Published in final edited form as:
*J Exp Psychol Hum Percept Perform*. 1982 April ; 8(2): 297–314.

# Some Effects of Laboratory Training on Identification and Discrimination of Voicing Contrasts in Stop Consonants

**David B. Pisoni**, **Richard N. Aslin**, **Alan J. Percy**, and **Beth L. Hennessy**
Indiana University

## Abstract

For many years there has been a consensus that early linguistic experience exerts a profound and often permanent effect on the perceptual abilities underlying the identification and discrimination of stop consonants. It has also been concluded that selective modification of the perception of stop consonants cannot be accomplished easily and quickly in the laboratory with simple discrimination training techniques. In the present article we report the results of three experiments that examined the perception of a three-way voicing contrast by naive monolingual speakers of English. Laboratory training procedures were implemented with a small computer in a real-time environment to examine the perception of voiced, voiceless unaspirated, and voiceless aspirated stops differing in voice onset time. Three perceptual categories were present for most subjects after only a few minutes of exposure to the novel contrast. Subsequent perceptual tests revealed reliable and consistent labeling and categorical-like discrimination functions for all three voicing categories, even though one of the contrasts is not phonologically distinctive in English. The present results demonstrate that the perceptual mechanisms used by adults in categorizing stop consonants can be modified easily with simple laboratory techniques in a short period of time.

Over the last 15 years numerous studies employing synthetically produced speech stimuli have investigated *voice-onset-time* (VOT) perception in human adults, human infants, chinchillas, and monkeys. These developmental and cross-species comparisons have been undertaken to better understand the potential interactions between genetic predispositions and experiential factors in perceptual categorization of speech signals. The results of these diverse studies have shown the combined influence of two factors operating in speech perception. First, linguistic experience has been shown to have a substantial effect on speech perception, particularly in human adults exposed to different language-learning environments. Subjects identify and discriminate speech sounds with reference to the linguistic categories of their language. Second, basic sensory and psychophysical constraints on auditory system function seem to affect perception of speech and nonspeech control signals in similar ways. For example, the perception of voicing in stop consonants apparently requires the analysis of a temporal relation between laryngeal and supralaryngeal events. Basic constraints on auditory perception may play an important role in defining the inventory of acoustic correlates for distinctive features used in speech (Stevens, 1972). These considerations suggest that experiential and genetic factors are both implicated in the process by which speech signals are perceived by human listeners.

The results of the earliest cross-language experiments on the perception and production of VOT by Lisker and Abramson (1964, 1967) quantitatively confirmed that the linguistic environment exerts a profound influence on the ability to produce and perceive voicing differences in initial stops. They examined the voicing and aspiration differences among

---

Requests for reprints should be sent to David B. Pisoni, Department of Psychology, Indiana University, Bloomington, Indiana 47405..

stops produced by native speakers of 11 diverse languages and were able to identify three primary modes of voicing: (a) a lead mode in which voicing onset precedes the release from stop closure, (b) a short-lag mode in which voicing onset is more or less simultaneous with release from stop closure, and (c) a long-lag mode in which voicing onset occurs substantially after the release. In addition to measurements of VOT in the production of stop contrasts, Lisker and Abramson (1967) and Abramson and Lisker (1970) also carried out perceptual experiments using synthetically produced speech stimuli that differed in VOT. The results of these experiments demonstrated that in general, subjects from different linguistic backgrounds identify and discriminate these synthetic stimuli in terms of the distinctive phonological categories of their language.

As shown in Figure 1, the cross-language identification functions obtained by Lisker and Abramson (1967) displayed perceptual boundaries at either one or two locations along the VOT continuum, corresponding to the presence of two or three voicing categories. The discrimination functions showed discontinuities along the stimulus continuum with peaks located at the crossover points separating perceptual categories in identification. The correspondence of heightened discrimination at the category boundaries combined with relatively poor discrimination within perceptual categories suggested that subjects could discriminate between stimuli only as well as they could differentially identify them. These results suggested that phonological categories are determined, in large part, by linguistic experience (see also Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973).

The subjects in these early perceptual experiments, as well as those in more recent studies, appeared to have a great deal of difficulty in identifying and subsequently discriminating between stimuli that were *not* phonologically distinctive in their native language. The failure of adults to perceive non-native distinctions in voicing has been interpreted by some investigators as support for the view that linguistic experience exerts a profound effect on an individual's ability to discriminate speech stimuli. Indeed, in a recent chapter, Eimas (1978) has suggested that the neural mechanisms mediating VOT perception might atrophy or degenerate if stimulation is not forthcoming during an early period of language development (see also Aslin & Pisoni, 1980, Eimas & Tartter, 1979 for further discussion of early experience in speech perception).

These conclusions concerning the role of linguistic experience in speech discrimination have been widely accepted in the literature on speech perception, despite the existence of several studies demonstrating that subjects can discriminate small differences between speech sounds that were identified as belonging to the same phonological category (see Pisoni & Lazarus, 1974; Pisoni & Tash, 1974; Streeter, 1976a, 1976b). When the experimental conditions are modified to reduce uncertainty or when the subject's attention is explicitly directed to the acoustic differences between the test signals rather than to their phonetic qualities, subjects can accurately discriminate very small differences in VOT (see also Carney, Widin, & Viemeister, 1977). These recent findings, then, challenge a general conclusion prevalent in the literature for over 20 years—namely, that subjects do not *typically* discriminate between speech sounds unless they are used distinctively in their native language. However, this conclusion is based on a small number of laboratory training studies that have been conducted in the past to determine whether subjects can learn to discriminate and identify nonphonemic contrasts in VOT. In discussing these training studies, Strange and Jenkins (1978) concluded that the use of laboratory training techniques with adult subjects was generally ineffective in promoting enhanced discrimination of phonetic contrasts that were not employed phonemically in the subject's native language.

These conclusions about the effects of linguistic experience are also consistent with the views of Eimas (1975) based on his work with young infants. Eimas (1975) suggested that "the course of development of phonetic competence is one characterized by a loss of abilities over time if specific experience is not forthcoming" (p. 346). Thus, as with the adult, if phonetic differences are not used distinctively in the language-learning environment of the infant, sensitivity to the relevant acoustic attributes of these speech sounds may be attenuated, and the child may fail to develop the specific mechanisms needed to discriminate the differences between these sounds (but see Streeter & Landauer, 1976).

One of the most interesting aspects of the recent work on infant speech perception is the possibility of finding significant effects of early auditory experience on the developmental course and final sensitivity of the perceptual mechanisms employed in processing speech signals (see Aslin & Pisoni, 1980). The extensive literature on the role of early experience in visual system development indicates that early environmental experience can modify the selectivity of cortical cells in kittens (see Blakemore, 1974, and Daniels & Pettigrew, 1976, for reviews). The analogy to the findings on visual system development has already been drawn by Eimas in reviewing the earlier infant work (Eimas, 1978; Eimas & Tartter, 1979). He has argued that the lack of experience with particular phonetic contrasts in the local environment during language acquisition has the effect of modifying the appropriate *phonetic feature detectors* by reducing their sensitivity to specific acoustic cues in the speech signal. Thus, some detectors that were originally designed to process certain phonetic distinctions in speech may be "captured" or "subsumed" by other detectors after exposure to particular acoustic signals in the language-learning environment. These detectors might therefore assume the specificity for only those attributes present in the stimuli to which they have been exposed. As a consequence, then, the poor discrimination observed for some phonetic contrasts might actually be due to the modification of low-level sensory mechanisms employed in discrimination of these acoustic attributes (see however, Remez, 1979, for a critique of detector conceptualizations in speech perception).

Our own recent findings have demonstrated that 5- and 6-month-old infants from English-speaking environments *can* discriminate both lead and lag contrasts along the VOT continuum (Aslin, Pisoni, Hennessy, & Percy, 1981). Thus, we became quite interested in reexamining the ability of adults to identify and discriminate VOT contrasts that are not phonemically distinctive in their native language. In considering the previous attempts to use laboratory training procedures, we were also interested in determining why the earlier efforts appeared to be so uniformly unsuccessful in producing substantial changes in identification and discrimination of VOT. As part of the present study, we also wanted to determine precisely how much training and experience would be required for adult subjects to reacquire a nondistinctive perceptual category in voicing: whether it could be accomplished effectively in the laboratory in just a few hours, or whether it would require substantially more experience and training to produce changes in both identification and discrimination performance. We were also interested in determining how specific the training procedures would have to be to reveal an additional perceptual category in voicing.

## Experiment 1

The purpose of this first experiment was twofold. First, we wanted to determine whether a large group of naive subjects could identify three perceptual categories along a VOT continuum without any formal training or systematic feedback during identification. Second, we wanted to determine if successful three-category labeling would selectively modify subsequent discrimination performance. Our interest in this problem stemmed from studies showing that discrimination could be predicted from identification under the strong categorical perception assumption—namely, that subjects can discriminate between two stimuli only to the extent that they can identify the stimuli differently on an absolute basis. If labeling or perceptual categorization is the primary factor responsible for the observed differences in discriminability found in previous speech perception experiments, it is important to determine if discrimination of VOT is also affected by prior labeling experience and, furthermore, whether such discrimination performance could be modified selectively in the laboratory in a short period of time.

## Method

**Subjects—**Two groups of naive undergraduate students at Indiana University were recruited as paid subjects through an advertisement in the student newspaper. All subjects were right-handed monolingual speakers of English and reported no history of a hearing or speech disorder in a pretest questionnaire. They were paid at a flat rate of $2.50 per hour for each testing session. Group A consisted of 17 females and 3 males, with a mean age of 20.2 years. Group B consisted of 20 females and 5 males, with a mean age of 20.1 years.

**Stimuli—**The stimuli for this and all subsequent experiments in this report consisted of a set of 15 synthetic labial stop consonant-vowel (CV) syllables that were generated on the cascade-parallel software synthesizer designed by Klatt (1980). The 15 stimuli differed in 10 msec steps of VOT from −70 msec to +70 msec. The values used for synthesis of these stimuli were chosen from measurements of natural speech originally made by Klatt (1980) as well as from our own measurements from spectrograms of the speech of a male talker, a trained phonetician, producing various voicing contrasts in CV syllables. The stimuli consisted of a 255-msec steady-state pattern with formant (F) values and bandwidths (BW) appropriate for the vowel [a] (Fl = 700 Hz, BW1 = 90 Hz; F2 = 1,200 Hz, BW2 = 90 Hz; F3 = 2,600 Hz, BW3 = 130 Hz; F4 = 3,300 Hz, BW4 = 400 Hz; F5 = 3,700 Hz, BW5 = 500 Hz). The formant transitions into the vowel were 40 msec in duration and had starting frequencies appropriate for a bilabial stop in stressed syllable initial position (Fl = 438 Hz, F2 = 1,025 Hz, F3 = 2,425 Hz). Voicing lead was simulated by passing a combination of the sinusoidal voicing source and normal voicing through F1 that was set at 180 Hz with a bandwidth of 150 Hz. The amplitude values of the voicing source were chosen to match natural productions measured from broadband spectrograms and average amplitude contours. A 10-msec release burst was generated by passing a turbulent noise source which has a broadband (5 kHz) flat spectrum through the bypass channel of the parallel branch of the synthesizer. The amplitude of the release burst was chosen on both theoretical and empirical grounds after a long series of listening tests. Finally, the aspiration associated with voiceless stops was generated by passing a noise source through the cascade branch to simulate the turbulence produced at the glottis. The amplitude of aspiration was again chosen to match measurements of spectrograms of natural speech. During the period of aspiration, the bandwidth of F1 was also widened to 300 Hz. To simulate breathiness at the end of the syllable, the bandwidth of F1 was widened linearly from 90 Hz to 180 Hz, and some aspiration noise was added to the final 35 msec. The pitch contour had a slight rise at the onset of the release of the consonant from 120 Hz to 125 Hz and then fell linearly to 100 Hz over the remaining steady-state portion of the vowel.

**Procedure—**All experimental events involving the presentation of stimuli and collection of responses were controlled on-line in real-time by a PDP-11 computer. Subjects participated in small groups in a quiet room equipped with six individual cubicles interfaced to the computer. The test stimuli were converted to analog form via a 12-bit D/A converter, low-pass filtered at 4.8 kHz and then presented to subjects binaurally through Telephonies TDH-39 matched and calibrated headphones. All stimuli were presented at a comfortable listening level of about 80 dB SPL for the steady-state portion of the vowel. The same voltage levels were maintained throughout all of the experiments.

The present experiment consisted of two 1-hour sessions, conducted on separate days. On each day, subjects in Group A identified the test stimuli by using either two or three response categories. Half of the subjects carried out two-category identification on Day 1 followed by three-category identification on Day 2. The order was reversed for the remaining subjects.

In the two-category condition, subjects simply identified the stimuli into two categories corresponding to English [ba] and [pa] by depressing an appropriate response key. In the three-category condition, subjects were required to use three response categories, corresponding to [b], [p] and [$p^h$]. Immediately before testing began in the three-category condition, subjects listened to 10 tokens of the −70, 0, and 70 msec VOT stimuli arranged in a blocked format (i.e., −70, 0, 70, −70, 0, 70, etc.) to familiarize them with the stimulus contrasts and the appropriate responses. However, subjects only listened to these stimuli; they were not required to respond to them overtly. No attempt was made at this time to “train” subjects in any explicit way by informing them of the consistency of their labeling responses through immediate feedback.

Identification testing on each day consisted of the presentation of two blocks of 150 trials each (i.e., 20 tokens of each of the 15 stimuli). Stimuli were presented one at a time in a random order. All timing and sequencing of trials in the experiment were paced to the slowest subject in a given session.

Subjects in Group B received either a two- or three-category identification test, identical to the procedure for Group A, followed by an ABX discrimination test. Ten subjects were assigned to the two-category identification condition, and 15 subjects were assigned to the three-category condition. On each of two separate days, subjects first identified the VOT stimuli into two or three categories and then discriminated-differences between pairs of stimuli in an ABX format.

As in Group A, no formal training procedures or feedback was used in two-category identification. On each day subjects first received one block of 150 trials for identification. This was followed, after a short break, by the ABX test, which consisted of one block of 208 trials. All 13 two-step pairs of stimuli from the VOX continuum were presented four times each in all possible ABX arrangements. Stimuli within an ABX triad were separated by 500 msec. Subjects were instructed to determine whether the third sound in each triad was most like the first or second sound and to enter their response accordingly on the response box. Subjects received no training during identification nor were they provided with any feedback during the ABX discrimination test.

Subjects in the three-category group received ten tokens of the −70, 0, and 70 msec VOT stimuli in a blocked format before identification testing began to familiarize them with the test stimuli and the three response categories. All other details were the same as in Group A. Timing and sequencing of test trials in both identification and discrimination were paced to the slowest subject in a given session.

## Results and Discussion

The average two- and three-category labeling results for subjects in Group A are shown in Figure 2. As expected, subjects showed very reliable and consistent two-category identification functions for phonologically distinctive English voicing categories. More interesting, both groups of subjects also were able to reliably identify stimuli into a third perceptual category, a category with values in the voicing lead region of the VOT continuum that are not phonologically distinctive in English. Only 2 out of the 20 subjects tested failed to use three responses at all. Although there was some variability in the labeling data of the remaining subjects when looked at individually, there was also a surprising amount of consistency among most of the subjects as shown by these group data.1

The average identification and ABX discrimination functions for Group B are shown in Figure 3. The two- and three-category labeling functions shown in the left-hand panel of each figure are similar to those obtained for Group A. Although the average two-category data shown here are consistent and representative of individual subjects, the average three-category data are less consistent. However, this group inconsistency is the result of greater variance in the minus VOT boundary region.

Examination of the average ABX discrimination functions shown in the right-hand panels of Figure 3 reveals the presence of two distinct peaks in discrimination regardless of prior labeling experience. The larger peak occurs in the voicing lag region of the continuum at roughly 20 msec, whereas a smaller peak can be observed in the voicing lead region at roughly −20 msec. The main effect for stimulus comparison was highly significant, $F(12, 275) = 21.44$, $p < .0001$. Although the overall level of ABX discrimination is significantly higher for subjects in the three-category labeling condition, $F(1, 275) = 10.50$, $p < .001$, the shape of the two ABX discrimination functions is nearly identical. This observation was supported by the presence of two significant main effects and the absence of a significant interaction between these variables.

To establish that the two peaks in the ABX discrimination function represent performance that is reliably different from chance expectation, we computed the difference between two proportions, a procedure that is equivalent to chi-square (Walker & Lev, 1953). These analyses were carried out separately for each labeling group, and they revealed reliable differences. The small peak in the voicing lead region corresponding to the ABX comparison straddling the −20 msec VOT value (i.e., the −30 vs. −10 msec pair) was significantly different from chance: $\chi^2(1) = 4.84$, $p < .05$, and $\chi^2(1) = 10.38$, $p < .05$, respectively, for the two- and three-category groups. The larger peak in the voicing lag region corresponding to the ABX comparison straddling the +20 msec VOT value (i.e., the 10 vs. 30 msec pair) was also significantly different from chance: $\chi^2(1) = 76.62$, $p < .01$, and $\chi^2(1) = 139.36$, $p < .01$, respectively, for the two groups.

Note that subjects in the two-category labeling condition showed evidence of discriminating stimuli in the voicing lead region of the stimulus continuum despite the fact that these stimuli were all classified into the same perceptual category. Such a finding is not surprising given previous demonstrations of within-category discrimination in speech perception (Carney, et al., 1977; Pisoni & Lazarus, 1974; Pisoni & Tash, 1974). However, we want to emphasize here that no special efforts were made to direct the subjects' attention to the voicing differences between VOT test stimuli in this region of the stimulus continuum or to modify the discrimination task to improve their sensitivity.

---

[1] To quantify these observations in more precise terms, we carried out a number of additional analyses on the identification data collected in each labeling condition. For each subject's data we fitted a normal ogive with the procedures outlined in Woodworth (1938) and computed the mean and standard deviation of the function. This provided us with numerical values of the 50% crossover points in the identification functions and their respective slopes. In all cases, the slopes of the functions were steeper for the boundaries separating the voiced and voiceless stops than the boundaries separating the prevoiced and voiceless unaspirated stops. To obtain a quantitative measure of response consistency in identification, we also calculated the amount of informational uncertainty (H) for each stimulus value along the continuum with the methods described by Attneave (1959). These H values were then converted to informational redundancy values (1 – Rel H) so that complete response diversity (i.e., two or three equiprobable. responses) would register as 0 and complete response consistency would equal 1. In all cases, these measures showed that subjects were more consistent in the two-category conditions than in the three-category conditions. These quantitative measures reflected quite accurately the qualitative observations that can be made by examining individual subjects' identification performance. For purposes of this discussion, another perceptual category was assumed to be present when consistent and reliable identification performance at levels well above change expectation could be observed in a subject's labeling functions.

Identification functions for individual subjects are shown in Figures 4 and 5. Many of the individual subjects in the three-category labeling condition also showed two peaks in their ABX discrimination functions at values corresponding to the boundaries separating the three voicing categories; The individual ABX discrimination data in Figure 5 show a fairly close correspondence between identification and ABX discrimination for some subjects, whereas other subjects are considerably more variable in discrimination. Despite the variability observed among individual subjects, the results indicate that a majority of naive unselected subjects can identify and discriminate an additional perceptual category in voicing.

Some investigators have argued that the observed relation between identification and discrimination is a consequence of the subject's initial labeling of the test stimuli prior to measurement of discrimination (Lane, 1965). The ABX data obtained in our two-category condition suggest that this may not be correct, since subjects did not identify stimuli in the voicing lead region differentially. Nevertheless, given the apparent ease with which naive subjects can use a new label for a nonphonemic VOT category, it is worthwhile to examine discrimination in the absence of any explicit labeling experience. The next experiment is directed to this question.

## Experiment 2

In this experiment we measured discrimination in the absence of any prior labeling experience. In addition, one group of subjects received immediate feedback during discrimination testing to determine if such feedback would enhance discrimination performance, particularly in the minus region of the VOT continuum corresponding to a nonphonemic contrast in voicing.

### Method

**Subjects—**Twenty additional naive subjects who met the same requirements as in Experiment 1 were recruited for this experiment. There were 13 females and 7 males, with a mean age of 20 years.

**Stimuli—**The 15 synthetic stimuli differing in VOT that were used in Experiment 1 were also used in this experiment.

**Procedure—**The procedure for ABX discrimination was, identical to that used in the previous experiment, except that half of the subjects received immediate feedback for correct responses after each trial. Ten subjects were tested with feedback, whereas subjects assigned to the other group were tested without feedback. Subjects in each group participated in a single testing session lasting about 1 hour. Two blocks of 208 ABX test trials were presented in each session.

### Results and Discussion

The average ABX discrimination functions are shown in Figure 6. Both groups displayed two peaks in their ABX discrimination functions, a relatively large peak in the lag region and a somewhat smaller peak in the lead region of the continuum. An analysis of variance on the ABX discrimination scores showed no overall statistical difference for the main effect of the feedback versus nonfeedback manipulation. As anticipated, the main effect for stimulus comparison was highly significant, $F(12, 216) = 22.51$, $p < .001$; the interaction of this variable with the feedback manipulation was not significant.

The ABX data for individual subjects is shown in Figures 7 and 8. The presence of two peaks in ABX discrimination was observed in the data for all subjects tested in both conditions of the experiment, although the individual peaks did not always correspond precisely to +/− msec VOT. To determine if both peaks in the ABX discrimination function represent performance that is reliably different from chance, we computed the difference between two proportions for the −30 versus −10 msec pairs and the 10 versus 30 msec pairs combined over the two groups. As in the previous experiment, both peaks were significantly different from chance expectation: $\chi^2(1) = 48.17$, $p < .01$, and $\chi^2(1) = 155.01$, $p < .O1$, respectively, for the lead and lag comparisons.

The ABX discrimination functions are similar to those obtained in the previous experiment in which the labeling task preceded discrimination testing. Thus, the present findings demonstrate that the peaks in discrimination are not entirely the consequence of prior labeling experience brought about by having subjects identify stimuli into categories before discrimination testing.2 Instead, subjects are able to respond differentially to the psychophysical properties of the stimuli without recourse to perceptual categorization. These ABX discrimination data suggest the presence of relatively narrow regions along the VOT continuum characterized by high discriminability (i.e., boundaries separated by somewhat broader regions of lower discriminability (i.e., perceptual categories). The present findings provide further support for the recent proposals that some underlying sensory or psychophysical factor, rather than perceptual categorization per se, is responsible for the observed discontinuities typically found in VOT discrimination at selected regions along the stimulus continuum (see Miller, Wier, Pastore, Kelly, & Pooling, 1976; Pisoni, 1977).

The present ABX discrimination data also suggest that within the relatively short testing period, the presence of immediate feedback does not significantly enhance ABX discrimination performance along the entire VOT continuum. Thus, if these discrimination peaks are based on the sensory or psychophysical aspects of the stimuli, it would appear that these sensory mechanism(s) are not easily modified. Nevertheless, the basic sensory mechanism(s) needed for discrimination are clearly present, thus providing naive speakers of English with the underlying abilities to identify and reliably use a third perceptual category in a nonphonemic region of the VOT continuum. Whether individual subjects would also show enhanced discrimination performance after a more extensive training regimen (with immediate feedback), and whether such improvements would be reflected in much sharper identification boundaries, remains a matter of further research at the present time (see however, Carney et al, 1977). Experiment 3 pursued this issue.

## Experiment 3

In this final experiment, we were interested in reducing intersubject variability while increasing response consistency in perceptual categorization. To accomplish this in a relatively short time, we used a discrimination-training procedure with immediate feedback after exposure to salient exemplars of the three perceptual categories. The training sequences presented only three stimuli, one token of each of the three voicing types, arranged in a specified order. After the training phase was completed, subjects who met a predetermined, 85% correct performance criterion in identification were selected for subsequent testing in which identification and ABX discrimination data were collected. The purpose of this experiment, therefore, was to determine if subjects who received a brief period of training would show steeper slopes in identification at both voicing boundaries and heightened peaks in ABX discrimination.

---

[2] It is of course quite possible that subjects still used phonetic labels in the ABX discrimination task, since these particular synthetic stimuli are very good appropriations to the natural speech they were modeled from. Nevertheless, our results indicate the presence of a discrimination peak in 'the lead region of the VOT continuum which corresponds to the boundary between voiced and voiceless unaspirated stops.

## Method

**Subjects—**Twelve additional naive subjects were recruited for this experiment in the same way as in the previous two experiments. This group consisted of 10 females and 2 males, with a mean age of 21.2 years. All subjects met the same requirements as in the previous experiments and were paid at a flat rate of $3 per hour for their services.

**Stimuli—**The same 15 synthetic speech stimuli used in Experiments 1 and 2 were used in this experiment.

**Procedure—**The experiment involved four 1-hour testing sessions that were conducted on consecutive days. On Day 1 all subjects were trained to identify the −70, 0, and 70 msec VOT stimuli into three categories. As in the previous experiments involving three-category identification, subjects first listened to the three category exemplars presented in sequential order 10 times (i.e., −70, 0, 70; −70, 0, 70, etc.) to familiarize themselves with the stimuli and required responses. When this phase was completed, subjects received a block of 240 trials for identification. This block of trials consisted of 80 replications of each of the three exemplar stimuli presented in a random order. Immediate feedback indicating the correct response was provided after each trial. After completing this phase of the experiment on Day 1, subjects who met a predetermined criterion of 85% correct for each stimulus were invited back for the remaining 3 days of testing.

On Day 2 the subjects initially received a 75-trial warm-up sequence containing 25 replications of each of the three training stimuli in a random order with feedback, duplicating the procedure from Day 1. When this was completed, subjects were presented with two blocks of 150 trials of the full stimulus continuum (i.e., −70 through 70 msec) in a random order for identification testing. No feedback was provided during identification testing (i.e., generalization) on the full series.

Testing on Days 3 and 4 consisted of one block of 150 trials for identification followed by one block of 208 trials for ABX discrimination. No feedback was in effect during identification testing although immediate feedback was provided for correct responses after each trial in the ABX test.

## Results and Discussion

Of the original 12 subjects recruited, 6 passed the 85% criterion on Day 1 and were invited back for the remaining sessions. Subjects who failed to meet this criterion all responded to the three training stimuli at levels well above chance, although they did not reach the required performance level. These results were anticipated, since our earlier experiments demonstrated some variability among individual subjects in VOT identification.

The average identification functions for the six criterion subjects are shown in the left-hand panel of Figure 9. These are the data collected on Day 2 of testing. As expected, these six subjects showed a high level of consistency in labeling stimuli in the voicing lead region of the continuum, despite receiving only a very modest number of training trials on the three VOT exemplars (−70, 0, 70 msec). Moreover, the very steep slopes in the group-identification function indicate the presence of three discrete and well-defined perceptual categories. In addition, the slope in the minus VOT region is steeper in this experiment than in Experiment 1, where no specific training procedures were used. The group data shown in this panel of the figure closely mirrors the performance of the six subjects as shown by an examination of the individual data in Figure 103.

Thee average identification and ABX discrimination data collected on Days 3 and 4 are shown in the right-hand panel of Figure 9. The corresponding individual subject data for Days 3 and 4 are also shown in the right-hand panel of Figure 10. As observed earlier in Experiments 1 and 2, the present ABX discrimination functions show two peaks, corresponding to the boundary between categories and troughs within perceptual categories.

## General Discussion

The overall results of these three experiments demonstrate clearly that naive listeners can perceive differences in the minus region of the VOT continuum. The present findings differ from the results reported in earlier investigations of VOT perception, which indicated that prior linguistic experience substantially diminishes perceptual sensitivity to nonphonemic voicing contrasts in adults. Our results also differ from previous investigations that were concerned with the use of laboratory training procedures in speech perception. Given appropriate experimental procedures, the present results have demonstrated clearly that naive subjects can perceive an additional perceptual contrast easily in the laboratory after a short training period. Since this additional perceptual contrast in voicing corresponds to a distinction that is nonphonemic in English, it seems appropriate to examine the reasons why several of the earlier training studies in speech perception failed to selectively modify the identification and discrimination of VOT in adult listeners.

The one previous exception to this general pattern of results is a seemingly obscure study reported by Lane and Moore (1962), who examined the identification and discrimination of voicing by an aphasic patient. Before training, the patient was unable to discriminate differences between [d] and [t] when presented in isolation or in the context of minimal word pairs. Using synthetic speech stimuli differing in F1 cutback between [do] and [to], Lane and Moore (1962) measured identification and discrimination both before and after a brief training interval. Before training began, the aphasic subject was unable to reliably sort the [do]—[to] stimuli into two well-defined perceptual categories. Moreover, he was unable to discriminate differences between pairs of these stimuli in an ABX format at levels above chance for any of the comparisons along the stimulus continuum. The 15-min training session alternated presentation of the two endpoint stimuli with F1 cutback values of 0 and 60 msec, respectively. The aphasic subject received immediate feedback for correct responses to these two training stimuli after each trial. When the training phase was completed, the aphasic listener carried out identification and ABX discrimination tests again with the full series of synthetic stimuli.

The results of these tests showed a very dramatic increase in differential labeling of the entire set of synthetic stimuli in the identification task as well as a marked improvement in ABX discrimination performance, particularly for pairs of stimuli selected from different perceptual categories. We suspect that previous investigators were unaware of this early training study by Lane and Moore (1962) and the specific procedures used in identification training. These same training procedures were, however, used in earlier studies carried out by Cross and Lane (1962), Lane (1965), and more recently, Pisoni (1976, 1977, 1979), all of whom obtained substantial changes in labeling of nonspeech signals after relatively short periods of training.

Nevertheless, if we set aside for the moment both the early results of Lane and Moore (1962) on the *reacquisition* of the voicing contrast in an aphasic and the previous nonspeech experiments noted above, it appears that all previous attempts to modify voicing perception

---

[3] The quantitative analyses we carried out on the identification data in terms of slope and the informational uncertainty index support these qualitative observations. In each case, the slopes for the prevoiced category were steeper and the overall uncertainty values higher than in the previous experiments using unselected subjects.

in adult subjects, except for the present experiments, have failed to produce very consistent or reliable effects on subsequent measures of identification and/or discrimination (see review by Strange & Jenkins, 1978).

Why have previous researchers been so uniformly unsuccessful at selectively modifying the perception of VOT in adults? Is there something peculiar about the specific speech stimuli used, or might the differences be a consequence of the specific methodologies employed? It is possible that the present findings are due to differences in the synthetic VOT stimuli, since we generated new stimuli based on measurements of natural speech, and all previous studies have used the original Lisker and Abramson (1967) VOT stimuli. However, we feel that such a criticism is unlikely to account for all of the results reported here. As we noted earlier, Carney et al. (1977) used the original Lisker and Abramson VOT stimuli and found substantial differences in identification and discrimination performance. However, they obtained these results by modifying in very substantial ways the experimental procedures typically used in speech perception experiments. If the Carney et al. data were not available in the literature to demonstrate the importance of the specific experimental methodology, then stimulus differences might account for the discrepancy between our results and previous investigations. However, given that these earlier results were obtained with the original Lisker and Abramson stimuli, we find it difficult to argue that our results are due to stimulus differences. Thus, our emphasis in the following discussion is concerned primarily with methodological differences.

Let us turn first to an examination of the earliest cross-language perceptual experiments on VOT carried out by Lisker and Abramson (1967). They found that subjects could readily identify synthetic VOT stimuli into the phonological categories of their native language. These experiments required native speakers of the language to name the initial stop consonant by identifying it with one or another words in their language. As far as we know, no tests were conducted to ascertain whether these subjects could reliably *identify* more perceptual categories than were used distinctively in their language.

Although subjects in the Lisker and Abramson (1967) cross-language experiments might have been able to recognize additional categories by having more response choices available to them in identification, results of oddity discrimination tests indicated that English subjects apparently could not discriminate these differences (see Abramson & Lisker, 1967). When discrimination is measured in the “oddity” paradigm, subjects are strongly encouraged to adopt a “context-coding” mode of response (Durlach & Braida, 1969). That is, the stimuli are immediately receded into a more durable phonological form for maintenance in short-term memory to solve the discrimination problem (see Pisoni, 1973, 1975). Such a “context-coding” mode of perception is also favored by the high uncertainty conditions of the oddity discrimination task brought about by the use of a roving standard from trial to trial, which effectively mixes “easy” trials with “hard” trials. Finally, in the Lisker and Abramson experiments, as well as many of the earlier training studies, immediate feedback was not provided during identification training and testing or discrimination testing. Strange (1972) provided immediate verbal feedback during oddity discrimination after each trial, but she did not use this procedure in identification training.

Under conditions such as these, untrained listeners have great difficulty in determining precisely which acoustic attributes of the signal they are supposed to attend to in carrying out the required tasks. Thus, subjects may fail to discriminate fine phonetic differences within a perceptual category if they adopt a lax criterion for detecting small differences between speech sounds; the present results weigh strongly against the alternative that a true capacity limitation on processing sensory input is responsible for the poor discrimination in these earlier studies. Thus, we suspect that the particular combination of experimental tasks and their order of presentation to subjects may have been responsible for the observed relations between identification and discrimination found by Lisker and Abramson in their well-known cross-language investigations and by other investigators in the past.

In an experiment specifically designed to study the learning of a new contrast in voicing, Lisker (1970) attempted to train native speakers of Russian to distinguish between voiceless unaspirated and voiceless aspirated stops, a voicing contrast that is distinctive to English speakers but not to Russian speakers. Lisker used a training procedure that was superficially similar to the one used earlier by Lane and Moore (1962), except that no immediate feedback was provided to subjects after training trials. Although the Russian subjects learned to identify the endpoint stimuli (i.e., 10 and 60 msec VOT) in this task slightly better than chance, their performance was not the same for both stimuli. Thus, although the majority of these subjects could differentiate the training stimuli and apparently could use two discrete labeling responses, their performance on this task was not always consistent or reliable. Since immediate feedback for correct responses in identification was not provided after each training trial, the subjects probably had difficulty in determining which specific criterial attributes of the stimuli they were to attend to selectively.

When the training phase was completed, all of Lisker's (1970) subjects were given the full series, including all the intermediate stimuli in what appeared to be a scaling or magnitude estimation task. The stimuli were presented one at a time, and subjects were required to assign a numerical rating on each trial. Compared to a group of English subjects who were also run in the same scaling procedure, the Russians failed to show sharp or consistent identification functions for these stimuli. Instead, their responses showed a more gradual or “continuous” change from one stimulus to the next along the test continuum from 10 to 60 msec. Lisker interpreted these results as evidence that his Russian listeners were not generally able to recognize the voicing boundary found in English. However, the outcome of this study is ambiguous, since the Russian subjects may not have been able to selectively attend to or focus on the relevant acoustic attributes that distinguish aspirated and unaspirated voiceless stops in the absence of relevant feedback during the initial identification training phase. Taken together with the scaling procedure used in this study, subjects may have adopted a perceptual strategy of focusing on several different properties of the stimuli over the course of the experiment.

Another attempt to modify voicing perception in adults was carried out by Strange (1972); she tried to train a small number of college-age students to identify and discriminate differences in VOT in the lead region of the continuum where the Thai voiced/voiceless unaspirate boundary occurs. In the first study, four subjects received training in the oddity discrimination paradigm with “right”–“wrong” feedback provided verbally by the experimenter after each trial. When the training phase was completed, subjects carried out the oddity discrimination task without feedback. In comparison to the present data, all four subjects showed slightly improved overall oddity discrimination performance on the VOT stimuli during the posttest. However, no improvement was observed for discrimination of pairs of stimuli straddling the Thai labeling boundary at −20 msec VOT. The greatest increase in discrimination occured for stimuli adjacent to the voicing boundary in English. Based on these results, Strange (1972) concluded that her subjects did not “learn” to discriminate the VOT dimension as native Thai-speaking subjects typically do. Moreover, Strange (1972) concluded that

> there is no prepotency for adult native English speakers to discriminate differences in the region of the Thai prevoiced-voiced boundary that can be easily realized by mere practice with feedback, (p. 40)

In the second study, Strange (1972) trained three subjects to identify the members of a truncated apical series of VOT stimuli (i.e., −100 to 10 msec) into two perceptual categories. Initial training involved presentation of the endpoint stimuli in alternation without immediate feedback. However, the experimenter told the subjects the number of errors they made after each block of trials. Oddity testing was carried out after labeling, and the results showed some evidence for a slight increase in discrimination at the boundary between these two new perceptual categories. However, identification and discrimination tests using a labial VOT series failed to show transfer of training from one VOT series to another.[4] Nevertheless, subjects in this experiment were able to reliably identify members of the truncated apical-place series into two categories and, moreover, this labeling experience was carried over to discrimination of the same series.

Strange (1972) also carried out a third study using a scaling procedure similar to the one used earlier by Lisker (1970). Subjects were required to rate each stimulus along a scale between two endpoint reference stimuli. This procedure was adopted as a way of training subjects to perceive the VOT dimension as an acoustic continuum rather than directing their attention to discrete labeling responses as in the identification task. After training in the scaling task, subjects also carried out oddity discrimination. Although the results of this study were complicated by high subject variability in both tasks, there was some weak evidence that training with the scaling procedure did produce effects on perception of VOT. Posttest results for some subjects showed a shift in the scaling responses toward more gradual or continuous functions. The oddity discrimination results were more inconsistent. Some subjects showed an overall improvement in discrimination, whereas others did not. As in Strange's second experiment, no consistent transfer effects from one VOT series to another were observed (see also Footnote 4). Based on the outcome of these three experiments, Strange and Jenkins (1978) offered the following conclusions about the effects of laboratory training in speech perception:

> The results of these three studies show that, in general, changing the perception of VOT dimensions by adult English speakers is not easily accomplished by techniques that involved several hours of practice spread over several sessions. Although performance on each of the kinds of tests did change somewhat with experience, only the identification training task (which involved practice with general feedback only) produced categorial results approaching those found for native speakers of Thai. (p. 154).

Although the previous studies of Lisker (1970) and Strange (1972) indicated that modification of speech sound perception may not be obtained easily in a short period of time with simple laboratory procedures, more recent investigations have provided more positive results using much more extensive training procedures and highly experienced listeners. Carney et al. (1977) modified several aspects of the standard procedures used to measure VOT discrimination and observed very substantial improvements in within-category performance. In addition, they also showed that with long-term practice and the use of immediate feedback, subjects could learn to identify various VOT stimuli into arbitrarily defined categories *depending simply* on the experimenter's prior criterion for category membership. Although the results of the Carney et al. study are important in demonstrating the existence of noncategorical perception of VOT, these findings were obtained by

---

[4] Although Strange and Jenkins (1978) have concluded that “tests of identification and discrimination of comparable labial VOT series showed no generalization of improvement in the perception of the same VOT differences in another phonemic context” (p. 152), we have recently found strong evidence for transfer of trajning in identification of VOT stimuli. In a study completed after the present article was prepared, McClasky, Pisoni, and Carrell (1980) showed consistent transfer of training of identification of VOT stimuli from labial to alveolar and from alveolar to labial places of articulation with only a small amount of laboratory training experience comparable to that used in the present investigation. These transfer of training results provide additional support for the conclusions outlined in the present article.

introducing very substantial changes in the experimental procedures typically used in most speech perception experiments (see also Edman, Soli & Widin, 1978). First, a “same” – “different” AX discrimination procedure was substituted for the more traditional oddity or ABX testing paradigm. From earlier work, it is likely that this reduced the memory load requirements and encouraged subjects to operate in a trace-coding (auditory) mode rather than a context-coding (phonetic) mode (see Ades, 1977; Durlach & Braida, 1969; Pisoni, 1971; Pisoni, 1973). Second, the particular arrangement of the stimuli in the experiment resulted in a low-uncertainty task for the observers, since the traditional roving standard procedure was replaced with a fixed standard during a block of trials. Moreover, and perhaps most important, only a very small number of highly practiced subjects were used in this study; these subjects all had extensive experience in previous psychophysical experiments. Carney et al. (1977) tested their three subjects in more than a dozen sessions distributed over a period of several weeks. Thus, it is not at all surprising that discrimination performance was so much better when compared with results typically obtained in earlier training studies. Extensive experience in listening to these particular stimuli during training, combined with immediate feedback and the use of a low uncertainty testing paradigm, appear to be several methodological factors responsible for the very marked improvement in discrimination reported by Carney et al.

When the results of our experiments are considered in light of these previous findings, it is apparent that numerous factors contributed to the poor performance observed by earlier investigators. Nevertheless, it has generally been assumed that the failure to “learn” to perceive a new voicing contrast was somehow related to a permanent change in the perceptual or sensory mechanisms of the listener. We believe that such widely held conclusions are unjustified in light of the results reported in the present article. There is little solid empirical evidence that the underlying sensory or perceptual apparatus, as yet undefined, has been “retuned” or modified in any permanent manner as a result of selective early experience. Our results suggest that the perceptual selectivity observed in almost all of the previpus studies on VOT perception is primarily a consequence of attentive processes brought about by exposure to a specific subset of distinctive acoustic attributes as used in the phonological system of the listener's native language.

Although investigators such as Strange and Jenkins (1978) have minimized the importance of the recent findings demonstrating that adults can discriminate very fine phonetic details by arguing that these subjects do not “typically” discriminate these differences, we find these arguments less than convincing. Descriptions of what subjects “typically” do in speech perception experiments are of obvious importance in understanding language acquisition and speech processing. However, accurate descriptions of what subjects “can” do in these experiments are equally important in determining the underlying sensory and cognitive basis for perceptual categorization of speech signals. Our results suggest that the adult perceptual system, as used in processing voicing information in stops, is quite capable of responding quickly to a cateogory that is not phonologically distinctive in the listener's native language. Thus, the underlying sensory mechanism(s) are not lost, and the attentional strategies used in speech perception are far from being as rigid as many investigators have assumed in the past.

## Acknowledgments

The preparation of this article was supported, in part, by National Institute of Child Health and Human Development Grant HD-11915 and National Institute of Mental Health Grant MH-24027 to Indiana University at Bloomington. The article was written while the first author was a Guggenheim Fellow at the Research Laboratory of Electronics, Massachusetts Institute of Technology.

We thank Diane Kewley-Port for her help and assistance in preparing the synthetic, stimuli, Wendy Crawford for help in running subjects and drawing graphs, Jerry C. Forshee for his expert advice with regard to the computer facilities used to carry out this work, and

## References

Abramson, AS.; Lisker, L. Proceedings of the 6th International Congress of Phonetic Sciences 1967. Academia; Prague: 1970. Discriminability along the voicing continuum: Cross language tests.

Ades AE. Vowels, consonants, speech, and nonspeech. Psychological Review. 1977; 84:524–530.

Aslin, RN.; Pisoni, DB. Some developmental processes in speech perception. In: Yeai-Komshian, G.; Kavanagh, J.; Ferguson, CA., editors. Child phonology: Perception and production. Vol. Vol. 1. Academic Press; New York: 1980.

Aslin RN, Pisoni DB, Hennessy B, Perey AJ. Discrimination of Voice-onset-time by human infants: New findings concerning phonetic development. Child Development. 1981; 52:1135–1145. [PubMed: 7318516]

Attneave, F. Applications of information theory to psychology. Holt, Rinehart & Winston; New York: 1959.

Blakemore, C. Developmental factors in the formation of feature extracting neurons. In: Schmidt, FO.; Worden, FG., editors. The neurosciences: Third study program. MIT Press; Cambridge, Mass.: 1974.

Caramazza A, Yeni-Komshian GH, Zurif EB, Carbone E. The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. Journal of the Acoustical Society of America. 1973; 54:421–428. [PubMed: 4759015]

Carney AE, Wildin GP, Viemeister NF. Non-categorical perception of stop consonants differing in VOT. Journal of the Acoustical Society of America. 1977; 62:961–970. [PubMed: 908791]

Cross DV, Lane HL. On the discriminative control of concurrent responses: The relations among response frequency, latency and topography in auditory generalization. Journal of the Experimental Analysis of Behavior. 1962; 5:487–496. [PubMed: 14023994]

Daniels, JD.; Pettigrew, JD. Development of neuronal response in the visual system of cats. In: Gottlieb, G., editor. Neural and behavioral specificity: Studies on the development of behavior and the nervous system. Vol. Vol. 3. Academic Press; New York: 1976.

Durlach NI, Braida LD. Intensity perception: I. Preliminary theory of intensity resolution. Journal of the Acoustical Society of America. 1969; 46:372–383. [PubMed: 5804107]

Edman TR, Soli SD, Widin GP. Learning and generalization of intra-phonemic VOT discrimination. Journal of the Acoustical Society of America. 1978; 63:SI, 19.

Eimas PD. Auditory and phonetic coding of the cues for speech: Discrimination of the [r-1] distinction by young infants. Perception & Psychophysics. 1975; 18:341–347.

Eimas, PD. Developmental aspects of speech perception. In: Held, R.; Leibowitz, H.; Teuber, HL., editors. Handbook of sensory physiology: VIII. Perception. Springer-Verlag; New York: 1978.

Eimas, PD.; Tartter, VC. On the development of speech perception: Mechanisms and analogies. In: Lipsitt, L.; Reese, H., editors. Advances in child development and behavior. Vol. Vol. 13. Academic Press; New York: 1979.

Klatt DH. Software for a cascade/parallel formant synthesizer. Journal of the Acoustical Society of America. 1980; 67:971–995.

Lane HL. The motor theory of speech perception: A critical review. Psychological Review. 1965; 72:275–309. [PubMed: 14348425]

Lane HL, Moore DJ. Reconditioning a consonant discrimination in an aphasic: An experimental case history. Journal of Speech and Hearing Disorders. 1962; 27:232–243. [PubMed: 14462012]

Lisker, L. Status Report on Speech Research (SR-24). Haskins Laboratories; New Haven, Conn: 1970. On learning a new contrast.

Lisker, L. Status Report on Speech Research SR-24. Haskins Laboratories; New Haven, Conn: 1970. On learning a new contrast.

Lisker, L.; Abramson, AS. The voicing dimension: Some experiments in comparative phonetics. Proceedings of the 6th International Congress of Phonetic Sciences; Prague: Academia; 1967.

Miller JD, Wier CC, Pastore R, Kelly WJ, Dooling RJ. Discrimination and labelling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. Journal of the Acoustical Society of America. 1916; 60:410–417. [PubMed: 993463]

McClasky, CL.; Pisoni, DB.; Carrell, TD. Research on Speech Perception Progress Report No. 6. Indiana University; Bloomington: 1980. Effects of transfer of training on identification of a new linguistic contrast in voicing.

Pisoni, DB. Status Report on Speech Research SR--27. Haskins Laboratories; New Haven, Conn.: 1971. On the nature of categorical perception of speech sounds. Supplement

Pisoni DB. Auditory and phonetic memory codes in the discrimination of consonants and vowels. Perception & Psychophysics. 1973; 13:253–260.

Pisoni DB. Auditory short-term memory and vowel perception. Memory & Cognition. 1975; 3:7–18.

Pisoni, DB. Research on Speech Perception Progress Report No. 3. Indiana University; Bloomington: 1976. Some effects of discrimination training on the identification and discrimination of rapid spectral changes.

Pisoni DB. Identification and discrimination of the relative onset of two component tones: Implications for voicing perception in stops. Journal of the Acoustical Society of America. 1977; 67:1352–1361. [PubMed: 881488]

Pisoni, DB. Perception of speech vs. nonspeech: Evidence for different modes of processing. Proceedings of the Ninth International Congress of Phonetic Sciences; Copenhagen: Institute of Phonetics, University of Copenhagen; 1979.

Pisoni DB, Lazarus JH. Categorical and non-categorical modes of speech perception along the voicing continuum. Journal of the Acoustical Society of America. 1974; 55:328–333. [PubMed: 4821837]

Pisoni DB, Tash JB. Reaction times to comparisons within and across phonetic categories. Perception & Psychophysics. 1974; 15:285–290.

Remez RE. Adaptation of the category boundary between speech and nonspeech: A case against feature detectors. Cognitive Psychology. 1979; 11:38–57. [PubMed: 761448]

Stevens, KN. The quantal nature of speech: Evidence from articulatory-acoustic data. In: David, EE., Jr.; Denes, PB., editors. Human communication: A unified view. McGraw-Hill; New York: 1972.

Strange, W. Unpublished doctoral dissertation. University of Minnesota; 1972. The effects of training on the perception of synthetic speech sounds: Voice onset time.

Strange, W.; Jenkins, JJ. Role of linguistic experience in the perception of speech. In: Walk, RD.; Pick, HL., editors. Perception and experience. Plenum Press; New York: 1978.

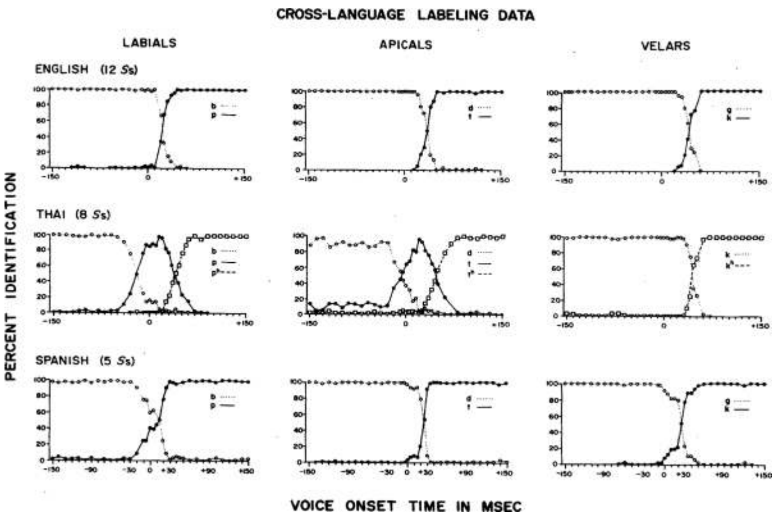
Streeter LA. Kikuyu labial and apical stop discrimination. Journal of Phonetics. 1976; 4:43–49. a.

Streeter LA. Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. Nature. 1976; 259:39–41. b. [PubMed: 1256541]

Streeter LA, Landauer TK. Effects of learning English as a second language on the acquisition of a new phonemic contrast. Journal of the Acoustical Society of America. 1976; 59:448–451. [PubMed: 1249332]

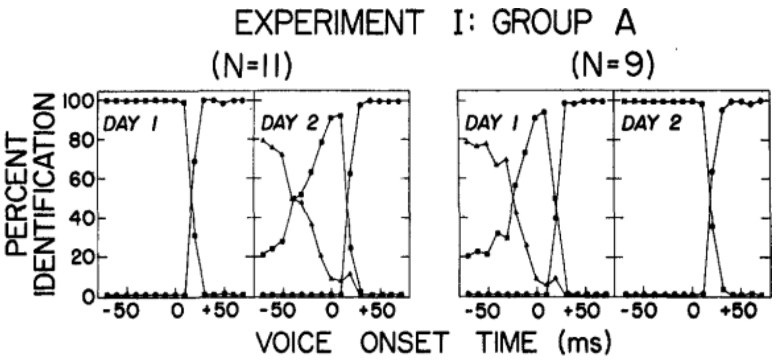
Walker, HM.; Lev, J. Statistical inference. Holt; New York: 1953.

Woodworth, RS. Experimental psychology. Holt; New York: 1938.



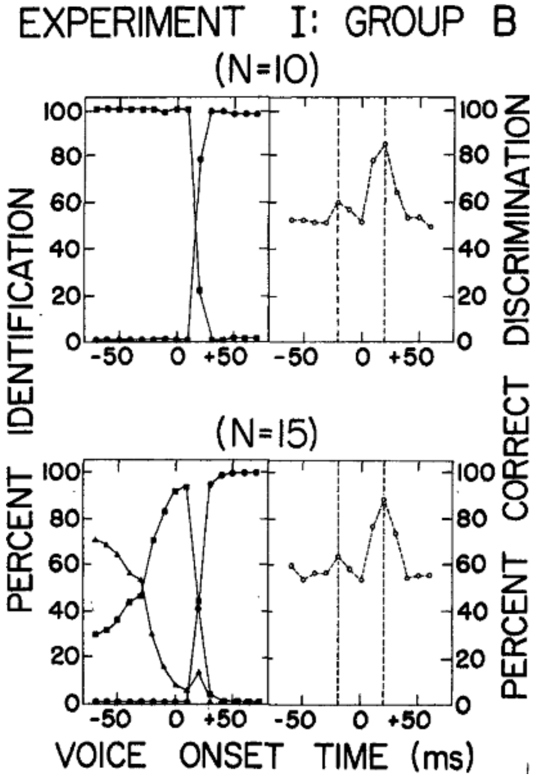
**Figrue 1.**

Cross-language identification data reported from Lisker and Abramon (1967) for labial, apical and velar stops differing in voice from −150 msec to 150 msec.

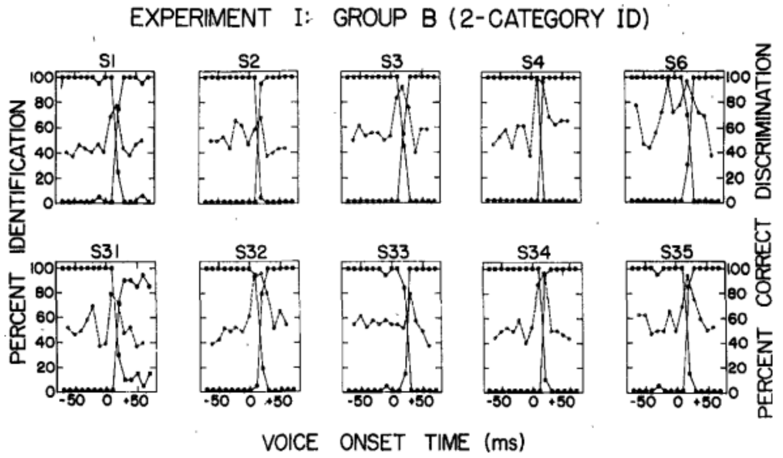


**Figrue 2.**

Average identification functions for two- and three-category labeling of voice onset time for Group A in Experiment 1. (ms = millisecond.)

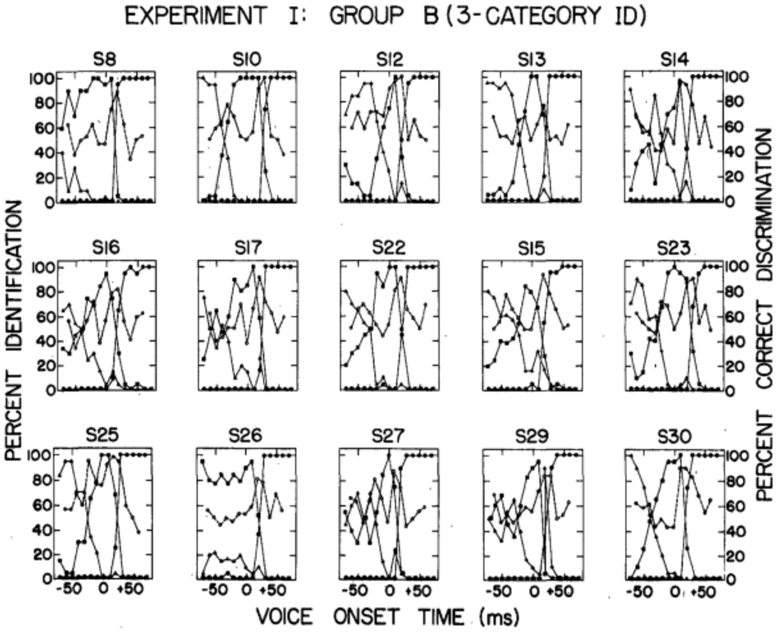


**Figrue 3.**
Average identification and ABX discrimination functions for two category (upper panel) and three category (lower panel) labeling obtained for Group B in Experiment 1.



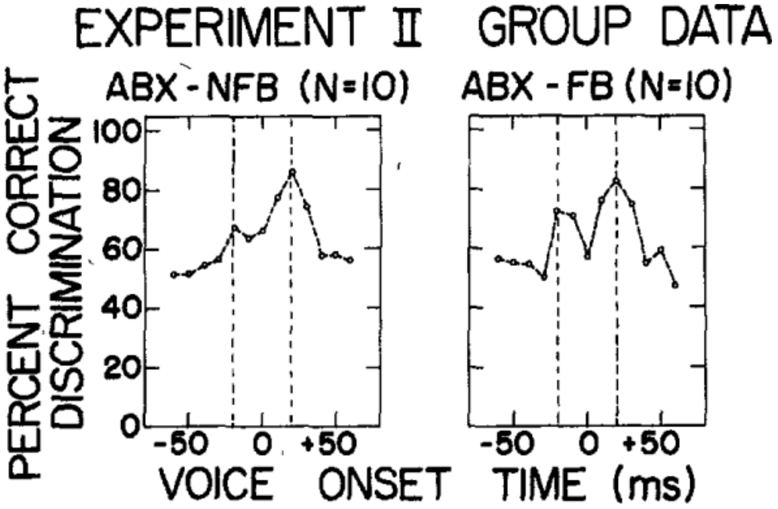
**Figrue 4.**

Individual subject data for two category identification (ID) obtained from Group B (two-category ID) in Experiment 1. (The ABX discrimination data for each subject is also plotted on the identification function.)



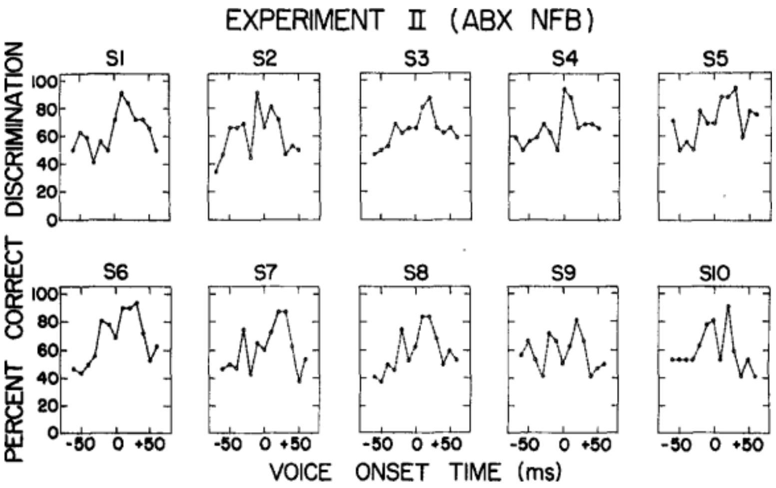
**Figrue 5.**

Individual subject data for three-category identification (ID) obtained from Group B (three-category ID) in Experiment 1. (The ABX discrimination data for each subject is also plotted on the identification function.)



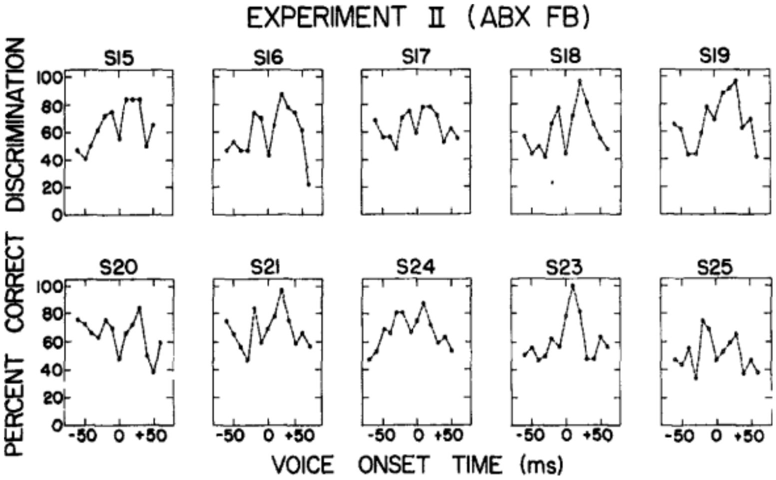
**Figrue 6.**

Average ABX discrimination functions obtained for subjects in Experiment 2. (The no-feedback group [NFB] is shown on the left, the feedback group [FB] is shown on the right. Vertical lines have been drawn at values of +/−20 msec to facilitate visual comparisons.)



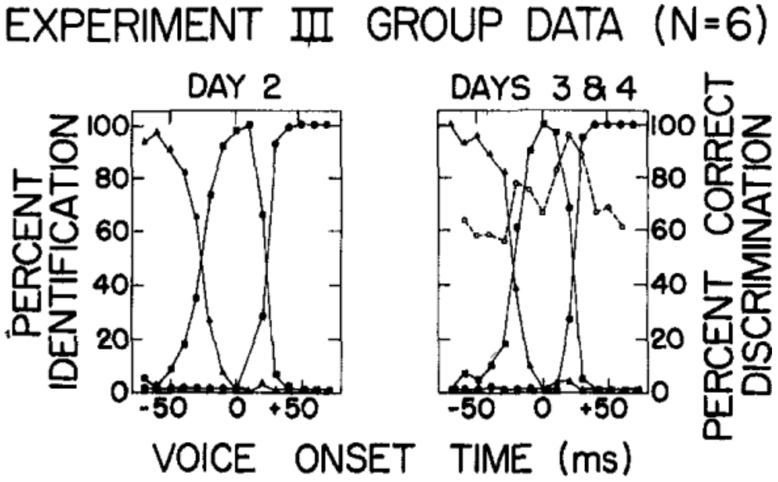
**Figrue 7.**

Individiual subject data obtained in ABX discrimination in Experiment 2 for the no feedback (NFB) condition.



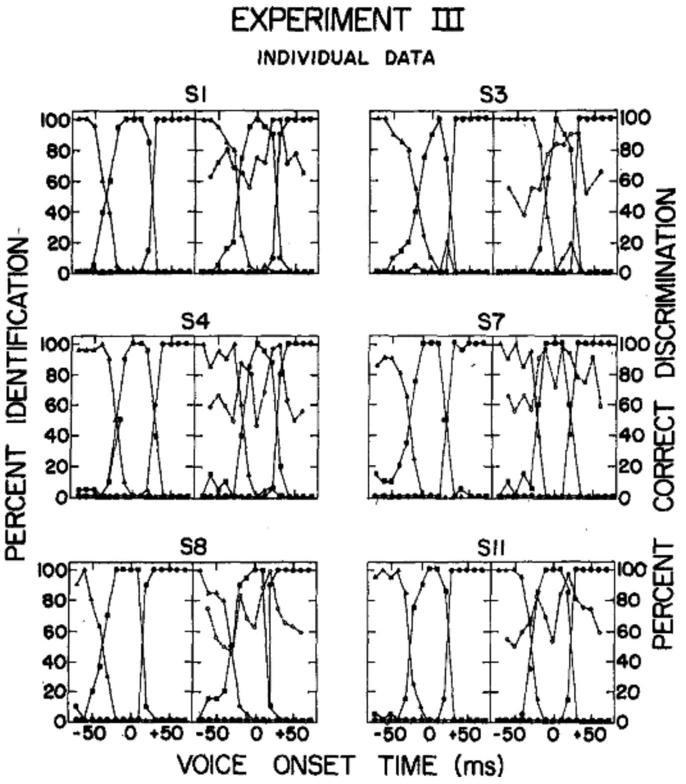
**Figrue 8.**

Individual subject data obtained in ABX discrimination in Experiment 2 for the feedback (FB) condition.



**Figrue 9.**

Average identification and ABX discrimination functions for subjects meeting an 85% identification criterion in Experiment 3. (Data on the left shows the average identification function on Day 2 of testing; the data on the right shows both average identification and ABX discrimination combined over Days 3 and 4).



**Figrue 10.**

Individual subject data for identification obtained on Day 2 and identification and ABX discrimination combined over Days 3 and 4 of the experiment.