COGNITION

Cognition 87 (2003) B47–B57

www.elsevier.com/locate/cognit

Brief article

A perceptual interference account of acquisition difficulties for non-native phonemes

Paul Iverson[a,*], Patricia K. Kuhl[b], Reiko Akahane-Yamada[c], Eugen Diesch[d], Yoh'ich Tohkura[e], Andreas Kettermann[f], Claudia Siebert[f]

[a]Department of Phonetics and Linguistics, University College London, 4 Stephenson Way, London NW1 2HE, UK
[b]Center for Mind, Brain, and Learning, University of Washington, Seattle, WA, USA
[c]ATR Human Information Science Laboratories, Kyoto, Japan
[d]Department of Cognitive and Behavioral Neuroscience, Central Institute of Mental Health, Mannheim, Germany
[e]NTT Science and Core Technology Laboratory Group, Atsugi-shi, Kanagawa, Japan
[f]Institute of Psychology, Technical University of Berlin, Berlin, Germany

Received 17 May 2002; accepted 11 October 2002

Abstract

This article presents an account of how early language experience can impede the acquisition of non-native phonemes during adulthood. The hypothesis is that early language experience alters relatively low-level perceptual processing, and that these changes interfere with the formation and adaptability of higher-level linguistic representations. Supporting data are presented from an experiment that tested the perception of English /r/ and /l/ by Japanese, German, and American adults. The underlying perceptual spaces for these phonemes were mapped using multidimensional scaling and compared to native-language categorization judgments. The results demonstrate that Japanese adults are most sensitive to an acoustic cue, F2, that is irrelevant to the English /r/-/l/ categorization. German adults, in contrast, have relatively high sensitivity to more critical acoustic cues. The results show how language-specific perceptual processing can alter the relative salience of within- and between-category acoustic variation, and thereby interfere with second language acquisition.

Keywords: Phonetic perception; Language acquisition; Perceptual magnet effect

1. Introduction

Exposure to speech during childhood alters neural organization such that individuals,

* Corresponding author.
E-mail address: paul@phon.ucl.ac.uk (P. Iverson).

0010-0277/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0010-0277(02)00198-1

born capable of learning any language, develop perceptual and cognitive processes that are specialized for their native language. The changes in neural organization are evident when an individual tries to learn a second language as an adult. The second-language speech can be difficult to segment into words and phonemes, different second-language phonemes can sound as if they are the same, and the motor articulations of the second language can be difficult to reproduce.

What causes the transition from a language-general to a language-specific pattern of perception? Early research (e.g. Eimas, Siqueland, Jusczyk, & Vigorito, 1971) demonstrated that infants are born with perceptual sensitivities for speech that parallel some aspects of adult categorical perception (i.e. higher perceptual resolution near phoneme boundaries than within categories; Liberman, Harris, Hoffman, & Griffith, 1957), and can exhibit these perceptual sensitivities even for phonemes that are not used in their parents' language (e.g. Streeter, 1976). Based on these findings, researchers theorized that newborn infants are innately endowed with a universal set of *phonetic feature detectors* that encode speech into linguistic units (Eimas & Corbit, 1973). It was thought that these feature detectors atrophy during development if they are not used, such that adults retain only the feature detectors that were stimulated by their native language (Eimas, 1975).

This early conception of perceptual development has been proven false in at least two ways. First, it now seems that the initial perceptual abilities of infants reflect auditory processing, not innate linguistic structures; animals can exhibit the same perceptual sensitivities (e.g. Kuhl & Miller, 1975). Second, the atrophy hypothesis has been falsified (e.g. Werker & Tees, 1984); adults maintain the ability to distinguish some non-native phonemes to which they have had little exposure (Best, McRoberts, & Sithole, 1988), and lose the ability to distinguish some non-native phonemes that they have heard in the allophonic variation of their native language (see MacKain, 1982). Instead of atrophy, it appears that the perceptual abilities of infants are actively changed to reduce perceptual sensitivity within native phoneme categories (Kuhl, 1992, 1998, 2000; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker, 1994). These perceptual changes can make it difficult for adults to distinguish non-native phonemes, depending on the degree to which the native and non-native phonemes conflict (Best, 1994; Flege, 1995; Harnsberger, 2001).

It is uncertain which levels of processing are altered by language experience. The predominant view has been that auditory processing is unaffected by language exposure, and that the observed changes in perception are due to higher-level linguistic processes, such as phonological encoding (Best, 1994) or speech-specific selective attention (Pisoni, Lively, & Logan, 1994). The central fact supporting this view is that adults can acquire, with varying success, even the most difficult non-native phonemes if they receive training (see Pisoni et al., 1994; Rvachew & Jamieson, 1995). In addition, adults can, at least within some experimental tasks, exhibit the ability to discriminate acoustic differences between non-native phonemes that they cannot correctly categorize linguistically (Werker & Tees, 1984). It thus seems clear that adults do not permanently lose the auditory resolution necessary to distinguish non-native speech sounds.

More recent evidence, however, suggests that language exposure may indeed affect auditory processing. Neurophysiological studies with adults (Näätänen et al., 1997; Sharma & Dorman, 2000; Winkler et al., 1999; Zhang, Kuhl, Imada, Kotani, & Pruitt,

2001) have found that language-specific perceptual sensitivities are present in the mismatch-negativity (MMN) event-related potential as well as its magnetic equivalent. These measures are thought to reflect the resolution of pre-attentive auditory processing, and the localization of these language-specific effects to the auditory cortex has been supported using MEG and fMRI (Guenther, Nieto-Castanon, Tourville, & Ghosh, 2000; see Näätänen, 2001 for a review). Developmental studies, using behavioral measures (Kuhl et al., 1992) as well as MMN (Cheour et al., 1998; Kuhl & Coffey-Corina, 2001), have demonstrated that infants exhibit language-specific perceptual sensitivities for phonetic units between 6 and 12 months of age, prior to the age that word meanings are thought to be acquired. Kuhl (1994, 1998, 2000) has proposed that the mapping between acoustics and perception becomes warped based on the statistical distribution of speech sounds in the infant's ambient language, reducing perceptual sensitivity near distribution peaks or prototypes (Iverson & Kuhl, 1995, 1996, 2000; Kuhl, 1991; Kuhl et al., 1992) and facilitating later stages of language acquisition. The changes in perception due to language experience are almost certainly speech-specific, and thus can be considered, by definition, to be phonetic rather than purely auditory (see Diesch, Biermann, & Luce, 1998; Näätänen, 2001; Sharma & Dorman, 2000; Winkler et al., 1999). The current evidence thus suggests that the perceptual changes due to language experience occur at an early phonetic or late auditory level, prior to the recognition or categorization of speech in terms of higher-level linguistic units.

The aim of the present study was to reconcile these views, to understand how language experience could alter relatively low levels of processing without permanently blocking the ability to acquire non-native phonemes during adulthood. The experiments specifically examined the English /r/-/l/ contrast. Three groups of listeners with varying linguistic backgrounds were tested: Japanese speakers, who have a well-documented difficulty in acquiring English /r/ and /l/ (e.g. Goto, 1971; Miyawaki et al., 1975); German speakers, who have no marked difficulty acquiring this phonetic contrast, but do not have a phoneme like English /r/; and American English speakers, for whom these are native phonemes. Multidimensional scaling (MDS) and signal detection theory were used to map the perceptual spaces underlying these phonemes. Identification and goodness judgments were used to characterize how these sounds related to each listener's native phoneme categories. The results were analyzed to see how the perceptual spaces for English /r/ and /l/ were altered by native-language exposure and to hypothesize how the structure of the perceptual spaces could affect the acquisition of English /r/ and /l/.

2. Method

2.1. Participants

Twenty-four native speakers of Japanese were tested in Tokyo, 12 native speakers of German were tested in Berlin, and 19 native speakers of English[1] were tested in Seattle.

[1] These native English speakers completed only the discrimination task. The similarity, goodness, and identification data for English speakers reported here were collected in a previous study (Iverson & Kuhl, 1996).

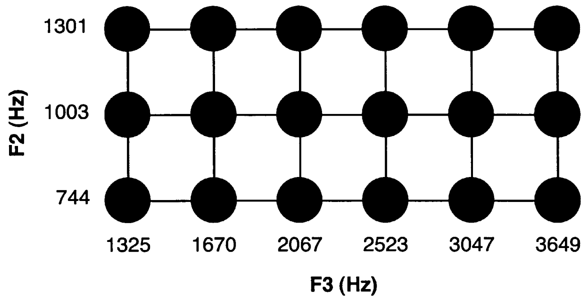

Fig. 1. The formant frequencies for the English /ra/ and /la/ stimuli used in this study (from Iverson & Kuhl, 1996). The stimuli varied in terms of the second (F2) and third (F3) formants during the initial consonant. The formant frequencies were spaced equally on the mel scale (Stevens et al., 1937).

All participants had been raised in monolingual homes, and had not learned other languages prior to attending school. All Japanese and German speakers had received English language instruction in school, for an average duration of 7.2 years for German speakers and 7.5 years for Japanese speakers. No Japanese speakers had lived abroad. German speakers had spent an average of 2.6 months visiting English-speaking countries.

2.2. Stimuli

The stimuli were 18 /ra/ and /la/ tokens from a previous study (Iverson & Kuhl, 1996). They were synthesized (Klatt & Klatt, 1990) to model natural citation speech recordings of an adult female native American English speaker. As shown in Fig. 1, the stimuli varied in the frequencies of the second (F2) and third (F3) formants during the consonant closure, to create a two-dimensional stimulus grid with the frequencies equally spaced on the mel scale (Stevens, Volkmann, & Newman, 1937). The stimuli were identical in all other respects. During the closure, F1 was 351 Hz and F4 was 4512 Hz; the bandwidths were 200, 100, 150, and 100 Hz, to match the formant amplitudes of the natural recordings. During the vowel, the formant frequencies for F1–F4 were 796, 1221, 2973, and 4512; the bandwidths were 200, 100, 150, and 400.

2.3. Procedure

2.3.1. Identification and goodness

Participants identified each stimulus in terms of their own native-language phonemes, and rated whether the stimulus was a good exemplar of that category on a scale from 1 (bad) to 7 (good). Each session had a practice block of 18 trials (one trial for each stimulus), and an experimental block of 36 trials (two trials for each stimulus).

2.3.2. Similarity scaling

Participants rated the acoustic similarity of stimulus pairs (350 ms ISI) on a scale from 1 (dissimilar) to 7 (similar). Each session had a practice block of 36 trials (randomly selected pairs), and an experimental block of 306 trials (every possible pair of the 18 stimuli, in both presentation orders, with no stimulus being paired with itself). The results were analyzed using MDS (Kruskal, 1964) to map the perceptual space underlying these stimuli. The procedure placed the stimuli in a two-dimensional Euclidean space such that perceptually similar stimuli were placed close together and dissimilar stimuli were placed far apart, maximizing the fit between distances and similarity using an inverse monotonic function.

2.3.3. Discrimination

Participants heard pairs of stimuli (250 ms ISI) and judged whether they were the same or different. Half were *same* pairs, containing two repetitions of the same stimulus. Half were *different* pairs, containing stimuli that varied in F3 along the stimulus series (all had the same F2, 1003 Hz). Subjects completed a practice block of 20 trials (two same and two different for each stimulus pair), and an experimental block of 480 trials (48 same and 48 different for each stimulus pair). The results were analyzed using a differencing model of signal detection theory (Macmillan & Creelman, 1991) to calculate perceptual sensitivity (i.e. the ability to detect a stimulus difference) for each stimulus pair.

3. Results

As shown in Fig. 2, the underlying perceptual spaces for these stimuli were affected by language experience. The MDS analyses modeled at least 90% of the variance in each set. The order of stimuli along each dimension matched their formant frequencies (i.e. the horizontal dimension corresponded to F3 and the vertical to F2), but the distances between neighboring stimuli were distorted relative to the equally spaced stimulus grid. American listeners were most sensitive to differences in F3 that distinguished /r/ and /l/ (stretching the MDS space in the middle of the F3 dimension), and were less sensitive to acoustic differences among excellent exemplars of these phonetic categories (shrinking the MDS space near the ends of the F3 dimension). In contrast, Japanese listeners were more sensitive to variation in F2 than F3, and had no marked stretching of the perceptual space in the middle of the F3 dimension. German listeners had MDS solutions for these stimuli that were more related to those of American listeners; there was stretching of the perceptual space in the middle of the F3 series and shrinking near the lowest and highest values of F3.

The discrimination results (Fig. 3) agreed with the MDS solutions. American listeners had higher sensitivity near the /r/-/l/ boundary than within each category. Japanese listeners had increasing sensitivity as F3 increased, with no peak in sensitivity at the English /r/-/l/ boundary. Statistical comparisons of the American and Japanese data revealed that Japanese listeners had significantly higher sensitivity within the English /r/ ($t(41) = -3.3$, $P = 0.002$) and /l/ ($t(41) = -5.1$, $P < 0.001$) categories; American listeners had significantly higher sensitivity at the phoneme boundary ($t(41) = 3.5$, $P = 0.001$). German listeners had sensitivity patterns that were similar to those of Americans. Statistical comparisons of the American and German data revealed that German listeners had significantly higher sensitivity within the English /r/ category ($t(29) = -2.1$, $P = 0.046$), but sensitivity was not significantly different within the /l/ category ($t(29) = -1.3$, $P = 0.214$), or at the boundary ($t(29) = -0.76$, $P = 0.450$).

The identification and goodness results (Fig. 2) provide clues about the causes of these differences in perceptual sensitivity. Japanese adults assimilated these stimuli into their /r/ category, but the strength of the assimilation varied with F2 frequency (stimuli with lower F2 frequencies began to sound like /w/). Their greater sensitivity to F2 may have been caused by this category assimilation (see Best, 1994), although there was no clear evidence for a perceptual magnet effect (Iverson & Kuhl, 1996; Kuhl, 1991) because no stimuli were good exemplars of Japanese phonemes. German adults heard these stimuli as good exemplars of their /l/ and as poor exemplars of their uvular fricative. Their category

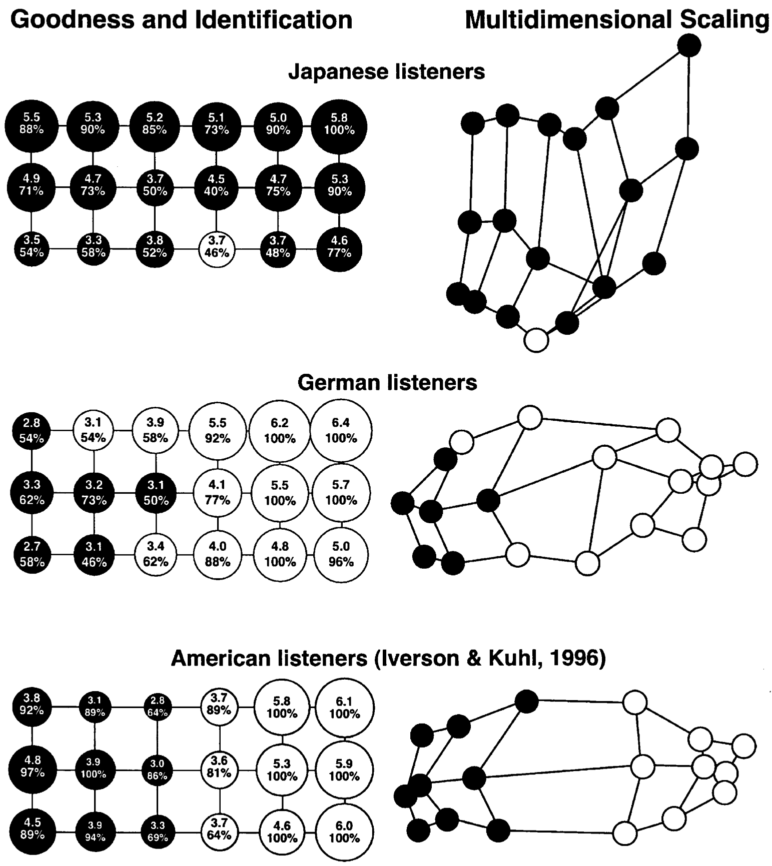

Fig. 2. Goodness, identification, and MDS solutions for Japanese, German, and American listeners. In the goodness and identification graphs, the size of the circle indicates the average goodness rating (larger circle for higher goodness), and the shading indicates the most frequently chosen phonetic category for that stimulus in terms of the listeners' native language (black for the respective /r/ sounds in Japanese, German, and English; white for /l/ sounds in German and English, and /w/ in Japanese). The numbers within the circles list the average goodness ratings and the identification percentages for the predominant phonetic category. The MDS solutions are geometric representations of the average similarity ratings for these stimuli. The lines between stimuli reflect their spacing in the stimulus grid (see Fig. 1), and the length of the lines reflects perceptual sensitivities for these acoustic differences (perceptually similar stimuli are placed close together; perceptually dissimilar stimuli are placed far apart).

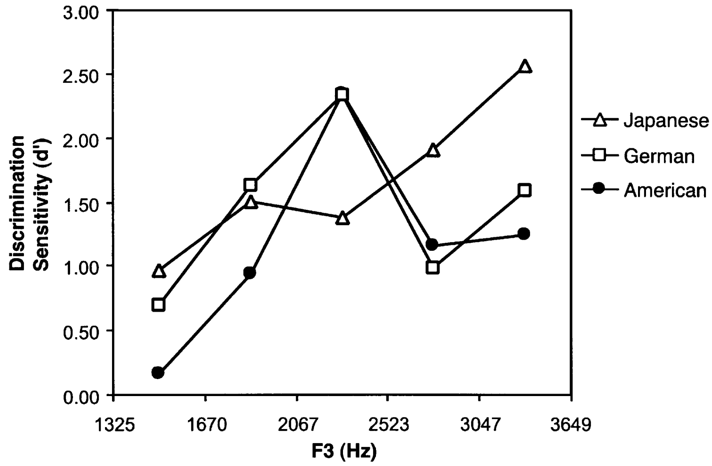

Fig. 3. Discrimination sensitivity along the middle horizontal vector of the stimulus grid (see Fig. 1) for Japanese, German, and American listeners. The stimuli varied in F3 and had the same F2 frequency.

boundary was less sharp than for American listeners, and they likewise had broader stretching of the perceptual space in the middle of the F3 dimension. There was evidence for a perceptual magnet effect for German /l/; the best exemplar had a higher F2 than did those of American English listeners, corresponding with a shrinking of the perceptual space at higher F2 values.

4. General discussion

It had long been known (e.g. Goto, 1971; Miyawaki et al., 1975) that Japanese adults are relatively insensitive to F3 differences near the English /r/-/l/ boundary, and the present study replicates this fact. However, the results also demonstrate that it is not the case that Japanese adults are entirely unable to discern differences among these phonemes; Japanese adults are more sensitive to variation along a dimension, F2, that is mostly irrelevant to the English /r/-/l/ categorization, and can have high within-category sensitivity to F3 differences, particularly for /l/. Japanese adults thus have a distorted perceptual space for these phonemes, but not a total lack of perceptual sensitivity (i.e. the critical acoustic variation is not filtered out; see Trubetzkoy, 1958/1969). In contrast, the perceptual space for German adults approximates that of American English speakers, being stretched near the English /r/-/l/ boundary and shrunk within each category.

The perceptual spaces of Japanese adults are thus miss-tuned for acquiring the English /r/-/l/ contrast, making acoustic variation that is irrelevant to categorization more salient than the critical differences in F3. These perceptual spaces are hypothesized to interfere with acquisition in at least two ways. First, Japanese adults could be prone to form erroneous category representations for /r/ and /l/, by relying on acoustic cues, such as F2, that are perceptually salient but not reliable or robust for categorization. In fact, Yamada (1995) has found that Japanese adults form these types of erroneous category representations, giving more weight to secondary acoustic cues than to F3. Second, high sensitivity to irrelevant acoustic differences could create problems akin to Garner interference (Garner, 1974), requiring focused attention and longer processing times to detect the critical F3 differences, even for Japanese adults who have formed correct category representations based on F3. These hypotheses extend Kuhl's (1992, 1994, 2000) Native Language Magnet Theory to further detail how perceptual warpings affect speech perception.

It is telling that the most successful training procedures for teaching English /r/ and /l/ to Japanese adults have involved multi-talker high-variability stimulus sets (e.g. Logan, Lively, & Pisoni, 1991). Training procedures involving smaller stimulus sets are easier to learn, but do not readily generalize to new stimuli (Strange & Dittmann, 1984). Training with larger stimulus sets may generalize better, because the variability provides information about which cues are most robust and trains individuals to ignore irrelevant variation. Likewise, enhancement of acoustic cues, in infant-directed (Kuhl et al., 1997) or signal-processed (e.g. McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002; Zhang et al., 2001) speech, may improve early stages of learning by making the most robust acoustic cues more salient.

The results from German listeners are compatible with this interpretation of the Japanese results. German listeners may not have a marked difficulty acquiring English /r/ and /l/ because their perceptual sensitivities, although subtly different from those of American English speakers, do not interfere with learning this categorization (see Hallé, Best, & Levitt, 1999 for related data on French speakers). German listeners retain a sensitivity peak for F3 differences near the English /r/-/l/ boundary that all individuals likely have at birth (Eimas, 1975), and do not acquire increased sensitivity to F2 variation. They thus find the acoustic differences that are critical to the English /r/-/l/ categorization to be more salient than irrelevant differences, which facilitates category learning.

The present results do not provide additional proof that perceptual changes resulting from language experience occur at an auditory or early-phonetic level, but they help show how such an account is plausible. Previous theorists have suggested that the perceptual changes that occur as an adult learns a second language result from modifications of higher-level linguistic processes (e.g. Best, 1994; Flege, 1995; Pisoni et al., 1994). The current theory is in agreement with this view that higher-level processes are changed, but further asserts that lower-level perceptual processes can interfere with the adaptability of these higher-level processes. It is likely too that lower-level perceptual processes are modified during adult second language acquisition (see Tremblay, Kraus, Carrell, & McGee, 1997), due, for example, to top-down effects of category learning (Guenther, Husain, Cohen, & Shinn-Cunningham, 1999; Guenther et al., 2000) or changes in the strength of connections between peripheral and cortical neurons (McCandliss et al., 2002).

This theory is compatible with the hypothesis (e.g. Flege, 1995; Kuhl, 1994, 1998, 2000; McCandliss et al., 2002) that the decline of language acquisition abilities from infancy through adulthood results from the interference of previous experience on perception. The changes in perceptual processing due to language experience may be self-reinforcing, because initial exposure to language will alter how all subsequent speech sounds are perceived. That is, a loss of perceptual sensitivity for non-native phoneme contrasts may be difficult to reverse in adulthood, because perceptual resolution would have become reduced for the types of acoustic variation that are most critical for training. Adults may thus become *neurally committed* (Kuhl, 2000) to a particular network structure for analyzing language, due more to this type of self-reinforcing perceptual interference than to any age-related biological limitations.

Acknowledgements

This research was supported by grants to P. Kuhl (NIH HD37954 and HD35465), by the William P. and Ruth Gerberding University Professorship Fund, and by the Talaris Research Institute and the Apex Foundation funded by Bruce and Jolene McCaw. We are grateful to Andrew Faulkner for comments on this manuscript.

References

Best, C. T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: the transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.

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: Human Perception and Performance*, *14* (3), 345–360.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, *1* (5), 351–353.

Diesch, E., Biermann, S., & Luce, T. (1998). The magnetic mismatch field elicited by words and phonological non-words. *NeuroReport*, *16*, 455–460.

Eimas, P. D. (1975). Auditory and phonetic coding of the cues for speech: discrimination of the [r-l] distinction by young infants. *Perception & Psychophysics*, *18* (5), 341–347.

Eimas, P. D., & Corbit, J. D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, *4*, 99–109.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171* (968), 303–306.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (Ed.), *Speech perception and language experience: issues in cross-language research* (pp. 233–277). Baltimore. MD: York Press.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Lawrence Erlbaum Associates distributed by Halsted Press, New York.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia*, *9*, 317–323.

Guenther, F. H., Husain, F. T., Cohen, M. A., & Shinn-Cunningham, B. G. (1999). Effects of categorization and discrimination training on auditory perceptual space. *Journal of the Acoustical Society of America*, *106*, 2900–2912.

Guenther, F. H., Nieto-Castanon, A., Tourville, J. A., & Ghosh, S. S. (2000). *The representation of prototypical and non-prototypical vowels in peri-sylvian cortical areas*. Poster presented at the 30th annual meeting of the Society for Neuroscience, New Orleans, LA.

Hallé, P. A., Best, C. T., & Levitt, A. (1999). Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics*, *27*, 281–306.

Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *Journal of the Acoustical Society of America*, *110* (1), 489–503.

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America*, *97*, 553–562.

Iverson, P., & Kuhl, P. K. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *Journal of the Acoustical Society of America*, *99*, 1130–1140.

Iverson, P., & Kuhl, P. K. (2000). Perceptual magnet and phoneme boundary effects in speech perception: do they arise from a common mechanism? *Perception & Psychophysics*, *62*, 874–886.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*, 820–857.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrica*, *29*, 1–27.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*, 93–107.

Kuhl, P. K. (1992). Psychoacoustics and speech perception: internal standards, perceptual anchors, and prototypes. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 293–332). Washington, DC: American Psychological Association.

Kuhl, P. K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology*, *4*, 812–822.

Kuhl, P. K. (1998). The development of speech and language. In T. C. Carew, R. Menzel & C. J. Shatz (Eds.), *Mechanistic relationships between development and learning* (pp. 53–73). New York: Wiley.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences USA*, *97*, 11850–11857.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*, 684–686.

Kuhl, P. K., & Coffey-Corina, S. (2001). *Language and the developing brain: changes in ERPs as a function of linguistic experience*. Paper presented at the annual meeting of the Cognitive Neuroscience Society, New York.

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science*, *190*, 69–72.

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.

Liberman, A., 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.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: a first report. *Journal of the Acoustical Society of America*, *89*, 874–886.

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

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: a user's guide*. New York: Cambridge University Press.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the r-l contrast to Japanese adults: predictions of a hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, and Behavioral Neuroscience*, *2*, 89–108.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J., & Fujimura, O. (1975). An effect of language experience: the discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception & Psychophysics*, *18*, 331–340.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, *38*, 1–21.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*, 432–434.

Pisoni, D. B., Lively, S. E., & Logan, J. S. (1994). Perceptual learning of nonnative speech contrasts: implications for theories of speech perception. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: the transition from speech sounds to spoken words* (pp. 121–166). Cambridge, MA: MIT Press.

Rvachew, S., & Jamieson, D. G. (1995). Learning new speech contrasts: evidence from adults learning a second language and children with speech disorders. In W. Strange (Ed.), *Speech perception and language experience: issues in cross-language research* (pp. 411–432). Baltimore, MD: York Press.

Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *Journal of the Acoustical Society of America*, *107*, 2697–2703.

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America*, *8*, 185–190.

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics*, *36*, 131–145.

Streeter, L. A. (1976). Language perception of 2-month old infants shows effects of both innate mechanisms and experience. *Nature*, *259*, 39–41.

Tremblay, K., Kraus, N., Carrell, T. D., & McGee, T. (1997). Central auditory system plasticity: generalization to novel stimuli following listening training. *Journal of the Acoustical Society of America*, *102*, 3762–3773.

Trubetzkoy, N. S. (1969). *Principles of phonology* (C. A. M. Baltaxe, Trans.). Berkeley, CA: University of California Press. (Original work published 1958)

Werker, J. F. (1994). Cross-language speech perception: development change does not involve loss. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: the transition from speech sounds to spoken words* (pp. 93–120). Cambridge, MA: MIT Press.

Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*, 1866–1878.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csepe, V., Ilmoniemi, R. J., & Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, *36*, 638–642.

Yamada, R. A. (1995). Age and acquisition of second language speech sounds: perception of American English /r/ and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech perception and language experience: issues in cross-language research* (pp. 305–320). Baltimore, MD: York Press.

Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Pruitt, J. (2001). Brain plasticity in behavioral and neuromagnetic measures: a perceptual training study. *Journal of the Acoustical Society of America*, *110*, 2687.