Research Report

Development of Phonological Constancy

Toddlers' Perception of Native- and Jamaican-Accented Words

Catherine T. Best,[1,2] Michael D. Tyler,[1] Tiffany N. Gooding,[2,3] Corey B. Orlando,[3] and Chelsea A. Quann[3]

[1]MARCS Auditory Laboratories, University of Western Sydney; [2]Haskins Laboratories, New Haven, Connecticut; and [3]Wesleyan University

ABSTRACT—*Efficient word recognition depends on detecting critical phonetic differences among similar-sounding words, or sensitivity to phonological distinctiveness, an ability evident at 19 months of age but unreliable at 14 to 15 months of age. However, little is known about phonological constancy, the equally crucial ability to recognize a word's identity across natural phonetic variations, such as those in cross-dialect pronunciation differences. We show that 15- and 19-month-old children recognize familiar words spoken in their native dialect, but that only the older children recognize familiar words in a dissimilar nonnative dialect, providing evidence for emergence of phonological constancy by 19 months. These results are compatible with a perceptual-attunement account of developmental change in early word recognition, but not with statistical-learning or phonological accounts. Thus, the complementary skills of phonological constancy and distinctiveness both appear at around 19 months of age, together providing the child with a fundamental insight that permits rapid vocabulary growth and later reading acquisition.*

Efficient word learning requires a child to grasp that two complementary principles of phonetic variation define a word's spoken form. One is *phonological distinctiveness*, by which critical differences between phonetic segments can distinguish a word from similar-sounding words or nonwords (e.g., *cake* from *take* or *pake*). The other is *phonological constancy*, by which a

Address correspondence to Catherine T. Best, MARCS Auditory Laboratories, University of Western Sydney, Locked Bag 1797, Penrith South DC, NSW 1797, Australia, e-mail: c.best@uws.edu.au.

word remains true to itself despite phonetic variations that leave phonological structure intact, for example, across speech registers or regional accents.

Research on phonological distinctiveness has shown that children reject minimal-pair "mispronunciations" of words, for example, "vaby" for *baby*, more readily and consistently at 18 to 19 months of age than at 11 to 17 months of age (Hallé & de Boysson-Bardies, 1996; Stager & Werker, 1997; Swingley & Aslin, 2002). Children at the younger age do show some sensitivity in certain contexts, however (Fennell & Werker, 2003; Swingley & Aslin, 2002), and these findings have sparked debate over the factors that drive developmental change in word recognition. Although mispronunciation tests provide important insights about children's appreciation of phonological distinctiveness, they cannot resolve the theoretical debate because segmental substitutions (e.g., saying "vaby" for *baby*) conflate phonetic, phonological, and lexical transformations, and thus fail to address the crucial complementary skill of phonological constancy. Dialect variation offers an excellent natural tool for probing phonological constancy, because the phonetic details of pronunciation can differ substantially between native and nonnative dialects yet leave a word's phonological structure and lexical identity intact.

We exploited English-dialect differences in word pronunciation to evaluate development of phonological constancy with respect to three relevant theoretical accounts. The *phonological hypothesis* posits that early word representations are holistic phonological patterns that are deficient in phonetic detail and lack distinct phonemic segments. Those underspecified representations give way to more fully specified, segmental representations by about 18 months of age (Brown & Matthews, 1997; Metsala & Walley, 1998). Because native- and nonnative-dialect pronunciations are globally comparable but differ in specific

phonetic details, we extrapolate that familiar words in either dialect should be holistically recognizable to younger toddlers, but that the unfamiliar phonetic specifications of phonemes in nonnative-dialect pronunciations should cause difficulties for older toddlers. Alternatively, basic statistical-learning accounts posit that children compute the statistical properties of spoken input (Maye, Werker, & Gerken, 2002; Saffran, Aslin, & Newport, 1996), forming phonetically detailed word representations (e.g., Swingley, 2007). Such accounts generally assume that these statistics are tracked over the output of domain-general sensory analyzers (e.g., Anderson, Morgan, & White, 2003; Jusczyk, 1997), and phonemes emerge as nexuses of experienced phonetic patterns recurring across stored exemplars (e.g., Pierrehumbert, 2003). Extrapolating from these principles, word recognition should be more reliable for native-dialect pronunciations than for pronunciations that deviate from experienced speech patterns. Even if emerging phonemes become "attractors" for phonetically similar novel inputs (Anderson et al., 2003; see also Saffran, 2002), some differential in recognition should still hold for previously unencountered, notably deviant pronunciations from native speech.

The third view posits a qualitative shift in children's word recognition around the time of the vocabulary spurt (e.g., Nazzi & Bertoncini, 2003). Specifically, we propose that 18-month-old children use their earlier perceptual attunement to lower-order phonetic patterns in native speech (Best, 1994) to begin to discern that higher-order phonological organization within those patterns (see Best, 1993) identifies a word's constant underlying form across surface variations. By this perceptual-attunement account, phonological constancy emerges as children begin to detect higher-order invariants within native speech, which they recognize across novel forms, for example, even markedly differing nonnative pronunciations. Our view is derived from the perceptual assimilation model (PAM; Best, 1995), which is founded on the principles of articulatory phonology (e.g., Browman & Goldstein, 1992). According to PAM, perception of unfamiliar speech reflects perceivers' detection of information about articulatory gestures, that is, of the active movements of one or more speech articulators (lips, tongue tip, tongue body, tongue root, velum, glottis) to achieve constrictions of varying degrees (closed, critical, narrow, mid, wide) at specific locations within the vocal tract (e.g., at the lips, upper front teeth, alveolar ridge; Browman & Goldstein, 1992). Those articulatory gestures are posited to serve as the common metric for development of language-environment-attuned speech production and perception (see also Best & McRoberts, 2003; Goldstein & Fowler, 2003; Studdert-Kennedy & Goldstein, 2003). From this perspective, young toddlers (14–15 months old) perceive familiar words as dialect-specific phonetic patterns and so should recognize them in the native dialect but have difficulty recognizing nonnative pronunciations. Older toddlers (18–19 months old), however, should detect their phonological structure and recognize them across both dialects.

METHOD

We took advantage of toddlers' tendency to preferentially listen to familiar over unfamiliar words (Hallé & de Boysson-Bardies, 1994, 1996) to probe the three predictions regarding development of phonological constancy. We gave 15-month-old children (n = 20; mean age = 14 months 28 days; 9 females, 11 males) and 19-month-old children (n = 20; mean age = 19 months 15 days; 10 females, 10 males) two word-familiarity preference tests, one in their native dialect, American English of Connecticut, and the other in a markedly different dialect they had not previously experienced. For the nonnative dialect, we selected Jamaican Mesolect, a Patois-influenced English dialect that deviates extensively from American English in its phonetic realization of consonants, vowels, and stress patterns (Patrick, 1999; Wassink, 2006).

Two word sets were developed: 12 familiar stress-initial disyllabic words produced by more than 50% of 13- to 16-month-old children (Rescorla, Alley, & Christine, 2001) and 12 unfamiliar vowel- and stress-matched disyllabic words that occur in English adult corpora at frequencies of less than 2 per million words (Kucera & Francis, 1967). Stressed syllables contained the following dialectally differing vowels: /ɪ/, /ɛ/, and /ʊ/ (pronounced [ɪ], [ɛ], and [ʊ] in Connecticut American English hid, head, and hood, but pronounced in Jamaican Mesolect dialect as [i], [e], and [u], which are similar to the vowels in shortened versions of Connecticut American English heed, hayed, and who'd), and /æ/, /ɔ/, and /ɑ/ (pronounced [æ], [ɔ], and [a] in Connecticut American English had, hawed, and hod, but all pronounced in Jamaican Mesolect dialect as [a], which is similar to the vowel in Connecticut American English hod).

Two female speakers, one from Connecticut and one from Montego Bay, Jamaica, were selected on the basis of similar voice quality and fundamental frequency (f_0). Each conversed briefly with another native-dialect speaker, then recorded the 24 target words six times each in random order in the carrier "Say ___ again." Words were excised from the sentences, and one token was selected for each word by each speaker for final stimuli, based on cross-speaker similarities in duration and f_0 contours. These paired tokens were equated for duration and loudness, and onsets and offsets were ramped over 10 ms.

Each child completed two eight-trial tests, one with all words spoken in Connecticut American English and one with all words spoken in Jamaican Mesolect, using a conditioned-fixation version of Hallé and de Boysson-Bardies's (1996) familiar-word-preference task. Within each test, four trials included words from the familiar set and four trials included words from the unfamiliar set, all in the designated dialect. Familiar-word trials and unfamiliar-word trials alternated within each test. On each trial of a given test, words were selected randomly from the set of 12 familiar or 12 unfamiliar words, as appropriate, and presented continuously as long as the child remained fixated on a colored checkerboard on an LCD monitor (65–70 dB sound-

pressure level; interstimulus interval = 750 ms). Audio ceased when the child looked away, and the trial ended when the child looked away for 2 s. The checkerboard flashed until the child's gaze was recaptured; when the child's gaze was recaptured, the checkerboard stabilized, and a new trial began. Dialect-test order (Connecticut American English or Jamaican Mesolect test first) and the word set represented on the starting trial (familiar or unfamiliar) were counterbalanced across children at each age.

RESULTS

The three sets of theoretical predictions were evaluated via planned simple contrasts and interaction contrasts on the sum of fixation times for familiar-word trials versus the sum of fixation times for unfamiliar-word trials. We ran a three-way analysis of variance on the between-subjects factor of age and the within-subjects factors of dialect and word familiarity. The left panel of Figure 1 displays idealized predictions for each hypothesis. For simplicity of presentation, familiar-word preferences are displayed as the ratio of fixation times across familiar-word trials divided by fixation times across unfamiliar-word trials. The right panel of Figure 1 displays the actual results, also as the ratio of familiar fixation to unfamiliar fixation.

As the figure shows, the pattern of word familiarity preferences across dialects and ages is consistent with the perceptual-attunement account. The results failed to uphold the predictions of the phonological hypothesis. That is, the younger children failed to show the predicted familiarity preference across dialects, $F(1, 38) = 0.671, p > .48$, and the older children failed to show the predicted familiarity preference only for Connecticut American English, $F(1, 38) = 0.039, p > .84$. The results also failed to uphold the statistical-learning prediction of familiarity preference for Connecticut American English but not Jamaican Mesolect for both ages, $F(1, 38) = 0.723, p = .40$. Nor was there support for an alternative version of statistical learning according to which age differences in experience might yield a larger between-dialect difference in familiar-word preference at 19 months than at 15 months, $F(1, 38) = 0.328, p = .56$. As predicted by the perceptual-attunement account, however, 19-month-old children showed a familiarity preference across dialects, $F(1, 38) = 5.89, p < .02$. The simple interaction contrast for the second perceptual-attunement prediction of a familiar-word preference in Connecticut American English but not Jamaican Mesolect at 15 months of age was not significant. However, this prediction was supported by t tests on the familiar-to-unfamiliar ratios at this age, which were significantly greater than 1.0 (familiarity preference) for Connecticut American English, $t(19) = 2.24, p < .019$, but not for Jamaican Mesolect, $t(19) = 0.907, p > .187$ (one-tailed, Bonferroni corrected).

DISCUSSION

These results offer important new insights into early word perception. At 15 months of age, recognition of familiar words is restricted to native pronunciations, a finding that is inconsistent with the phonological hypothesis, but compatible with the statistical-learning and perceptual-attunement accounts. Both views posit native dialect-specific phonetic signatures for familiar early words, hampering recognition of unfamiliar pronunciations.

In contrast, 19-month-old children accept Jamaican Mesolect pronunciations of familiar words. This finding complements their previously reported rejection of native single-segment
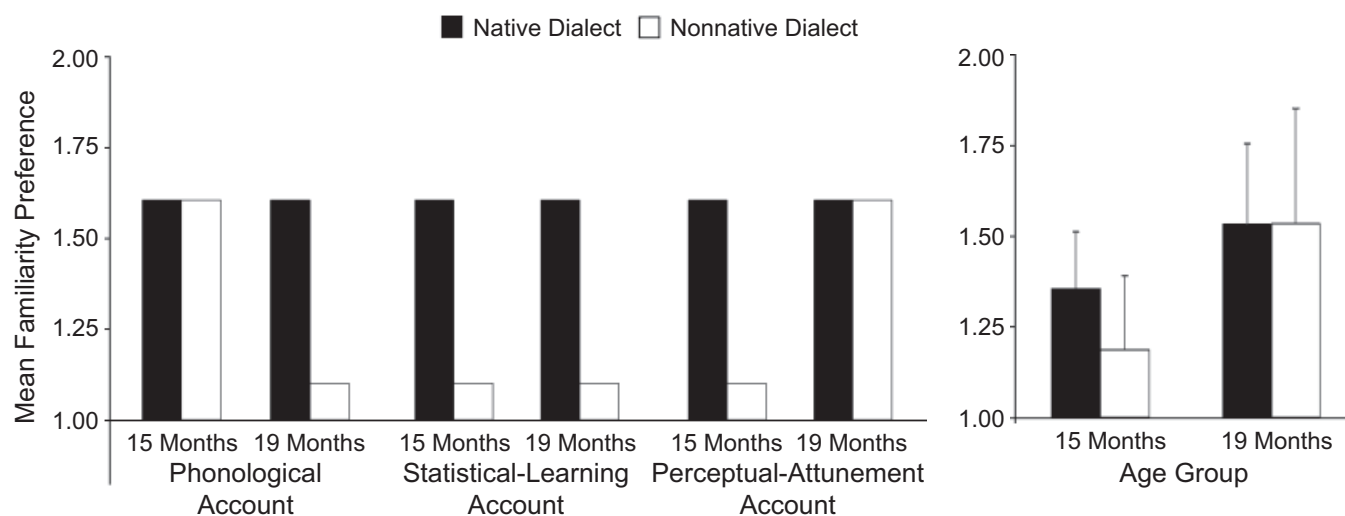

Fig. 1. Idealized and observed listening preferences for the native and nonnative dialects at ages 15 and 19 months. The graphs show mean familiarity preference, the ratio of fixation times across familiar-word trials to fixation times across unfamiliar-word trials. Idealized results according to three theoretical accounts are presented in the left panel. Observed listening preferences are presented in the right panel; the error bars represent standard errors of the means.

mispronunciations, and cannot be explained by the statistical account. Native-dialect statistical representations should be strengthened by additional experience, which would accentuate the deviance of unfamiliar pronunciations and further limit recognition. However, performance at 19 months of age is compatible with perceptual attunement, which suggests that these children have begun to discover phonological constancy across varying pronunciations.

Coemergence of phonological distinctiveness and phonological constancy around 19 months of age marks a dawning awareness that the two types of phonetic variation together define a word's identity. This insight is crucial to children's discovery of the fundamental phonological generalization that words are composed of discrete subunits, which emerges midway through the 2nd year, contemporaneously with other core linguistic skills (morphological regularities, syntactic operations), as exemplified by rhyming word play. This emergence is not mere coincidence, and instead reflects a nascent grasp of the particulate principle of language (Studdert-Kennedy & Goldstein, 2003) that carries a child into the world of linguistic structure, making possible both an infinitely expandable vocabulary and reading acquisition (Liberman, Shankweiler, & Liberman, 1989).

REFERENCES

Anderson, J.L., Morgan, J.L., & White, K.S. (2003). A statistical basis for speech sound discrimination. *Language and Speech*, *46*, 155–182.

Best, C. (1993). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 289–304). Dordrecht, The Netherlands: Kluwer Academic.

Best, C.T. (1994). Learning to perceive the sound pattern of English. In C. Rovee-Collier & L.P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 9, pp. 217–304). Norwood, NJ: Ablex.

Best, C.T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange & J.J. Jenkins (Eds.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

Best, C.T., & McRoberts, G.W. (2003). Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Language and Speech*, *46*, 183–216.

Browman, C.P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155–180.

Brown, C., & Matthews, J. (1997). The role of feature geometry in the development of phonemic contrasts. In S. Hannahs & M. Young-Scholten (Eds.), *Focus on phonological acquisition* (pp. 67–112). Amsterdam: John Benjamins.

Fennell, C.T., & Werker, J.F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, *46*, 245–264.

Goldstein, L.G., & Fowler, C.A. (2003). Articulatory phonology: A phonology for public language use. In N.O. Schiller & A.S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 159–207). Berlin: Mouton de Gruyter.

Hallé, P.A., & de Boysson-Bardies, B. (1994). Emergence of an early receptive lexicon: Infants' recognition of words. *Infant Behavior and Development*, *17*, 119–129.

Hallé, P.A., & de Boysson-Bardies, B. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, *19*, 463–481.

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

Kucera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Liberman, I.Y., Shankweiler, D., & Liberman, A.M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I.Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1–33). Ann Arbor: University of Michigan Press.

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

Metsala, J.L., & Walley, A.C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J.L. Metsala & L.C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Erlbaum.

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: Two modes of word acquisition? *Developmental Science*, *6*, 136–142.

Patrick, P.L. (1999). *Urban Jamaican Creole: Variation in the Mesolect*. Amsterdam: Benjamins.

Pierrehumbert, J.B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, *46*, 115–154.

Rescorla, L., Alley, A., & Christine, J.B. (2001). Word frequencies in toddlers' lexicons. *Journal of Speech, Language, and Hearing Research*, *44*, 598–609.

Saffran, J.R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*, 172–196.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Stager, C.L., & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381–382.

Studdert-Kennedy, M., & Goldstein, L. (2003). Launching language: The gestural origin of discrete infinity. In M.H. Christiansen & S. Kirby (Eds.), *Language evolution* (pp. 235–254). Oxford, England: Oxford University Press.

Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology*, *43*, 454–464.

Swingley, D., & Aslin, R.N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*, 480–484.

Wassink, A.B. (2006). A geometric representation of spectral and temporal vowel features: Quantification of vowel overlap in three linguistic varieties. *Journal of the Acoustical Society of America*, *119*, 2334–2350.