COGNITION

Cognition 82 (2002) B101–B111

www.elsevier.com/locate/cognit

Brief article

Infant sensitivity to distributional information can affect phonetic discrimination

Jessica Maye[a,*], Janet F. Werker[b], LouAnn Gerken[c,d]

[a]*Department of Brain and Cognitive Sciences, Meliora Hall, University of Rochester, Rochester, NY 14627, USA*
[b]*Department of Psychology, University of British Columbia, Vancouver, BC, Canada*
[c]*Department of Speech and Hearing Sciences, University of Arizona, Tucson, AZ, USA*
[d]*Department of Linguistics, University of Arizona, Tucson, AZ, USA*

Received 29 May 2001; received in revised form 21 August 2001; accepted 30 August 2001

Abstract

For nearly two decades it has been known that infants' perception of speech sounds is affected by native language input during the first year of life. However, definitive evidence of a mechanism to explain these developmental changes in speech perception has remained elusive. The present study provides the first evidence for such a mechanism, showing that the statistical distribution of phonetic variation in the speech signal influences whether 6- and 8-month-old infants discriminate a pair of speech sounds. We familiarized infants with speech sounds from a phonetic continuum, exhibiting either a bimodal or unimodal frequency distribution. During the test phase, only infants in the bimodal condition discriminated tokens from the endpoints of the continuum. These results demonstrate that infants are sensitive to the statistical distribution of speech sounds in the input language, and that this sensitivity influences speech perception. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Infant sensitivity; Distributional information; Phonetic discrimination

1. Introduction

Infants are extraordinarily good at discriminating speech sounds. Young infants can discriminate virtually every phonetic contrast on which they have been tested (for review see Aslin, Jusczyk, & Pisoni, 1998). In fact, young infants discriminate

* Corresponding author. Tel.: +1-716-275-7192; fax: +1-716-442-9216.
E-mail address: jmaye@bcs.rochester.edu (J. Maye).

PII: S0010-0277(01)00157-3

certain contrasts better than adults from the same language community (Werker, Gilbert, Humphrey, & Tees, 1981). In contrast, adults' perception of speech sounds is constrained by the phonetic organization of their native language. In particular, it is easier for adults to discriminate contrasts between speech sounds that distinguish word meanings in their native language than contrasts that do not (Liberman, Harris, Hoffman, & Griffith, 1957).

By the end of their first year, infants' pattern of speech sound discrimination resembles that of adults from their language community (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984). In many cases, this means that infants *stop* discriminating previously discriminable contrasts (Best, McRoberts, LaFleur, & Silver-Isenstadt, 1995; Polka & Werker, 1994; Werker & Tees, 1984).[1] Although it is somewhat counterintuitive to characterize decline in performance as an indication of cognitive development, these changes in perception reflect the efficacious focus of infants' attention to only those acoustic dimensions that are relevant for their native language.[2] For this reason, developmental changes in speech perception are among the first evidence that an infant has begun acquiring a native language.

The early age at which infants' speech perception exhibits influence from the native language is remarkable. Yet, in the 17 years since these early language effects were first documented, there has been no demonstration of a definitive mechanism to account for this perceptual development. Some researchers have suggested that the development occurs as a result of word learning (Best, 1995; Jusczyk, 1985; Lalonde & Werker, 1995; MacKain, 1982; Werker & Pegg, 1992). For example, an infant who learns the meanings of two words whose pronunciations differ only by a single sound, e.g. *bear* vs. *pear*, may discover that this difference is important to attend to. The plausibility of important sound differences being highlighted via meaning differences is attested, at least for adults, because this is the standard method by which field linguists discover the inventory of sounds that are used contrastively in a language (Pike, 1947). However, this account cannot entirely explain native language effects on infant speech perception, because the initial changes in perception *precede* infants' ability to distinguish between similar word pairs. Although language-specific discrimination of minimally-different nonsense syllables (e.g. [ba] vs. [pa]) is evident before the age of 12 months, infants have not been shown to discriminate minimally-different meaningful words (e.g. *bear* vs. *pear*) before 17 months of age (Stager & Werker, 1997; Swingley & Aslin, 2000; Werker, Fennell, Corcoran, & Stager, in press). It is therefore unlikely that word learning is a primary component in the initial restructuring of phonetic perception during infancy.

A second account for changes in speech perception draws on infants' sensitivity to

[1] For certain phonetic contrasts discrimination appears to improve with exposure to a language (Aslin, Pisoni, Hennessy, & Perey, 1981; Polka, Colantonio, & Sundara, 2001).

[2] Certain non-native phonetic distinctions are apparently immune to this effect, however. Best, McRoberts, and Sithole (1988) and Best et al. (1995) showed that Zulu click sounds, which are unlike any phones of English, are well discriminated by both English-learning infants and English-speaking adults, suggesting that reductions in discriminability are mediated by the relationship of the non-native contrast to native language phonetic categories.

distributional properties of their language (Guenther & Gjaja, 1996; Jusczyk, Bertoncini, Bijeljac-Babic, Kennedy, & Mehler, 1990; Kuhl, 1993). Infants are remarkably attuned to stochastic patterns in language. By 6 months, perception of vowel sounds is affected by the phonetic distribution of the native language, such that vowels that are prototypical in the language exhibit a strong perceptual pull on less prototypical vowels, inhibiting discrimination (Kuhl et al., 1992). By 9 months infants can discriminate between speech sound sequences that occur frequently in their language and those that occur with lower probability (Jusczyk, Luce, & Charles-Luce, 1994). The timing of these abilities, which rely on exposure to the input language, conspicuously parallels the age at which language experience affects the perception of non-native speech sounds, which begins to occur by 6 months for vowels (Kuhl et al., 1992; Polka & Werker, 1994) and by 10 months for consonants (Werker & Tees, 1984). Furthermore, infants' sensitivity to distributional regularities is so robust that it is evident after only brief (e.g. 2 min) exposure (Gómez & Gerken, 1999; Saffran, Aslin, & Newport, 1996).

Infants' sensitivity to distributional patterns provides a potential account for the early effect of the native language on speech perception, because the distribution of speech sounds an infant is exposed to from native language input reflects the phonetic categories that are used in the language. For any given phonetic category in a language, e.g. [b], actual tokens of the category vary considerably on several acoustic dimensions. However, despite this variation, along certain acoustic dimensions (e.g. voice onset time (VOT)) most tokens of [b] that are heard in the language are more similar to each other than to tokens from some neighboring phonetic category, e.g. [p].[3] These stochastic properties of the input indicate which acoustic dimensions are most informative for differentiating the phonetic categories of a language: for properties that are highly informative for differentiating two sounds, tokens from the two categories will form a bimodal distribution of values on that dimension, such that most fall into one of two clusters, separated by a sparsely populated region (see Fig. 1). If a given acoustic property is non-contrastive, however (that is, it does not differentiate between two categories), speech sound tokens will fall into a single (potentially wider) cluster, forming a unimodal distribution. Previous research has demonstrated that non-human animals utilize distributional information for the purpose of discriminating phonetic categories (Kluender, Lotto, Holt, & Bloedel, 1998); therefore, it is likely that infants are also sensitive to this cue to category structure.

Our goal was to determine whether infants marshal their keen sensitivity to stochastic patterns to track the distribution of speech sounds in a language. If so, they should be able to use this information to determine the linguistic relevance of various acoustic properties: unimodally distributed regions of sound indicate that an acoustic property is uninformative for distinguishing speech sounds in a given language, and therefore that property need not be attended to, while bimodal distributions signal the linguistic importance of a contrast.

[3] For evidence that phonetic categories are reflected in the distribution of sounds produced in a language see Lisker and Abramson (1964), Magloire and Green (1999), and Sundberg and Lacerda (1999).

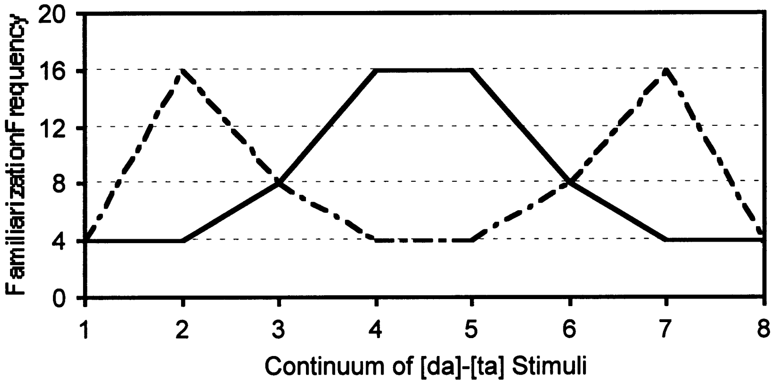

Fig. 1. Bimodal vs. Unimodal distributions of [da]–[ta] stimuli during familiarization. The continuum of speech sounds is shown on the abscissa, with Token 1 corresponding to the endpoint [da] stimulus, and Token 8 the endpoint [ta] stimulus. The ordinate axis plots the number of times each stimulus occurred during the familiarization phase. The presentation frequency for infants in the Bimodal group is shown by the dotted line, and for the Unimodal group by the solid line.

To test this, we exposed infants to novel speech stimuli, arranged according to systematically different distributions. We created a continuum of speech sounds based on a phonetic contrast that infants between 6 and 8 months of age have been shown to discriminate: voiced unaspirated vs. voiceless unaspirated stop consonants (Pegg & Werker, 1997). We then exposed infants to the full continuum of stimuli arranged in one of two distributional patterns (see Fig. 1). One group was presented with a bimodal frequency distribution, such that stimuli near the endpoints of the continuum occurred more frequently than the center stimuli. The other group was presented with a unimodal distribution for the same stimuli, such that stimuli from the center of the continuum occurred most frequently. We predicted that infants exposed to a bimodal distribution would form a two-category representation of these sounds, while infants exposed to a unimodal distribution would form a one-category representation, and that they would be able to do this without any information about whether the sounds expressed the same or different meanings in this mini-lexicon. If this prediction is correct, after familiarization to stimuli exhibiting these distributions, infants exposed to a bimodal distribution should be better able to discriminate the contrast than infants exposed to a unimodal distribution.

Infants at 8 months of age have been shown to be sensitive to statistical information (Jusczyk et al., 1994; Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999) and capable of learning distributional relationships between linguistic units after short-term experimental exposure (Gómez & Gerken, 1999; Jusczyk, Houston, & Newsome, 1999; Saffran, Aslin, & Newport, 1996). At 8 months infants should therefore be capable of performing the necessary computations for using distributional information to learn phonetic categories. However, at this age the native language is already beginning to affect speech perception (Kuhl et al., 1992; Werker & Tees, 1984). If evidence of development is already present at 8 months, the mechanism responsible for that development should be present prior to 8 months. For this reason we tested both 6- and 8-month-olds, although for the younger age group there is no experimental evidence of acquisition of native language consonants or of statistical learning.

2. Experiment

2.1. Methods

2.1.1. Participants

Infants from English-speaking homes were recruited based on parental interest in research participation. We excluded infants for whom English did not account for at least 75% of their total language exposure. Twenty-four 6-month-olds (age range 6 months, 0 days to 7 months, 8 days; mean 6 months, 16 days) and 24 8-month-olds (age range 7 months, 13 days to 9 months, 11 days; mean 8 months, 2 days) were included in the study. An additional 12 infants were tested but excluded from analyses for the following reasons: failure to meet language requirements ($n = 5$), crying ($n = 2$), inattention to the visual stimulus ($n = 2$), equipment failure ($n = 1$), experimenter error ($n = 1$), and parental interference ($n = 1$). Parental consent was obtained prior to testing, in accordance with hospital and university standards for the ethical treatment of human subjects. The infants were randomly divided between two familiarization conditions (discussed below), with 12 infants from each age group per condition. In all groups, infants were balanced for gender.

2.1.2. Stimuli

The experimental contrast was between voiced vs. voiceless unaspirated alveolar stops [d] and [t]. The voiceless unaspirated [t] occurs after "s" in English, in words like *stop*.[4] The syllables [da] and [ta] were produced by a female American-English speaker ([s] was excised from [sta] to create [ta]), and then digitally edited and re-synthesized to form an eight-point continuum from [da] to [ta], using Kay Elemetrics Analysis and Synthesis Laboratory. All [da] and [ta] stimuli were 465 ms in duration.

The phonetic difference between [da] and [ta] was in the presence of prevoicing (VOT ≥ -90 ms) for [da], as well as in the trajectories of the first two formants, from vowel onset to vowel center. For [da], the difference between formant frequencies at vowel onset vs. vowel center was greater than for [ta], resulting in a steeper onset trajectory for [da].

To de-emphasize the experimentally relevant acoustic dimensions, making the task more like natural language acquisition, we included four tokens each of filler syllables [ma] and [la] during familiarization, in addition to the eight [da]–[ta] stimuli. The filler stimuli ranged from 459 to 472 ms in duration, and were created

[4] In English, syllable-initial /t/ is pronounced with aspiration: [t^h]. It is somewhat difficult for adult English speakers to discriminate syllable-initial unaspirated [t] from voiced [d], and English-learning infants lose sensitivity to this contrast by 10–12 months (Pegg & Werker, 1997).

in the same manner as the experimental stimuli, with the exception that they did not form a continuum or correspond to any particular statistical distribution.

2.1.3. Procedure

We utilized a variation of the preferential looking procedure (Jusczyk & Aslin, 1995). After familiarization with speech stimuli, infants were presented with an auditory test stimulus from a speaker located above a monitor displaying a visual pattern, and looking times to the auditory/visual stimulus pairing were measured. Differential looking times to two different types of auditory stimuli indicate that infants discriminate them.

The experiment began with a familiarization phase, during which infants heard six blocks of 24 syllables each. Each block was composed of 16 stimuli from the [da]–[ta] continuum (the entire continuum, presented according to either a unimodal or bimodal frequency distribution, as shown in Fig. 1), and the eight filler syllables. Syllables were presented in random order, with an interstimulus interval (ISI) of 500 ms. The total length of familiarization was 2.3 min.

We tested phonetic discrimination using a paradigm developed by Best and Jones (1998), in which there are two types of test trials; on half of the test trials (*Non-Alternating* trials) a single stimulus is repeated, while on the other test trials (*Alternating* trials) infants hear an alternation between two different stimuli. Because discrimination of these two types of test trials hinges on infants' ability to perceive the alternation on Alternating trials, differential looking times to Alternating vs. Non-Alternating trials demonstrate discrimination of the stimuli composing the Alternating trials. Previous studies using this type of procedure to assess phonetic discrimination have found longer looking times on Alternating trials (Best & Jones, 1998). In the statistical learning literature, though, infants generally display a novelty preference (Saffran, Aslin, & Newport, 1996; Saffran et al., 1999). Since in the present study the test phase directly follows a familiarization phase in which infants hear a string of varied syllables, a novelty preference would be evident in longer looking times on Non-Alternating trials. However, any significant difference in looking times for the two types of test trials would indicate discrimination of the test stimuli.

Infants were presented with eight test trials. Test trials were ordered such that every other trial was Alternating (and vice versa), with order counterbalanced in each group. On each test trial, infants heard a string of eight stimuli with an ISI of 1 s, for a total length of 11 s per trial. On half of the Non-Alternating trials the repeated stimulus was Token 3 from the experimental continuum, while on the other Non-Alternating trials it was Token 6. On the remaining four test trials, infants heard an alternation between the two endpoint stimuli, Tokens 1 and 8. The test stimuli were chosen such that they had each occurred equally frequently in both familiarization conditions (Tokens 3 and 6 occurred eight times each, and Tokens 1 and 8 occurred four times each). Thus, any differences observed between infants from the two familiarization conditions could not be attributed to differences in exposure to the particular test stimuli. To reduce the possibility that the frequency distribution of stimuli presented during the *test phase* would counteract any learning that had occurred during familiarization, the test phase was designed to exemplify a flat distribution: Tokens 1, 3, 6, and 8 were each presented a total of 16 times over the course of the test phase.

The experiment was conducted in a sound-attenuated room. Infants were seated on their parent's lap, 115 cm in front of a Panasonic CT-27XF37C television monitor and a BOSE model 101 speaker. The monitor and speaker were connected to a PowerMacintosh G3 computer in an adjoining control room, and presentation of auditory and visual stimuli was controlled by the program Habit, developed by L.B. Cohen (University of Texas, Austin, TX). The infants' gaze was monitored over closed-circuit TV in the control room via a Panasonic PV-5770-K Omnimovie SVHS video camera positioned below the television monitor, and the entire experiment was videotaped to check the reliability of looking-time measurements.

On each trial a salient visual stimulus (a flashing ball) was presented to draw the infant's attention to the television monitor. Once the infant oriented towards the screen, the experimenter began a trial. On each trial, the auditory stimulus was presented at a volume between 65 and 70 dB, paired with an unrelated visual stimulus (a colorful picture of flowers during familiarization, a black-and-white checkerboard pattern during the test). Throughout the experiment the parent listened to masking music through Peltor workstyle headphones, so that they could not inadvertently influence their infant's looking behavior.

2.2. Results and discussion

Mean looking times for infants in each condition for the two types of test trials are provided in Table 1. A mixed-design ANOVA (2 Familiarization Conditions × 2 Age Groups × 2 Test Trial Types) revealed a main effect of Familiarization Condition ($F[1,44] = 4.913$, $P < 0.05$), with infants in the Bimodal condition showing longer overall looking times ($M_{6\text{months}} = 6.03$ s, $M_{8\text{months}} = 5.80$ s) than infants in the Unimodal condition ($M_{6\text{months}} = 4.81$ s, $M_{8\text{months}} = 5.09$ s), but no significant effect of Age Group ($F[1,44] = 0.033$, NS), or Trial Type ($F[1,44] = 2.787$, NS). The interaction of Trial Type and Familiarization Condition was marginally significant ($F[1,44] = 3.629$, $P = 0.063$). Because of the lack of an age effect, data from the two age groups were pooled in subsequent analyses. Planned pairwise comparisons confirmed the specific hypothesis that a difference in Test Trial Type (Alternating vs. Non-Alternating) would be present for infants in the Bimodal ($t[23] = 2.273$, $P < 0.04$), but not the Unimodal condition ($t[23] = 0.197$, NS).

Table 1
Mean (SE) looking times for infants in each age group and familiarization condition on Alternating and Non-Alternating trials

	Alternating trials (s)	Non-Alternating trials (s)
6 months Unimodal	4.85 (0.47)	4.53 (0.51)
8 months Unimodal	4.98 (0.63)	5.20 (0.56)
6 months Bimodal	5.66 (0.44)	6.41 (0.32)
8 months Bimodal	5.45 (0.52)	6.15 (0.56)

The frequency distribution of familiarization stimuli significantly affected behavior during the test phase. At both ages, infants from the Bimodal condition looked longer on Non-Alternating test trials than on Alternating trials, while infants from the Unimodal condition showed no preference, indicating that only infants in the Bimodal condition discriminated the test stimuli. Because the infants were familiarized to strings of varied syllables, the Bimodal infants' longer looking times on Non-Alternating trials exemplify a novelty preference.

3. General discussion

These results demonstrate that infants are sensitive to the frequency distribution of speech sounds in the input, and that this sensitivity is present during the age range in which developmental changes in speech perception are observed. This finding suggests that attention to the statistical distribution of speech sounds in the input is one factor driving the development of speech perception over the first year of life.

An interesting corollary is that infants familiarized to a unimodal distribution did not discriminate the test stimuli, although previous research showed this contrast to be discriminable at these ages (Pegg & Werker, 1997). This finding suggests that exposure to a unimodal distribution has the effect of *reducing* discrimination. These results parallel the behavior of infants who have amassed months of exposure to the native language: many non-native contrasts that are discriminable during the first 6 months are no longer discriminated later in infancy (Best et al., 1995; Werker & Tees, 1984). Of course, in the present study this reduced discrimination for infants in the unimodal condition was both more rapidly induced (after only 2.3 min of exposure) and presumably less robust than the reduced discrimination of non-native contrasts resulting from native language input, where information about phonetic distributions is less concentrated than in this artificial lexicon. It is likely that distributional information also has the converse effect on perception. For certain native language contrasts discrimination appears to improve between infancy and adulthood (Aslin et al., 1981; Polka et al., 2001). A distribution-sensitive mechanism predicts that discrimination can be enhanced by exposure to a multimodal distribution of speech sounds.

Sensitivity to probabilistic patterns in the input is not restricted to infancy, but contributes to learning throughout the lifetime (Maye, 2000; Maye & Gerken, 2000; Saffran, 2001; Saffran, Newport, & Aslin, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997), and is present in both humans and non-human animals (Hauser, Newport, & Aslin, 2001; Kluender et al., 1998). In addition to its probable role in speech perception, this sensitivity contributes to word segmentation (Saffran, 2001; Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996) and the acquisition of constraints on speech sound sequences (Jusczyk et al., 1994; Zamuner, 2001) and grammatical structure (Gómez & Gerken, 1999; Saffran, 2001), as well as to non-linguistic tasks, such as visual discrimination of textures (Chubb, Econopouly, & Landy, 1994; Julesz, Gilbert, Shepp, & Frisch, 1973) and segmentation of tone sequences (Saffran et al., 1999).

The broad generality of a statistical learning mechanism suggests that it may also be present in younger infants. If this is the case, then performance in the present experiment should not be restricted to these age groups; rather, we would expect younger infants to perform similarly to 6- and 8-month-olds. If this prediction bears out, the question arises of why infant speech perception does not exhibit native language effects until well into the first year of life? We suggest that the latency of changes in speech perception is governed by the quality of infants' statistical representations for speech sounds in the input. In the real world, as opposed to this artificial paradigm, it may take 10 months of exposure to a language for an infant to build up dense enough representations of consonants to differentiate between unimodal and multimodal distributions. The fact that changes in vowel perception occur somewhat earlier (around 6 months) supports this explanation: since there are fewer vowel categories in any language than consonant categories (e.g. Standard American English has ten vowels and 24 consonants), and vowels make up the majority of the speech signal, we would expect distributional representations to develop earlier for vowels than consonants.

In summary, this study demonstrates that infants of 6 and 8 months of age can harness a domain-general learning mechanism, stochastic learning, to facilitate acquisition of a domain-specific system. Infants are able to use distributional information in input speech to detect phonetic category structure. Evidence of the operation of this powerful learning mechanism in the service of phonetic categorization provides our first insights into just what the mechanism might be that allows infants, prior to acquisition of words or grammatical structure, to use the acoustic variability inherent in speech to discern the phonetic organization of the native language.

Acknowledgements

This research was supported by a University of Arizona SBSRI graduate student research grant to J. Maye, a Canada Research Chair and Canadian NSERC research grant (RGP 1103) to J. Werker, and an NSF grant (#SBR9696072) to L. Gerken. The research was conducted in part while the first author was a participant in the Junior Scholars in Residence program sponsored by the Peter Wall Institute of Advanced Studies at the University of British Colombia. We are grateful to Karla Kaun, Fiona Scott, and Elizabeth Job for their assistance in recruiting and testing participants, and to Richard Aslin, Elissa Newport, Athena Vouloumanos, Daniel Weiss, and three anonymous reviewers for their comments on earlier versions of this paper.

References

Aslin, R. N., Jusczyk, P. W., & Pisoni, D. B. (1998). Speech and auditory processing during infancy: constraints on and precursors to language. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology: cognition, perception, and language* (pp. 147–254), 2. New York: Wiley.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: new findings and implications for the effects of early experience. *Child Development*, *52*, 1135–1145.

Best, C. T. (1995). Learning to perceive the sound pattern of English. In C. Rovee-Collier & L. Lipsitt, *Advances in infancy research* (pp. 217–304). Norwood, NJ: Ablex.

Best, C. T., & Jones, C. (1998). Stimulus-alternation preference procedure to test infant speech discrimination. *Infant Behavior and Development*, *21* (Special Issue), 295.

Best, C. T., McRoberts, G. W., LaFleur, R., & Silver-Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two nonnative speech contrasts. *Infant Behavior and Development*, *18*, 339–350.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology*, *14*, 345–360.

Chubb, C., Econopouly, J., & Landy, M. S. (1994). Histogram contrast analysis and the visual segregation of IID textures. *Journal of the Optical Society of America A*, *11*, 2350–2374.

Gómez, R. L., & Gerken, L. (1999). Artificial language learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109–135.

Guenther, F. H., & Gjaja, M. N. (1996). The perceptual magnet effect as an emergent property of neural map formation. *Journal of the Acoustical Society of America*, *100*, 1111–1121.

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition*, *78*, B53–B64.

Julesz, B., Gilbert, E. N., Shepp, L. A., & Frisch, H. L. (1973). Inability of humans to discriminate between visual textures that agree in second-order statistics: revisited. *Perception*, *2*, 391–405.

Jusczyk, P. W. (1985). On characterizing the development of speech perception. In J. Mehler & R. Fox, *Neonate cognition: beyond the blooming, buzzing, confusion* (pp. 199–229). Hillsdale, NJ: Erlbaum.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.

Jusczyk, P. W., Bertoncini, J., Bijeljac-Babic, R., Kennedy, L. J., & Mehler, J. (1990). The role of attention in speech perception by young infants. *Cognitive Development*, *5*, 265–286.

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., Luce, P., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, *33*, 630–645.

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. (1993). Early linguistic experience and phonetic perception: implications for theories of developmental speech perception. *Journal of Phonetics*, *21*, 125–139.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*, 606–608.

Lalonde, C. E., & Werker, J. F. (1995). Cognitive influences on cross-language speech perception in infancy. *Infant Behavior and Development*, *18*, 459–475.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 358–368.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, *20*, 384–482.

MacKain, K. S. (1982). Assessing the role of experience on infants' speech discrimination. *Journal of Child Language*, *9*, 527–542.

Magloire, J., & Green, K. P. (1999). A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica*, *56*, 158–185.

Maye, J. (2000). *Learning speech sound categories on the basis of distributional information*. Unpublished doctoral dissertation, University of Arizona, Tucson.

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

Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America*, *102*, 3742–3753.

Pike, K. L. (1947). *Phonemics: a technique for reducing languages to writing*. Ann Arbor, MI: University of Michigan Press.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/~/ð/ discrimination: evidence for a new developmental pattern. *Journal of the Acoustical Society of America*, *109*, 2190–2201.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 421–435.

Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language*, *44*, 493–515.

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

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101–105.

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.

Sundberg, U., & Lacerda, F. (1999). Voice onset time in speech to infants and adults. *Phonetica*, *56*, 186–199.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147–166.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (in press). Infants' ability to learn phonetically similar words: effects of age and vocabulary size. *Infancy*.

Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*, 349–355.

Werker, J. F., & Pegg, J. E. (1992). Infant speech perception and phonological acquisition. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: models, research, implications* (pp. 285–311). Timonium, MD: York Press.

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

Zamuner, T. S. (2001). *Input-based phonological acquisition*. Unpublished doctoral dissertation, University of Arizona, Tucson.