WARPING AUDITORY PERCEPTUAL SPACE WITH CATEGORIZATION AND DISCRIMINATION TRAINING

Frank H. Guenther[1,2] and Fatima T. Husain[1]
[1]*Department of Cognitive and Neural Systems, Boston University*
[2]*Research Laboratory of Electronics, Massachusetts Institute of Technology*

ABSTRACT

This paper investigates the effects of two different kinds of training on auditory perceptual space. It is first shown that categorization training, in which subjects learn to identify stimuli within a particular frequency range as members of the same category, leads to a decrease in sensitivity (as measured by d') to stimuli in that category. This phenomenon is an example of acquired similarity for a category-relevant dimension. Discrimination training with the same set of stimuli was shown to have the opposite effect: subjects became more sensitive to differences in the stimuli presented during training. Further experiments investigated some of the conditions that are necessary to generate the acquired similarity found in the first experiment. The results of these experiments are interpreted using a neural network model of auditory map formation in the central nervous system.

1. INTRODUCTION

It is well-known that our perceptual spaces for some auditory stimuli, such as phonemes, are warped. That is, the perceptual distance between two stimuli, as evidenced by a subject's ability to discriminate them, is not always a straightforward function of their distance measured along physical dimensions such as frequency or time. English stop consonants, for example, have long been known to undergo a process of categorical perception (see [14] for a review). Researchers have also shown that the perceptual space for synthetic vowels and semivowels is warped [7, 8]. Kuhl [7] referred to this warping as a "perceptual magnet effect", distinguishing it from categorical perception. Roughly speaking, the effect is characterized by a warping of perceptual space such that acoustic patterns near phonemic category prototypes are perceived as closer together than equally spaced acoustic patterns that are further away from phonemic category prototypes.

Some instances of auditory space warping appear to arise from learning rather than from built-in properties of the auditory system. One example is the language specificity of the warping of auditory space for vowels as measured in studies of the perceptual magnet effect. In a study of 6-month-old English and Swedish infants presented with English and Swedish vowel stimuli, Kuhl et al. [8] found that infants had more difficulty discriminating between stimuli falling near a prototypical vowel from their native language than stimuli falling near a prototypical vowel in the non-native language.

Liberman [10] identified two possible learning processes that might underly this phenomenon. The first, acquired distinctiveness, is defined as an increase in perceptual sensitivity for items that are repeatedly categorized differently in some learning situation. Liberman [10] reported evidence for acquired distinctiveness in detecting duration differences for speech sounds vs. non-speech sounds, and later studies provided further examples of acquired distinctiveness for non-speech stimulus sets [3,9]. The second possible learning process identified by Liberman was acquired similarity, also referred to by some authors as acquired equivalence. In acquired similarity, sounds that were originally distinguishable from each other become less distinguishable after repeatedly being categorized together. If acquired similarity is playing a role the learned instances of categorical perception or the perceptual magnet effect, it must involve category-relevant dimensions: the very notion of "nearer to the category boundary" that is commonly used to describe these phenomena implies that we are talking about category-relevant dimensions, such as formant frequencies for vowels. Although attempts have been made [3], acquired similarity for a category-relevant dimension has apparently not been shown experimentally [11, pp. 18-19]. One goal of the current study was to address the following question: is it possible to induce a warping of auditory space that takes the form of reduced discriminability for category-relevant dimensions of frequently encountered stimuli, as appears to be the case in the perceptual magnet effect, by training adult subjects on a categorization task using stimuli that are not categorical prior to training?

2. EXPERIMENTS

Four experiments were performed, utilizing 10 subjects each. Subjects were male and female adults between the ages of 18 and 50 with no history of speech, language or hearing disorders. The stimuli for all experiments were auditory stimuli consisting of narrow-band filtered samples of white noise with different center frequencies. These stimuli did not resemble speech sounds.

All experiments consisted of four phases: a calibration phase in which a subject's detection threshold for auditory stimuli like those to be used later in the experiment was detemined, a pre-test phase to determine baseline sensitivity to the auditory stimuli, a training phase which differed between experiments, and a post-test phase to measure any change in sensitivity that may have resulted from training. The pre- and post-tests followed an AX Same-Different paradigm, and sensitivity was estimated using a collapsed d' measure. Sensitivity was measured in two different regions of frequency space: a control region and a training region. The stimuli in each region consisted of a "milestone" stimulus at the center of the region and six stimuli located 1, 1.5, and 2 JND above and below the milestone in frequency space. The training region stimuli were used during the training phase of the experiment. Control region stimuli were not encountered during training.

2.1. Experiment I

The main goal of the first experiment was to investigate whether it is possible to induce a decrease in discriminability along a category-relevant dimension of a set of non-speech auditory stimuli that were repeatedly encountered during a training session. This would constitute a demonstration of acquired similarity along a category-relevant dimension.

2.1.1. Methods. In the training phase of Experiment I, subjects were trained to identify sounds that belonged to the training region from a list of sounds. More specifically, subjects were told that they were to learn to identify sounds from a category, referred to as the "prototype category" and corresponding to the training region of frequency space, and that during training they would have to choose the prototype category sound from a list of sounds that included only one member of the prototype category. Since the subjects were learning to treat the training region sounds as members of the same category, we will refer to this type of training as categorization training. The training phase lasted approximately 45 minutes.

2.1.2. Results. Figure 1 compares the change in sensitivity for the control and training regions. The change in sensitivity was calculated as the percentage increase or decrease in d' from pre-test to post-test. This figure indicates that the change in sensitivity for the training set of stimuli was significantly more negative (t=-5.14, $p<0.01$) than the change in sensitivity for the control region. These results indicate that it is possible to induce acquired similarity along a category-relevant dimension if an appropriate training regime is utilized.

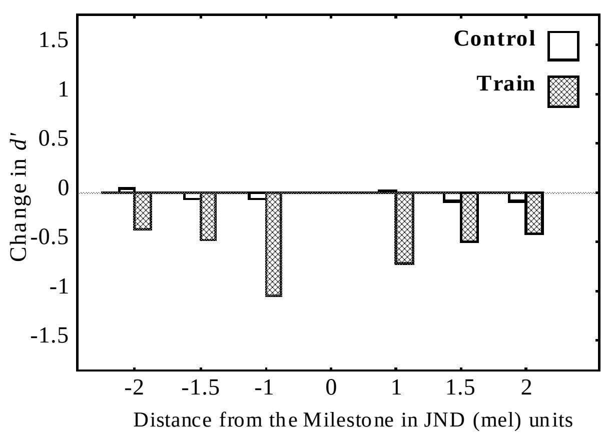

Figure 1. Change in sensitivity after training for the control and training regions in Experiment I. Subjects showed a significant decrease in sensitivity for the stimuli in the training region but not in the control region.

2.2. Experiment II

The second experiment tested whether a training regimen different from that used in Experiment I could result in a different effect on the subjects' sensitivity to the training stimuli even though a similar distribution of sounds is presented during training.

2.2.1. Methods. In Experiment II, a discrimination training paradigm was used in which subjects were repeatedly asked to report whether they thought two sounds chosen from the training region were the same or different. Subjects were given feedback concerning the correctness of their responses. The training phase lasted approximately 45 minutes.

2.2.2. Results. Figure 2 shows the change in sensitivity after training for the training and control regions. The training region showed a significantly greater increase in sensitivity than the control region (t=3.23, $p<0.05$). This indicates that the same distribution of training stimuli that led to a decrease in sensitivity for the training region in Experiment I can lead to an increase in sensitivity if the training regime is changed to a discrimination training task. This is a case of acquired distinctiveness along a category-relevant dimension (see also [3]).

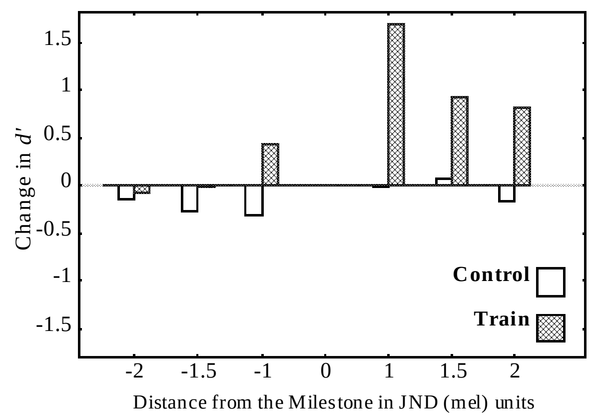

Figure 2. Change in sensitivity after training for the control and training regions in Experiment II.

2.3. Experiment III

The third experiment was designed to elaborate on the training conditions required to induce the perceptual magnet-like effect of acquired similarity along a category-relevant dimension that was demonstrated in Experiment I. The specific question this experiment sought to answer was whether training with only a single exemplar from a category is sufficient to induce decreased sensitivity in its immediate region of acoustic space. It is possible that a listener must experience many exemplars from the same category in order to induce acquired similarity. This scenario makes sense if one takes the view that acquired similarity is a case of learning to "ignore" differences between exemplars of the same category; if subjects hear only one exemplar of a category, there are no differences between category exemplars to learn to ignore.

2.3.1 Methods. This experiment involved a categorization training regime that differed from that of Experiment I in only one respect: instead of hearing different exemplars from the training region when performing the training task, subjects always heard the same exemplar, the milestone.

2.3.2. Results. The change in sensitivity for the training region was not significantly different from the change in sensitivity for the control region (*t*=0.30, p=0.38). In other words, using only a single exemplar from the training region during training did not lead to a significant decrease in discrimination performance for the training region as compared to the control region. This suggests that a single category exemplar is not sufficient to induce acquired similiarity in the neighborhood of the category exemplar, or at minimum that a single exemplar does not induce as much acquired similarity as multiple exemplars.

2.4. Experiment IV

Several investigators have suggested that the brain's representation of sounds can be broken into two different memory modes: a continuous auditory memory mode that consists of a reasonably accurate representation of a sound that decays relatively rapidly after the stimulus goes away or is interrupted by a new auditory stimulus, and a more "discretized" or "categorical" mode that can be maintained in memory for a longer period of time, e.g. for comparison to a second stimulus in a discrimination task with a relatively large interstimulus interval (ISI). When investigating speech sounds, Pisoni [13] referred to the different memory forms as auditory mode and phonetic mode. In a model of sound intensity discrimination, Durlach and Braida [2] delineated two memory modes that they termed sensory-trace mode and context-coding mode; these modes are roughly analogous to Pisoni's auditory mode and phonetic mode, respectively. The Durlach and Braida model has been extended to explain experimental results involving speech stimuli [12]. Since we are not dealing with speech stimuli in the current experiments, we will use the terms sensory-trace mode and context-coding mode here.

The purpose of Experiment IV was to determine whether the acquired similarity induced in Experiment I could be better characterized as a result of changes in the sensory-trace mode or the context-coding mode of auditory memory. It is usually assumed that increasing the ISI and/or adding a brief noise burst between two stimuli interferes with the sensory-trace mode of memory more than the context-coding mode [14, 16]. Given the relatively long ISI of Experiment I and the use of a noise burst between the two stimuli in a discrimination trial, one might reasonably conclude that the effect measured in that experiment primarily involved the context-coding mode of auditory memory.

2.4.1 Methods. The methods for Experiment IV were the same as those of Experiment I except that the ISI for the same-different testing procedure was reduced from 1 second to 250 milliseconds and no interfering noise burst was used. These conditions should favor a sensory-trace mode of memory over a context-coding mode.

2.4.2. Results. The change in sensitivity in the training region was not significantly different from the change in the control region (*t*=-0.63, p=0.28). Since decreasing the ISI and removing the noise burst presumably favors a sensory-trace memory mode over a context-coding memory mode, this result suggests that the acquired similarity seen in Experiment I was primarily associated with the context-coding mode of auditory short term memory. This result is consistent with the hypothesis that a shorter ISI can diminish the categorical nature of the responses made by an observer [12, 13, 14, 16].

3. DISCUSSION

The results of the first two experiments indicate that, depending on the training regime, it is possible to induce either an increase or a decrease in the discriminability of a set of auditory stimuli. The first experiment indicated that categorization training, in which subjects were asked to identify sounds belonging to a small region of frequency space as members of the same category, led to a decrease in the discriminability of stimuli within this small range. That is, subjects exhibited acquired similarity along the category-relevant dimension of center frequency of the narrow-band noise stimuli. The third and fourth experiments helped elucidate some of the necessary conditions for attaining this acquired similarity. In Experiment III, the small range of frequencies corresponding to the learned category in Experiment I was shrunk down to a single exemplar during training. This eliminated the acquired similarity seen in Experiment I, suggesting that a listener needs to be exposed to different examples of a category during training, not just a single exemplar, in order to decrease the listener's ability to discriminate between stimuli falling near the center of the category. In Experiment IV, a testing regime that favors a hypothesized sensory-trace mode of auditory memory over a context-coding mode weakened the acquired similarity effect of training, suggesting that categorization training primarily affects the context-coding mode of memory processing.

Bauer et al. [1] and Guenther and Gjaja [5] have proposed models of the perceptual magnet effect that attribute it to neural map formation properties in auditory brain areas such as the primary auditory cortex (Brodmann's Areas 41/42; Heschl's gyrus). According to both of these models, the learning process during which infants develop phonemic categories involves a change in the distribution of firing preferences of cells in auditory cortex; it is this change in the auditory neural map for vowel-like sounds that underlies the perceptual magnet effect. Because both models posit that the magnet effect results from neural map formation properties that are not specific to speech stimuli, they predict that exposing a listener to new, non-speech auditory stimuli within a training regime that appropriately mimics the learning of phonemic categories by an infant should lead to a similar change in the distribution of firing preferences of cells coding these stimuli in auditory cortex. This change in the auditory neural map should in turn result in a measurable "perceptual magnet-like" effect for these auditory stimuli: we should see a decreased ability for subjects to discriminate the training stimuli. This prediction was tested in Experiment I and supported by that experiment's results: subjects showed a decrease in the ability to discriminate stimuli from a heavily experienced training region.

In keeping with these neural models of the magnet effect, the results of Experiments I and II can be interpreted in terms of changes in the distribution of cells in auditory neural maps. Figure 3 illustrates the hypothesized changes in the neural map in auditory cortex as a result of categorization training (left; Experiment I) and discrimination training (right; Experiment II). The top and bottom panels schematize the auditory map as a function of acoustic space before and after training, and the middle panel schematizes the distribution of training stimuli in acoustic space. The x and y axes of all plots correspond to two acoustic dimensions, such as the first two formant frequencies. The z axis corresponds to the number of cells in the map devoted to each region of frequency space (top and bottom plots) or the number of training stimuli from that region of frequency space (middle plots). In categorization training, heavy exposure to a set of training sounds leads to fewer cells coding these sounds in the auditory map, and the resulting smaller cortical representation diminishes a listener's ability to differentiate sounds in this region of acoustic space. This is how the Bauer et al. model, with an appropriate parameter choice that leads to a negative magnification factor for the cortical representation, accounts for the perceptual magnet effect. The right side of Figure 3 corresponds to a discrimination training situation, as in Experiment II. Here, more cells in the map become tuned to the most frequently encountered training stimuli, and the resulting larger cortical representation increases the listener's ability to differentiate sounds in this region of acoustic space. This learning situation corresponds to the "classical" formulation of a self-organizing feature map in the computational neuroscience literature, in which increased exposure to a set of stimuli leads to a larger cortical representation for those stimuli [4, 6, 15], and can also be accounted for by using a positive magnification factor in the Bauer et al. model. We are currently testing predictions of the hypothesis illustrated in Figure 3 using functional magnetic resonance imaging techniques.

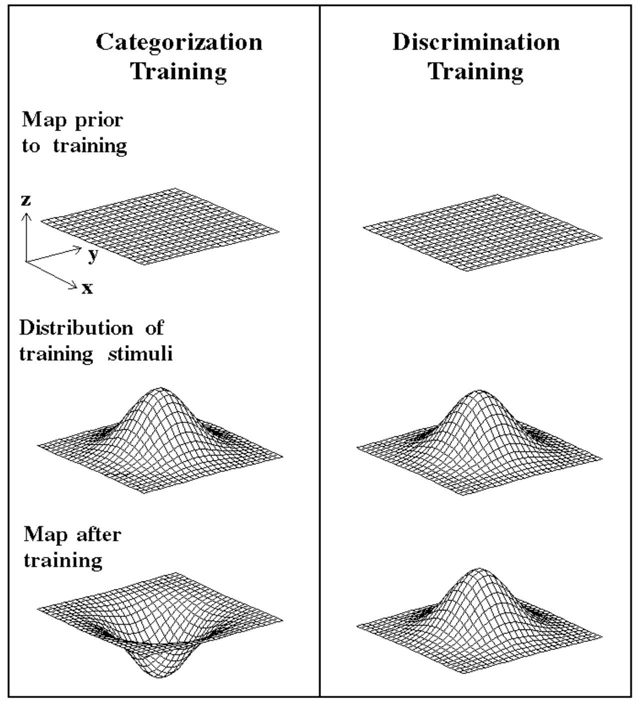

Figure 3. Hypothesized changes in the neural map in auditory cortex as a result of categorization training (left) and discrimination training (right). Categorization training leads to a decrease in the number of cells coding the most frequently encountered training stimuli, whereas discrimination training leads to an increase in the number of cells coding the most frequently encountered training stimuli.

ACKNOWLEDGMENTS

This research was supported by NIDCD grant 1R29DC02852.

REFERENCES

[1] Bauer, H.-U., Der, R., and Herrmann, M. 1996. Controlling the magnification factor of self-organizing feature maps. *Neural Computation*, 8, 757-771.

[2] Durlach, N.I., and Braida, L.D. 1969. Intensity perception. I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America*, 46(2) pt. 2, pp.372-383.

[3] Goldstone, R. 1994. Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology*, 123, 178-200.

[4] Grossberg, S. 1976. Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics*, 23, 121-134.

[5] Guenther, F.H., and 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.

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

[7] 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.

[8] Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., and Lindblom, B. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606-608.

[9] Lane, H. 1965. The motor theory of speech perception: A critical review. *Psychological Review*, 72, 275-309.

[10] Liberman, A.M. 1957. Some results of research on speech perception. *Journal of the Acoustical Society of America*, 29, 117-123.

[11] Liberman, A.M. 1996. *Speech: A special code*. Cambridge, MA: MIT Press.

[12] MacMillan, N.A., Goldberg, R.F., and Braida, L.D. 1988. Resolution for speech sounds: Basic sensitivity and context memory on vowel and consonant continua. *Journal of the Acoustical Society of America*, 84, 1262-1280.

[13] Pisoni, D.B. 1973. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, 13, 253-260.

[14] Repp, B. 1984. Categorical perception: Issues, methods, and findings. In Lass, N.J. (ed.), *Speech and language: Advances in basic research and practice* (Vol. 10). New York: Academic Press.

[15] von der Malsburg, C. 1973. Self-organization of orientation sensitive cells in the striata cortex. *Kybernetik*, 14, 85-100.

[16] Werker, J.F. and Pegg, J.E. 1992. Infant speech perception and phonological acquisition. In Ferguson, C.A., Menn, L., and Stoel-Gammon, C. (eds.), *Phonological Development: Models, research, implications*. Parkton, MD: York Press.