Cognitive, Affective, & Behavioral Neuroscience
2002, 2 (2), 89-108

Success and failure in teaching the [r]–[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception

BRUCE D. McCANDLISS and JULIE A. FIEZ
Center for the Neural Basis of Cognition, Pittsburgh, Pennsylvania and University of Pittsburgh, Pittsburgh, Pennsylvania

ATHANASSIOS PROTOPAPAS
Scientific Learning Corporation, Oakland, California

MARY CONWAY
Center for the Neural Basis of Cognition, Pittsburgh, Pennsylvania and University of Pittsburgh, Pittsburgh, Pennsylvania

and

JAMES L. McCLELLAND
Center for the Neural Basis of Cognition, Pittsburgh, Pennsylvania and Carnegie Mellon University, Pittsburgh, Pennsylvania

A Hebbian model of learning predicts that adults may be able to acquire a nonnative speech contrast if they are trained with stimuli that are exaggerated to make them perceptually distinct. To test these ideas, we asked Japanese adults to identify contrasting [r]–[l] stimuli (e.g., *rock–lock*) in two training conditions. In the adaptive condition, the [r]–[l] contrast was exaggerated at first and then adjusted to maintain accurate identification. In the fixed condition, a fixed pair of stimuli were used that were distinguishable by native English speakers but difficult for the Japanese learners to discriminate. To examine whether feedback contributes to learning, we ran separate groups with and without feedback in the fixed and the adaptive conditions. Without feedback, 3 days of adaptive training produced substantial improvements, but 3 days of fixed training produced no benefit relative to control, consistent with the Hebbian account. With feedback, both fixed and adaptive training led to robust improvements, and the benefit of training transferred to a second continuum (e.g., *road–load*). The results are consistent with Hebbian models that are augmented to be sensitive to feedback.

Adults are good at learning many things, but there are some exceptions. It can be very difficult for adults to learn to perceive certain speech contrasts that their native language does not use. A classic case in point is the persistent difficulty that native speakers of Japanese face in distinguishing the English liquids [r] and [l].

It is clear that Japanese natives exposed to English in adulthood have great difficulty identifying or discriminating stimuli forming minimal pairs that contrast [r] versus [l], such as *road* and *load* (Miyawaki et al., 1975), even after years of exposure to English. There is, however, a growing body of evidence that some ability to learn nonnative contrasts remains into adulthood. Several groups have shown that training can lead to improvement in identification and discrimination of nonnative contrasts (Jamieson & Morosan, 1986, 1989; Morosan & Jamieson, 1989; Pisoni, Aslin, Perey, & Hennessy, 1982; Pruitt, 1995; Tees & Werker, 1984), including [r] versus [l] (Logan, Lively, & Pisoni, 1991; Strange & Dittmann, 1984). Although the effects have often been short-lived and/or specific to the training stimuli used, there are some data that show broader and/or longer lasting improvements in adults (Lively, Logan, & Pisoni, 1993; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994). Also, identification and production of [r] and [l] is better in Japanese who have lived in an English-speaking society for 20 years, as compared with Japanese with only 1 or 2 years of such experience (Flege, Takagi, & Mann, 1996), although even Japanese with extensive exposure to

The work reported in this article was supported by grants from the National Science Foundation (LIS-9720348), the National Institute of Mental Health (MH47566), and the J. S. McDonnell Foundation (96-34). We thank Helen Neville for asking the question that led to the Hebbian formulation of the basis of the stabilization of speech representations and Susan E. Holmes for helpful input in the early formation of this research. Correspondence concerning this article should be addressed to J. L. McClelland, Center for the Neural Basis of Cognition, 115 Mellon Institute, Pittsburgh, PA 15213 (e-mail: jlm@cnbc. cmu.edu).

English performed more poorly than native English speakers. Other studies have shown that extensive laboratory training can lead to substantial, generalizable improvement. In one study (Bradlow, Pisoni, Yamada, & Tohkura, 1997), 45 h of training on a range of different minimal pair stimuli from five different speakers resulted in an average increase of 19% in the accuracy of forced-choice identification of the correct [r] or [l] member of each minimal pair. Generalization was nearly as good as posttraining performance on the training materials, and the benefits of training persisted for at least 6 months, affecting production as well as identification of words containing [r] and [l] (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999). Overall, existing data certainly suggest that the mechanisms of language perception are not completely fixed in adulthood; clearly, though, they seem relatively inflexible, on the basis of comparisons of adult language learners with younger learners or with adults learning other skills. Understanding why speech perception can be so resistant to change in adulthood may help us find ways to help second-language learners overcome this resistance and may lead to new insights into the nature of the mechanisms of learning and natural language acquisition.

Our work explores the idea that one factor in these persistent learning difficulties may be a paradoxical characteristic of a domain-general learning mechanism, which may work to maintain preestablished perceptual representations in the face of exposure to new experiences. Specifically, we will consider the hypothesis that Hebbian synaptic plasticity may tend to reinforce perceptual categories that have been established early in life, contributing to the difficulty of new learning in adulthood. This Hebbian approach leads to predictions concerning what kinds of retraining regimes might be used to produce effective remediation, predictions that will be tested in the reported experiments.

A premise of the Hebbian account is the observation that language experience has an influence on how we perceive acoustic information, affecting which stimuli we perceive as similar and which as different. For example, adult native English speakers clearly distinguish [r] and [l], but native Japanese speakers lose the ability to perceive this distinction (Miyawaki et al., 1975). These effects begin to occur very early in life (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1983, 1984). A range of mechanisms and metaphors have been introduced to capture the role of experience in shaping the perception of speech, and several of these ideas have been used to address the difficulties second-language learners face in learning nonnative language contrasts. One influential account suggests that experience with instances of a speech sound category are represented in the form of an abstract cognitive prototype and, furthermore, that this perceptual prototype acts as a *perceptual magnet*, influencing perception of similar stimuli (Kuhl, 1991). Others have proposed that collections of representations of specific spoken experiences, rather than abstract prototypes, could serve as the basis for a magnet-like effect (Lacerda, 1995). Another account (Guenther & Gjaja, 1996) draws on concepts from neuroscience studies of perceptual systems to suggest that language experiences directly influence organization of a topographic map for acoustic perception and shows how such experience-based changes in the perceptual map can influence how nonnative speech stimuli are perceived.

The structure of the native phonological space has been demonstrated to greatly affect perception of nonnative speech sounds, so that some of these sounds are perceived as (good or poor) exemplars of a native speech sound category, whereas others are perceived as not belonging to any native speech category or even to the domain of speech. In the perceptual assimilation model (PAM) developed by Best and colleagues (Best, 1995; Best, McRoberts, & Sithole, 1988), nonnative sounds that are similar to sounds in the learner's native language are assimilated to representations of the characteristic articulatory programs used to produce the related sounds from the learner's native language. The focus of the PAM is on predicting the difficulty of discriminating specific nonnative phonetic contrasts, and the model's predictions have been generally quite successful (Best, 1995; cf. Guion, Flege, Akahane-Yamada, & Pruitt, 2000).

Beyond static nonnative phonetic perception and contrast discrimination ability, the extent of *learning* to discriminate nonnative phonetic contrasts is a principal focus of the speech-learning model (SLM; Flege, 1995). The SLM postulates that "the mechanisms and processes used in learning the [native language] sound system . . . remain intact over the life span" (Flege, 1995, p. 239) and that these mechanisms are continually at work in second-language learning. On the basis of these principles, the SLM strives to predict long-term phonetic development for specific nonnative phonetic contrasts. Several SLM hypotheses relate to a person's age of initial important use of the nonnative language and to the extent of reliance on the nonnative language versus continued use of the native language. Perceptual similarity between native and nonnative phonetic categories also plays an important role in the SLM. Experience with nonnative speech sounds can gradually lead to the formation of new phonetic categories according to this model, but the rate at which this occurs will be a function of how strongly the sounds in the new language are assimilated into the established phonetic category structure. The strength of the assimilation is greater if the native sounds are more strongly entrenched before exposure to the new language and is also greater when the sounds of the nonnative language are close in perceptual space to existing sounds of the native language. In the SLM, experience still shapes the phonetic category landscape at all ages, but greater entrenchment carries with it greater resistance to change. At the extreme, category formation for a nonnative language sound may be blocked by the mechanism of "equivalence classification"—that is, if both members of a nonnative contrastive pair are perceptually equivalent with respect to the native phonetic categories, new categories may never form.

The PAM and the SLM provide explanations for the great differences in the ease with which nonnative contrast pairs can be initially discriminated and ultimately learned, depending on one's native language. However, given a particular nonnative phonetic contrast, with its associated perceptual difficulty and learnability, neither model proposes an actual mechanism through which experience with nonnative sounds leads to change. Our own ideas address this issue, building on the starting place provided by these important models. We take the explicit step of suggesting (with Guenther & Gjaja, 1996) that the neural substrate of a phonetic percept may be a pattern of neural activity within some region or regions of the cerebral cortex. Given this, we suggest that experience with a language may result in a situation in which a recurring categorical pattern of neural activity is elicited by a range of acoustically distinct inputs—in some cases, failing to capture aspects of speech input that distinguish nonnative phonemes.

We are now in a position to see how Hebbian learning, when applied to the learning of perceptual categories in speech, might provide a framework within which the mechanisms that reinforce the tendency to hear a range of stimuli as being the same might impede the ability to learn new perceptual distinctions in adulthood. The basic idea follows quite directly from the essential assumption underlying Hebb's proposal, which we paraphrase from Hebb (1949) as follows: *When one neuron, A, participates in firing another neuron, B, the strength of the effect of A on the firing of B is increased*. Such a learning rule should tend to strengthen whatever response a neural system makes to an incoming stimulus. Applying this idea to speech perception, when a pattern of neural activity at a peripheral level elicits a cortical pattern corresponding to a percept, the strength of the connections from the peripheral neurons to the neurons constituting the percept, and among the neurons constituting the percept, is increased. Thus, if it is true that, when a Japanese listener hears either an [r] or an [l], one and the same perceptual representation is activated, Hebbian learning should work to maintain this tendency. Computer simulations of a model illustrating how this may occur have been presented (McClelland, Thomas, McCandliss, & Fiez, 1999). Here, we consider the consequences of these ideas for retraining language perception.

If the Hebbian account is correct, it should be possible for Japanese adults to learn the [r]–[l] distinction, if only we can find contrasting [r]- and [l]-like stimuli that will elicit distinct perceptual representations. In this case, Hebbian learning would reinforce the tendency of these inputs to elicit such distinct representations, potentially leading to rapid progress in perceptual learning. One way to do this is to create [r] and [l] stimuli based on natural speech, in which the portion of the speech signal corresponding to the [r] and the [l] phonemes and adjacent transitions has been acoustically altered to exaggerate the differences between them. By using highly exaggerated stimuli that most of our Japanese native subjects do hear as different, we expect Hebbian learning to reinforce the resulting distinct percepts. Assuming that the tendency to hear the exaggerated stimuli as distinct generalizes to other, very similar stimuli, it should then be possible to gradually reduce the degree of exaggeration, establishing distinct percepts for stimuli that are more and more similar until the ability to distinguish them extends into the natural range. The Hebbian account also leads us to predict that if we were to attempt to train Japanese adults with stimuli that they could not initially discriminate, Hebbian learning would tend to maintain the tendency to perceive the stimuli as similar, and no learning should occur. Predictions similar to the ones just outlined can be derived from many of the other models discussed above, on the basis of the idea that stimuli that are strongly assimilated to existing perceptual representations will be difficult to distinguish and this difficulty will impede acquisition.

A further prediction arises from the fact that perceptual learning under our Hebbian account does not depend on feedback about the accuracy of one's perceptions. Indeed, an appealing aspect of a Hebbian account is that it provides a basis for understanding how language representations may be affected by experience very early in life, when there may be no explicit task or intent to learn (Saffran, Aslin, & Newport, 1996) but the child is simply hearing ambient speech in its environment. This form of learning may play a role during the perinatal period, during which the child is establishing perceptual representations consistent with the native language environment (Kuhl, 1993; Kuhl & Iverson, 1995). In any case, the Hebbian account predicts that the use of stimuli that exaggerate the contrast between [r] and [l] will allow nonnative speakers to learn to discriminate these stimuli, even in the absence of any feedback. This prediction appears to be contrary to a tacit assumption of previous researchers who have investigated the ability to learn novel speech categories, even though the matter has not been systematically investigated (Logan & Pruitt, 1995). To our knowledge, all previous studies of learning nonnative contrasts have provided subjects with feedback during at least some part of the training procedure. Thus, it is of considerable interest to determine what can be learned in the absence of feedback. For this reason, a crucial aspect of the design of our experiment is that we examined conditions in which subjects received no feedback during training. Again, we note that this prediction may also be consistent with other approaches. We would note, however, that our explicit consideration of the applicability of the Hebbian learning rule brings the issue of learning without feedback into focus.

To address these issues, we used two contrasting training conditions with two groups of Japanese adult subjects whose initial ability to discriminate minimal pairs beginning with [r] or [l] was very poor, relative to English speak-

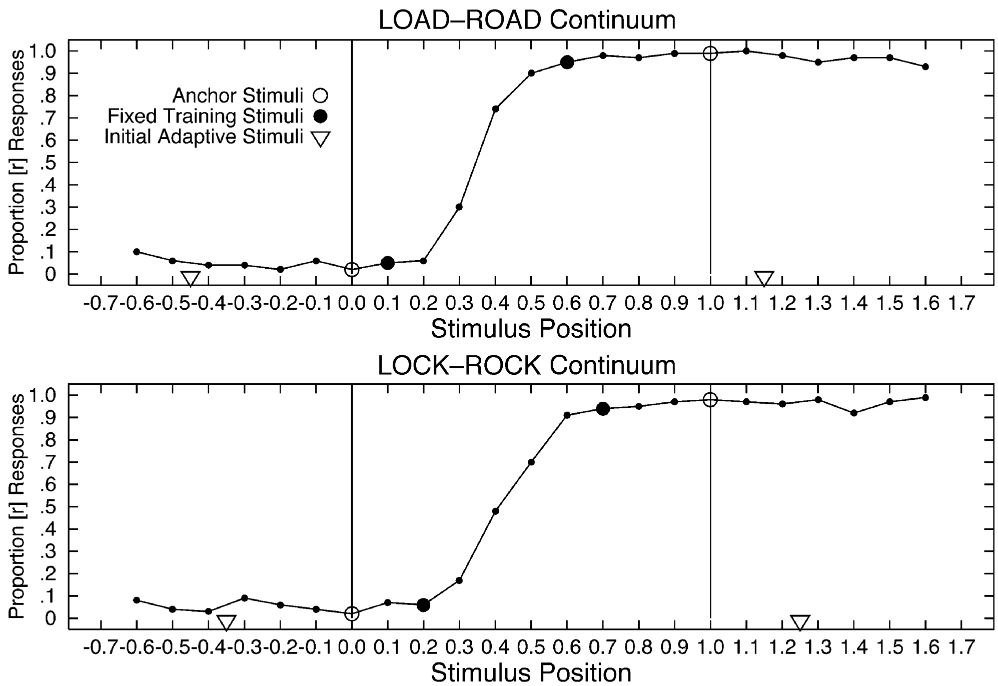

Figure 1. Mean categorization functions of 12 native English speakers for synthesized speech stimuli from each of the two continua used in the experiments. The *x*-axis represents the stimulus index variable *x*, which indicates the position on the stimulus in relation to the anchor stimuli (see the Method section for details). Percentages of trials eliciting [r] responses are plotted on the *y*-axis for each stimulus. Large empty circles represent the anchor stimuli resynthesized from the recorded base stimuli (see the Method section). Data points between the anchor stimuli are responses to stimuli interpolated between these anchors, and data points in the peripheral regions represent responses to extrapolated speech stimuli. Stimuli used for the fixed training condition are indicated with large filled circles. Triangles point to the positions of the initial stimuli used in the adaptive training condition.

ers. In these conditions, the two groups of subjects received no feedback on the accuracy of their responses. We examined whether the subjects in each condition could learn to respond differentially to the members of these pairs. The two contrasted conditions are called the *adaptive* and the *fixed* training conditions. In the adaptive condition, subjects were trained with initially exaggerated stimuli (see Figure 1) that were equidistant from the point at which perception of our native speaker group crossed over from [l] to [r]. Then, the stimuli were adjusted throughout the course of learning, in an effort to ensure that the subject received perceptually distinguishable stimuli throughout. On each training trial, the subject responded with a keypress to identify the first phoneme of the word as an [r] or an [l]. Success on eight stimuli in a row resulted in replacement of one of the training stimuli with a slightly less exaggerated stimulus, moving one small step inward along the [r]–[l] continuum. A single incorrect response resulted in replacing one of the stimuli with the item one step further from the crossover point. Adjustments were balanced to keep the stimuli as nearly equidistant from the crossover point as possible, stopping at the endpoints indicated in the figure. In the fixed condition, a second group of subjects received the same number of training trials and had the same task of attempting to identify the stimuli. However, the stimuli were a fixed pair of contrasting speech tokens (also indicated in Figure 1) that were reliably identified by English speakers but that the Japanese subjects could not discriminate with an accuracy greater than 70%.

Our adaptive training procedure is similar to the procedure used by Jamieson and Morosan (1986) and Morosan and Jamieson (1989) to train adult francophones on the English [ð]–[θ] contrast, although these investigators used feedback throughout training and provided no contrast to any other training condition. Other investigators (Merzenich et al., 1996; Tallal et al., 1996) have used exaggerated stimuli and adaptive training in their protocol for remediating language-processing deficits in children with specific language impairment, and the reported success of their intervention provided part of the initial motivation of our own efforts. They found a greater benefit for their experimental group over a control group that received training with nonexaggerated stimuli. Like Jamieson and Morosan, they used feedback in many of their training conditions. Also, it should be noted that their interventions were rather broad-ranging, in that they involved training on discrimination of rapid auditory transitions, training on discrimination of a wide range of speech contrasts, and extensive training on tasks requiring comprehension of meaningful connected speech. Thus, further evaluation of the effects of using exaggerated stimuli and an adaptive training regime in a study focused on speech identification seems warranted.

It should be noted that the acoustic exaggeration we used was different in an important way from that used in previous studies. Instead of increasing the salience of a given acoustic difference by making the critical stimulus region louder or longer or making it stand out by removing its acoustic context (cf. Pruitt, Kawahara, Akahane-Yamada, & Kubo, 1998), our method was designed to directly increase the magnitude of the natural distinction between English [r] and [l]. Specifically, we enhanced naturally occurring differences between tokens of [r] and [l] uttered in identical contexts (see the Method section).

In order to assess the effects of training, it was necessary to provide both a pre- and a posttraining test. We assessed both identification of stimuli as tokens of [r] or [l] and discrimination of pairs of stimuli, using a same–different (AX) discrimination task (see the Method section for details). No feedback was given during the test, but it is possible (and indeed, predicted by our Hebbian account, if some of the stimuli used are distinct enough) that the pretest itself might have led to some improvement in performance. To control for such effects, an additional group of subjects received the pre- and posttest without any training in between. The subjects in this group received the pretest battery and then returned, after a 3-day delay, to take the posttest.

With the above aspects of the experiment in view, we will review the predictions arising from our Hebbian account. Insofar as the stimuli used in the adaptive condition are perceived as distinct, and under the assumption that what is learned about a particular stimulus tends to generalize to very similar stimuli, the Hebbian approach would lead us to predict that subjects would succeed in learning the [r]–[l] distinction in the adaptive condition. Although no feedback is given during training or testing, none should be necessary for learning, according to the Hebbian account. This should result in improvements both in identification and in discrimination of [r] versus [l] stimuli after training. On the other hand, in the fixed condition, insofar as the subjects initially perceive the stimuli used in this condition as the same, we would predict that Hebbian learning will, if anything, reinforce the tendency to hear them as the same, thereby preventing progress in learning the [r]–[l] distinction, leading to the prediction that there will be little or no improvement in identification or discrimination after training.

A further prediction that arises from an unelaborated version of the Hebbian account is that learning should be the same with feedback as without it. If it is only the neural response that is elicited immediately by an input that drives learning, we would still expect subjects in the adaptive condition to learn and subjects in the fixed condition to fail to learn, whether or not they are given feedback on the accuracy of their responses. However, given the adaptive significance of sensitivity to outcomes, not to mention the huge literature on the role of positive and negative reinforcement in learning, it would be surprising if feedback played no role, and thus there may be reasons to suspect that this unelaborated version of the Hebbian account would turn out to be incorrect. However, there is relatively little data comparing learning with versus without feedback, and thus a test of the hypothesis should lead to useful information that may constrain the further elaboration of an adequate theory.

To assess the role of feedback, two additional groups of subjects were run, using the same conditions described above for the adaptive and the fixed training conditions without feedback, but this time with visual feedback provided immediately after each identification response during the training phase of the experiment. In all other ways, the training and testing of the subjects in the two feedback groups were the same as they were for the subjects in the corresponding no-feedback conditions.

An important issue, with both practical and theoretical implications, is the extent to which what is learned about one minimal pair of stimuli exemplifying the [r]–[l] contrast in one specific phonological context (e.g., *rock–lock*) generalizes to other stimuli exemplifying the same contrast. This issue has been an important focus of a considerable body of prior research (Bradlow et al., 1999; Bradlow et al., 1997; Lively et al., 1993; Lively et al., 1994). As a small step toward addressing this issue as it arises within the context of the present study, we trained each of our subjects on one of two continua (*rock–lock* or *road–load*) and tested pre- and posttest performance not only on the trained continuum, but also on the untrained, transfer continuum, with each continuum used equally often for training and transfer within each group. It should be clear that this test of generalization is a limited one. The stimuli are spoken by the same individual, so generalization across speakers is not assessed. Also, the contrast is in word-initial, pre-vocalic position in both cases, so generalization to other positions within the word is not assessed. Nevertheless, it is important to know whether the effects of training generalize even to this limited extent and, if so, what training conditions foster generalization.

METHOD

Stimulus Materials

Two synthetic speech continua were generated, one ranging from *rock* to *lock*, the other ranging from *road* to *load*. The stimuli for each continuum were based on recorded tokens of the two words spoken by a male native English speaker, sampled at 11.025 kHz with 16-bit quantization. Care was taken in selecting the base tokens so that the pairs were acoustically matched in amplitude, fundamental frequency, and formant frequencies over their vowel portions, so that cross-splicings would be undetectable. The base stimuli were time aligned on the basis of the formant transitions from the consonant to the vowel. For each stimulus, an onset portion was defined as including the steady-state portion of the initial consonant and the transitions, and a body portion was defined as including the acoustically matched vowel–consonant portion of the syllable. The waveforms of the base stimuli were subjected to 14-pole linear predictive coding (LPC) analysis on 27.21-msec Hamming-windowed speech frames every 9.07 msec, and the LPC coefficients were used to derive log area ratio coefficients of an equivalent vocal tract model (Rabiner & Schafer, 1978, p. 444). Estimates of the pitch tracks of the base stimuli were derived using an auto-correlation algorithm and then manually corrected and adjusted to match between the two base stimuli used for each continuum.

A series of speech sounds ranging from [r] to [l] was generated for each continuum, using vocal tract model coefficient values determined by weighted averaging of coefficient values from the base stimuli, using the equation $v(x) = v(l) + x * [v(r) - v(l)]$, where $v(r)$ stands for the coefficients from the [*r*] base and $v(l)$ stands for the coefficients at the corresponding time point from the [l] base, with x ranging from 0.00 to 1.00 in steps of 0.05. The stimuli constructed with values of $v(l)$ equal to 0.0 or 1.0 (circled in Figure 1) are called anchor stimuli. Exaggerated stimuli were created by extrapolating outside the range of the two base stimuli by using the same formula with values of x from -0.60 to -0.05 for exaggerated [l]-like sounds and from 1.05 to 1.60 for exaggerated [r]-like sounds. In every case, only the onset portions of the stimuli were combined for the interpolation or extrapolation; the body coefficients were always taken from the [r] base stimulus.

The resulting vocal tract model coefficient sets were converted to LPC coefficients, and all the stimuli were synthesized by direct application of the recursive LPC filter equations with pitch-synchronous interpolation of parameters throughout.

Twelve University of Pittsburgh undergraduates whose first language was English were used to determine native identification functions for both continua. Stimuli from a given continuum were randomly ordered within separate blocks. For each continuum, each subject heard 8 presentations of 23 stimuli constructed as above, with values of x ranging from -0.6 to 1.6 in steps of 0.10, and was asked to indicate whether each began with [r] or [l] (stimuli were constructed at 0.05 intervals; however, we obtained normative data only at 0.1 intervals). The subjects heard an additional 20 presentations each of the 11 stimuli from 0.0 to 1.0. Some of the extreme stimuli sounded unnatural to most listeners. Typically, the onsets would be disproportionately loud, or the stimuli would be described as having "weird metallic overtones." This may be inevitable given the need to create marked exaggerations. What is important is that the resulting group identification functions showed that each continuum could be divided into three regions: one in which all the stimuli were reliably ($p > 90\%$) identified as beginning with [l], one in which all the stimuli were reliably identified ($p > 90\%$) as beginning with [r], and one over which the stimuli showed a sigmoidal transition from reliable identification as [l] to reliable identification as [r]. For each continuum, a pair of stimuli were chosen to be the fixed training stimuli. Those were the most similar pair of stimuli with the property that one was reliably identified by the native English subjects as [r] and the other was reliably identified by the native English subjects as [l] (Figure 1).

Subjects

Forty adult Japanese subjects residing in or near Pittsburgh were used overall, 8 in each of the four training conditions and 8 in the control condition. Every subject met the following criteria: He or she (1) had been born in Japan and had lived there at least until 18 years of age and (2) had performed at or below 70% accuracy in discriminating the fixed training stimuli on at least one of the two continua during the same–different judgment task in the pretest. The subjects were pseudorandomly assigned to the different conditions, within the constraints that pretest performance had to be less than 70% on the continuum to be used in training and overall pretest performance of all of the training groups and of the control group should be roughly equivalent.

Training and Testing Procedures

Before and after training (or for control subjects, before and after a 3-day test–retest delay interval), each subject was given a test battery to assess ability to identify and discriminate stimuli from each of the two continua. Each test was administered separately for each of the two continua, with the *rock–lock* continuum always tested before *road–load*. The identification task, always administered first in the battery, required the subject to indicate, by pressing one of two buttons, whether a word began with [r] or [l], without feedback. Before the test on each continuum, the subjects were given three labeled instances of each of the two anchor stimuli. For each continuum, the subject received 132 trials in a random order, consisting of 12 instances of the anchor stimuli and 12 instances of the interpolated stimuli spaced 0.1 units apart along the continuum. Next came the discrimination test, which required the subjects to indicate whether a pair of stimuli were the same or different. Two variants of the discrimination task were used. In the first variant, the subjects received 48 trials that used only the two anchor stimuli. Each AX pair consisted of both [r] anchors, both [l] anchors, or one [r] and one [l] anchor. This variant proved relatively easy for all the subjects, and performance on this variant was not further analyzed. In the second variant, called the *expand* test, eight stimulus pairs from each continuum were used, with the members of each pair centered around the midpoint of the transition from [l] to [r] in English speakers (0.35 for *load–road*, 0.45 for *lock–rock*) and spanned separations ranging from 0.1 (most difficult) to 1.5 (least difficult) in equal intervals, with a separation of 1.0 corresponding to the separation between the anchor stimuli on each continuum. For each continuum, each subject received 192 trials in random order, consisting of 24 trials based on each pair. In half of the trials, the two different members were used as stimuli, so that the correct response was *different*; otherwise, one of the two stimuli was presented twice, so that the correct response was *same*. For each continuum, the fixed training stimuli served as one of the eight pairs, and performance on these stimuli was used as the basis for subject selection, as indicated above. All of the above tests were administered both in the pretest and in the posttest. One additional discrimination test was administered in the posttest only. This test, called the *slide* test, made use of stimuli with a separation of 0.3, centered at 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, and 0.85 on each continuum. As with the expand test, half of the stimulus pairs were *same* pairs, and the other half were *different*. A total of 128 trials, consisting of 16 trials with each pair, was administered in random order for each continuum.

Each subject in the study took home a portable computer to complete three training sessions on separate days. Each session took approximately 20 min to complete. Within each training session, the subjects were presented with 480 training trials, in which one of two synthesized speech tokens (an [r] token or an [l] token) was selected at random, subject to the constraint that each group of 20 trials contained an equal number of [r] and [l] stimuli. The subjects indicated whether the word began with [r] or [l] by pressing one of two buttons. Each response was simply recorded by the computer in the no-feedback conditions. In the feedback conditions, correct responses were followed immediately by a row of three green check marks, whereas errors were followed by a row of three red x characters. In the fixed training condition, the same 2 fixed training stimuli were used in all training trials (0.1 vs. 0.6 for *load–road*, and 0.2 vs. 0.7 for *lock–rock*). For the adaptive training condition, a set of 38 stimulus pairs was created for each continuum. The pairs were constructed around the midpoint between the fixed training stimuli, and the distance between the members of each stimulus pair ranged from 0.1 (Level 0, hardest) to 1.95 (Level 37, easiest). The pairs were constructed by starting with the Level 0 pair, then alternating between decreasing the x value of the [l]-side stimulus and increasing the x value of the [r]-side stimulus by 0.05. All the subjects in the adaptive condition initially began training at Level 30 (-0.45 vs. 1.15 for the *road–load* continuum and -0.35 vs. 1.25 for the *rock–lock* continuum; see Figure 1). Whenever a subject made an error, the stimuli were switched to the next easier pair. Whenever a subject made eight consecutive correct responses, the stimuli were switched to the next harder pair. At the beginning of each training session and again after every block of 20 training stimuli, there were 2 probe trials, 1 with each of the 2 fixed training stimuli for the continuum in use for the given subject, for a total of 50 probe trials per session. Performance on the probe stimuli was not considered in adjusting the training stimuli, but the subjects in the feedback conditions did receive feedback on their responses to these stimuli.

After completion of the posttest, as described above, 4 of the subjects in each training group (2 trained with *rock–lock* and 2 trained with *road–load*) were assigned to continue in the identical training condition for an additional three sessions. After the additional training, these subjects received an additional test, called the *final posttest*, identical to the test previously administered at the end of the first 3 days of training.

RESULTS

We will begin by considering the effects of training without feedback, focusing on the key contrast between the adaptive and the fixed training conditions. We then will consider whether the results were the same when feedback was used.

Effects of Training Without Feedback

In accordance with the predictions of the Hebbian account, the subjects in the adaptive training condition demonstrated substantial gains in identifying stimuli on the continuum used in training (Figure 2, upper left), even without feedback. The pretraining identification function was fairly flat, indicating poor identification, consistent with the poor initial discrimination of these stimuli, which was a condition of the subjects' inclusion in the study. After training in the adaptive condition, the group identification function took on a sigmoid shape. Stimuli on the [l] end of the acoustic continuum were reliably identified as beginning with [l], and there was a clear transition leading to a region over which the stimuli were reliably identified as beginning with [r]. Assessment of the statistical significance of training effects was based on the difference between the slopes of best-fitting logistic regression functions applied to each subject's pre- and posttraining identification data,[1] using one-tailed tests unless otherwise indicated. The slope differences were significantly greater than 0 in the adaptive condition [$t(7) = 4.90, p < .001$], and every subject in this condition showed an increase in the slope. An improvement from the pretest to the posttest was also exhibited by the control group, which received no training between the pre- and the posttest [$t(7) = 2.88, p = .012$][2] (Figure 3, top). Therefore, it became important to contrast the slope difference found in the control group with the value obtained in the adaptive group. Importantly, the adaptive group showed greater change from the pre- to the posttest than did the control group, as indicated by a larger slope difference for adaptive subjects than for control subjects [$t(14) = 3.10, p = .004$].

Even though the adaptive training condition did lead to the emergence of fairly sharp identification functions, these

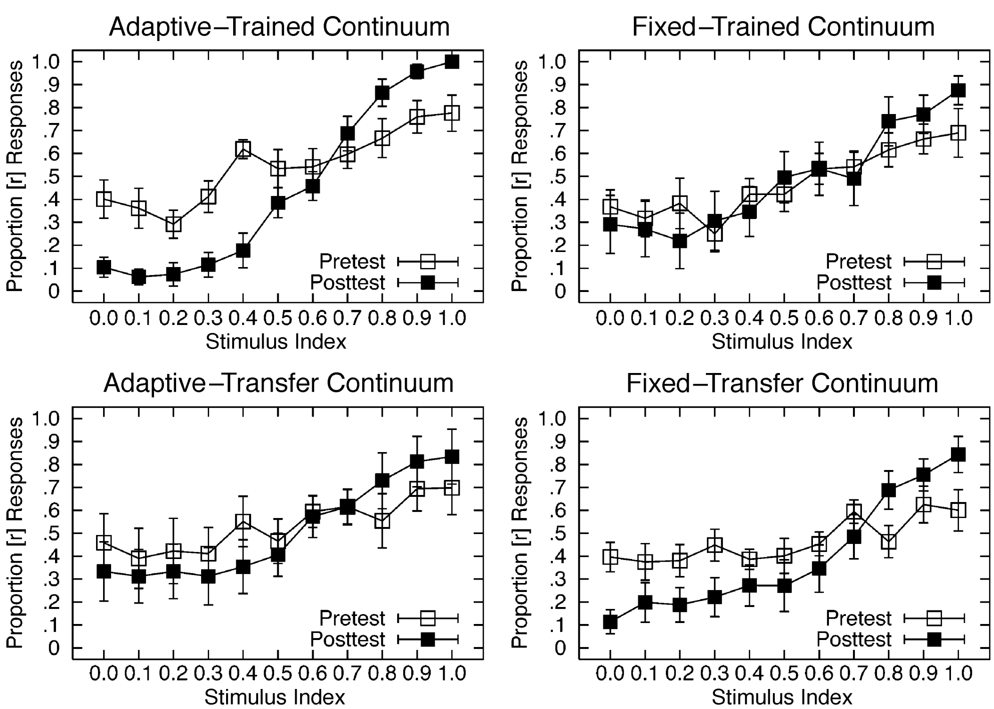

Figure 2. Mean categorization functions (with standard error bars) for two groups of Japanese subjects ($n = 8$) before and after three 20-min training sessions in the adaptive condition (left) or the fixed condition (right) without feedback. Pre- and posttest performance is shown on the continuum used in training (upper panels)and on the other continuum used to assess transfer (bottom panels).

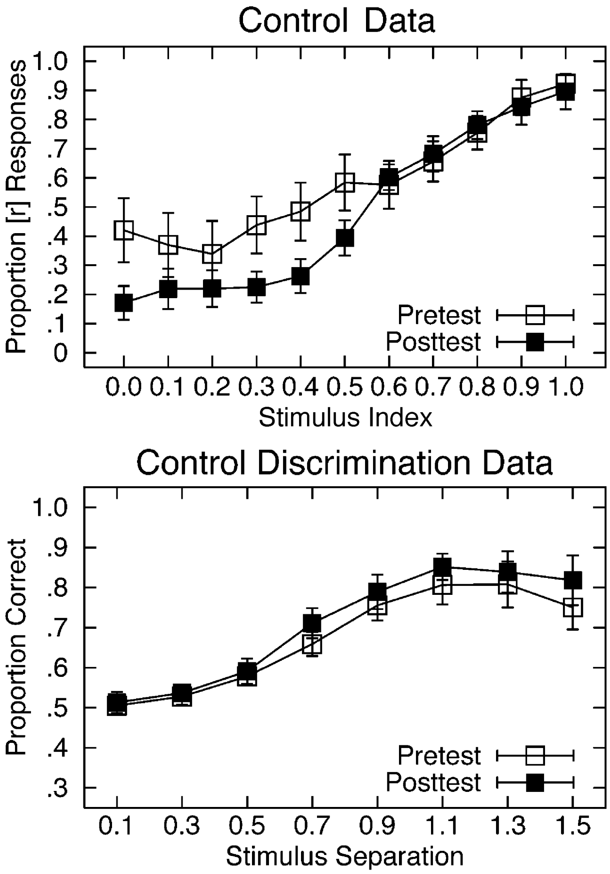

Figure 3. Top: mean categorization functions (with standard error bars) for a group of Japanese control subjects (*n* = 8) before and after a 3-day control period, during which no training was given. Bottom: mean discrimination performance graphed against the separation of the discriminative stimuli (with standard error bars) for the same subjects under the same conditions. Data are collapsed over the two continua, both of which were tested in the pretest and the posttest.

functions were shifted relative to those of native speakers. For the native speakers, we calculated the crossover point for each continuum and obtained a mean value of 0.391 (SD = 0.064). For the Japanese subjects in the adaptive condition, the mean locus of the crossover point was 0.586 (SD = 0.141). The difference was highly significant [$t(29) = 5.390, p < .001$]. Thus, the adaptive subjects had learned to distinguish the [l] and the [r] categories, but their boundary appeared to be in a different place than that of native English speakers.

Figure 4 provides examples of individual subjects' patterns of progress during training in the adaptive training condition. The separation of the stimuli at each point in training is indicated, and a separation of 1.0 corresponds to the separation of the natural speech tokens (the anchor stimuli in Figure 1). Even though stimuli were initially exaggerated, nearly all the subjects made some errors at first, causing the training algorithm to present even more exaggerated stimuli, until the limit of our constructed set of exaggerated stimuli was reached. After working with the most exaggerated stimuli for a brief period, some subjects, such as the one shown at the top of the figure, showed rapid progress within the first training session and continued to progress to the point where they were reliably labeling stimuli well within the range of the actual spoken items. Other subjects—for example, the one whose data are shown in the middle of the figure—exhibited slower and/or more erratic progress. We also had 3 subjects (such as the one shown at the bottom of the figure) who did not progress beyond the most exaggerated stimuli within the three training sessions. Two of these subjects showed improvement, as measured by percentage of correct identifications of the most extreme stimuli, and the 3rd apparently mapped the stimuli to the wrong responses during training but exhibited differentiation of the stimuli in his responses (i.e., his accuracy was far *less* than chance). All 3 of these subjects showed clear improvement in the posttest. These subjects all undertook three further sessions of training after the posttest, and all 3 progressed to less exaggerated stimuli during these additional training sessions (the subject who had the stimuli mapped to the wrong responses remapped them to the correct responses in the second 3 days of training). Thus, even the subjects with the poorest initial ability to respond differentially to [r] and [l] stimuli made progress in the adaptive training condition.

We turn now to the group of subjects who received training with the fixed stimuli, where the Hebbian account predicts there should be no effect of training, given that these stimuli are initially difficult for our Japanese adult subjects to discriminate. As was predicted, this group showed a smaller slope difference between the pre- and the posttest for the stimuli used in training than was found for the adaptive group [$t(14) = 2.01, p = .032$; Figure 2, upper right]. Although this group did exhibit a slight improvement from the pretest to the posttest [$t(7) = 2.31, p = .021$], this effect was not reliably different from the improvement seen in the control group [$t(14) = 0.67, p = .256$].

Effects of training on discrimination. According to the Hebbian hypothesis, our adaptive training regime, using stimuli that subjects tend to perceive as different, not only should increase the ability to apply correct labels to stimuli, but also should enhance discrimination between them. Such a training effect should not be obtained if subjects are trained in the fixed condition, using stimuli they tend to perceive as the same. Here, we will consider evidence bearing on these predictions from the expand and the slide tests. In the expand test, we obtained clear evidence that training improved discrimination in the adaptive condition (Figure 5, top left) but little sign of a training effect in the fixed condition (Figure 5, top right). An analysis of variance performed on the pre- versus posttest data from the adaptive condition produced a main effect of test [pretest vs. posttest, $F(1,7) = 45.835, p < .001$], as well as a main effect of the stimulus separation variable [$F(7,49) = 26.147, p < .001$]. Although a pretest–posttest difference

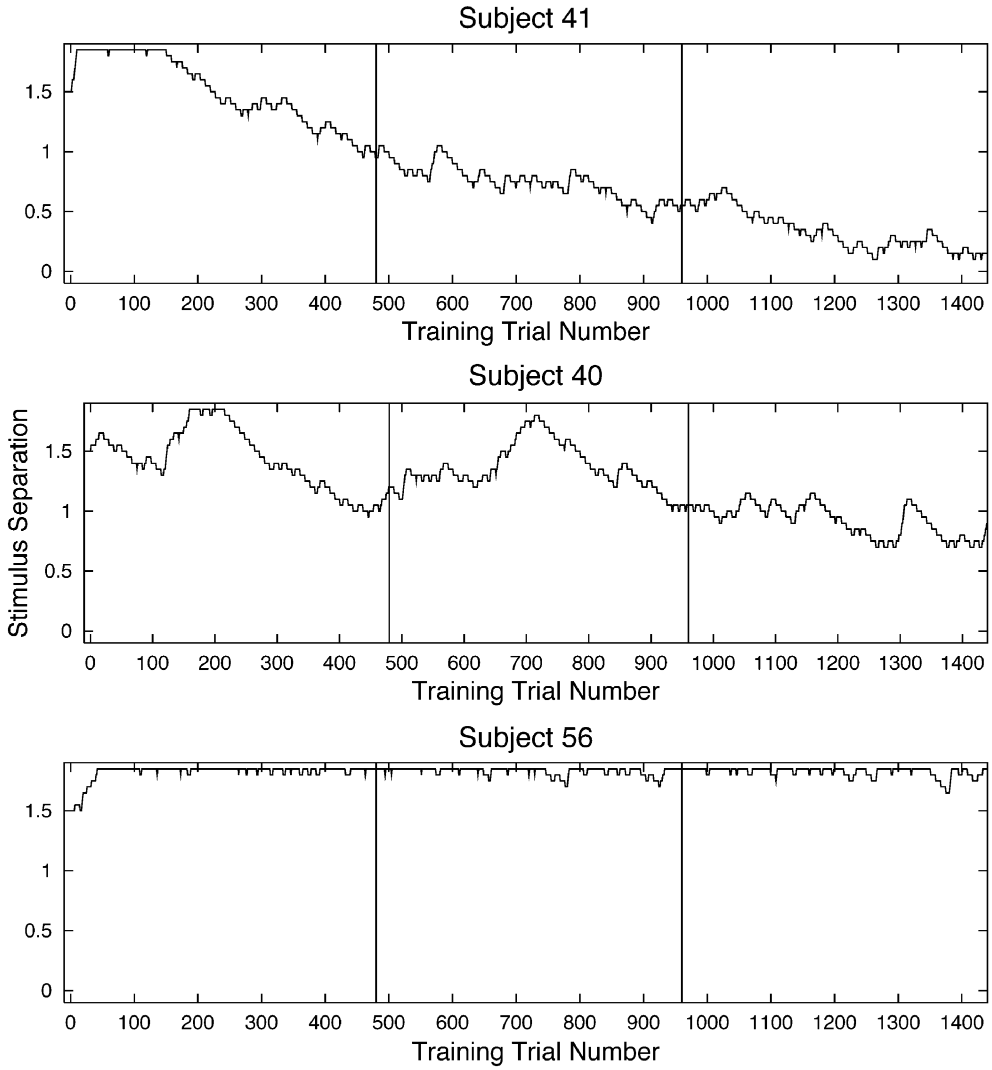

Figure 4. Three examples of single-subject performance illustrating the range of progress shown by different subjects through the separation levels of the stimuli in the adaptive training condition without feedback over the 1,440 training trials. Vertical lines indicate boundaries between daily training sessions. A separation of 1.0 corresponds to the distance between the natural speech tokens beginning with [r] or [l]. Separation was increased by 0.05 after each incorrect response and decreased by the same amount after each run of eight consecutive correct responses. See the Method section for details. Note that although Subject 56 did not progress beyond the most extreme separation level, there was some evidence of learning, since this subject progressed from an accuracy level of 76% on Day 1 to 82% on Day 3, a reliable increase (*p* = .011).

does not appear for the smallest separations, the interaction of pre- versus posttest with separation was not significant [$F(7,49) = 1.603$, $p = .157$]. A corresponding analysis on the data from the fixed condition produced only a main effect of separation [$F(7,49) = 34.143$, $p < .001$] and no evidence of a pretest–posttest difference. Interestingly, the control condition (Figure 3, bottom panel) produced a slight but clear improvement from the first to the second test [$F(1,7) = 9.697$, $p = .017$], along with a main effect of separation [$F(7,49) = 36.274$, $p < .001$]. Although the magnitude of the pre- to posttest improvement did not differ reliably between the fixed and the control conditions, the finding of a significant pre- versus posttest effect in the control condition and no such effect after fixed training hints at the possibility that fixed training may have tended to reverse a slight improvement in discrimination owing to the experience with the pretest.

In the slide test, if training is effective in inducing a perceptual distinction between [r] and [l] stimuli, we would expect to find an accuracy increase for stimuli that straddle the crossover point in the posttraining identification functions. There was evidence of such an effect in the adaptive condition, but no evidence of a training effect in the fixed condition. The slide test was not administered during the pretest, so statistical testing focused on posttest performance, comparing performance of the subjects with either adaptive or fixed training with the control subjects. An analysis of variance on the adaptive training versus control data revealed a main effect of training [adaptive vs. control, $F(1,14) = 4.740$, $p = .047$] and a main effect of position [$F(7,98) = 8.756$, $p < .001$] but no interaction [$F(7,98) = 1.397$, $p = .215$]. Even though the interaction was not significant, post hoc one-tailed *t* tests were carried out to determine at which points the discrimination curves differed between the adaptive and the control conditions. The differences were reliable at midpoint positions 0.45 ($p = .026$) and 0.55 ($p = .002$) and approached reliability at 0.65 ($p = .053$). The positions of the reliable or near-reliable differences (0.45–0.65) corresponded to that portion of the [r]–[l] continuum near the crossover point ($M = 0.586$, $SD = 0.141$, as was reported above) for [l] versus [r] identification responses in the posttest identification data for these subjects (Figure 2, top left). In contrast to these findings, an analysis comparing fixed training versus control yielded only a main effect of position [$F(7,98) = 5.795$, $p < .001$], and there was no significant difference between the fixed and the control performance at any point (of the eight one-tailed *p* values, the smallest was .233). In summary, adaptive training led to reliable improvements in discrimination at intermediate points along the [r]–[l] continuum, but not elsewhere, and fixed training produced no evidence of such an improvement.

Generalization of the effects of training to the transfer continuum. Because the subjects were tested on a second continuum not used in training, we were able to consider whether there was any evidence of transfer. Considering first the identification results, data for the transfer continuum for the subjects from both the adaptive and the fixed training conditions are shown in the lower panels of Figure 2. An analysis of variance indicated a slight improvement from the pre- to the posttest [$F(1,14) = 7.685$, $p = .015$], but the size of the improvement did not differ between the two groups [$F(1,14) = 2.483$, $p = .137$], and the subjects who received training did not show a greater improvement than that of the control subjects [$t(22) = 0.676$, $p = .253$]. With respect to the discrimi-

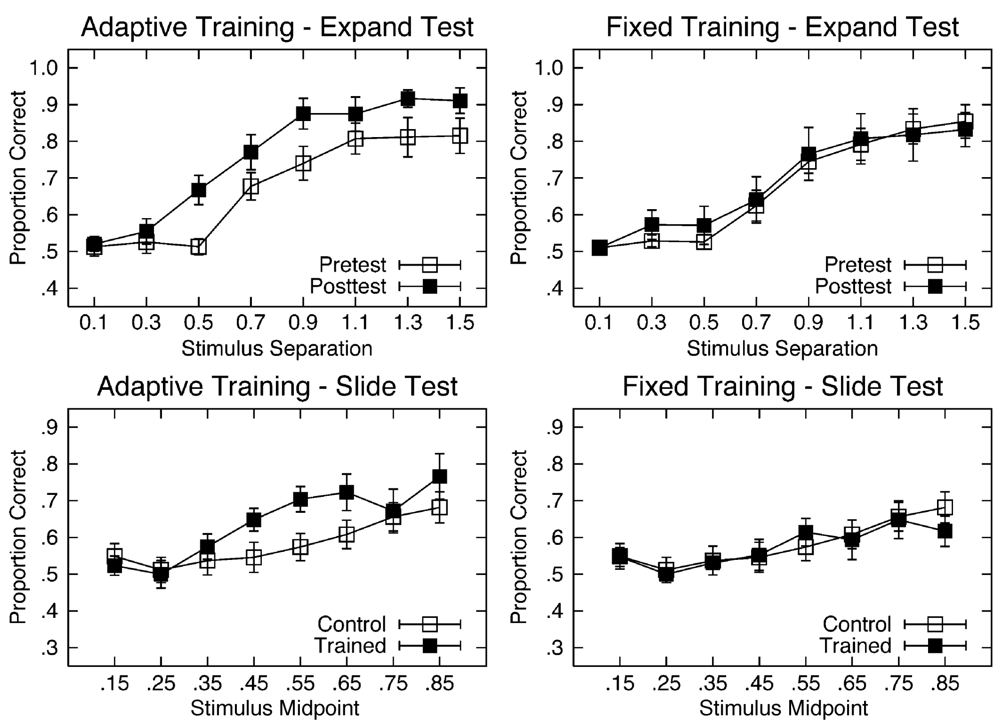

Figure 5. Effects of training without feedback on discrimination of the [r]–[l] contrast. All data come from performance on the continuum used in training. Top: mean discrimination performance (with standard error bars) as a function of the separation of the discriminative stimuli for two groups of Japanese subjects before and after training in the adaptive condition (left) or the fixed condition (right) without feedback. The separation is measured relative to the separation of the anchor stimuli, which are 1.0 units apart. Bottom: mean discrimination performance for stimuli separated by 0.3 units as a function of the position of the midpoint between the two stimuli along the continuum between the two anchor stimuli. The panels contrast performance after adaptive training (left) or fixed training (right) with performance by control subjects on the posttest.

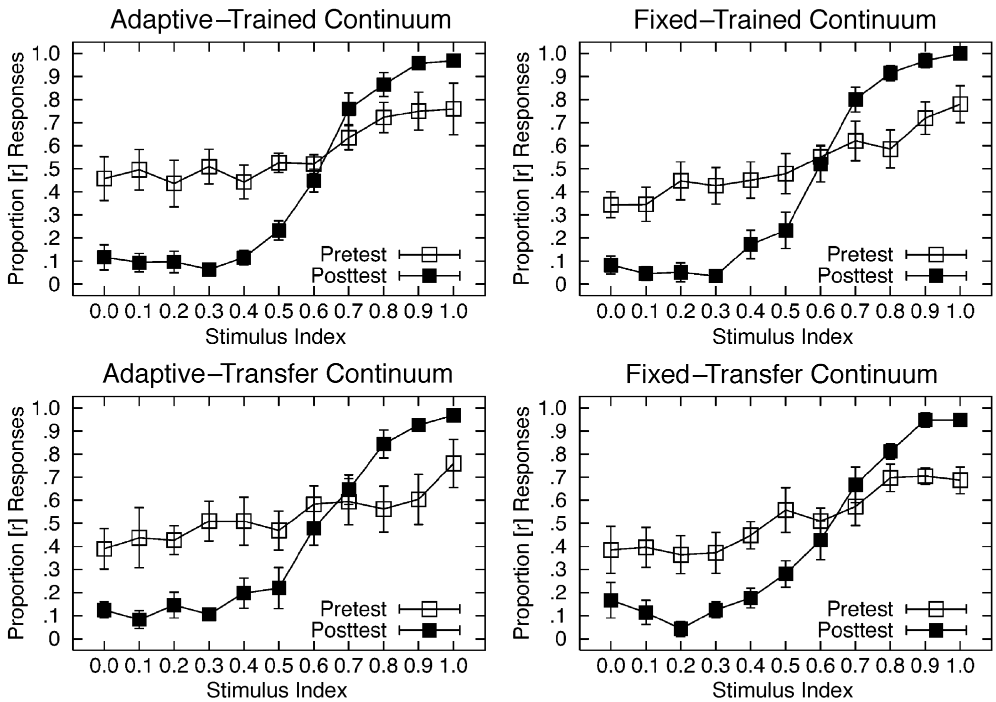

Figure 6. Mean categorization functions (with standard error bars) for two groups of Japanese subjects (n = 8) before and after three 20-min training sessions in the adaptive condition (left) or the fixed condition (right) with feedback. Pre- and posttest performance is shown on the continuum used in training (upper panels) and on the other continuum used to assess transfer. Assignment of *road–load* and *rock–lock* as training and transfer continua was counterbalanced.

nation results, there were no reliable effects of either fixed or adaptive training on the transfer continuum in either discrimination test.

Effects of Training With Feedback

As was previously discussed, Hebbian learning makes no provision for feedback to alter the outcome of the learning process. Without further elaboration, then, a Hebbian account would predict that the effects of adaptive and fixed training would be the same with and without feedback. We begin our consideration of this matter by examining the identification performance of the subjects who received feedback in the adaptive and nonadaptive training regimes. As can be seen in Figure 6 (top panels), both the fixed and the adaptive group showed a clear increase in the slope of their identification functions from the pretest to the posttest [$t(7) = 2.48, p = 0.21$ for the adaptive group and $t(7) = 3.41, p = .006$ for the fixed training group], and both groups showed larger gains than did the no-training control group [$t(14) = 1.94, p = .036$ for adaptive vs. control and $t(14) = 2.70, p = .009$ for fixed vs. control]. There was no significant difference in the gains shown by the two feedback groups [$t(14) = -0.17, p > .5$]. Furthermore, a direct comparison of the fixed training group trained with feedback versus the fixed/no-feedback group indicated a reliably larger gain from training with feedback than from training without it [$t(14) = 2.24, p = .021$]. This finding clearly contradicts the prediction of the Hebbian approach that feedback should not make any difference. There was no significant difference, however, in the gain shown by the two adaptive groups [$t(14) = 0.677, p = .255$]. Once again, the boundary between categories exhibited by the two feedback groups (0.623, SD = 0.051 for the adaptive group and 0.589, SD = 0.087 for the fixed group) differed from the boundary seen in native English speakers [0.391, SD = 0.064, as was previously noted; $t(29) = 9.211, p < .001$ for the adaptive group vs. native speakers and $t(29) = 6.868, p < .001$ for the fixed group vs. native speakers].

Effects of training on discrimination. In line with the identification data, training with feedback led to clear gains in discrimination between [r] and [l] stimuli, for both the fixed and the adaptive training regime (Figure 7). Considering first the performance in the expand test, an analysis of variance performed on the pre- vs. posttest data from the adaptive condition produced a main effect of training [posttest vs. pretest, $F(1,7) = 7.835, p = .026$], as well as a main effect of the stimulus separation variable [$F(7,49) = 58.214, p < .001$], and there was a trend toward a reliable interaction [$F(7,49) = 2.032, p = .070$]. A corresponding analysis of the data from the fixed condition produced a main effect of separation [$F(7,49) = 34.143, p < .001$] and

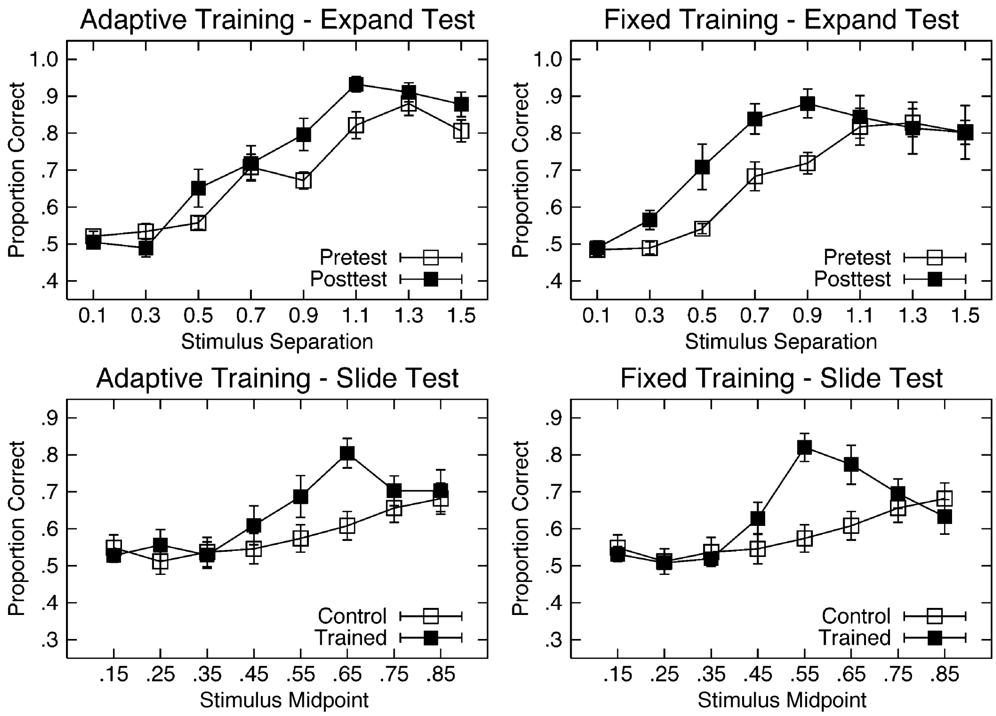

Figure 7. Effects of training with feedback on discrimination of the [r]–[l] contrast. All data come from performance on the continuum used in training. Top: mean discrimination performance (with standard error bars) as a function of the separation of the discriminative stimuli for Japanese subjects before and after training in the adaptive condition (left) or the fixed condition (right) with feedback. The separation is measured relative to the separation of the anchor stimuli, which are 1.0 units apart. Bottom: mean discrimination performance for stimuli separated by 0.3 units as a function of their midpoint relative to the continuum between the two anchor stimuli. The panels contrast performance after adaptive training (left) or fixed training (right) with performance by control subjects on the posttest.

a clear interaction of separation with training [$F(7,49) = 2.866, p = .014$], but no main effect of training [$F(1,7) = 3.268, p = .114$].

In the slide test, the effect of training can be seen to be concentrated at intermediate positions along the [r]–[l] continuum. An analysis of variance on the adaptive training versus control data revealed a main effect of training [adaptive vs. control, $F(1,14) = 4.719, p = .048$] and a main effect of position [$F(7,98) = 8.713, p < .001$], with a trend toward an interaction [$F(7,98) = 1.910, p = .076$]. A corresponding analysis comparing fixed training versus control yielded a main effect of training [$F(1,14) = 5.097, p = .040$], a main effect of position [$F(7,98) = 12.538, p < .001$], and a highly reliable interaction [$F(7,98) = 5.451, p < .001$]. The peak in the discrimination function landed at 0.55 for the adaptive condition and at 0.65 for the fixed condition, in both cases falling near the crossover point in the identification functions ($M = 0.589$, $SD = 0.087$ for the adaptive group and $M = 0.623$, $SD = 0.064$ for the fixed group). Such a pattern indicates that the feedback subjects were able to establish distinct perceptual categories corresponding to [r] and [l], albeit with a different placement of the boundary between the categories than we found in native speakers.

Generalization after training with feedback. In the identification data for the transfer continuum (Figure 6, bottom), there was a large overall improvement from the pre- to the posttest [$F(1,14) = 19.042, p < .001$] and no difference in the size of the improvement for the subjects receiving fixed versus adaptive training [$F(1,14) = 0.475, p = .502$]. A one-tailed t test supported the hypothesis that the improvement shown by the two groups of feedback subjects combined was significantly greater than the pre- to posttest improvement exhibited by the control subjects [$t(1,22) = 1.95, p = .032$]. As with the trained continuum, the crossover points of the identification functions were higher for our Japanese subjects than they were for native speakers [$M = 0.621$, $SD = 0.110$ for the adaptive training group and $M = 0.621$, $SD = 0.077$ for the fixed group, both significantly different from native speakers, $t(29) = 7.284, p < .001$ and $t(29) = 8.480, p < .001$, respectively]. There was also some slight evidence of transfer in the discrimination data after fixed training (Figure 8). In the expand test, there was a main effect of training in an analysis of variance carried out on discrimination performance in the fixed training condition [$F(1,7) = 7.373, p = .030$]. In the slide test, there was a trend after fixed training toward increased discrimination for stimuli bracketing 0.45,

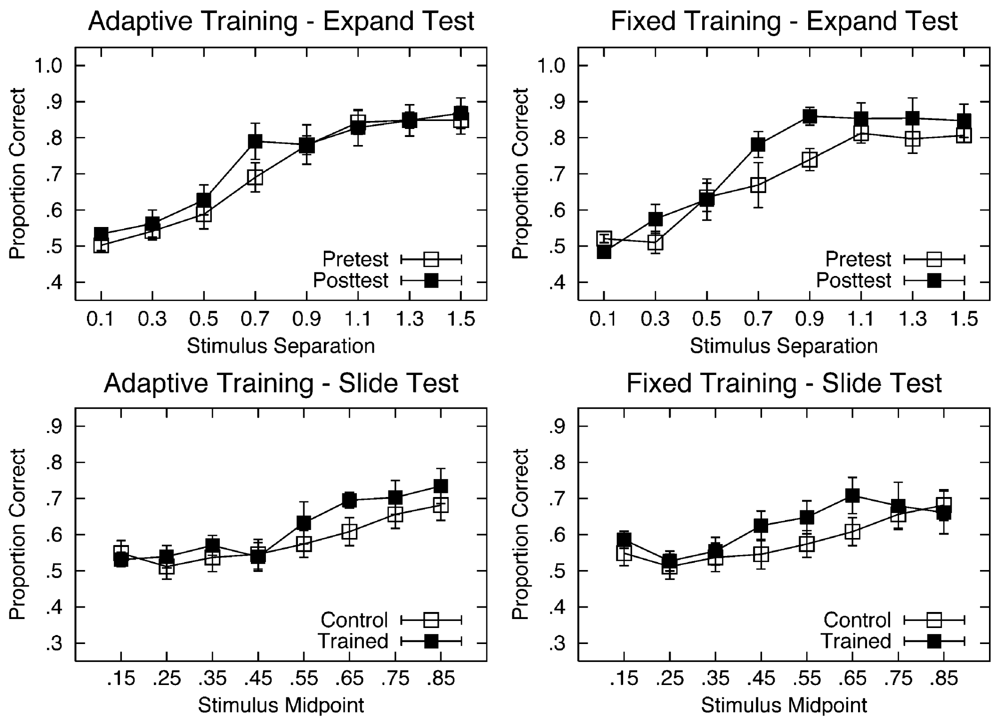

Figure 8. Transfer of training with feedback on discrimination of the [r]–[l] contrast. All data come from performance on the transfer continuum, rather than the continuum used in training. Top: mean discrimination performance (with standard error bars) as a function of the separation of the discriminative stimuli for Japanese subjects before and after training with feedback in the adaptive condition (left) and the fixed condition (right). The separation is measured relative to the separation of the anchor stimuli, which are 1.0 units apart. Bottom: mean discrimination performance for stimuli separated by 0.3 units as a function of the position of the midpoint between the two stimuli along the continuum between the two anchor stimuli. The panels contrast performance after adaptive training (left) and fixed training (right) with performance by control subjects on the posttest.

0.55, and 0.65, in the region of the crossover point for these subjects, but neither the effect of training (fixed vs. control) nor the interaction of training with position was reliable. There was no sign of a statistically reliable transfer effect after adaptive training in the expand test, and although there were some signs of improvement in the slide test, the effect was not statistically reliable.

Progress in the Different Training Conditions

The evidence presented thus far is quite clear in establishing that the availability of feedback can have a dramatic effect on learning, at least when subjects are trained with fixed stimuli that are initially difficult for them to discriminate. To provide further evidence relevant to this issue, we compared the performances of all four groups on the probe trials that employed the fixed training stimuli. Group average performance on these probe trials as a function of training (as reflected in successive sets of 10 test trials covering successive groups of 100 training trials) are shown in Figure 9 for all four groups. Note that the subjects in the two feedback groups received feedback on these probe trials, whereas the subjects in the two no-feedback groups did not receive feedback. Note further that for the subjects in the two fixed training groups, such trials were in no way different from the ongoing training trials. For the subjects in the two adaptive groups, such trials could differ in that the probe stimuli might not be the same as the training stimuli in use at that point in training.

It is evident from the data in this figure that the subjects receiving fixed training with feedback showed very rapid gains and the best performance overall, whereas the subjects in the fixed-training/no-feedback group showed little improvement. The subjects in the two adaptive conditions produced intermediate results. A three-way analysis of variance with session (first, second, or third), feedback condition (feedback vs. no feedback), and training condition (fixed vs. adaptive)[3] revealed main effects of session [$F(2,56) = 23.883$, $p < .001$] and feedback [$F(1,28) = 10.846$, $p = .003$], and the interaction of training condition with feedback approached significance [$F(1,28) = 3.639$, $p = .067$]. In a follow-up two-way analysis of variance focusing on the two fixed training conditions, there was a main effect of feedback [$F(1,14) = 11.273$, $p = .005$] and session [$F(2,28) = 7.014$, $p = .003$]. The interaction was not significant [$F(2,28) = 1.157$, $p = .329$]. In contrast, in a parallel analysis focusing on the two adaptive con-

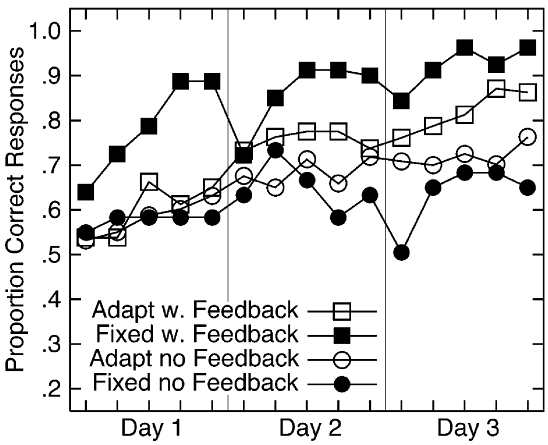

Figure 9. Mean percentage correct on probe trials for subjects in each of the four training conditions over the course of training. Each datapoint is based on 10 probe trials per subject and encompasses 100 training trials. Data from 2 subjects in the fixed/no-feedback condition have been excluded owing to a data recording error on Day 3. Their data were typical of the group on Days 1 and 2.

ditions, there was a main effect only of session [$F(2,28) = 17.351$, $p < .001$], but no effect of feedback [$F(1,14) = 1.200$, $p = .292$] and no interaction [$F(2,28) = 0.848$, $p = .439$].

Given that the advantage of fixed training with feedback began to appear at the very beginning of training, it is worth considering the possibility that the subjects in this condition were slightly better initially in identifying and/or discriminating the stimuli than were the subjects in the other conditions. To address this issue, we first examined performance on the fixed training stimuli from the identification component of the pretest. The fixed–feedback subjects were 59% correct in identifying these stimuli in the pretest, identical to the overall mean of the four groups, and an analysis of variance revealed no reliable initial differences among the groups. We also examined discrimination as measured in the expand task and found that the mean discrimination accuracy of the fixed–feedback group was 67% correct, just below the overall mean of 68% correct over all four training groups. For the actual fixed training stimuli, which constituted one of the expand test pairs, the fixed–feedback group achieved an accuracy of 54% correct, one percentage point above the mean of 53% correct for this pair over all four groups. An overall analysis of variance on the expand pretest results revealed no reliable effect of group and no interaction of group with the degree of separation of the members of the expand test pair. Overall, it appears that the subjects in the fixed–feedback group were not distinguishable from the other groups before the beginning of training. Thus, their advantage on the first set of probe trials seems most likely to reflect the onset of an advantage for fixed training with feedback within the first 100 trials of the experiment.

Effects of Further Training

The effects of 3 additional days of training were assessed by considering the performance of the 4 subjects, in each experimental condition, who received a second round of 3 days of training under identical training conditions. A comparison of the performances of this subset of subjects on the 3-day posttest versus the final posttest is shown for each of the four training conditions in Figure 10. For all groups, the 3-day posttest identification functions were quite similar to the posttest identification functions shown previously (see Figures 2 and 6), and an analysis of variance performed on the 3-day posttest data (with the slope of the best-fitting logistic regression function as the dependent variable) revealed no main effects or interactions involving a factor that distinguished those who received continued training and those who did not. Thus, it appears that each subgroup that was given further training was representative of the larger group from which it was derived.

Because of the small number of subjects in each subgroup, statistical comparisons between the groups are too weak to provide useful information, and there is no control available for assessing the extent to which any further improvements shown by these subjects can be attributed to the effects of the 3-day posttest, rather than to the effects of the additional training per se. Nevertheless, the data may be worth examining for signs that effects of training without feedback that were absent after 3 days of training might begin to appear with additional training experience. Before turning to these matters, we note that in the feedback conditions (bottom panels), further training had only a small effect on the identification functions over and above the effects already produced by the first 3 days of training.

The results in Figure 10 suggest that a second 3 days of training may have led to some improvement in identification performance of at least some of the subjects in the fixed–no-feedback condition. Indeed, 3 of the 4 subjects in this group showed a slope increase from the 3-day posttest to the final test (the remaining subject produced a flat identification function and stayed at chance across all 6 days of training). The results cannot unequivocally be taken as evidence of an effect of the training procedure per se in this condition, because there is no control for the effects of the 3-day posttest on performance on the final test. However, the fact that this group showed no improvement on the transfer continuum does provide some reason to suspect that the improvement seen on the trained continuum was not due solely to the effects of the 3-day posttest. With so few subjects, statistical confirmation of this difference is lacking, so once again we must stress that the data do not unequivocally indicate that there was a real training effect. At the same time, given the prediction that fixed training without feedback should produce no benefit, it seems important to acknowledge the possibility that extended training without feedback may produce a beneficial effect at least for some subjects.

The results in Figure 10 also provide some indication that with three extra training sessions, adaptive training with-

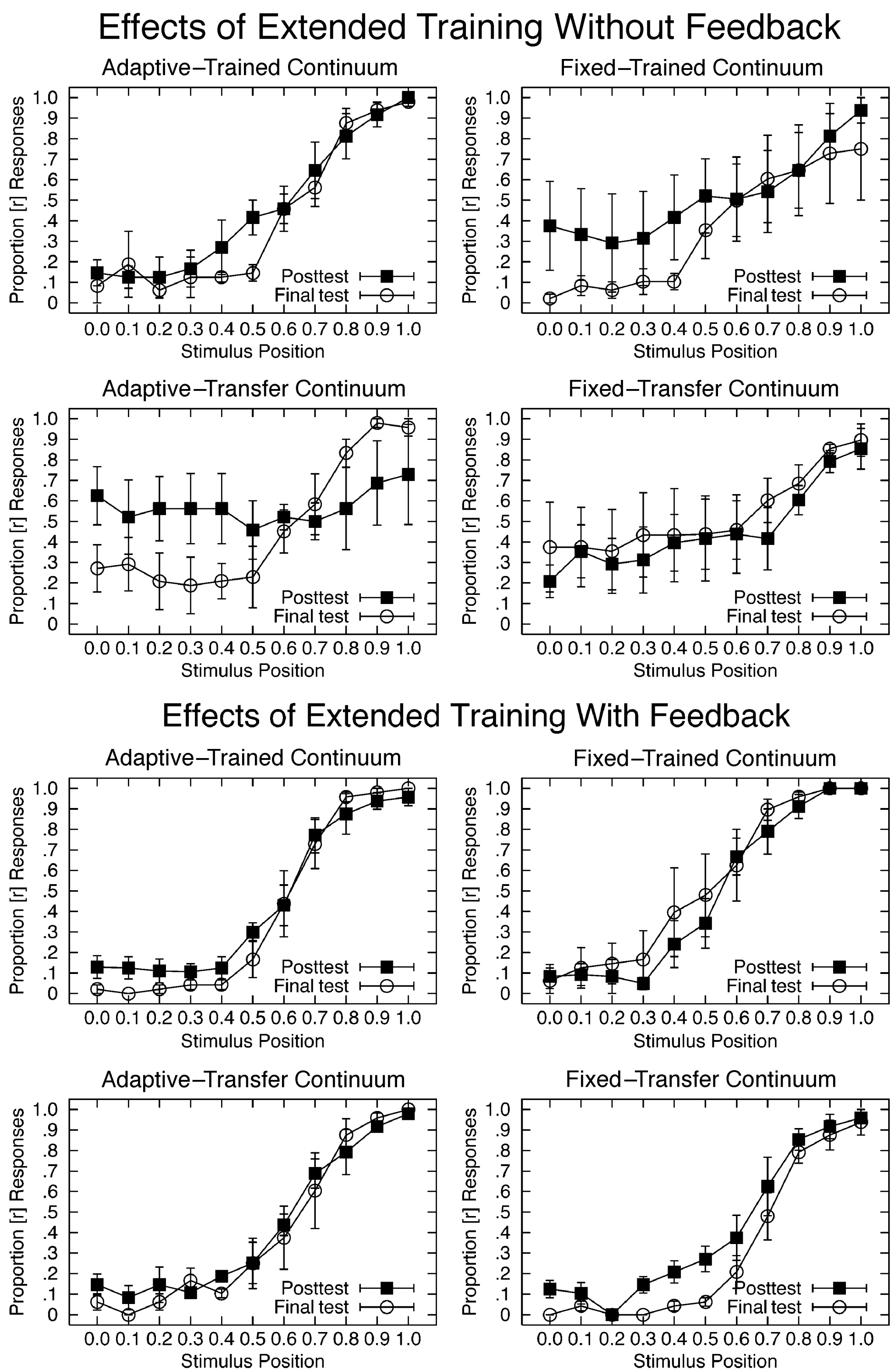

Figure 10. Performance on the posttest after 3 days of training and on the final test after a total of 6 days of training for 4 subjects in each training condition.

out feedback led to a transfer effect for at least some subjects. In the final posttest, the adaptive–no-feedback subgroup showed a much steeper group identification function on the transfer stimuli, relative both to their own function after 3 days of training and to the function shown on the transfer continuum by the subjects in the fixed–no-feedback condition. Three of the 4 subjects in the adaptive–no-feedback subgroup showed steeper identification functions after the second round of training, whereas 1 subject showed a slight reduction in slope. Again, firm conclusions are not possible, owing to the absence of a control for the effects of the 3-day posttest, but the data are at least suggestive that a total of 6 days of adaptive training without feedback can produce a transfer effect.

The discrimination data were somewhat noisier than the identification data, and there were no clear indications that further training led to further improvements in discrimination.

DISCUSSION

Our experiments were designed to test predictions arising from a Hebbian account of the mechanisms underlying the acquisition and stabilization of speech perception. According to the Hebbian account, the acquisition of one's native language can lead to the formation of a strong tendency to treat certain nonnative stimuli as the same. This tendency may be self-reinforcing, leading to its maintenance even when it is counterproductive, perhaps accounting in part for the difficulty Japanese adults have in acquiring the English contrast between [r] and [l]. The basis for this is the idea that the mechanisms of learning may tend to reinforce whatever reaction the brain has to a particular stimulus. Two predictions generated from the Hebbian account were borne out in the no-feedback conditions of our experiment.

First, the Hebbian account predicts that subjects who receive exposure to stimuli they perceive as the same should fail to benefit from training, even with hundreds of trials of exposure. In accord with this prediction, the group receiving training without feedback, with fixed stimuli that were difficult for them to discriminate, showed no evidence of learning after 3 days of training, once the slight gain attributable to exposure to the pretest was taken into account. Second, the Hebbian account predicts that subjects who receive exposure to stimuli that they *can* discriminate should benefit from that exposure, even if they receive no feedback on their accuracy in identifying these stimuli. In support of this, we found that the group receiving adaptive training, without feedback, in which initially exaggerated stimuli were used showed considerable gains in both identification and discrimination, when compared with a no-training control and when compared with the group receiving fixed training with difficult stimuli.

The findings are consistent with the idea that the self-reinforcing nature of Hebbian learning might play a role in determining how experience shapes and maintains the perceptual categorization of speech. Further support is provided by simulations (McClelland et al., 1999; Thomas & McClelland, 1997) using a Hebbian learning rule within a network based on Kohonen's self-organizing map architecture (Kohonen, 1982). The network embodies assumptions similar to those of Kuhl's native language magnet model (Iverson & Kuhl, 1995; Kuhl, 1993) and Flege's SLM (Flege, 1995), and a network model similar to ours in some ways has previously been used to simulate language-specific perceptual magnet effects (Guenther & Gjaja, 1996). In our simulations, "English" networks were exposed from their initialization to an English-like training environment, including two patterns analogous to [r] and [l], whereas "Japanese" networks were exposed instead to a Japanese-like environment that included just one pattern between the [r] and the [l] analogues. The English networks developed distinct representations for the [r] and the [l] analogues, whereas the Japanese networks developed a single representation to which they mapped both the [r] and the [l] inputs. We found that this representation persisted indefinitely in the Japanese networks if they were switched over to the English-like training environment, illustrating how Hebbian learning can maintain a single representation for two different inputs, once that representation has been established. However, as with our subjects, the Japanese networks could be remediated by using exaggerated training stimuli to produce distinct representations for the [r] and the [l] analogues. Once elicited, these distinct representations were strengthened by the operation of the learning mechanism, producing rapid change in perception, similar to what we have seen in our experiment.

As was previously noted, the data provided some suggestion of a beneficial effect of a second set of 3 days of fixed training without feedback in 3 of the 4 subjects who continued training beyond the first 3 days. Such a finding runs counter to the Hebbian account as we have stated it thus far, but the prediction of absolutely no learning in the no-feedback/fixed-training condition depends on the idea that the representations elicited by the two alternative stimuli are absolutely identical. If they tend to be even slightly different (e.g., if they involve the activation of largely, but not completely, overlapping populations of neurons), Hebbian learning may strengthen the connections supporting the nonoverlapping aspects of the representations promoting gradual differentiation. Just such a gradual differentiation effect may underlie the improvement seen in Japanese adults who have been exposed to English as adults for extended periods (Flege et al., 1996). We have seen in our modeling work (McClelland et al., 1999) that this can lead to a gradual process of differentiation. In the model, such differentiation is facilitated when the only items being processed are the two overlapping but not identical inputs. We will return to this point below when we consider the differences between naturalistic learning conditions and the conditions of our experiments.

Although two predictions of the Hebbian account were supported by our experiment, it is clear from the data from the feedback conditions that the Hebbian account is, at best, incomplete. In contrast to the fixed training condition without feedback, fixed training with feedback produced very robust learning. Indeed, the subjects in this condition showed learning as good as and, by some measures, slightly better than the subjects who received adaptive training with feedback. The finding that learning was so effective in the fixed training condition with feedback may be somewhat surprising, given the apparent benefits of an adaptive training regime in the studies reported in Merzenich et al. (1996) and Tallal et al. (1996). It should be noted, however, that there are very large differences between our study and theirs, and we would not want to suggest that the use of feedback would always eliminate an advantage for adaptive training, relative to fixed training. In particular, our subjects were all highly motivated adults who may have been able to tolerate more frustration during early stages of fixed training than the language-impaired children used in Merzenich et al.'s and Tallal et al.'s studies.

In any case, our finding that feedback does make a very large difference, at least in the fixed training conditions of our experiment, makes it clear that Hebb's idea about the basis of changes in the strengths of connections between neurons does not provide a full account of the factors that influence perceptual learning. One response to this situation might be to call into question the Hebbian approach and to consider whether it should be abandoned in favor of other approaches. Within the domain of connectionist or neural models, one alternative approach to accounting for our findings could be constructed on the basis of error-driven, instead of Hebbian, learning (Rescorla & Wagner, 1972; Rosenblatt, 1962; Rumelhart, Hinton, & Williams, 1986). Although we certainly cannot rule such approaches out, we nevertheless believe that the Hebbian approach has considerable merit. It provides a mechanism that can be used to allow a system to organize in response to experience, even when no feedback is available. This type of learning appears to us to be highly relevant to the initial shaping and stabilization of the neural mechanisms of perception in other modalities, including vision (Linsker, 1986a, 1986b, 1986c; Miller, Keller, & Stryker, 1989), as well as for the formation of speech categories in the absence of feedback or category labels, as was discussed above. We therefore favor proposals in which Hebbian learning plays a role, and either is combined with error-correcting learning methods (O'Reilly, 1996, 1998) or is modulated by outcome information, as in reinforcement learning (Barto, 1994). A challenge for a model based on these ideas will be to explain why fixed training with feedback can lead to learning that is at least as good as or better than adaptive training under some conditions. This is a matter we intend to consider in ongoing modeling studies.

Our research has been motivated by the Hebbian account of learning we have discussed throughout this article, but we do not wish to suggest that our findings provide unique support for a Hebbian approach to understanding the acquisition of a new phonological distinction in adulthood. Quite apart from this approach, there are other ideas and considerations that may be relevant to understanding the perceptual learning process and the role of feedback in our experiments. For example, it might be suggested that both exaggeration and feedback facilitate learning by calling a subject's attention to the cues that distinguish the training stimuli. This suggestion provides one way of offering a unified account of our identification findings, since both exaggeration and feedback would be operating to produce the same effect. One can easily see how exaggeration could help orient attention, since exaggeration would make the differences between the stimuli highly salient. Ways in which feedback can operate in directing attention to distinguishing cues have been explored by several investigators. For example, Kruschke (1992) has introduced mechanisms that allow the explicit adjustment, on the basis of feedback, of attention weights assigned to particular stimulus dimensions. A second idea is that feedback may allow subjects to associate the correct category label to the fixed training stimuli. Sawusch and Gagnon (1995) have suggested that labels associated with representations of training stimuli may allow the subject to generate an internal referent for each of the two categories, with which the current input can be compared for purposes of identification. This idea can provide the basis for a unified account of the identification results as well, in that one might suggest that the effect of exaggeration is to increase the likelihood that the subject will be able to generate consistent labels even in the absence of feedback. We have not elaborated upon either of these accounts, since we are not specifically advocating either one, and we emphasize that these ideas are far from exhaustive and may not be mutually exclusive. We present them primarily as examples indicating that there may be more than one way to account for the findings we have reported.

Three days of training without feedback did not appear to produce transfer of training between the continua used in our experiments, but evidence of transfer began to emerge after 6 days of adaptive training. Furthermore, training with feedback produced clear evidence of transfer, even after 3 days. In the fixed training condition with feedback, there was clear evidence that transfer occurred in both identification and discrimination. This evidence of transfer is important in ruling out the possibility that what the subjects learned was merely some idiosyncratic property of the particular stimuli used in training. We emphasize, however, that our findings cannot be taken as evidence that what our subjects have learned is a very general phonological discrimination that would apply to all instances of [r] and [l] spoken in all contexts by all speakers. Previous investigations (Bradlow et al., 1999; Bradlow et al., 1997; Lively et al., 1993; Lively et al., 1994) do suggest that training can lead to broad generalization, but only if the training is correspondingly broad. Morosan and Jamieson (1989) reported generalization of the benefits of training with synthetic stimuli on a word-initial [ð]–[θ] distinction to natural tokens of word-initial [ð]–[θ] contrasts produced by several speakers in a range of vowel contexts, but the effect of training did not generalize to medial or final positions. It seems likely that the exact extent of generalization will vary with the particular contrast used.

With these points in mind, we can consider the question of whether our findings actually reflect the effects of training on the mechanisms that adult speakers ordinarily use for the processing of speech. The fact that the training tended to generalize from one continuum to another is consistent with this, and the fact that generalization was somewhat less than perfect is not unexpected, given the previous findings just cited. One further point relevant to this issue is the observation that, at least with feedback, training appears to produce evidence of categorical perception, a phenomenon that has sometimes been described as special to the mechanisms of spoken language processing (Liberman, Cooper, Shankweiler, & Studdert-Kenedy, 1967). Evidence for categorical perception is provided by the fact that when training was successful (i.e., in the adaptive/no-feedback condition and in both conditions with feedback), the identification functions became fairly steep and sigmoidal in shape and by the fact that training produced an increase in discrimination around the crossover point of the posttraining identification functions. However, even with this evidence, it is possible to adopt an approach like that advocated by Sawusch and Gagnon (1995), in which perceptual category formation (as evidenced by a steep identification function, enhanced discrimination at a category boundary, and generalization) can arise with complex auditory stimuli that are not perceived as speech. Thus, we cannot rule out the possibility that the categories our subjects learned were auditory, rather than strictly phonological. Furthermore, even if they were phonological, it would be difficult to know whether the effects of training reflected reorganization of the same perceptual mechanisms as those used in the perception of a native language or whether, instead, they reflected the recruitment of additional neural hardware that had not previously been used for speech discrimination but that exhibited properties similar to those exhibited by the established speech perception machinery. Functional brain imaging studies of second-language learners may provide some evidence bearing on these issues, and some of the members of our group are pursuing such investigations.

There is at least one way in which the perceptual categories acquired by our Japanese subjects differed from the [r]–[l] categories of native English speakers. The Japanese subjects placed the boundary between their two categories farther toward the [r] end of the continuum than did our native English speaking subjects; in other words, they tended to place intermediate stimuli in the same category with English [l], rather than English [r]. One interpretation of these findings arises in the context of evidence in other studies that English [r] and [l] may not be equidistant from the closest Japanese phoneme (an apical-alveolar tap; Guion et al., 2000). Some findings (Guion et al., 2000, Experiment 2; Takagi, 1993; but see Guion et al., 2000, Experiment 1) suggest that the Japanese sound may be more similar to English [l] than it is to English [r]. In this context, one possibility is that our subjects learned a new category for English [r], while continuing to hear members of the English [l] category as exemplars of their existing Japanese category. The fact that the boundary was shifted toward the [r] end of the continuum, even though the training stimuli were symmetrically placed around the English crossover point, would be consistent with the possibility that their preexisting category continued to exert a relatively strong attractor or magnet-like effect, assimilating some of the intermediate stimuli. Only stimuli with strong [r]-like content could escape the attractor and be reliably heard as members of the newly emerging perceptual category.

The substantial learning in just three 20-min sessions seen in three of our training conditions underscores the point that adult language learners maintain considerable plasticity in their ability to learn perceptual speech contrasts. This rapid learning contrasts with the very slow improvement shown by adult Japanese second-language learners receiving the [r] and [l] stimuli in spoken English (Flege et al., 1996). There are two obvious factors that may contribute to the difference between our experimental conditions and naturalistic situations. First, in some of our conditions, we made use of highly exaggerated contrasts, and even the natural tokens used to construct our stimulus continua may have been somewhat exaggerated or, at least, very clear, as compared with natural tokens encountered in daily experience with running speech. Although mothers do exaggerate speech to infants in ways that might facilitate the initial establishment of phonetic categories (Kuhl et al., 1997), it seems unlikely that exaggeration is regularly provided in naturalistic speech to adults. Second, in some of our conditions, we provided feedback, and it is unlikely that feedback is often available in naturalistic conditions. Indeed, in training without either exaggeration or feedback, we find little or no sign of learning over the first three training sessions. These two factors alone may be enough to account for the difficulty that second-language learners face in acquiring the [r]–[l] contrast under naturalistic conditions. However, there is another characteristic of our experiment that may have contributed to the rapid learning that we observed in most of the conditions and even to the eventual emergence of a slight learning effect in the fixed/no-feedback condition. This characteristic is the use of a single contrasting stimulus pair, such as *rock* and *lock*, spoken by a single individual. In our modeling work (McClelland et al., 1999), we have found that when we expose our networks to a single pair of stimuli, their representations eventually separate, even if initially they are highly overlapping. This separation does not occur for such pairs if the set of training stimuli includes several additional stimuli. This difference in the response of the model in these two situations reflects the fact that the competing stimuli divide up the units available for representing speech sounds into clusters. When there are many stimuli competing for units, stimuli that are highly overlapping are forced into the same cluster, but when there are only two stimuli, they will eventually divide up the space between them even if they are highly overlapping. Slight differences in their representations eventually become amplified, leading to the gradual emergence of distinct representations.

The focusing of experience on a single contrasting stimulus pair may also be part of the explanation for the differences between our findings and those of Lively et al.

(1993; Lively et al., 1994) and Bradlow et al. (1999; Bradlow et al., 1997). These studies used fixed stimuli with feedback—conditions that led to very rapid learning in our experiment— but they employed a large number of different minimal pairs all mixed together in their training set. Although their approach did lead to robust generalization, it also led to learning that was rather slow and still incomplete even after 45 h of training. The contrast between our findings and theirs raises two possibilities that may be worth examining in future investigations. The first is that the diversity of the [r]–[l] contrast, as it is exhibited in different contexts and with different speakers, may contribute to the difficulty second-language learners have in acquiring this contrast from natural experience. The second is that focusing training on a single contrast and/or a single speaker at a time may lead to faster and/or asymptotically better learning. Some initial success has been obtained with such an approach (Protopapas & Calhoun, in press; cf. Pruitt, 1995), but the matter deserves additional investigation.

REFERENCES

Barto, A. G. (1994). Reinforcement learning control. *Current Opinion in Neurobiology*, **4**, 888-893.

Best, C. [T.] (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Baltimore: York.

Best, C. T., McRoberts, G. W., & Sithole, N. (1988). The phonological basis of perceptual loss for nonnative contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 345-360.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in speech perception and production. *Perception & Psychophysics*, **61**, 977-985.

Bradlow, A. R., Pisoni, D. B., Yamada, R. A., & Tohkura, Y. (1997). Training the Japanese listener to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, **101**, 2299-2310.

Flege, J. [E.] (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-273). Baltimore: York.

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *Journal of the Acoustical Society of America*, **99**, 1161-1173.

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.

Guion, S. G., Flege, J. E., Akahane-Yamada, R. K., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, **107**, 2711-2724.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Iverson, P., & Kuhl, P. (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.

Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English ð–θ contrast by francophones. *Perception & Psychophysics*, **40**, 205-215.

Jamieson, D. G., & Morosan, D. E. (1989). Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology*, **43**, 88-96.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, **43**, 59-69.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, **99**, 22-44.

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. (1993). Developmental speech perception: Implications for models of language impairment. In P. Tallal, A. M. Galaburda, R. R. Llinas, & C. von Euler (Eds.), *Temporal information processing in the nervous system: Special reference to dyslexia and dysphasia* (Annals of the New York Academy of Sciences, Vol. 682, pp. 248-263). New York: New York Academy of Sciences.

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., & Iverson, P. (1995). Linguistic experience and the "perceptual magnet effect." In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121-154). Baltimore: York.

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*, **225**, 606-608.

Lacerda, F. (1995). The perceptual-magnet effect: An emergent consequence of exemplar-based phonetic memory. In K. Elenius & R. Brandrud (Eds.), *Proceedings of the 13th International Congress of Phonetic Sciences* (Vol. 2, pp. 140-147). Stockholm: Kungliga Tekniska Högskolan and Stockholm University.

Liberman, A. M., Cooper, F. S., Shankweiler, D. S., & Studdert-Kenedy, M. (1967). Perception of the speech code. *Psychological Review*, **74**, 431-461.

Linsker, R. (1986a). From basic network principles to neural architecture: I. Emergence of spatial-opponent cells. *Proceedings of the National Academy of Sciences*, **83**, 7508-7512.

Linsker, R. (1986b). From basic network principles to neural architecture: II. Emergence of orientation-selective cells. *Proceedings of the National Academy of Sciences*, **83**, 8390-8394.

Linsker, R. (1986c). From basic network principles to neural architecture: III. Emergence of orientation columns. *Proceedings of the National Academy of Sciences*, **83**, 8779-8783.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, **94**, 1242-1255.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/and /l/: III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, **96**, 2076-2087.

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.

Logan, J. S., & Pruitt, J. S. (1995). Methodological issues in training listeners to perceive nonnative phonemes: Linguistic experience and the "perceptual magnet effect." In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121-154). Baltimore: York.

McClelland, J. L., Thomas, A., McCandliss, B. D., & Fiez, J. A. (1999). Understanding failures of learning: Hebbian learning, competition for representational space, and some preliminary experimental data. In J. Reggia, E. Ruppin, & D. Glanzman (Eds.), *Brain, behavioral, and cognitive disorders: The neurocomputational perspective* (pp. 75-80). Oxford: Elsevier.

Merzenich, M. M., Jenkins, W. M., Johnson, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science*, **271**, 77-81.

Miller, K. D., Keller, J. B., & Stryker, M. P. (1989). Ocular dominance column development: Analysis and simulation. *Science*, **245**, 605-615.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, **18**, 331-340.

Morosan, D. E., & Jamieson, D. G. (1989). Evaluation of a technique for training new speech contrasts: Generalization across voices, but not word-position or task. *Journal of Speech & Hearing Research*, **32**, 501-511.

O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, **8**, 895-938.

O'Reilly, R. C. (1998). Six principles for biologically based computational models of cortical cognition. *Trends in Cognitive Sciences*, **2**, 455-462.

Pisoni, D. B., Aslin, R. N., Perey, A. J., & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 297-314.

Protopapas, A., & Calhoun, B. (2000). Adaptive phonetic training for second language learners. In P. Delcloque (Ed.), *Proceedings of the 2nd International Workshop on Integrating Speech Technology in Language Learning* (pp. 31-38). University of Abertay Dundee, U.K.

Pruitt, J. S. (1995). *The perception of Hindi dental and retroflex stop consonants by native speakers of Japanese and American English* (University Microfilms No. 9542085). Unpublished doctoral dissertation, University of South Florida.

Pruitt, J. S., Kawahara, H., Akahane-Yamada, R., & Kubo, R. (1998). Methods of enhancing speech stimuli for perceptual training: Exaggerated articulation, context truncation, and "STRAIGHT" re-synthesis. In R. Carlson, C. Dunger, B. Granstrom, & A. Oster (Eds.), *Speech technology in language learning* (pp. 105-108). Stockholm, Sweden.

Rabiner, L. R., & Schafer, R. W. (1978). *Digital processing of speech signals*. Englewood Cliffs, NJ: Prentice-Hall.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Block & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Rosenblatt, F. (1962). *Principles of neurodynamics*. New York: Spartan.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, **323**, 533-536.

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

Sawusch, J. R., & Gagnon, D. A. (1995). Auditory coding, cues, and coherence in phonetic perception. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 635-652.

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.

Takagi, N. (1993). *Perception of American English /r/ and /l/ by adult Japanese learners of English: A unified view*. Unpublished doctoral dissertation, University of California, Irvine.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagaraja, S. S., Schreiner, C., Jenkins, W. M., & Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, **271**, 81-84.

Tees, R. C., & Werker, J. F. (1984). Perceptual flexibility: Maintenance or recovery of the ability to discriminate nonnative speech sounds. *Canadian Journal of Psychology*, **38**, 579-590.

Thomas, A., & McClelland, J. L. (1997). How plasticity can prevent adaptation: Induction and remediation of perceptual consequences of early experience. *Society for Neuroscience Abstracts*, **23**, 234.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, **37**, 278-286.

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.

NOTES

1. Slope estimates were constrained to values less than or equal to 4.0 because the slope estimation procedure is oversensitive at the high end.

2. Note that in all statistical tests involving the control group, results have been collapsed over the two continua, since the distinction between *trained* and *transfer* is meaningless in this case.

3. For this analysis, it was necessary to estimate the Day 3 performance of 2 subjects in the fixed/no-feedback condition, since their data were lost owing to a recording error. The estimated data were constrained to preserve the mean and standard deviation of the Day 3 data exhibited by the rest of the fixed/no-feedback subjects.

(Manuscript received August 30, 2001;
revision accepted for publication April 9, 2002.)