Speech, Hearing and Language: work in progress

Volume 13

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

Paul IVERSON, Patricia K. KUHL, Reiko AKAHANE-YAMADA, Eugen DIESCH, Yoh'ich TOHKURA, Andreas KETTERMANN, and Claudia SIEBERT

Department of Phonetics and Linguistics
UNIVERSITY COLLEGE LONDON

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

Paul IVERSON, Patricia K. KUHL, Reiko AKAHANE-YAMADA, Eugen DIESCH, Yoh'ich TOHKURA, Andreas KETTERMANN, and Claudia SIEBERT

Abstract
This article presents a new 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 low-level changes interfere with the formation and adaptability of higher-level linguistic representations. Supporting data is 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 were 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.

Introduction
Exposure to speech during childhood alters neural organization such that individuals, born capable of learning any language, develop perceptual and cognitive processes that are specialized for their own native language. The changes in neural organization are particularly 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 phonemes in the second language 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., high sensitivity to acoustic differences that cross phoneme boundaries and low sensitivity to within-category differences; Liberman, Harris, Hoffman, & Griffith, 1957), and that infants can exhibit these perceptual sensitivities even for phoneme boundaries 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 eventually retain only the phonetic 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 (c.f., Diehl, 1981; Nittrouer, 2001). First, it now seems that the perceptual abilities of infants are due to auditory processing, not innate linguistic structures; animals have been shown to exhibit the same patterns of perceptual sensitivity (e.g., Kuhl & Miller, 1975). Second, the atrophy hypothesis has been falsified (e.g., Werker

& Tees, 1984); adults have been shown to 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 to which they have been exposed in the allophonic variation of their native language (see MacKain, 1982). Instead of atrophy, it appears that the initial perceptual abilities of infants are actively changed by language exposure, such that individuals acquire reduced perceptual sensitivity to within-category acoustic variation for native phonemes (Kuhl, 2000; Werker, 1994). These perceptual changes can interfere with the ability of adults to distinguish non-native phonemes (Kuhl, 1998, 2000), and the degree of this interference is related to the extent that 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 and completely lose the auditory resolution necessary to distinguish non-native speech sounds.

More recent evidence, however, suggests that language exposure may indeed affect auditory processing. Electrophysiological studies with adults (Näätänen et al., 1997; Sharma & Dorman, 2000; Winkler et al., 1999) have found that language-specific perceptual sensitivities are present in the mismatch-negativity (*MMN*) event-related potential and its magnetic equivalent (*MMNm*). Both measures are thought to reflect the resolution of pre-attentive auditory processing to stimulus differences; they can be obtained from individuals who are not paying attention to the stimuli (e.g., when a person is watching a movie, sleeping, or in a coma), and the sources of MMNm have been localized to the auditory cortex (see Näätänen, 2001 for a review). Developmental studies, using behavioral measures (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992) as well as MMN (Cheour et al., 1998), have demonstrated that infants exhibit language-specific perceptual sensitivities for vowels by 6-12 months of age, prior to the age that word meanings are thought to be acquired. Kuhl (1998; 2000) has theorized that the mapping between acoustics and perception becomes warped based on the statistical distribution of speech sounds in the infant's ambient language (see Guenther & Gjaja, 1996 for a possible mechanism), 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; Iverson & Kuhl, 2000; Sharma & Dorman, 2000; Winkler et al., 1999). Moreover, Näätänen (2001) has recently argued that MMNm can be produced by both auditory and early phonetic processing, and that the more phonetic processes tend to be left-lateralized, located near Wernicke's area. The current developmental and electrophysiological evidence thus both suggest 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 is to reconcile these current views of language acquisition, 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 non-native phonetic contrast; 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 related to each listeners' native phoneme categories, and to hypothesize how the structure of the perceptual spaces could affect the acquisition of English /r/ and /l/.

1. Method

1.1 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. All participants had been raised in monolingual homes; Japanese and German speakers had received English language instruction in school.

1.2 Stimuli

The stimuli were eighteen synthesized /ra/ and /la/ tokens that were modeled from clear citation speech recordings of an adult female native English speaker (see Iverson & Kuhl, 1996). As shown in Figure 1, the stimuli varied in the starting frequencies of the second (*F2*) and third (*F3*) formants, to create a 2-dimensional stimulus grid. The stimuli were identical in all other respects.

[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).

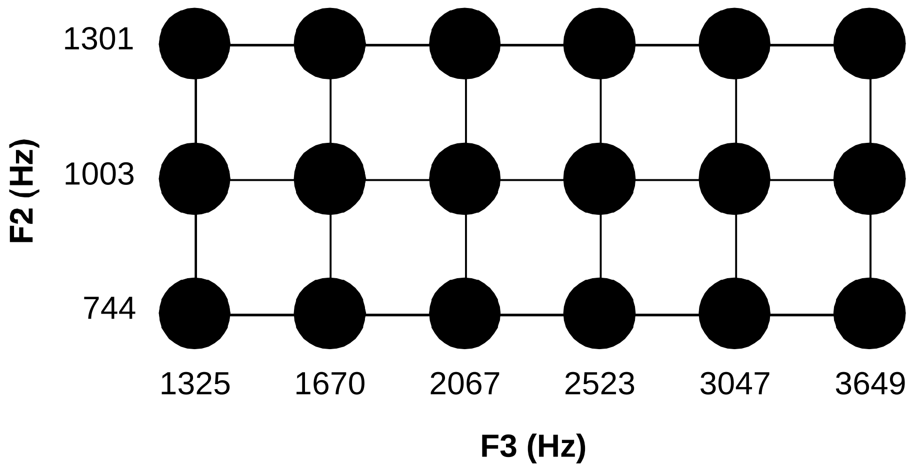

Figure 1: *Formant frequencies for the English /ra/ and /la/ stimuli used in this study (from Iverson & Kuhl, 1996). The stimuli varied in terms of the second (F2) and third (F3) formants during the initial consonant. The formant frequencies were spaced equally using the Mel scale (a perceptual scale based on the mapping between frequency and pitch).*

1.3 Procedure

Identification and goodness. Participants identified each stimulus in terms of their own native-language phonemes, and then rated whether the stimulus was a good exemplar of that category on a scale from 1 (bad) to 7 (good). Each session started with a practice block of 18 trials (1 trial for each stimulus, presented in a random order), and was followed by an experimental block of 36 trials (2 trials for each stimulus, presented in a random order).

Similarity scaling. Participants rated the acoustic similarity of each pair of the 18 stimuli on a scale from 1 (dissimilar) to 7 (similar). Each session started with a practice block of 36 trials (randomly selected pairs), and was followed by 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 nonmetric MDS (Kruskal, 1964), to geometrically model the perceptual space underlying these tokens.

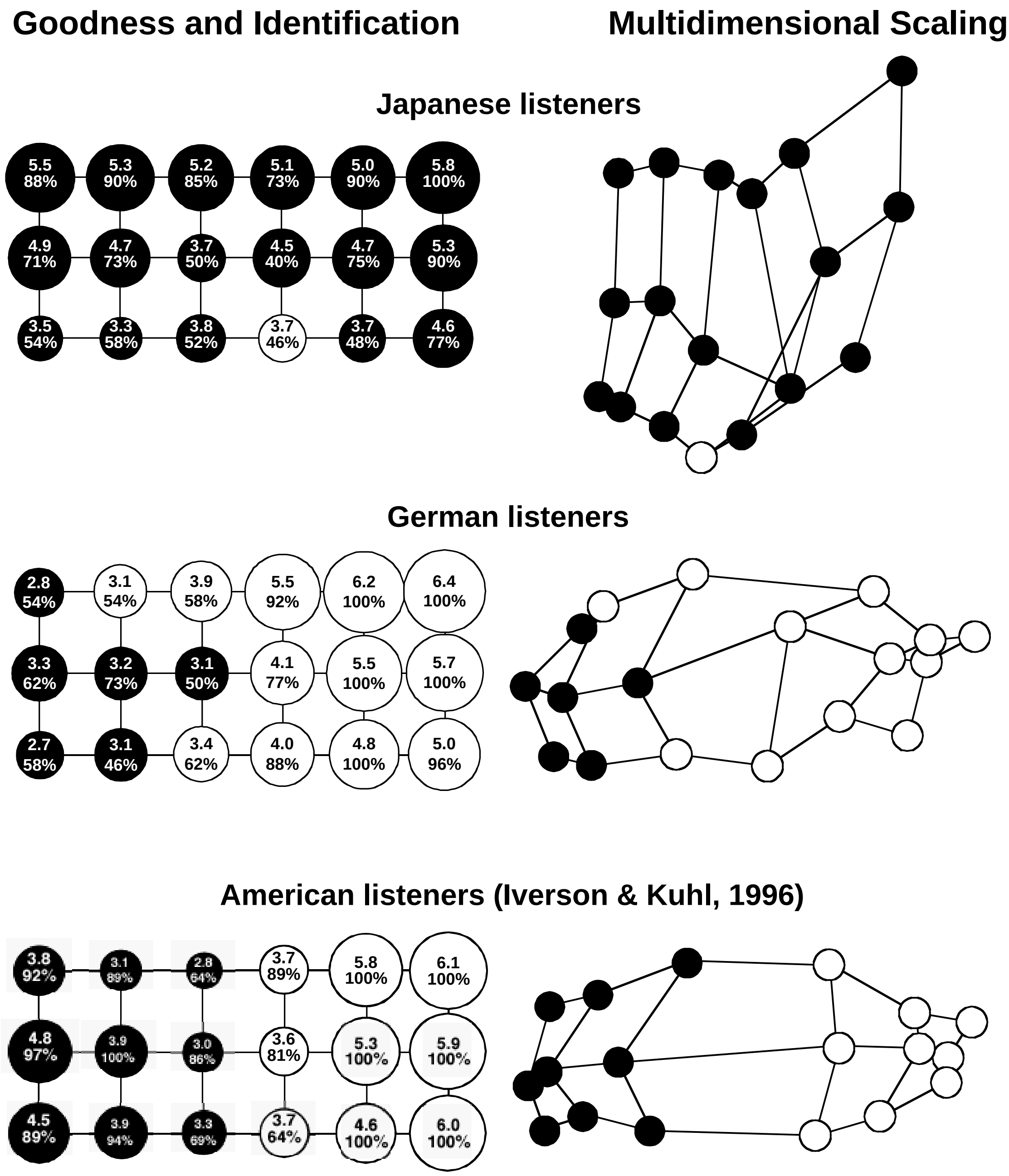

Figure 2: *Goodness, identification, and MDS solutions for Japanese, German, and American listeners. In the goodness and identification graphs, the size of the circle indicates the average goodness rating (larger circle for higher goodness), and the shading indicates the predominant phonetic category (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 Figure 1), and the length of the lines reflect perceptual sensitivities for these acoustic differences (perceptually similar stimuli are placed close together; perceptually dissimilar stimuli are placed far apart).*

Discrimination. Participants heard pairs of stimuli, and reported 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 stimuli had the same middle F2 value). Subjects first completed a practice block of 20 trials (2 same and 2 different trials for each stimulus pair), and an experimental block of 480 trials (48 same and 48 different trials for each stimulus pair). The results were analyzed using a differencing model of signal detection theory (Macmillan & Creelman, 1991) to calculate the perceptual distance between each stimulus pair.

Results

As shown in Figure 2, the underlying perceptual spaces for these stimuli were affected by language experience. The MDS analyses fit each set of similarity data into a 2-dimensional space, modeling at least 90% of the variance in each set. The basic ordering of stimuli in the MDS solutions matched their acoustic characteristics (i.e., the horizontal dimension corresponded to F3 and the vertical to F2), but all solutions were distorted relative to the equally spaced stimulus grid (see Figure 1). American listeners were most sensitive to differences in F3 that distinguished /r/ and /l/ (illustrated by the stretching of the MDS space in the middle of the horizontal dimension), and were less sensitive to acoustic differences among excellent exemplars of these phonetic categories (illustrated by the shrinking of the MDS space near stimuli that had received the highest goodness ratings). 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 series (sensitivity to F3 was greater as F3 increased). German listeners had MDS solutions for these stimuli that were more closely 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 (Figure 3) agreed with the MDS solutions. American listeners had high sensitivity in the middle of the F3 series (i.e., at the /r/-/l/ boundary), and poor sensitivity within the /r/ and /l/ categories. Japanese listeners had progressively increasing sensitivity as F3 increased, with no peak in sensitivity at the American English /r/-/l/ boundary. Statistical comparisons of the American and Japanese data revealed that Japanese listeners had significantly higher discrimination accuracy within the English /r/, $t(41) = -3.3$, $p = .002$, and /l/, $t(41) = -5.1$, $p < .001$, categories; American listeners had significantly higher discrimination accuracy at the phoneme boundary, $t(41) = 3.5$, $p = .001$. German listeners had discrimination sensitivity that was similar to American listeners, with the only difference being somewhat higher sensitivity within the /r/ category. Statistical comparisons of the American and German data revealed that German listeners had significantly higher discrimination accuracy within the English /r/ category, $t(29) = -2.1$, $p = .046$; discrimination accuracy was not significantly different within the /l/ category, $t(29) = -1.3$, $p = .214$, or at the phoneme boundary, $t(29) = -0.76$, $p = .45$.

The identification and goodness results (Figure 2) provide clues about the causes of these differences in perceptual sensitivity. Japanese adults mostly 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/); exposure to the Japanese /r/ and /w/ categories may have caused these listeners to have greater sensitivity to F2 among English /r/ and /l/ phonemes. German adults heard these stimuli as good exemplars of their /l/ and as poor exemplars of their uvular fricative. Their category boundary was less sharp than for American listeners, and this corresponded to a somewhat broader stretching of the perceptual space in the middle of the F3 series. There was thus evidence in both cases that these distortions in the underlying perceptual spaces were caused by exposure to the listeners' respective native language phonetic categories.

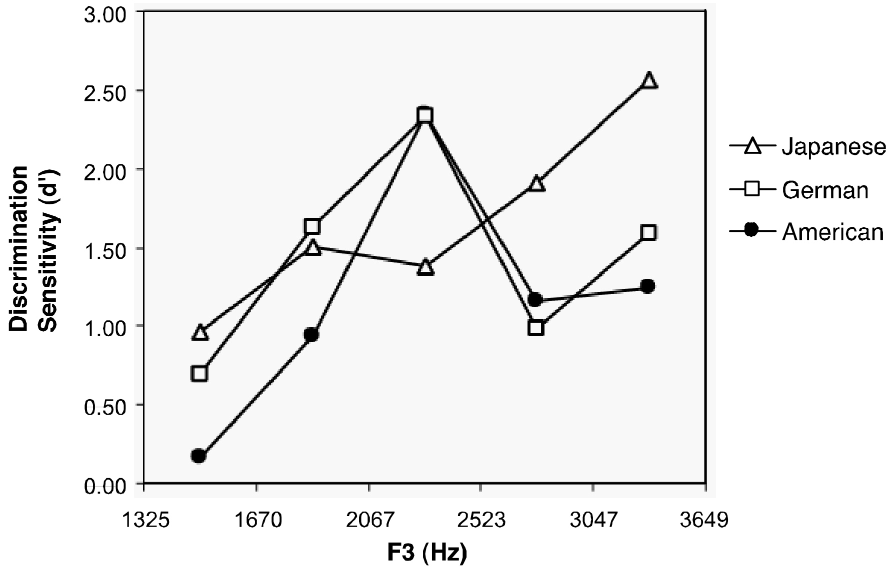

Figure 3: *F3 discrimination sensitivity along for Japanese, German, and American listeners. The stimuli were from the middle horizontal vector of the stimulus grid (i.e., the stimuli varied in F3 and had the same middle F2 frequency). The discrimination sensitivity measure, d', was calculated using signal detection theory, to remove potential effects of response biases.*

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 to acoustic variation. In contrast, the perceptual space for German adults approximates that of American English speakers, being stretched near the English /r/-/l/ boundary and shrunk near the best exemplars of each category.

The perceptual maps of Japanese adults are thus miss-tuned for acquiring the English /r/-/l/ contrast, making acoustic variation which is irrelevant to the English /r/-/l/ categorization more salient than the critical differences in F3. These types of perceptual maps 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 are not reliable or robust enough to accurately recognize these phonemes. 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, the high sensitivity to irrelevant acoustic variation could increase the processing load for making the English /r/-/l/ categorization, even for Japanese adults who have formed category representations based on F3. For example, high sensitivity to irrelevant acoustic differences could create problems akin to Garner interference (Garner, 1974), requiring more focused attention and longer processing times to attend to critical F3 differences even after correct category representations for these phonemes have been learned.

It is telling that the most successful training procedures for teaching English /r/ and /l/ to Japanese adults have involved multi-talker stimulus sets with a high degree of acoustic variability (e.g., Logan, Lively, & Pisoni, 1991). Training procedures involving smaller stimulus sets are easier for Japanese adults to learn, but it is typically difficult for participants to generalize their training to new stimuli (Strange & Dittmann, 1984). Training with larger stimulus sets may generalize better, because participants are less likely to form erroneous and artifactual category representations based on the most perceptually salient cues; such training provides more information about which cues are most robust. In addition, training with larger stimulus sets could teach listeners to more efficiently focus attention on the critical acoustic cues, reducing the processing load required to ignore irrelevant variation.

The results from German listeners are compatible with this interpretation of the Japanese data. German listeners may not have a marked difficulty acquiring English /r/ and /l/ because their perceptual sensitivities do not interfere with learning this categorization. German listeners retain a sensitivity peak for F3 differences near the English /r/-/l/ boundary that they likely had a birth (see 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 are irrelevant differences, which facilitates category learning and decreases attentional demands.

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. In addition, it is likely that that the lower-level perceptual processes are themselves modified during adult second language acquisition (see Tremblay, Kraus, Carrell, & McGee, 1997).

This theory is compatible with Kuhl's hypothesis (1998; 2000) that the critical period for language acquisition (i.e., the decline of language acquisition abilities from infancy through puberty) results more from the interference of previous experience than from age. 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, even though an adult learning a second-language could be exposed to the same acoustic distribution of speech sounds as an infant learning the same language, the auditory distribution of those sounds would be different for adults due to prior perceptual changes. 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 which would have normally been thought to be the 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. The decline in second-language acquisition abilities from childhood through puberty may likewise reflect a progressively stronger neural commitment to one's native language, as changes in perception interfere more with new experience.

Author Note

P. Iverson, Department of Phonetics and Linguistics, University College London; P. K. Kuhl, Department of Speech and Hearing Science, University of Washington; R. Akahane-Yamada, ATR Spoken Language Acquisition Project; E. Diesch, Department of Cognitive and Behavioral Neuroscience, Central Institute of Mental Health Mannheim; Y. Tohkura, NTT Communication Science Research Laboratories; A. Kettermann and C. Siebert, Institute of Psychology, Technical University of Berlin.

Research was supported by grants to PKK (NIH HD37954 and HD35465) and support from the William P. and Ruth Gerberding University Professorship Fund. We are grateful to Andrew Faulkner for comments on this manuscript.

Please direct correspondence to Paul Iverson, Department of Phonetics and Linguistics, University College London, 4 Stephenson Way, London NW1 2HE, England. Electronic mail may be sent to paul@phon.ucl.ac.uk.

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.

Diehl, R. L. (1981). Feature detectors for speech: a critical reappraisal. *Psychological Bulletin, 89*(1), 1-18.

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: York Press.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, Md.,: L. 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*(3), 317-323.

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*(2 Pt 1), 1111-1121.

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*(1), 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*(2), 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*(4), 874-886.

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*(2), 93-107.

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, 97*(22), 11850-11857.

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science, 190*(4209), 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*(5044), 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*(2), 874-886.

MacKain, K. S. (1982). Assessing the role of experience on infants' speech discrimination. *Journal of Child Language, 9*(3), 527-542.

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

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*(6615), 432-434.

Nittrouer, S. (2001). Challenging the notion of innate phonetic boundaries. *Journal of the Acoustical Society of America, 110*(3 Pt 1), 1598-1605.

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: York Press.

Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *Journal of the Acoustical Society of America, 107*(5 Pt 1), 2697-2703.

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics, 36*(2), 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*(6), 3762-3773.

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*(6), 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*(5), 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: York Press.