Journal of Experimental Psychology
Vol. 54, No. 5, 1957

THE DISCRIMINATION OF SPEECH SOUNDS WITHIN AND ACROSS PHONEME BOUNDARIES[1]

ALVIN M. LIBERMAN, KATHERINE SAFFORD HARRIS, HOWARD S. HOFFMAN, AND BELVER C. GRIFFITH[2]

Haskins Laboratories and the University of Connecticut

In listening to speech, one typically reduces the number and variety of the many sounds with which he is bombarded by casting them into one or another of the phoneme[3] categories that his language allows. Thus, a listener will identify as *b*, for example, quite a large number of acoustically different sounds. Although these differences are likely to be many and various, some of them will occur along an acoustic continuum that contains cues for a different phoneme, such as *d*. This is important for the present study because it provides a basis for the question to be examined here: whether or not, with similar acoustic differences, a listener can better discriminate between sounds that lie on opposite sides of a phoneme boundary than he can between sounds that fall within the same phoneme category.

There are grounds for expecting an affirmative answer to this question. The most obvious, perhaps, are to be found in the common experience that in learning a new language one often has difficulty in making all the appropriate sound discriminations. The evidence for this is impressionistic in the extreme, and there is little information that would permit a definition of the more specific aspects of the difficulty. In whatever degree this difficulty exists, however, a reasonable assumption is that some part of it arises from the fact that a person who is newly exposed to the sounds of a strange language finds it necessary to categorize familiar acoustic continua in unfamiliar ways. If his discriminations have, by previous training, been sharpened and dulled according to the position of the phoneme boundaries of his native language, and if the acoustic continua of the old language are categorized differently by the new one, then the learner might be expected to have difficulty perceiving the sounds of the new language until he had mastered some new discriminations and, perhaps, unlearned some old ones.

In more explicit psychological terms, an affirmative answer is to be expected on the basis that the situations being considered here clearly meet the conditions for acquired similarity and acquired distinctiveness. If either or both of these processes do, in fact, occur, then two speech sounds which a listener normally lumps into the same phoneme class would come to be less discriminable than sounds to which he habitually attaches different phoneme labels. Indeed, one might conceivably find in language some very common and

[1] This research was supported in part by the Carnegie Corporation of New York and in part by the Department of Defense in connection with Contract DA49-170-sc-1642.

[2] Now at Bell Telephone Laboratories, Murray Hill, N. J.

[3] The phoneme is most often taken to be the smallest unit of speech that can, by itself, distinguish one utterance from another as to meaning. Thus, the existence of the two words *bill* and *dill* makes it clear that *b* and *d* are different phonemes in English. It should be emphasized that the phoneme is not a single sound, but is, rather, a class which can and usually does include a great many sounds that differ from each other in various ways without causing any change in meaning.

easily accessible cases in which the effects of such processes as acquired similarity and acquired distinctiveness are as great as many years of practice can make them.

The present experiment was designed to investigate the relation between phonemic labeling and discrimination in one language and within one group of phonemes. For this purpose a synthesizer was used to generate speech-like sounds and to vary them in small steps along an acoustic continuum known to contain important cues for the perception of the voiced stops, *b*, *d*, and *g*. When listeners are asked to label these sounds as *b*, *d*, or *g*, they normally tend by their responses to divide the continuum into three sharply defined phoneme categories, the shifts from one response, or phoneme label, to another being very abrupt. It was the purpose of this experiment to determine how well these same sounds can be discriminated, and, in particular, to see whether the discrimination functions have sharp inflections that correspond in position to the abrupt shifts (i.e., phoneme boundaries) in the labeling responses. In addition, an attempt has been made to determine to what extent the relation between discrimination and labeling has here been reduced to its theoretical limit. For that purpose, the obtained discriminations have been compared with a function that is computed from the labeling data on the extreme assumption that the listener cannot hear any differences among these sounds beyond those that are revealed by his use of the phonemic labels.

Method

Apparatus.—A special-purpose instrument, called a pattern playback, was used to generate the stimuli of this experiment. This instrument, which has been described in earlier papers (**1**, **2**, **3**), converts hand-painted spectrograms into sound, thus making it possible to synthesize speech-like auditory patterns and to control the various aspects of the pattern quite precisely.

Stimuli.—Figure 1 illustrates the spectrograms used to produce the stimuli. The stimulus variable is the direction and extent of the second-formant transition, this variable having been found previously (**6**) to be important for the perceived distinctions among *b*, *d*, and *g*. In the stimulus pattern at the extreme left of the top row of the figure, the second formant rises from a point 840 cps below its steady-state level, and in the pattern at the extreme right of the bottom row it falls from a point 720 cps above the steady state. Between these two extremes, the starting point of the transition varies in steps of 120 cps. For convenience these stimuli will be referred to by number, from 1 through 14, as indicated in Fig. 1.

The rising transition of the first formant had been found previously to be a marker for the class of voiced stops (**4**) and, as can be seen in Fig. 1, this first-formant transition is constant in all the stimuli. In the steady-state part of the pattern the first formant centers at 360 cps and the second at 2160 cps. Formants at these frequencies produce a synthetic approximation to the vowel *e* (as in *gate*).

The spectrograms were converted into sound on the pattern playback and recorded on magnetic tape. By copying, cutting, and splicing the magnetic tape, all the stimulus arrangements described below were made.

Stimulus presentation and S*s.*—The stimuli were presented to *S*s in two ways; singly, to determine how *S*s would label them as *b*, *d*, or *g*, and in an ABX arrangement to determine to what extent *S*s could discriminate them on any basis at all.

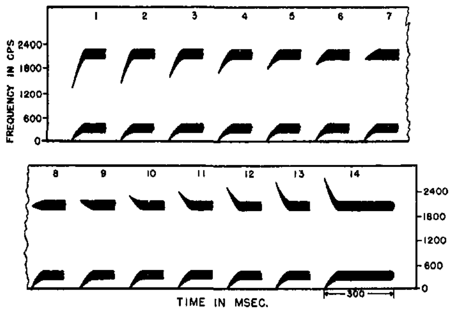

Fig. 1. Illustrations of the spectrographic patterns from which the stimuli of the experiment were produced. Pattern 14, at the lower right, is complete in all respects.

For the labeling part of the experiment, six magnetic tapes were prepared, each of which contained a different randomization of the entire series of 14 stimuli. There was a 6-sec. interval between stimuli. These tapes were presented to *Ss* with instructions to judge each stimulus as *b*, *d*, or *g*.

To test discrimination by the ABX procedure, the stimuli were arranged in triads, each of which consisted of an A stimulus, a B stimulus, and a third stimulus, X, which was identical either to A or to B. The *Ss* were instructed to determine whether X was the same as A or the same as B. It was strongly emphasized to each *S* that he was to make his judgment on the basis of any cues he could hear. The measure of discriminability is, of course, the percentage of the time that the X stimulus is correctly matched to A or to B.

It was not the purpose of this experiment to obtain actual DL's, but only to measure relative discriminability at every step on the continuum. For that purpose, the A and B stimuli of the ABX triad were made up by pairing each stimulus with the stimulus one step, two steps, and three steps removed from it. Each stimulus was paired in this fashion with the stimuli lying to its right on the continuum shown in Fig. 1. Thus, Stimulus 1 was paired with Stimulus 2 to form the A and B stimuli, respectively, of one ABX triad. Stimulus 1 was paired with Stimulus 3 in another triad, and with Stimulus 4 in a third. Similarly, Stimulus 2 was paired with Stimulus 3, with Stimulus 4, and with Stimulus 5, and so on for the remaining stimuli on the continuum. There was no stimulus one step to the right of Stimulus 14, so 14 does not appear as an A stimulus in the one-step series. For analogous reasons, Stimuli 13 and 14 do not appear as A stimuli in the two-step series, and Stimuli 12, 13, and 14 are missing from the three-step series.

From the above discussion, it follows that the number of A and B combinations was 36. For each combination of A and B stimuli, there were two triads—one in which X was identical to A and one in which X was identical to B. The total number of triads was 72. These triads were arranged into six tapes of 12 triads each. Within each ABX triad, the stimuli were spaced at 1-sec. intervals; successive triads were separated by 10 sec.

There were two groups of *Ss* in the experiment. Since the procedures for the two groups were slightly different they will be described separately. Group I consisted of five paid volunteers, all undergraduate students at the University of Connecticut, who had never heard any synthetic speech prior to this experiment. For the first 17 sessions, *Ss* were given only the ABX discrimination task to do, without being told that these were synthetic speech sounds. Beginning with Session 18, they were informed that the sounds on the discrimination tapes were synthetic approximations to speech. At this point the labeling tapes were introduced, and for the next three sessions *Ss* were given four labeling tapes per session with instructions to identify each stimulus as a speech sound. During this period they were given no discrimination task. At Session 22 and thereafter *Ss* were asked to identify the stimuli on the labeling tapes as *b*, *d*, or *g*, and the discrimination task was resumed. In these sessions, the discrimination task was always undertaken after *Ss* had finished judging the stimuli on the labeling tapes.

For several reasons the discrimination data obtained before and after the introduction of the labeling tapes have been combined. First, an examination of the results showed that there was no obvious change in the discrimination judgments following the introduction of the labeling tapes and the instruction to judge the stimuli on those tapes as *b*, *d*, or *g*. Second, in the sessions in which the labeling tapes were first introduced, *Ss* reported that they had previously heard the sounds on the discrimination tapes as speech. Moreover, when they were asked to identify the speech sounds they heard, they responded mostly with *b*, *d*, and *g*.

When the discrimination data are combined, there are, for each *S* in Group I, a total of 21 judgments of each ABX triad. Since there were two triads for each combination of A and B stimuli, the total number of judgments of each A and B combination by each *S* was 42. For each of the stimuli on the labeling tapes there are, for each *S*, 32 judgments. All of the labeling judgments were obtained after Session 21—i.e., after *Ss* had been asked specifically to identify the stimuli on the labeling tapes as *b*, *d*, or *g*.

Two of the five *Ss* in Group I were eliminated because they failed to apply the phoneme labels consistently. Since these *Ss* did not clearly divide the stimulus continuum into phoneme categories, one cannot compare their discrimination of speech sounds within and across phoneme boundaries. It should be noted here that the stimuli of this experiment have previously been presented to large groups of listeners, and that it is quite unusual to find as many as two out of five who are as unreliable in their responses as the two *Ss* who were eliminated from this experiment.

Group II consisted of four workers at the Haskins Laboratories. These *Ss* had previously had extensive experience in listening to synthetic speech sounds. The procedure for Group II was

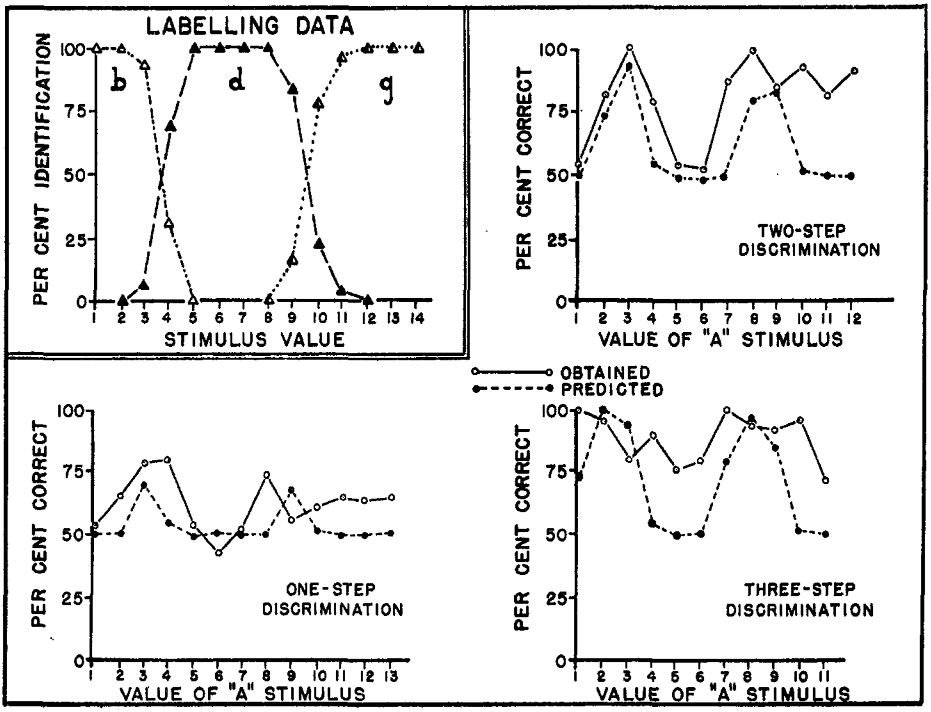

FIG. 2. Labeling and discrimination data for a single *S*, CD. The values given on the ordinates in terms of percentages are based on 32 and 42 judgments for the labeling and discrimination data, respectively.

like the procedure for Group I after Session 21. Thus, in each session, *S*s identified the stimuli on one labeling tape, and then judged three of the discrimination tapes—a total of 36 ABX triads. (This plan could not be followed exactly in the last few sessions due to scheduling difficulties.) In all, each *S* identified each of the stimuli on the labeling tapes 25 times, and he judged each ABX triad a total of 13 times. Since there were two triads for each A and B combination, each *S* judged each A-B combination a total of 26 times.

Of those *S*s whose data will be shown individually in the following section on results, CD and RV are from Group I, while LG and KH are from Group II.

RESULTS

In order to describe the method of presenting the data, and also to indicate in general terms the outcome of this study, the results that were obtained with one *S*, CD, will be presented first. At the upper left of Fig. 2 are plotted the labeling responses made by this *S* when the 14 stimuli were presented to him one at a time and in random order for judgment as *b*, *d*, or *g*. It can be seen that Stimuli 1, 2, and 3 were identified primarily as *b*, Stimuli 5, 6, 7, 8, and 9 as *d*, and Stimuli 11, 12, 13, and 14 as *g*. The shifts from one response to another are very abrupt, which is to say that the phoneme boundaries for this *S* are very sharp and stable.

The discrimination data obtained with the same *S*, CD, are also shown in Fig. 2. Only the "obtained" data points indicated by open circles and connected by solid lines will be considered at this time. Each point represents the percentage of correct responses for all ABX presentations (both ABA and ABB) when the A stimulus had the value shown on the abscissa and the B stimulus was one, two, or three steps removed (for the one-, two-, and three-step curves, respectively). Thus, for example, the first point on the one-step curve shows that this *S* correctly discriminated Stimulus 1 and Stimulus 2 (one step higher on our stimulus scale) 54% of the time.

A comparison of the discrimination functions with the labeling functions indicates that, other things equal, this *S* does, indeed, discriminate better between stimuli that lie on either side of a phoneme boundary than he does between stimuli that fall within the same phoneme category. For example, it can be seen from the labeling curves that Stimuli 1 and 3 were both identified as *b* almost all of the time, and the two-step discrimination curves show that this *S* correctly discriminated these stimuli only 55% of the time. Two steps beyond Stimulus 3, which was almost always identified as *b*, is Stimulus 5, which was always identified as *d*. The discrimination curves show that these two stimuli, consistently labeled as different phonemes, were correctly discriminated 100% of the time.

It must, of course, be supposed that if a listener can always identify two stimuli as different (as in the case of Stimulus 3 and Stimulus 5), he will surely be able to discriminate them. On the other hand, it does not necessarily follow from the fact that he identifies two stimuli as the same phoneme (as with Stimulus 1 and Stimulus 3) that he cannot discriminate them. One might think that he would hear *two* types of *b*—i.e., that he would hear what the linguist calls allophonic variations. In the example cited, *S* does not.

Clearly, the data obtained with this *S* are not all so neat and striking as the particular examples chosen, and some of the other *S*s were more variable, especially in their responses to the discrimination task, than the one *S*, CD, whose responses have been shown in Fig. 2. It is, nevertheless, reasonably apparent from an inspection of the data of all *S*s that the discriminations tend to be relatively more acute in the vicinity of phoneme boundaries than in the middle of phoneme categories. Before presenting these other data, however, it is desirable to provide a basis on which all the results can be evaluated. For that purpose the working hypothesis will be stretched to a theoretical limit and its quantitative implications will be developed.

Make the extreme assumption that *S* can discriminate the stimuli only to the extent that he can identify them as different phonemes. Then suppose that in the discrimination task—i.e., when *S* is presented with the stimuli in ABX fashion and asked to say whether X is like A or like B—he can only assign the phonemic labels *b*, *d*, and *g* to the individual stimuli, and that he has no other basis for discriminating among the stimulus members of the various ABX triads. One can, then, use *S*'s responses in the phonemic labeling part of the experiment as a basis for calculating the frequency with which he will correctly discriminate in any given ABX arrangement of the stimuli. To do this, one must first refer to the labeling part of the experiment and take account of the relative frequency with which *S* identified each of the 14 stimuli as *b*, *d*, or *g*. It is, of course, possible to go from these data to calculations of the probabilities that a given ABX triad will be heard as any one of the possible sequences of the three phonemes. One then needs only to determine for each triad which of the possible phonemic sequences will lead to responses that would be counted as correct discriminations.

By reference to the phonemic labeling part of the experiment we first determine for each A and B stimulus of the various ABX triads the relative frequency with which it was identified as *b*, *d*, and as *g*. These relative frequencies will be used as estimates of the probabilities of hearing the various stimuli of the ABX triads as *b*, as *d*, and as *g*. Let us call the probabilities of hearing these phones p_b, p_d, and p_g in the case of an A stimulus and p_b', p_d', and p_g' in the case of a B stimulus. We assume next that the various stimuli within each triad are perceived independently of each other. It follows, then, that the probability (for a given ABA triad) of hearing a particular sequence, such as *b*, *d*, *g*, is $p_b p_d' p_g$. Since there are three alternative responses, there are $3 \times 3 \times 3$, or 27, such phoneme sequences possible.

For any given triad the 27 possible phonemic sequences can be divided into three classes according to whether they lead *S* to make responses that would be counted as correct discriminations, as incorrect discriminations, or as discriminations that would on the average be correct half of the time and incorrect half of the time. In the case of any ABA type of triad, for example, *S* will be correct for any sequence in which the first and third stimulus members of the triad are heard as the same phoneme and the second member is heard as a different phoneme. He will be incorrect (again for the ABA type of triad) whenever the second and third members are heard as the same phoneme and the first is heard as a different phoneme. He will be correct half of the time and incorrect half of the time with two types of phoneme sequences: (*a*) all those in which he hears the first and second stimuli as the same phoneme and (*b*) those sequences in which he hears three different phonemes.[4]

These considerations can be expressed quantitatively in the following way. Let $P\ Corr_{(ABA)}$ be the proportion of the time that the listener is correct on a number of presentations of the same ABA sequence, P_R be the proportion of the time that the listener heard a sequence which would lead to correct discrimination, P_W be the proportion of the time that the listener heard a sequence which would lead to incorrect discrimination, and P_I be the proportion of the time that the listener heard a sequence which would, with equal likelihood, lead to correct and incorrect discriminations. Then

$$P\ Corr_{(ABA)} = 1P_R + 0P_W + .5P_I.$$

If the probabilities of the particular sequences described above are substituted appropriately for the general expressions P_R, P_W, and P_I, and if the resulting equation is then manipulated algebraically, we obtain

$$P\ Corr_{(ABA)} = \frac{.5 + p_b^2 + p_d^2 + p_g^2 - p_b p_b' - p_d p_d' - p_g p_g'}{2}.$$

So far, we have been concerned only with the case in which the presented sequence is ABA; analogous considerations lead to a similar equation for ABB, although the particular sequences which are correct and incorrect are different.

$$P\ Corr_{(ABB)} = \frac{.5 + p_b'^2 + p_d'^2 + p_g'^2 - p_b p_b' - p_d p_d' - p_g p_g'}{2}.$$

The ABA and ABB sequences were presented equally often; therefore, an average $P\ Corr$ is given by

$$\bar{P}\ Corr = \frac{.5 + (p_b - p_b')^2 + (p_d - p_d')^2 + (p_g - p_g')^2}{4}.$$

The solid points connected by dashed lines in Fig. 2 represent the discrimination values derived from CD's labeling curves on the assumptions outlined above. It is seen in the case of this *S* that the predicted functions do indeed take care of some, although by no means all, of the

[4] When *S* hears a sequence of three different phonemes, he can be expected to make the correct discrimination half of the time only if we assume that he perceives each phoneme as equally like the other two. It is possible that the typical listener does not perceive *b*, *d*, and *g* in precisely this way. However, the labeling data of this experiment are such as to produce essentially zero probabilities of hearing such sequences in the ABX triads. We have, therefore, not yet attempted to determine how the similarities among *b*, *d*, and *g* are perceived, because a correction, no matter how large, would have negligible effect on the results being reported here.

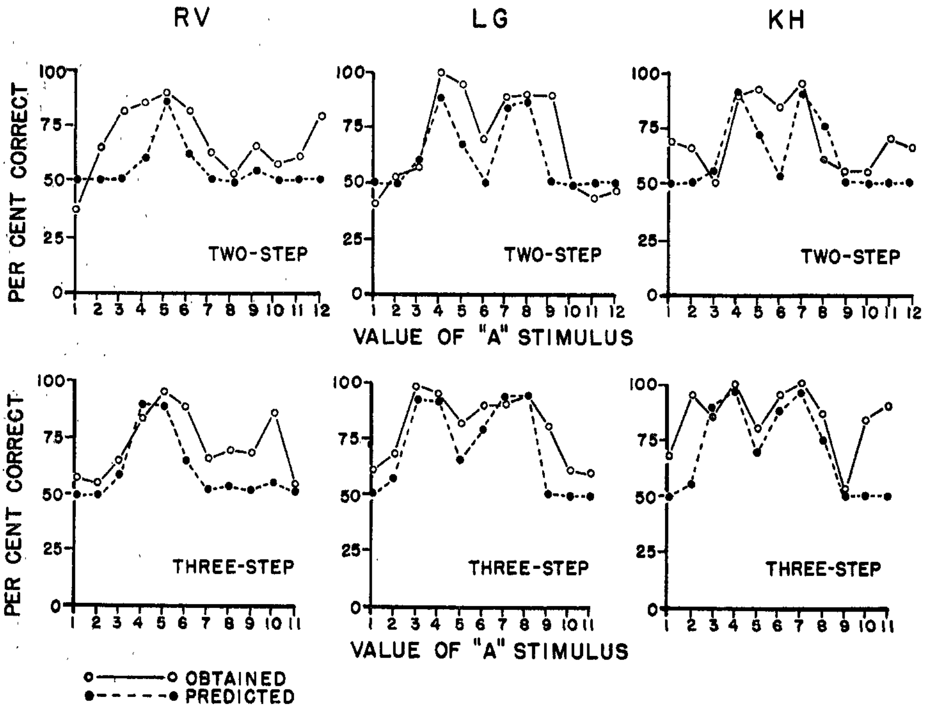

Fig. 3. Predicted and obtained discrimination values at two and three steps for RV, LG, and KH. Each of the obtained values is based on 42 judgments in the case of RV, and on 26 judgments for LG and KH.

variations in the obtained discriminations. The assumptions predict the points of high and low discrimination reasonably well, but they lead one to expect a general level of discrimination slightly lower than that obtained.

In Fig. 3 are data obtained with three other *S*s. These data are not presented as a way of indicating the results for all *S*s but only to show additional details of the results, including in particular a sample of the individual differences among *S*s. It is seen that the position and number of the peaks in the predicted discrimination functions vary somewhat from one *S* to another, reflecting differences in the way *S*s had assigned phoneme labels to the stimuli when they were presented for identification. It is also apparent that the obtained discriminations follow the inter-*S* differences in the predicted functions fairly well.

The simplest way to summarize the data for all *S*s is to make a scatter plot of obtained values against predicted values. Such plots are shown in Fig. 4, 5, and 6, for one-, two-, and three-step discriminations, respectively.[5] For the two- and three-step data, regression lines have been fitted by the method of least squares, and these are shown for each set of points. The regression for the one-step data has not been determined because so

[5] It should be noted here that the reliabilities of all points in the scatter plots are not equal, since, as was pointed out under the section on procedure, the two groups of *S*s made different numbers of judgments of the various ABX triads.

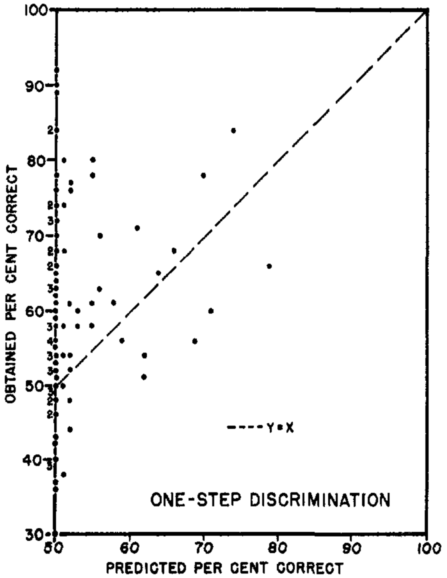

FIG. 4. Scatter plot of predicted vs. obtained values on the one-step discrimination for all *Ss*. The small numerals have been placed on the graph to indicate, where necessary, the number of values that occupy the same position of the coordinates.

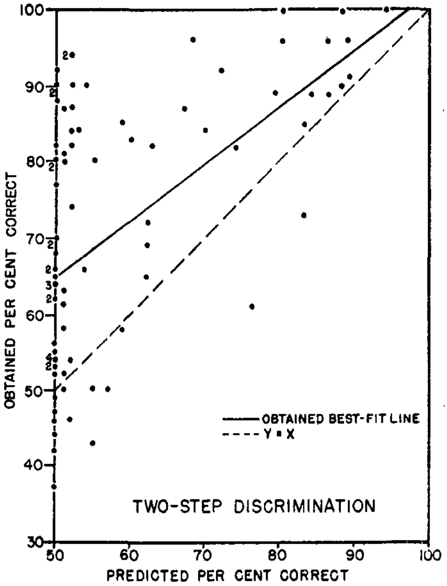

FIG. 5. Scatter plot of predicted vs. obtained values on the two-step discrimination for all *Ss*. The small numerals have the same meaning as in Fig. 4.

few predicted points lie above 50% that a meaningful fit cannot be obtained. The one-step data do not, therefore, provide a good test of our assumptions. For the two- and three-step data we should, of course, suppose that if the obtained data were essentially as predicted, give or take a little experimental error, the regression lines would be described by the equation "$x = y$."

The relationship between predicted and obtained values has been measured by computing tau,[6] a nonparametric measure of correlation. The correlations are +.14, +.43, and +.43 for the one-, two-, and three-step discriminations, respectively. Significance levels, in the same order,

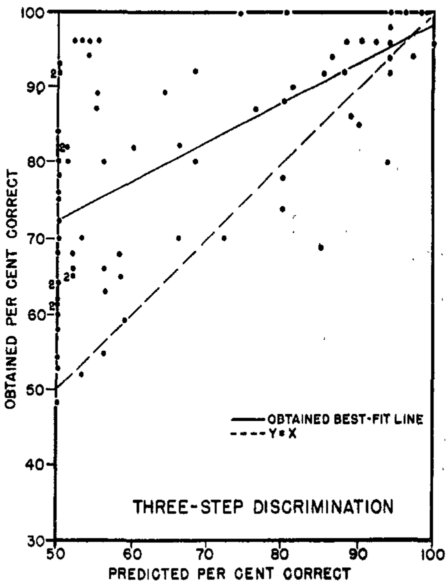

FIG. 6. Scatter plot of predicted vs. obtained values on the three-step discrimination for all *Ss*. The small numerals have the same meaning as in Fig. 4.

[6] We have used a nonparametric measure because our data fail to meet the assumption of homoscedasticity. For a description of tau, see Kendall (**5**).

are $P = <.08$, $P = <.001$, and $P = <.001$. Thus, it is seen that the correlations are highly significant for the two- and three-step discriminations, but in the case of the one-step discrimination the relationship could have arisen by chance. The failure to obtain significance for the one-step discrimination is not surprising since, as was pointed out earlier, most of the one-step points are predicted to lie at 50%.

Although there is a significant relationship between obtained and predicted points for the two- and three-step discrimination data, it is apparent in both cases that the lines of best fit are systematically displaced upward. This indicates that while the assumptions predict fairly well the occurrence and location of the inflections in the discrimination curves, they apparently lead to an underestimation of the general level of discrimination. This was previously noted in the data of the individual *S*s as shown in Fig. 2 and 3, where it was seen that the obtained discrimination functions tended to fit quite well with the predicted curves except that they were in general at a slightly higher level.

It is difficult on the basis of the data now available to make an unequivocal interpretation of the difference in level between obtained and predicted discrimination functions. One possibility, of course, is that this discrepancy represents a margin of "true" discrimination —i.e., an ability of *S* to distinguish the speech sounds, not solely on the basis of the phonemic labels, but also more directly by the essential acoustic differences among the patterns. A very different possibility is that the discrepancy between obtained and predicted is the result of certain detailed aspects of the experimental procedure. For example, irrelevant discriminable aspects of the stimuli, such as accidental stray noise, could have provided *S*s with an extraneous basis for deciding, in regard to the ABX triads, whether X was A or B. Such additional stimuli would, of course, have had no effect on *S*s' responses in the phonemic labeling part of the experiment. The result would have been, then, to make the obtained discriminations somewhat better than one should have expected them to be. In general, the procedures of this experiment were not such as to control most effectively for these irrelevant discriminanda. It will be possible in future research to take greater precautions against their occurrence, and then to determine whether and to what extent the discrepancy between obtained and predicted is reduced.

Within the data of the present experiment, however, there is some evidence that the discrepancy between obtained and predicted discrimination functions is due, at least in part, to "true" discrimination. This evidence is to be found in the fact that the discrepancy would appear to be greater for the three-step than for the two-step function. One might expect to obtain this difference if *S* were truly discriminating the stimuli on the basis of their essential acoustic characteristics, since one would then presumably find as between the two- and three-step discriminations that the three-step discrimination was the easier. Other factors, such as the irrelevant discriminanda discussed in the preceding paragraph, would have affected the two- and three-step conditions equally, and would not have caused the departure from predicted values to be greater in the three- than in the two-step data.

Discussion

The results of this experiment cannot be assumed unequivocally to reflect the effects of learning on discrimination. There is, of course, the possibility that the inflections in the discrimination function are given innately, and that the phoneme boundaries have been placed so as to coincide with these discontinuities. This begins to seem unlikely when one considers that other languages have put phoneme boundaries at different places along the *b-d-g* continuum. In itself, this would appear to reduce the probability that the sharp inflections in the discrimination functions are innately given. One would have far more compelling evidence, however, if it were found that native speakers of such languages have their points of maximal and minimal differential sensitivity displaced along the continuum to correspond with the phoneme boundaries of their respective languages. Research to test this possibility is feasible.

In order to find out whether the effects here described represent acquired similarity, acquired distinctiveness, or a combination of both, it will, of course, be appropriate to obtain discrimination data on nonspeech stimuli that are otherwise identical with the synthetic approximations to speech. Complete identity is obviously impossible, unless one can get *S*s to hear the same sounds sometimes as speech and sometimes not. One can, however, reasonably expect to make revealing comparisons between discrimination data obtained with speech and nonspeech stimuli that vary along the same, simple stimulus dimensions. Sheer temporal duration, for example, is sometimes a cue for distinguishing speech sounds. (Duration of transitions distinguishes stop consonant from semivowel; duration of fricative noise is a cue for distinguishing among the classes fricative, affricate, and stop.) It will be possible to obtain discrimination functions for variations in duration when those variations cue the perceived differences among phonemes, and to obtain comparable data for variations in duration of nonspeech sounds. We should suppose that a comparison of discrimination functions such as these would help greatly to determine whether the typical listener's long training in speech perception has served selectively to sharpen or to dull his discrimination of speech sounds, or whether, perhaps, it has done both.

Summary

This experiment was designed to measure the relation between *S*s' phonemic identifications (as *b*, *d*, *g*) of certain synthetic speech sounds and the extent to which they can discriminate the sounds as being different in any way. The stimuli were two-formant approximations to consonant-vowel syllables. They varied in the extent and direction of the second-formant transitions, this variable having previously been found to be an important cue for the perceived distinctions among *b*, *d*, and *g*.

In one part of the experiment these sounds were presented singly and in random order for identification as *b*, *d*, or *g*. The responses obtained in this way tended, with most listeners, to divide the stimulus continuum into three sharply bounded phonemic categories, indicating that the perceived shifts from one phoneme to another were rather abrupt.

In a second part of the experiment an ABX procedure was used to measure *S*s' ability to discriminate these sounds. The results indicated that discrimination was better at the phoneme boundaries than in the middle of the phoneme category. That is, with acoustic differences equal, *S*s discriminated better between speech sounds to which they habitually attach different phonemic labels than they did between sounds which they normally put into the same phoneme class. To obtain a more nearly precise evaluation of this effect, the obtained discrimination curves were compared with those that would be predicted from the phoneme identification curves on a most extreme assumption: namely, that *S* can only hear those differences that are revealed by the way in which he attached phonemic labels to the stimuli. The discrimination curves that were produced on this assumption predicted fairly well the occurrence and location of points of high and low discriminability in the obtained functions, but *S*s tended in general to discriminate the stimuli somewhat better than expected from the extreme assumption.

References

1. Cooper, F. S. Spectrum analysis. *J. acoust. Soc. Amer.*, 1950, **22**, 761–762.
2. 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 Univer., 1953. Pp. 46–53.

3. Cooper, F. S., Liberman, A. M., & Borst, J. M. 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.
4. Delattre, P. C., Liberman, A. M., & Cooper, F. S. Acoustic loci and transitional cues for consonants. *J. acoust. Soc. Amer.*, 1955, **27**, 769–773.
5. Kendall, M. G. *Rank correlation methods.* London: C. Griffin, 1948.
6. Liberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychol. Monogr.*, 1954, **68**, No. 8 (Whole No. 379).

(Received November 6, 1956)