Reprinted from JOURNAL OF EXPERIMENTAL PSYCHOLOGY, Vol. 61, No. 5, May, 1961
Printed in U. S. A.

THE DISCRIMINATION OF RELATIVE ONSET-TIME OF THE COMPONENTS OF CERTAIN SPEECH AND NONSPEECH PATTERNS[1]

A. M. LIBERMAN, KATHERINE S. HARRIS, JO ANN KINNEY,[2] AND H. LANE[3]

Haskins Laboratories and University of Connecticut

In earlier experiments (Griffith, 1958; Liberman, Harris, Hoffman, & Griffith, 1957) with synthetic stop consonants it was found that discrimination of a constant acoustic difference was more acute across phoneme boundaries (between /b/ and /d/ and between /d/ and /g/) than in the middle of the phoneme categories. Indeed, the discrimination was relatively so good across the boundaries and so poor within the categories as to suggest that *S*s could only respond to these sounds categorically (i.e., as phonemes) and could hear no other differences among them. It was possible to test this suggestion quantitatively by determining whether the discriminability of the sounds was predictable from the frequency with which they had been identified as belonging in one or another phoneme class. In general, the locations of the peaks and valleys of the discrimination functions were well accounted for on this basis. The level of discrimination, on the other hand, was on the whole somewhat higher than predicted, but not by a very great amount. Thus, differential sensitivity was in this case only slightly better than absolute identification. Such a result must appear unusual when it is remembered that in the typical psychophysical experiment the number of discriminably different stimuli is found to be many times greater than the number which can be labeled or identified absolutely (see Chapanis, 1956; Garner, 1953; Miller, 1956; Pollack, 1952, 1953).

The outcome of the experiment with the synthetic stop consonants can be interpreted in either of two ways. On the one hand it is possible that the peaks and valleys in the discrimination function are given innately. Alternatively, it can be assumed that two sounds which a listener must normally respond to in the same way come, by some learning process, to be less discriminable than two equally different sounds to which he must typically attach different phoneme labels.

If it is assumed that the discrimination functions have been radically modified by learning, there remains a further question concerning the direction of the effect. One possibility is that the innately given discrimination was originally as poor as that obtained for the least discriminable pair of stimuli, and that discriminations in the vicinity of phoneme boundaries have been sharpened by something like the process of acquired distinctiveness. The contrary possibility is that in its virgin state discrimination was everywhere as good as that at the peaks of the functions, and that long experience in responding similarly to stimuli

---

[1] This research was supported by the Operational Applications Laboratory of the Air Force Cambridge Research Center in connection with Contract AF 19(604)-2285.

[2] Now at Medical Research Laboratory, United States Naval Submarine Base, New London, Connecticut.

[3] Now at the University of Michigan.

within the same phoneme category has, through acquired similarity, significantly reduced the discriminability of those sounds. It is also possible, of course, that the original discrimination lay between the two extremes. In that case the results would be taken to represent the operation of both processes: acquired distinctiveness near phoneme boundaries and acquired similarity within the phoneme category.

The experiment to be reported was designed to investigate the relation between phoneme labeling and discrimination in the case of an acoustic and linguistic distinction different from the ones studied earlier, and, hopefully, to throw some light on the issues discussed above. The phonemes to be used are the voiced and voiceless alveolar stops /d, t/.

One reason for investigating /d,t/ is that Liberman, Delattre, and Cooper (1958) have shown that many *S*s can consistently identify as /d/ or /t/ synthetic patterns that differ by only small amounts of the appropriate acoustic cue. We may, then, expect to find with this phonemic distinction, as we have earlier found with /b,d,g/ (Liberman et al., 1957; Griffith, 1958), that the discrimination function has a sharp peak at the phoneme boundary.

A second reason for selecting /d,t/ is that the discrimination of these phonemes can be made to depend on a very simple acoustic variable—the relative displacement in time of onset between the first and second formants.[4] It is possible to study discrimination of this simple time cue in nonspeech contexts, thus providing a baseline control in terms of which one might infer whether the effects obtained with the speech sounds are the results of learning, and, if so, whether they represent acquired distinctiveness, acquired similarity, or both. A variety of different nonspeech controls is, in this case, possible and necessary. One such control has been undertaken with this experiment.

METHOD

Apparatus.—All the stimuli of this experiment were generated by using the Pattern Playback to convert hand-painted spectrograms into sound. The Pattern Playback and the techniques for using it are described in earlier papers (Cooper, 1950, 1953; Cooper, Liberman, & Borst, 1951).

Stimuli.—The spectrograms used to produce the stimuli are shown in Fig. 1. In the top row are patterns that are heard by most *S*s as /do/ or /to/ when converted into sound by the Pattern Playback. These will be referred to as the "speech" stimuli. As the figure shows, the stimulus variable which distinguishes these patterns is the delay in the onset of the first formant relative to the second and third formants. (This variable had been found in an earlier study by Liberman et al. (1958) to be sufficient for the perceived distinction between /d/ and /t/.) The delays vary in 10-msec. steps, from 0-msec. delay, as in the pattern at the extreme left, to 60-msec. delay, as shown at the right. The other characteristics of the pattern are those appropriate to an alveolar stop consonant in initial position before the vowel /o/. For convenience, these stimuli will be referred to by number, from 0 through 60, as indicated in Fig. 1.

One of the purposes of this study was to compare the discriminability of the synthetic speech sounds described above with those of a control set which were as much like them as possible, and yet not perceived as speech. The stimuli that were used for this purpose are shown in the bottom row of Fig. 1. These patterns, which will be referred to as the "control" stimuli, were made simply by turning the speech spectrograms upside down before running them through the Pattern Playback. The stimuli that are produced by this maneuver differ from each other in exactly the same way that the speech stimuli differ—namely, by the relative time of onset of the formants—and, like the speech stimuli, they can be discriminated only on that basis. As can be seen from an inspection of the control patterns, one rather small change was made in addition to the frequency inversion. The relatively "thin" formant which is at the top of the pattern (Formant 3) in the speech stimuli and at the bottom (Formant 1) in the control, was altered in the control so as to have a falling transition in place of the rising transition that would have resulted from the simple inversion. This was to provide a further precaution against the possibility that the control stimuli would resemble speech. An additional difference between the speech and control stimuli is in the relative intensities of the three formants. In the speech stimuli the first and second formants were of essentially equal intensity, while the thin third formant was approximately 5 db. less intense. In the control stimuli the first and second formants have become the second and third and are, again, of equal intensity; however, the relatively thin formant, which has become the first formant, is now only 3 db. below the other two.

---

[4] A "formant" is a relatively high concentration of acoustic energy within a restricted region of the frequency spectrum. The formants are numbered according to their positions on the frequency scale, the lowest formant being called the "first," the next higher the "second," and so on.

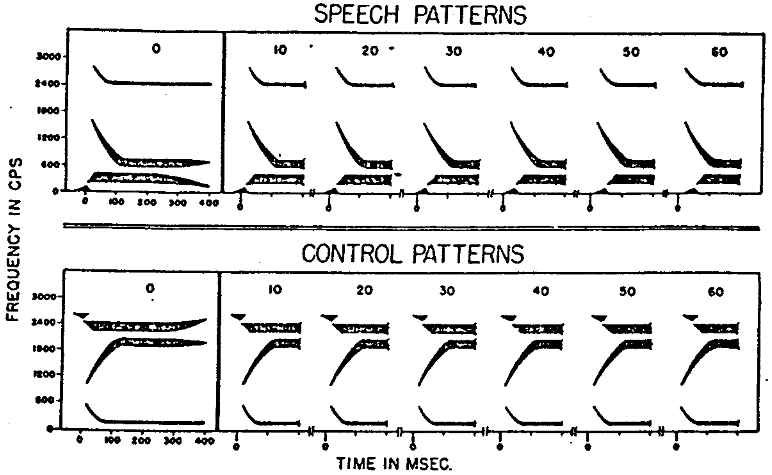

FIG. 1. Spectrographic patterns which were converted to sound by the Pattern Playback to form the stimuli of the experiment.

Careful listening to the control stimuli revealed that they did not sound at all like speech. Certainly, no phonemic distinction was conveyed by the stimulus variable.

All the patterns of Fig. 1 were converted into sound and recorded on magnetic tape. By copying, cutting, and splicing the magnetic tape, the various arrangements described below were prepared.

Stimulus presentation.—The purpose of this experiment made it necessary that we measure the discriminability of both the speech and control stimuli, and that we find the location and shape of the phoneme boundary. Discriminability of the speech and control stimuli was measured by the ABX technique. The *S*s were told that A and B would always be different, and that X would be identical with the one or the other. They were asked to judge, on whatever basis they could, whether X was like A or like B, and to guess if necessary.

The speech and control stimuli were grouped separately for ABX presentation, being arranged in magnetic tape sections that consisted of speech stimuli or control stimuli, but never both in the same tape section. The procedures for making the speech tapes and the control tapes were identical and are described below. This description has been written as if it were for only one of the two sets of stimuli (speech or control).

The various triads were made by pairing each stimulus with another stimulus having an onset delay that differed in the amount of 10, 20, or 30 msec. These pairs formed 1-, 2-, and 3-step intervals, respectively. Thus, for the 1-step intervals, Stimulus 0 was paired with Stimulus 10, Stimulus 10 with Stimulus 20, etc. The 2-step intervals were formed by pairing Stimulus 0 with Stimulus 20, Stimulus 10 with Stimulus 30, etc. The 3-step intervals were made in similar fashion. Since there were seven stimuli (in each of the speech and control sets), there were 6, 5, and 4 comparisons in the 1-, 2-, and 3-step series, respectively. Each stimulus comparison was arranged into four ABX (ABA, ABB, BAB, BAA) permutations for a total of 60 ABX triads (15 stimulus comparisons times 4 ABX permutations). Recorded copies of the 60 triads were made and distributed among 8 tape sections of 15 triads each. The 15 triads were ordered randomly with the restriction that one and only one of the four ABX permutations be represented in each tape section. The tape sections were prepared with 0.5 sec. between members of the stimulus triad and 4.0-sec. separation between triads.

In order to ascertain for each *S* the position and shape of the boundary between the "d" and "t" phonemes, each *S* was presented the same speech tapes that were used to measure discrimination, but in this case *S* was asked to identify each member of the triad as /d/ or /t/ and to guess if necessary. This aspect of the procedure will be referred to as "phoneme labeling."[4]

Subjects.—The *S*s were summer students at Columbia University. All were paid volunteers, and had no knowledge of the purpose of the experiment.

In order to determine whether discrimination is relatively sharper at the phoneme boundary, it is, of course, necessary that *S* assign the phoneme labels with a high degree of consistency, and that the region of uncertainty between /d/ and /t/ be narrow. We therefore preselected our *S*s. Twenty *S*s were each presented four speech tapes with phoneme labeling instructions. The 11 *S*s whose judgments provided the sharpest phoneme boundaries were selected for the experiment. The difference between these 11 *S*s and the 9 who were rejected was not very great. The judgments obtained during the selection session were not included in the data of the experiment.

The 11 *S*s served subsequently for approximately 30 min. per day, 5 days a week, for a total of 20 sessions. In Sessions 1 and 2 only labeling tapes were presented. Speech and control ABX tapes were introduced in Session 3. In each of the remaining sessions, *S*s heard at least one tape of each type.

In these 20 sessions, 4 speech ABX tapes were presented 6 times, and 4 were presented 5 times, thus providing 44 judgments of each comparison. Each of the 8 control tapes was presented 5 times, making a total of 40 judgments for each stimulus comparison.

Six of the tapes were presented for phoneme labeling four times, while two were presented three times. Because of the way the stimulus comparisons are arranged, the individual stimuli are not presented with equal frequency. The number of phoneme labeling judgments obtained from each *S* for the various stimuli was as follows: Stimulus 0, 135; Stimulus 10, 180; Stimulus 20, 225; Stimulus 30, 270; Stimulus 40, 225; Stimulus 50, 180; Stimulus 60, 135.

Results will also be presented for two *S*s who were run in a preliminary experiment conducted at the University of Connecticut. The procedure for these *S*s was the same as for the larger group, except that the number of ABX judgments was approximately 50% greater and the number of phoneme labeling judgments about 50% less than with the larger group.

RESULTS

In Fig. 2 are plotted the responses that were obtained when the triads of synthetic speech stimuli were presented for phoneme labeling, i.e., for identification as /d/ or /t/. Each function represents the data for one *S*, and is based on responses to all presentations of the stimuli. The values on the abscissa, which represent amount of first-formant cutback in milliseconds, are progressively displaced (by 10 msec. of cutback) for each function in order that all the data may be shown on a single graph. A reference for the displacement of the abscissa is provided by the circled points, which indicate the percentage of /dɔ/ responses for Stimulus 20.

The phoneme labeling curves are quite similar, both in position and in form. All the functions but one pass through the 50% point somewhere between Stimulus 20 and Stimulus 30; the remaining function crosses the midpoint between Stimulus 30 and Stimulus 40. For all *S*s the phoneme boundary between /d/ and /t/ is quite sharp; in every case a change of 10 msec. in the first-formant cutback is sufficient to shift the responses from 75% /d/ to 75% /t/, and there are no large inversions in the responses.

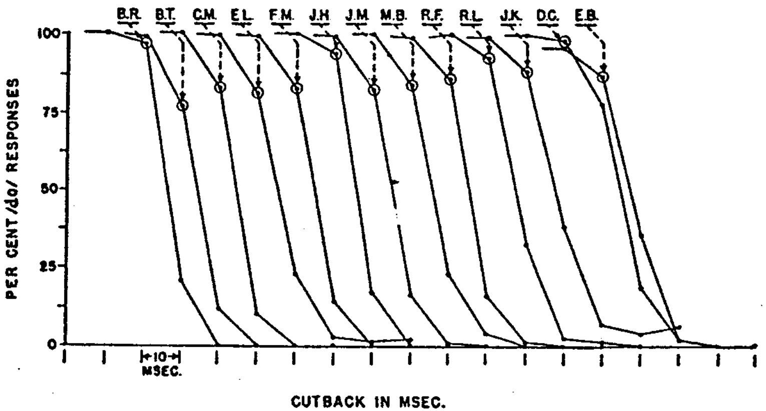

FIG. 2. Response functions for each of the 13 *S*s, showing how phoneme labels were assigned to the stimuli. The circled points represent, for each *S*, the percentage of /do/ responses to the stimulus that had 20 msec. of cutback. So that the functions might be better separated for the eye, each one has been started just beyond the point that would represent the results at 0 msec. of cutback. The /do/ responses at that point (100% for 10 *S*s, 99% for 2 *S*s, and 96% for 1 *S*) can, however, be inferred from the graphs.

One purpose of this experiment was to determine whether or not discrimination is better at the phoneme boundary than in the middle of the phoneme category. This question is answered by the graphs for individual *S*s in Fig. 3. These are based on data that were obtained when the stimuli were presented in ABX triads and *S*s were asked to judge whether X was identical with A or with B. The data of Subject BR may be used to illustrate the results. The labeling data of Fig. 2 show that for this *S* the /d/ responses shifted from 97% at Stimulus 20 to 18% at Stimulus 30, so that the phoneme boundary may be said to lie between these two stimuli. The upper left-hand graph of Fig. 3 (solid lines) shows the discrimination results for this *S* when the stimuli being compared differed by one step (10 msec.) on the stimulus scale. It can be seen that there is a peak in the discrimination function at the point which represents the comparison of Stimulus 20 and Stimulus 30, the two stimuli which straddle the phoneme boundary. In the 2-step case Stimulus 10 and Stimulus 30 fall on opposite sides of the phoneme boundary, as do Stimulus 20 and Stimulus 40, and it can be seen that discrimination is, indeed, better in these cases than in others. For the 3-step situation all stimulus pairs except Stimulus 30 and Stimulus 60 cross the phoneme boundary, and discrimination is at a high level; the 3-step pair in which the stimuli (30 and 60) lie on the same side of the boundary is seen to be considerably less discriminable.

---

[4] In the earlier study of /b,d,g/ (Liberman et al., 1957) the speech stimuli were presented separately for phoneme labeling, and not in triads as in the present experiment.

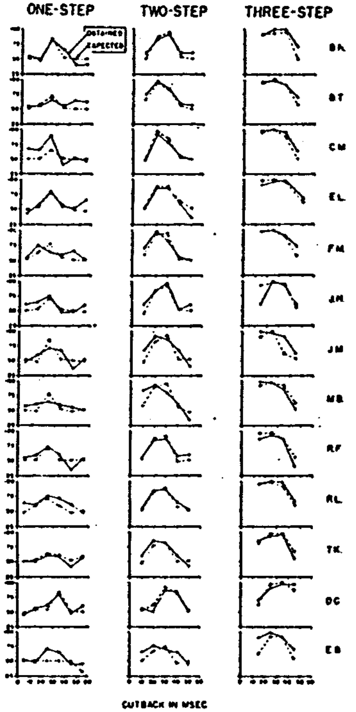

FIG. 3. Expected and obtained discrimination functions for the 1-, 2-, and 3-step differences among the synthetic speech stimuli.

The effects observed in the case of BR were obtained to a very large extent with all *S*s. In general, discrimination was better for a given stimulus difference when the two stimuli lay on opposite sides of a phoneme boundary than when the stimuli were within the same phoneme category. This effect is very apparent for the 2- and 3-step data, though perhaps not quite so obvious for the 1-step comparison.

In order to quantify the effect just described it is useful to compare the discrimination functions of Fig. 3 with those that would be predicted from the extreme assumption that *S*'s perception is completely categorical, i.e., that he can hear these sounds only as phonemes and can discriminate no other differences among them. It follows that one ought to be able to predict how often *S* will correctly discriminate any two stimuli by taking account of the frequency with which he attached one or another of the phoneme labels to each of them. The present procedure for deriving the discrimination functions is the same as that developed earlier (Liberman et al., 1957), but the calculations have been changed and simplified.[6]

An example will make the method of calculation clear. Suppose that the presented triad of stimuli is Stimulus 0–Stimulus 20–Stimulus 0. When these stimuli are presented with phoneme labeling instructions, *S* is asked to identify each one as /d/ or /t/. Neglecting for the moment how *S*s actually labeled these stimuli and considering only how they might have done so, we see that there are eight possible sequences of phoneme labels for any triad. For the purpose of predicting how *S* would respond in the ABX test of discrimination, these eight sequences can be divided into three classes. One class, which consists of two sequences (/d/ /t/ /d/ and /t/ /d/ /t/), would lead S to the correct response, "A." A second class, consisting of two sequences (/d/ /t/ /t/ and /t/ /d/ /d/), would lead *S* to the incorrect response, "B." The remaining four sequences (/t/ /t/ /t/, /d/ /d/ /d/, /d/ /d/ /t/, and /t/ /t/ /d/) would leave *S* no alternative but to guess, and he would, then, presumably respond "A" half the time and "B" half the time (probability of being correct = .5). Now suppose the triad (Stimulus 0–Stimulus 20–Stimulus 0) is presented to *S* four times for phoneme labeling, and suppose, further, that he responds three times with the sequence /d/ /d/ /d/ and once with /d/ /t/ /d/. The sequence /d/ /d/ /d/ would produce a .5 probability of being correct in the ABX discrimination; the sequence /d/ /t/ /d/ would produce a probability of 1.0. Multiplying the probability of .5 by the three sequences which are calculated to yield it, and adding this to a probability of 1.0 times the one sequence that would yield it, gives 2.5. Dividing this quantity by four, which is the number of presentations, one obtains .63 as the average predicted percent correct in the ABX discrimination of Stimulus 0 and Stimulus 20. Since the comparison of Stimulus 0 and Stimulus 20 will be made not only in the ABA order of the example taken here, but also the orders ABB, BAB, and BAA, the predicted percentage correct for the discrimination of these two stimuli is the average of the predictions obtained from all four ABX permutations.

---

[6] The simpler method can be used when the stimuli are presented in the same triads for labeling and discrimination. It yields the same result as the earlier method provided the various triad contexts have had no effect on *S*s' assignment of phoneme labels. For the data of several *S*s, discrimination functions were calculated according to the earlier method and were found to be almost identical with those produced by the simpler method.

The discrimination functions that are predicted from the assumption of completely categorical perception are shown in Fig. 3 by the dotted lines. A comparison with the obtained functions shows that on the whole the fit is a good one. For the 1- and 2-step data there appears to be no systematic displacement of the predicted and obtained curves; the differences occur equally in both directions. In the 3-step case the average of the obtained discriminations is higher on the whole than predicted, but the difference is very small.

To the extent that the obtained discrimination data fit the predicted, it can be said that the peaks in the obtained functions do, indeed, lie at the phoneme boundaries, a conclusion that otherwise could have been tested only roughly and by point-for-point comparison of the discrimination and labeling data of Fig. 2. Moreover, and more generally, the close fit supports the conclusion that in perceiving these speech sounds the listeners were able to extract no information beyond that which was given by the phoneme categories.

One asks about discrimination functions like those of Fig. 3 whether they reflect the effects of learning on perception. The discriminability of the "control" stimuli was measured to provide evidence on this point. These stimuli contained essentially the same stimulus variations as the synthetic speech patterns, but were otherwise sufficiently altered as not to be clearly perceptible as speech.

In Fig. 4 are the discrimination

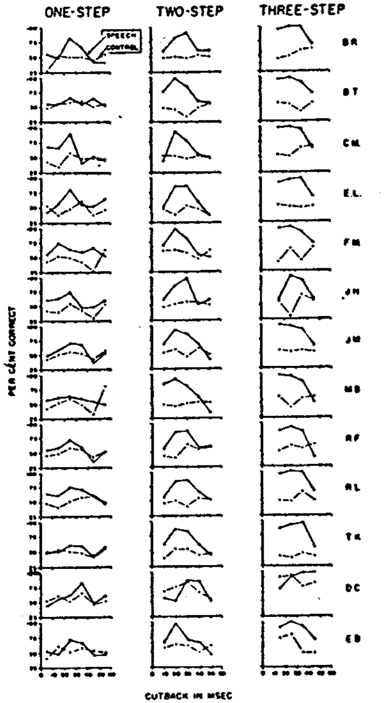

FIG. 4. Discrimination functions for the 1-, 2-, and 3-step differences among the nonspeech control stimuli. The discrimination functions that were obtained with the synthetic speech stimuli (and shown in Fig. 3 as "obtained" data) are reproduced here to facilitate comparison.

data obtained with the control stimuli. For purposes of easy comparison, the data obtained with the synthetic speech (and already shown in Fig. 3) are reproduced. Looking first at the control functions alone, one sees that there are no large or regularly occurring peaks at stimulus values corresponding to the position of the phoneme boundary or at any other points on the stimulus continuum. Comparison of the speech and control functions shows that the peaks of the speech discrimination functions for all *S*s tend in general to lie above the control functions. For the 1-step data this is not obviously true for all cases, but for the 2- and 3-step functions the speech discrimination is clearly superior to the control discrimination, and by a very wide margin.

Assuming the control to be a fair one, we conclude that the speech discrimination data reflect the effects of a large amount of learning. Inasmuch as the peaks of the speech discrimination lie above the control, we interpret the learning effect as acquired distinctiveness.

Acquired similarity would presumably occur, if at all, within the phoneme categories, where it would be reflected as a reduction in the level of speech discrimination below that of the control. One sees in Fig. 4 that the control data do not, in general, rise very far above the chance (50%) level. Thus, there is in this experiment no margin within which acquired similarity might have become manifest.

DISCUSSION

To determine whether the increased discrimination at the phoneme boundary is an effect of learning on perception, and, if so, whether it represents acquired distinctiveness, acquired similarity, or both, we compared the discriminability of the speech stimuli with a set of nonspeech controls. Ideally, these control stimuli should have been identical with the speech stimuli, and yet not perceived as speech. In practice, such controls can only be approximated, since the acoustic pattern in which the relevant physical variable is imbedded must necessarily be somewhat different if, in the one case, the stimulus is to be perceived as a speech sound and in the other not. The control stimuli used in this experiment represent only an approximation to the ideal. They are not the only possible approximation, nor are they necessarily the best.

It can be counted an advantage of the control stimuli used here that the physical differences which cued the distinction between /do/ and /to/ in the speech stimuli (i.e., the relative difference in time of onset of first and second formants) were in one sense perfectly preserved, and that they constituted in the control, as they had in the speech series, the only acoustic differences among the stimuli. The control stimuli were appropriate, too, in that many of the constant acoustic features of the speech stimuli, such as formant width, the complex nature of formant onset, etc., were closely duplicated. Many of these features of the rather complicated stimulus pattern might well affect the discrimination of onset time, and make it quite different from what it would be with the most nearly comparable simple stimuli.

In several other respects, however, the control stimuli were less than perfect. First, it should be noted that the frequency level of the middle formant was different in the control stimuli from what it was in the speech. More significant, perhaps, is the fact that in the control stimuli the formant whose time of onset varied was at a higher frequency than the other two formants, while in the speech stimuli it lay at a lower frequency than the other formants. This may be important since masking effects are greater from low frequencies to high, and it is possible, therefore, that the variations in formant onset were to some extent masked out in the control.

The point was made above that since the stimulus variations to be discriminated (the differences in onset-time) were imbedded in a complex pattern in the speech stimuli, it was appropriate that they be similarly imbedded in the control. It is nevertheless pertinent to inquire about the discriminability of essentially the same stimulus variable in a much simpler context—where the discrimination data might be considered to provide a psychophysical baseline—and to compare the results of such a study with those obtained here. A recent experiment by Hirsh (1959), in which *S*s judged which of two pure tones (of different frequency) started first, is relevant to this point.

One finds in Hirsh's results no indication of the sharp peaks so clearly evident in the discrimination functions of the present experiment. It is also of interest that the discrimination of onset time was better in the speech patterns than in Hirsh's stimuli. For speech stimuli that straddle phoneme boundaries (those that were predicted from the phoneme labeling results to yield the best discrimination), the difference in time of onset required to give 75% correct judgments was slightly less than 12 msec. The 75% point in Hirsh's experiment was about 17 msec. This difference appears the more significant when one takes into account that the stimulus variable was in a much simpler acoustic context in Hirsh's experiment, and that his *S*s were permitted to hear each stimulus pair as many times as they wished before making a judgment, while our *S*s had to respond after each presentation. These comparisons support the assumption that the discrimination peak at the phoneme boundary is a result of learning.

If we put aside the results of the control, we are left with speech discrimination data that fit very closely a set of curves derived on the extreme assumption that *S*s' discriminations were completely determined and limited by the phoneme labels they attached to the sounds. Such a result contrasts with the usual psychophysical finding, which is that many more stimuli can be discriminated than can be identified or labeled absolutely. It is difficult to conceive that discrimination and labeling would be equivalent, as they proved to be in this experiment, except on the assumption that the discrimination processes have been radically altered by long practice in attaching phoneme labels.[7]

SUMMARY

An experiment was designed (*a*) to measure the discriminability of certain acoustic differences when they are cues for the perceived distinction between phonemes and when they lie entirely within a single phoneme category, and (*b*) to compare such discrimination data with those obtained when essentially the same acoustic differences occur in sounds that are not perceived as speech.

The speech sounds were synthetic approximations to the syllables /do/ and /to/, which differed only in the relative time of onset of first and second formants, this feature having previously been found to be sufficient to cue the distinction between /d/ and /t/ phonemes. To find the location of the phoneme boundary on the scale of relative onset-times, the various stimuli were presented in several random orders, with instructions to identify each sound as /do/ or /to/. To measure discrimination, the stimuli were presented in ABX fashion, and *S*s were instructed to use whatever differences or similarities they could hear as a basis for deciding whether the third stimulus (X) was identical with the first (A) or the second (B).

It was found that discrimination was considerably better across a phoneme boundary than in the middle of a phoneme category. Indeed, this effect was so marked as to justify the assumption that the perception of these stimuli is essentially categorical—that is, that *S* can hear no differences among the stimuli beyond those that are revealed by the phoneme labels he applies to them. On that extreme assumption it was possible to use the frequencies with which the various stimuli were labeled /d/ or /t/ as a basis for predicting very accurately how well the stimuli could be discriminated.

A set of control stimuli—patterns of

---

[7] For a discussion of the mechanism which possibly underlies these learning effects, see Liberman (1957), pp. 121–123.

sound differing from each other in essentially the same way that the speech stimuli differed —was produced by inverting the speech patterns on the frequency scale. The discrimination data obtained with these inverted controls revealed no increase in discriminability at stimulus values in the region corresponding to the location of the phoneme boundary, and, in general, the discriminability of the control stimuli was much poorer than that of the speech. To the extent that the inverted stimuli are an appropriate control, it may be concluded that the sharpening of discrimination at the phoneme boundary is an effect of learning and, more specifically, that it represents a considerable amount of acquired distinctiveness.

REFERENCES

CHAPANIS, A. S., & HALSEY, R. M. Absolute judgments of spectrum colors. *J. Psychol.*, 1956, 42, 99–103.

COOPER, F. S. Spectrum analysis. *J. Acoust. Soc. Amer.*, 1950, 22, 761–762.

COOPER, F. S. Some instrumental aids to research on speech. In, *Report of the fourth annual round table meeting on linguistics and language teaching.* Washington, D. C.: Institute of Languages and Linguistics, Georgetown University, 1953. Pp. 46–53.

COOPER, F. S., LIBERMAN, A. M., & BORST, J. The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proc. Nat. Acad. Sci.*, 1951, 37, 318–325.

GARNER, W. R. An informational analysis of absolute judgments of loudness. *J. exp. Psychol.*, 1953, 46, 373–380.

GRIFFITH, B. C. A study of the relation between phoneme labeling and discriminability in the perception of synthetic stop consonants. Unpublished doctoral dissertation, University of Connecticut, 1958.

HIRSH, I. J. Auditory perception of temporal order. *J. Acoust. Soc. Amer.*, 1959, 31, 759–767.

LIBERMAN, A. M. Some results of research on speech perception. *J. Acoust. Soc. Amer.*, 1957, 29, 117–123.

LIBERMAN, A. M., DELATTRE, P. C., & COOPER, F. S. Some cues for the distinction between voiced and voiceless stops in initial position. *Lang. Speech*, 1958, 1, 153–157.

LIBERMAN, A. M., HARRIS, K. S., HOFFMAN, H. S., & GRIFFITH, B. C. The discrimination of speech sounds within and across phoneme boundaries. *J. exp. Psychol.*, 1957, 54, 358–368.

MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81–97.

POLLACK, I. The information of elementary auditory displays. Part I. *J. Acoust. Soc. Amer.*, 1952, 24, 745–749.

POLLACK, I. The information of elementary auditory displays. Part II. *J. Acoust. Soc. Amer.*, 1953, 25, 765–769.

(Received March 28, 1960)