regressions were also conducted to investigate how well CDI would predict the difference scores, but revealed no significant effects for the toddlers aged 2;0 ($R^2 = 0.008$, $F(1, 30) = 0.25$) or those aged 2;6 ($R^2 = 0.02$, $F(1, 22) = 0.35$).

GENERAL DISCUSSION

The experiment reported in this paper provides evidence that at age 2;0, toddlers experience some difficulty in generalizing dissimilar instances of the same novel words from training to test, when produced by two female speakers with different accents. In particular, when trained on novel words by a native English speaker, toddlers aged 2;0 are unable to then recognize the same words when produced by a Spanish-accented speaker. One account for this failure is that the speakers' productions are dissimilar, which could hinder token generalization and learning. However, notice that toddlers aged 2;0 succeed when trained by the accented speaker, lending little support to an account based on differences in phonetic instantiation, which is necessarily symmetrical. Instead, this asymmetrical pattern of results could have been driven by a speaker-specific effect, either a preference for the foreign-accented speaker, or an inability to learn from the native talker's training. Neither explanation is borne out by our data. First, overall fixation times during training and test were not significantly different between Generalization Orders. Second, in a control experiment with the same design, apparatus and procedure, sixteen toddlers aged 2;0 successfully generalized the same novel words from training to test, when produced by two different native English speakers ($F(1, 15) = 11.48$, $p = 0.004$, $\eta^2 = 0.66$), one of which was the native speaker used in the present work. This demonstrates that toddlers CAN successfully learn words when trained by this particular speaker, making it very unlikely that a speaker-specific effect in the present work impeded subsequent learning.

Another plausible explanation is that exposure to phonetic variability leads to more robust representations by promoting broader lexical categories (e.g. Houston, 2000; Lively, Logan & Pisoni, 1993; Rost & McMurray, 2009). For example, the Spanish-accented talker produced non-standard pronunciations of English sounds, such that the phonological structure involved would be more deviant or variable with respect to stored structure. Further, when compared with native speakers, foreign-accented speakers demonstrate a high level of variability in their speech, particularly in their utilization of vowel space (Jongman & Wade, 2007). This may

bolster learning, as variability present in training may facilitate token generalization by allowing listeners to disregard information identified as highly variable across tokens (e.g. Rost & McMurray, 2009). Thus, word representations may better accommodate phonetic variation after exposure to non-contrastive information in training, offering a distinct learning benefit for younger toddlers. On the other hand, being trained with less variable productions of a word might disrupt abstraction across dissimilar instances. So, when trained by the native speaker (whose words likely encompass less irrelevant phonetic variability), toddlers may have fewer opportunities to discover which dimensions are irrelevant to lexical identity. This may hinder generalization of words that deviate markedly from those heard previously, particularly given that toddlers do not know a priori which dimensions are relevant.

In contrast, older toddlers (at 2;6), successfully extract the invariant properties of words among speakers with different accents, regardless of the speaker in training. This suggests that while toddlers aged 2;0 may depend on increased phonetic variability to generalize dissimilar word instances, the representations of those aged 2;6 are more robust, possibly due to greater experience with variability in the input. Alternatively, they may be better able to contend with foreign-accented speech ‘on the fly’, during the test phase. Either interpretation fits well with previous findings, which emphasize the interaction of processing load and task demands on children’s word recognition and learning.

Indeed, at each stage of lexical development, younger children are more vulnerable to irrelevant phonetic information. Moreover, the type of representational access involved in the task interacts with processing load. For example, pattern recognition, lexical access and formation of new lexical items are all likely to elicit different levels of success depending on the sophistication of the learner. Thus, infants succeed in coping with voice and affect in word-to-passage segmentation tasks at 10.5, but not at 7.5 months of age (Houston & Jusczyk, 2000; Singh, 2008; Singh, Morgan & White, 2004). In contrast, it is not until after their first birthday that infants can recognize words across different accents under the same conditions (Schmale, Cristià, Seidl & Johnson, 2010; Schmale & Seidl, 2009). Nevertheless, this problem is not resolved at this point in development either. Once again toddlers experience difficulty in coping with an unfamiliar dialectal accent in lexical access at age 1;3, succeeding in preferring highly familiar words only at age 1;7 (Best *et al.*, 2009), but still struggling at age 2;0 to recognize a newly learned word when pronounced by a foreign-accented speaker. In other words, difficulty in coping with irrelevant phonetic information largely depends on the difficulty of other aspects of the task. It is clear that the developmental trajectory by which young children come to resolve this problem may require a more nuanced

characterization, especially considering that adults still encode information not relevant to word identity in word processing tasks (e.g. Goldinger, Pisoni & Logan, 1991).

In summary, these findings are the first to assess how toddlers accommodate foreign accent when learning new words. Future research will explore whether increased exposure to foreign-accented speech when learning new words facilitates generalization, how much exposure to foreign-accented speech is needed to produce benefits for early learners, and whether children exposed to regular forms of foreign-accented speech are better at disregarding information not relevant to word identity. This work will not only promote a better understanding of young children's interpretation of relevant and irrelevant phonetic information, but may also serve an important role in helping parents to consider the potential benefits of exposure to accented speech.

REFERENCES

- Ballem, K. D. & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language* **32**, 159–73.
- 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**, 539–42.
- Fagan, J. (1971). Infant recognition memory for a series of visual stimuli. *Journal of Experimental Child Psychology* **11**, 244–50.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development* **59**.
- Gelfer, M. P. (1993). A multidimensional scaling study of voice quality in females. *Phonetica* **50**, 15–27.
- Goldinger, S. D., Pisoni, D. B. & Logan, J. S. (1991). On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **17**, 152–62.
- Hallé, P. & de Boysson-Bardies, B. (1994). Emergence of an early lexicon: Infant's recognition of words. *Infant Behavior and Development* **17**, 119–29.
- Houston, D. M. (2000). The role of talker variability in infant word representations. Unpublished doctoral dissertation, Johns Hopkins University, Baltimore.
- 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**, 1570–82.
- Jongman, A. & Wade, T. (2007). Acoustic variability and perceptual learning. In O. S. Bohn and M. J. Munro (eds), *Language experience in second language speech learning*, 135–50. Amsterdam: John Benjamins Publishing Company.
- 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., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. (1992). Language experience alters phonetic perception in infants by 6 months of age. *Science* **255**, 606–608.
- 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**, 1242–55.

- Mani, N. & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language* **57**, 252–72.
- Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition* **98**, 13–30.
- Quam, C. & Swingley, D. (2010). Phonological knowledge guides two-year-olds' and adults' interpretation of salient pitch contours in word learning. *Journal of Memory and Language* **62**, 135–50.
- 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**, 651–66.
- 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*, 2063–64. New York: Acoustical Society of America.
- Rost, G. C. & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science* **12**, 339–49.
- 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*, 79–84. Cambridge, MA: MIT.
- Schmale, R., Cristiã, A., Seidl, A. & Johnson, E. K. (2010). Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy* **15**(6), 650–62.
- Schmale, R. & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science* **12**, 583–601.
- 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**, 1447–69.
- Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition* **106**, 833–70.
- 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**, 173–89.
- Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology* **15**, 626–36.
- Stager, C. L. & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* **388**, 381–82.
- Swingley, D. & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition* **76**, 147–66.
- Swingley, D. & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science* **13**, 480–84.
- Swingley, D. & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology* **54**, 99–132.
- Thiessen, E. D. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language* **56**, 16–34.
- Werker, J. F. & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development* **1**, 197–234.
- Werker, J. F., Fennell, C. T., Corcoran, K. M. & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy* **3**, 1–30.
- White, K. S. & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language* **59**, 114–32.