Current Directions in Psychological Science
http://cdp.sagepub.com/

How Do Infants Become Experts at Native-Speech Perception?

Janet F. Werker, H. Henny Yeung and Katherine A. Yoshida
Current Directions in Psychological Science 2012 21: 221
DOI: 10.1177/0963721412449459

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Current Directions in Psychological Science
21(4) 221–226

DOI: 10.1177/0963721412449459
http://cdps.sagepub.com
SAGE

How Do Infants Become Experts at Native-Speech Perception?

Janet F. Werker[1], H. Henny Yeung[2,3], and Katherine A. Yoshida[4]
[1]University of British Columbia; [2]Université Paris Descartes, Sorbonne Paris Cité; [3]Centre National de la Recherche Scientifique, Paris, France; and [4]New York University

Abstract

Infants begin life ready to learn any of the world's languages, but they quickly become speech-perception experts in their native language. Although this phenomenon has been well described, the mechanisms leading to native-language-listening expertise have not. In this article, we provide an in-depth review of one learning mechanism: distributional learning (DL), which has been shown to be important in phonetic category learning. DL is a domain-general statistical learning mechanism that involves tracking the relative frequency of phonetic tokens in speech input. Although DL is powerful, recent research has identified limitations to it as well. We conclude with a discussion of possible supplementary phonetic-learning mechanisms, which focuses on the surrounding context in which infants hear phonetic tokens and how it can augment DL and highlight important linguistic differences between perceptually similar stimuli.

Keywords

speech perception, infancy, category learning, phonetic

Phonetic Perception and Its Development

A parent points at a toy and says, "Look at this doll! That's your doll!" Learning words in this scenario requires many perceptual, pragmatic, and referential strategies on the part of an infant, but a first step is recognizing which sound properties distinguish word forms in his or her native language(s). This is not a simple task, given that phonetic input to infants is inherently variable. Consider the spoken tokens of "doll." To a Hindi speaker, the difference between the "d" sounds in "this doll" versus "your doll"—a phonetic contrast between a dental [d̪al] versus a retroflex [ɖal], respectively—would signal two possible word forms (either *lentils* or *branch*). In English, both of those "d" sounds signal just one possible word form—phonetically labeled as an alveolar [dal]. Figure 1 (top panel) illustrates these environments. In this article, we ask what developmental processes allow infants to learn native-language phonetic categories like these from the input they receive.

Early work suggested that infants begin life sensitive to all phonetic contrasts and that listening experience functions to maintain perceptual sensitivity only for native phonetic contrasts. For example, young English-learning infants easily discriminate the two Hindi "d" sounds mentioned above, whereas English-speaking adults find this much harder. By 10 months of age, English learners begin perceiving speech in accord with their native language, no longer discriminating these two "d" sounds, whereas Hindi-learning infants perceptually maintain this distinction (Werker & Tees, 1984).[1] Dozens of studies have reported this kind of perceptual attunement in consonant, vowel, and even tone perception within the first year of life (for reviews, see Curtin & Werker, 2007; Gervain & Mehler, 2010), even for bilingual infants, who maintain phonetic contrasts used in both of their native languages (Albareda-Castellot, Pons, & Sebastián-Gallés, 2010; Burns, Yoshida, Hill, & Werker, 2007).

Subsequent work has shown that phonetic development is actually much more complex than this single pattern of maintenance and decline (Best, 1995). Language experience also enhances the discrimination of some native phonetic contrasts (Kuhl et al., 2006) and can realign existing phonetic-category boundaries (Burns et al., 2007). Infants discriminate other contrasts only after early exposure to a language with these distinctions, a fact that suggests that listening experience may induce certain categories (Narayan, Werker, & Beddor, 2010; Sato, Kato, & Mazuka, 2012). Moreover, phonetic perception is not as categorical as was once imagined: Young infants also discriminate some within-category variation (McMurray & Aslin, 2005). Finally, perceptual change continues well beyond

Corresponding Author:
Janet F. Werker, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC V6T 1Z4, Canada
E-mail: jwerker@psych.ubc.ca

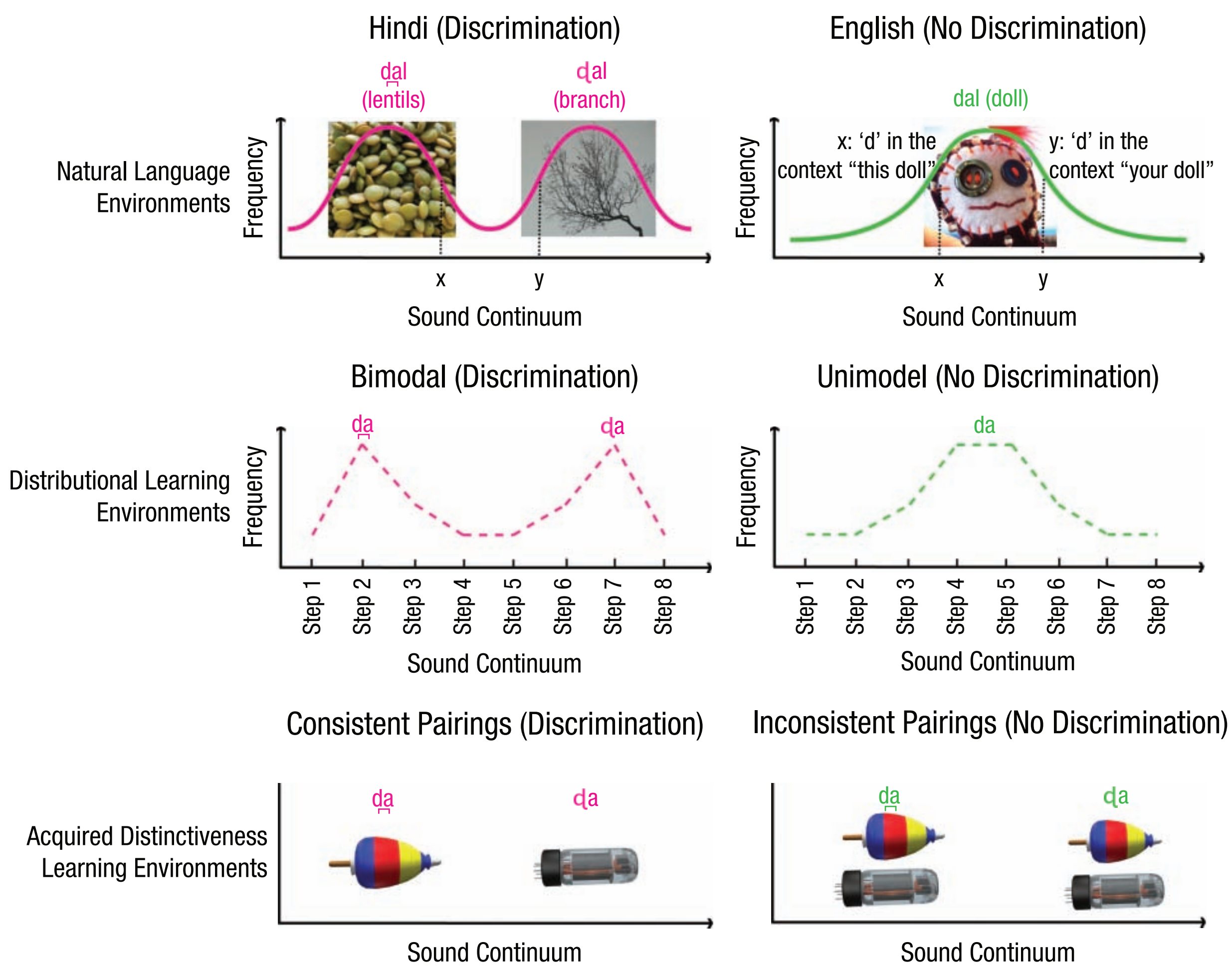

Fig. 1. An illustration of distributional learning and acquired distinctiveness. The top panel illustrates the variability around the pronunciation of "d" an infant might hear in a natural-language environment, depending on whether the infant is growing up learning Hindi (left) or English (right). The middle panel illustrates how this variability is modeled using relative frequencies of sounds that are created along an 8-step continuum from dental [d̪a] to retroflex [ɖa], as used in studies of distributional learning. The bottom panel illustrates how this variability is modeled in studies of acquired distinctiveness: Two consistent sound-object pairings are presented to highlight the distinction between [d̪a] and [ɖa] in comparison to inconsistent pairings. Image of lentils: FoodStories (2009). Retrieved from Flickr: http://www.flickr.com/photos/foodstories/4016993694/. Image of branch: Andy Magee (2011). Retrieved from Flickr: http://www.flickr.com/photos/amagee3/5394212150/. Image of doll: elasticcamel (2004). Retrieved from Flickr: http://www.flickr.com/photos/91524358@N00/2419004722/. All images used with permission.

the first year of life into later stages of development (Minagawa-Kawai, Mori, Naoi, & Kojima, 2007; Mugitani et al., 2009; Sundara, Polka, & Genesee, 2006).

What processes guide infants' phonetic learning? The input frequency of phonetic tokens is crucial. For example, Anderson, Morgan, and White (2003) showed that phonetic contrasts straddling frequently heard categories become language-specific earlier in development than those straddling less-frequently heard ones. Earlier theoretical models had suggested that token frequency "warps" acoustic-phonetic space, altering perceived similarity in that space (Jusczyk, 1993), or that frequently heard tokens attract acoustically similar ones, forming a "perceptual magnet" (Kuhl et al., 2008). In the following section, we review the literature on distributional learning (DL), which shares the token-counting approach of these previous models but also provides a mechanistic explanation of phonetic learning.

Distributional Learning (DL) From the Statistical Structure of Language Input

The term "statistical learning" refers to the notion that infants learn some language patterns from the statistical properties of language input, a fact that was first demonstrated by seminal work showing that infants can segment possible "words" in a language by tracking transitional probabilities between syllables (Saffran, Aslin, & Newport, 1996). Proponents of DL similarly suggest that native phonetic categories are identified at least in part by statistical information, but it is the relative frequencies of phonetic tokens in subregions of acoustic-phonetic space, not transitional probabilities, that are tracked. Consider, for example, an English-learning infant who hears the word "doll." Along one phonetic dimension, the infant will hear variation around a central tendency, creating a unimodal distribution. For a Hindi learner, this variation is distributed around two means, one for the dental "d" in [d̪al] and one for the retroflex "d" in [ɖal] (see Fig. 1, top and middle panels). Identifying these distributions could be used to infer the underlying phonetic structure of the native language(s).

Maye and Gerken (2000) demonstrated that adults are able to track phonetic distributions and use them to learn new perceptual categories (see Hayes-Harb, 2007, for a replication). To test DL in infants, Maye, Werker, and Gerken (2002) synthesized an 8-step continuum of "da" sounds[2] whose endpoints were within the same English category (likely still discriminable by these infants; see Pegg & Werker, 1997) but were in two possible categories from a different language. Two distributional environments were created from the continuum tokens. In the "bimodal" condition, the input contained more exemplars of steps 2 and 7, mimicking the variability infants might hear if they were raised in a language with two categories along that continuum. In the "unimodal" condition, the input contained more exemplars of steps 4 and 5, mimicking a language like English, which has only a single "da" category (see Fig. 1, middle panel). The 6- to 8-month-olds in the study were presented with these tokens in a semirandom order for just over 2 minutes and tested on their ability to discriminate the endpoints of the continuum (steps 1 and 8) immediately afterward. Infants in the bimodal condition discriminated the endpoints, but infants in the unimodal condition did not.

This work illustrated a mechanism for how perceptual sensitivity to pre-existing phonetic categories is maintained while sensitivity to nonnative categories declines. More recently, Maye, Weiss, and Aslin (2008) showed that DL from bimodal distributions can also enhance phonetic sensitivity to difficult phonetic contrasts. The researchers also showed that DL of a an acoustic/phonetic feature in one context (differences between [da] and [ta] in voice-onset time) also generalized to a new contrast in another context ([ka] and [ga]), suggesting that DL may occur at a more abstract, phonetic-feature level. This implies that the identification of two phonetic categories can also facilitate the learning of related contrasts, even if those latter distinctions are not as clearly demarcated by statistical distributions.

DL seems to be a basic mechanism of perceptual change in the phonetic domain. Like other statistical mechanisms, it is available not only to humans but also to other species (Pons, 2006). Moreover, it functions broadly across domains: For example, visual categories can be inferred from distributional information by both children (Duffy, Huttenlocher, & Crawford, 2006) and adults (Rosenthal, Fusi, & Hochstein, 2001). Recently, Cristià and her colleagues have further shown that DL is evident from at least 4-6 months of age and that it is robust to variability along two features simultaneously—an important demonstration, given that multiple redundant acoustic features cue most speech contrasts (Cristià, McGuire, Seidl, & Francis, 2011).

Limitations and Challenges to DL

Recent evidence suggests two important limitations to DL. First, the effectiveness of DL as a learning strategy has already begun to decline by 10 months of age, when most native phonetic categories have already emerged. For example, Maye et al. (2002, 2008) found that for 6- to 8-month old infants, 2.3 minutes of exposure to a bimodal or unimodal distribution was required to collapse phonetic categories, but by 10 months of age, infants required more than 4 minutes of exposure—still a small amount, but almost twice as much as before—to re-establish phonetic discrimination of a nonnative contrast in the process of decline (Yoshida, Pons, Maye, & Werker, 2010). In adults, even 20 minutes of exposure brings about smaller perceptual changes than those seen in infants (Hayes-Harb, 2007; Maye & Gerken, 2000). These findings mirror broader age-related changes in phonetic sensitivity, and they raise the possibility that DL may operate most effectively during a sensitive period in early development. This could be because phonetic systems are most open to input in infancy for maturational reasons (Werker & Tees, 2005), because other learning mechanisms become more important as infants develop (see the next section for more details), or because inferred distributions become more resilient to change as more speech input is accumulated.

Second, DL appears to interact with acoustic-phonetic salience. Although infants can use DL to learn many kinds of phonetic distinctions, infants show no evidence of learning certain particularly difficult categories (e.g., a Polish alveolar-palatal fricative; see Cristià et al., 2011). Future work will need to investigate whether DL is sufficient to induce phonetic categories that are not discriminable by infants without previous language experience (e.g., Narayan et al., 2010; Sato et al., 2010). Such studies must ask whether the difficulty in learning particular contrasts from DL training in the laboratory correlates with the relative timing of perceptual change for these same contrasts in natural languages.

An important challenge is to show that there are cues in speech input that would support DL in more naturalistic situations. Several analyses of parental speech input to infants have confirmed that, although imperfect, distributional cues are present that can signal both the number of native-language categories and possible phonetic distinctions in that language (Gauthier, Shi, & Xu, 2007; Vallabha, McClelland, Pons, Werker, & Amano, 2007; Werker et al., 2007). These analyses are supported by results from a compelling correlation study showing that distributional regularity in a mother's speech is predictive of the phonetic-category structure her infant learns (Cristià, 2011).

At the same time, other work has provided convincing empirical challenges to the notion that DL alone can explain phonetic-category learning. For example, some researchers have argued that computational modeling of DL is not viable in examinations of samples of natural conversational speech (Swingley, 2009) or without additional constraints (i.e., the ability of highly predictive distributional hypotheses to inhibit other hypothetical distributions; McMurray, Aslin, & Toscano, 2009). Collectively, these studies have suggested that DL must be supplemented by other learning strategies on the part of the infant.

Current Directions on Mechanisms of Phonetic Learning

Infants may supplement DL by relying on the fact that individual speech sounds frequently occur in unique perceptual contexts. As originally suggested by Lawrence (1949), who labeled this basic, domain-general learning mechanism *acquired distinctiveness* (AD), such contexts reprioritize the salience of perceptible cues in discrimination tasks. According to an AD-based explanation, the occurrence of value A in context X and value B in context Y is thought to highlight cues that distinguish value A from value B and hence facilitate discrimination (see Hayes-Harb, 2007; Kluender, Lotto, Holt, & Bloedel, 1998 for examples in the phonetic domain). Thus, in cases where distributional information might be muddled, hearing one vowel in context X (i.e., /i/ in 'see') contrasted with a similar vowel in context Y (i.e., /I/ in 'this') could enhance the perceptual distance between those vowels (i.e., /i/ versus /I/; see Swingley, 2009). Such auditory AD contexts have already been shown to highlight phonetic contrasts in 14-month-olds' word-form representations (Thiessen, 2011) and to improve the effectiveness of DL in adults (Feldman, Myers, White, Griffiths, & Morgan, 2011); moreover, they have been hypothesized to do the same for phonetic perception as infants learn native phonetic categories (Feldman et al., 2011; Swingley, 2009).

Of course, infants pay attention to much more than just the auditory stream. For example, redundant information in seen and heard speech can influence DL in infants (Teinonen, Aslin, Alku, & Csibra, 2008). Consider a clearer case of cross-modal AD, in which audio-visual co-occurrences are arbitrary. As young infants begin attending to word referents, the simple co-occurrence of distinct words and objects may further highlight the salience of the relevant phonetic cues. In the scenario described in the introduction of this article, a Hindi-learning infant might hear variability around a dental [d̪al] when his mother is talking about cooking lentils and variability around a retroflex [ɖal] when she points to a branch. An English-learning infant would hear that variability only around an alveolar [dal] when his or her mother talks about a doll.

Yeung and Werker (2009) investigated this exact AD scenario. English-learning 9-month-olds were presented with consistent pairings of a syllable beginning with a dental [d̪a] and a picture of one object as well as a syllable beginning with a retroflex [ɖa] and a picture of another object (see Fig. 1, bottom panel). Results showed that these infants were able to discriminate this nonnative contrast after seeing such contrastive object-speech pairings; however, infants who had received no training (a control group) or had seen inconsistent pairings between syllables and objects (i.e., [ɖa] presented with Object 1 on some familiarization trials and with Object 2 on others) could not.

This work all suggests that co-occurring contextual cues can facilitate the acquisition of native-language phonetic categories through learning from AD. However, it is important to note that the very same contexts in which infants learn from AD are also those that can be lexically informative. Such learning situations may allow infants to embed phonetic categorization in the broader context of language acquisition (i.e., word-learning) from very early in life. Specifically, it may be the case that as language-specific categories begin to solidify, and as infants come to treat speech as a means of communication, phonetic learning relies less on basic domain-general strategies and more on linguistically motivated ones. For example, there is some evidence that infants learn phonetic patterns better in face-to-face, contingent live interactions, in which the statistical characteristics of (audiovisual) speech input are presumably constant (Kuhl, Tsao, & Liu, 2003). Future work in infant phonetic learning must further investigate what specific learning mechanisms may supplement DL in such scenarios.

Several important issues must be explored by future work in this field. First, learning mechanisms supplementing DL, such as AD, must be further delineated. What are their characteristics? When do they become active in development, and how do different learning mechanisms interact in cases of conflict? Relatedly, are there clear distinctions between domain-general statistical mechanisms—such as DL or basic learning from co-occurring contexts—and more "sophisticated," lexically related mechanisms? And finally, are there sensitive periods in development during which different learning mechanisms are most effective? Answers to these questions will help provide a more complete picture of just how infants learn native-language speech patterns so early in development.

Recommended Reading

Cristià, A. (2011). (See References). A study showing that the distributional characteristics in maternal speech guide infant phonetic-category learning.

Gervain, J. & Mehler, J. (2010). (See References). A comprehensive, highly accessible overview of the many ways infant speech perception changes in the first year of life.

Swingley, D. (2009). (See References). A paper that provides a full discussion of challenges to distributional learning for readers who wish to learn more about computational approaches to the issue.

Yeung, H. H., & Werker, J. F. (2009). (See References). The original study identifying acquired distinctiveness as a possible phonetic-category-learning mechanism in infancy.

Yoshida, K. A., Pons, F., Maye, J., & Werker, J. F. (2010). (See References). A recent study that illustrates original research on distributional learning and how it becomes more difficult by 10 months of age.

Declaration of Conflicting Interests

The author declared no conflict of interest with respect to the authorship or the publication of this article.

Funding

Preparation of this article was supported in part by NSERC Grant 81103 and the Canada Research Chair Foundation.

Notes

1. Werker and Tees (1984) originally tested "t" sounds rather than "d" sounds, although subsequent work has tested "d" sounds as well.
2. The original work was done with a voicing continuum, not the Hindi retroflex-dental distinction, but subsequent work has tested that distinction as well.

References

Albareda-Castellot, B., Pons, F., & Sebastián-Gallés, N. (2010). The acquisition of phonetic categories in bilingual infants: New data from an anticipatory eye movement paradigm. *Developmental Science, 14*, 395–401. doi:10.1111/j.1467-7687.2010.00989.x

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. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language speech research* (pp. 171–204). Timonium, MD: York Press.

Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics, 28*, 455–474.

Cristià, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *Journal of the Acoustical Society of America, 129*, 3271–3280. doi:10.1121/1.3562562

Cristià, A., McGuire, G. L., Seidl, A., & Francis, A. L. (2011). Effects of the distribution of acoustic cues on infants' perception of sibilants. *Journal of Phonetics, 39*, 388–402. doi:10.1016/j.wocn.2011.02.004

Curtin, S., & Werker, J. F. (2007). The perceptual foundations of phonological development. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 579–599). Oxford, England: Oxford University Press.

Duffy, S., Huttenlocher, J., & Crawford, L. E. (2006). Children use categories to maximize accuracy in estimation. *Developmental Science, 9*, 597–603. doi:10.1111/j.1467-7687.2006.00538.x

Feldman, N., Myers, E., White, K., Griffiths, T., & Morgan, J. L. (2011). Learners use word-level statistics in phonetic category acquisition. In N. Danis, K. Mesh, & H. Sung (Eds.), *Proceedings of the 35th Annual Boston University Conference on Language Development* (pp. 197–209). Boston, MA: Cascadilla Press.

Gauthier, B., Shi, R., & Xu, Y. (2007). Simulating the acquisition of lexical tones from continuous dynamic input. *Journal of the Acoustical Society of America, 121*, EL190–EL195. doi:10.1121/1.2716160

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology, 61*, 191–218. doi:10.1146/annurev.psych.093008.100408

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research, 23*, 1–31.

Jusczyk, P. W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics, 21*, 3–28.

Kluender, K. R., Lotto, A. J., Holt, L. L., & Bloedel, S. L. (1998). Role of experience for language-specific functional mappings of vowel sounds. *Journal of the Acoustical Society of America, 104*, 3568–3582.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363*, 979–1000. doi:10.1098/rstb.2007.2154

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9*, F13–F21.

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, USA, 100*, 9096–9101.

Lawrence, D. H. (1949). Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology, 39*, 770–784.

Maye, J., & Gerken, L. (2000). Learning phonemes without minimal pairs. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development* (pp. 522–533). Somerville, MA: Cascadilla Press.

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science, 11*, 122–134.

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

McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition, 95*, B15–B26. doi:10.1016/j.cognition.2004.07.005

McMurray, B., Aslin, R. N., & Toscano, J. C. (2009). Statistical learning of phonetic categories: Insights from a computational approach. *Developmental Science, 12*, 369–378. doi:10.1111/j.1467-7687.2009.00822.x

Minagawa-Kawai, Y., Mori, K., Naoi, N., & Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *Journal of Neuroscience, 27*, 315–321. doi: 10.1523/jneurosci.1984-06.2007

Mugitani, R., Pons, F., Fais, L., Dietrich, C., Werker, J. F., & Amano, S. (2009). Perception of vowel length by Japanese- and English-learning infants. *Developmental Psychology, 45*, 236–247. doi:10.1037/a0014043

Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science, 13*, 407–420. doi:10.1111/j.1467-7687.2009.00898.x

Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *The Journal of the Acoustical Society of*

America, 102, 3742–3753. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9407666

Pons, F. (2006). The effects of distributional learning on rats' sensitivity to phonetic information. *Journal of Experimental Psychology: Animal Behavior Processes, 32*, 97–101. doi:10.1037/0097-7403.32.1.97

Rosenthal, O., Fusi, S., & Hochstein, S. (2001). Forming classes by stimulus frequency: Behavior and theory. *Proceedings of the National Academy of Sciences, USA, 98*, 4265–4270. doi:10.1073/pnas.071525998

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

Sato, Y., Kato, M., & Mazuka, R. (2012). Development of single/geminate obstruent discrimination by Japanese infants: Early integration of durational and nondurational cues. *Developmental Psychology, 48*, 18–34. doi:10.1037/a0025528

Sato, Y., Sogabe, Y., and Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology, 46*, 106–119.

Sundara, M., Polka, L., & Genesee, F. (2006). Language-experience facilitates discrimination of /d-th/ in monolingual and bilingual acquisition of English. *Cognition, 100*, 369–388. doi:10.1016/j.cognition.2005.04.007

Swingley, D. (2009). Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3617–3632. doi:10.1098/rstb.2009.0107

Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition, 108*, 850–855. doi:10.1016/j.cognition.2008.05.009

Thiessen, E. D. (2011). When variability matters more than meaning: The effect of lexical forms on use of phonemic contrasts. *Developmental Psychology, 47*, 1448–1458. doi:10.1037/a0024439

Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences, USA, 104*, 13273–13278.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition, 103*, 147–162.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development, 7*, 49–63.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology, 46*, 233–251. doi:10.1002/dev.20060

Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition, 113*, 234–243. doi:10.1016/j.cognition.2009.08.010

Yoshida, K. A., Pons, F., Maye, J., & Werker, J. F. (2010). Distributional phonetic learning at 10 months of age. *Infancy, 15*, 420–433. doi:10.1111/j.1532-7078.2009.00024.x