

SOME CUES FOR THE DISTINCTION BETWEEN VOICED  
AND VOICELESS STOPS IN INITIAL POSITION

A. M. LIBERMAN, P. C. DELATTRE AND F. S. COOPER

*(Haskins Laboratories, New York)*

*Reprinted from*

LANGUAGE AND SPEECH

Vol. 1, Part 3, July-September 1958, pp. 153-167.

# SOME CUES FOR THE DISTINCTION BETWEEN VOICED AND VOICELESS STOPS IN INITIAL POSITION\*

A. M. LIBERMAN,\*\* P. C. DELATTRE,\*\*\* AND F. S. COOPER

*Haskins Laboratories, New York*

Experiments with synthetic speech produced by the Pattern Playback indicated that the voiced stops in initial position could be made to sound like their voiceless counterparts by cutting back the beginning of the first-formant transition. Normally, the cutback of the first formant raises its starting frequency and also delays the time at which it begins relative to the other two formants. Both these changes appeared, in general, to be necessary. With certain combinations of stop and back vowel, however, a delay in the onset of the first formant was by itself sufficient to produce a strong voiceless effect. Substituting noise for harmonics in the transitions of the second and third formants (for the duration of the first-formant cutback) increased the impression of voicelessness over that obtained with cutback alone.

In judging stimuli that lay near phoneme boundaries, many listeners demonstrated what appeared to be a very high degree of acuity. It is possible that this is the result of long experience in the use of the language and thus represents an effect of learning on perception.

Phonetic observations and the results of spectrographic analysis point to several acoustic features that may underlie the discrimination of voiced and voiceless stops in initial position. Of these, the presence or absence of vocal cord vibration ought, perhaps, to be considered first, since it is the nominal basis for the designation of the two classes as "voiced" and "voiceless." On spectrograms this feature should be visible, and often is, as a line in the low-frequency region, appearing relatively earlier for the voiced than for the voiceless stops. Although this "voice bar" may be adequate as a phonetic basis for distinguishing the two classes of sounds there is no reason to believe that it is necessarily of overriding importance in perception.<sup>1</sup>

In the case of English, phoneticians agree about a second characteristic, aspiration, which is presumed to be present in initial stops uttered in the voiceless manner and absent in their voiced counterparts. It is difficult to guess what all the acoustic correlates of aspiration are likely to be, but it is reasonable to suppose that one of its acoustic manifestations will be the presence of noise rather than harmonics

\* *This work was supported in part by the Carnegie Corporation of New York and in part by the Department of Defense in connection with Contract No. DA49-170-sc-2564. A portion of this paper was presented at a meeting of the Acoustical Society of America in Ann Arbor, Michigan on October 24, 1957.*

\*\* *Also at the University of Connecticut.*

\*\*\* *Also at the University of Colorado.*

<sup>1</sup> *The distinction between voiced and voiceless stops (in initial position) is quite perceptible in whispered speech. In itself, this must mean that the presence or absence of vocal cord vibration cannot be the only cue.*

in the beginning parts of the formant transitions.<sup>2</sup> Observation of spectrograms indicates that this might be the case, but the details are often unclear.

Experimental work with synthetic speech suggests other possible cues. Thus there have been persistent indications that an important aspect of voicing is somewhere to be found in the first formant. In this connection, we have known for some time that we could produce a strong voiced stop only by starting the first formant very low on the frequency scale (Delattre, Liberman and Cooper, 1955). Starting the first formant anywhere else weakens the voicing impression, though it cannot be said to produce a strong impression of voicelessness. In the case of released stops in *final* position, Malécot has found some indication that he could convert voiced stops into voiceless ones by reducing the intensity of the first formant of the release, or by omitting it entirely.<sup>3</sup> Exploratory work by the authors of this paper indicated that Malécot's manoeuvre seemed to produce essentially the same effect when applied to stops in *initial* position; this is not at all surprising, since the release of the final stop is, in effect, the beginning part of a new syllable, albeit a short one.

The experiments to be described in this paper were intended primarily to investigate further this last-mentioned feature—the elimination of a portion of the first-formant transition—as a cue for the distinction between voiced and voiceless stops in initial position.<sup>4,5,6</sup> Investigations of the effects of the other two features described above, *viz.*, vocal cord vibration and noise in the transitions, have also been undertaken, and their results will be presented.

<sup>2</sup> Dr. Fischer-Jørgensen (1954) has presented a lengthy discussion of this and related matters.

<sup>3</sup> Personal communication from Dr. André Malécot.

<sup>4</sup> Our interest in this cue was heightened as a result of a conversation with Dr. Gunnar Fant in which he informed us that he had observed this effect in spectrograms, and suggested that there were reasons for supposing that it would occur. Since this article was written Dr. Fant has published these and related observations. See especially: G. Fant, *Acoustic Theory of Speech Production*. Royal Institute of Technology (Division of Telegraphy-Telephony) Report No. 10, 1958, Stockholm; and G. Fant, *Modern instruments and methods for acoustic studies of speech*. Acta Polytechnica Scandinavia, Ph. Series No. 1, 1958.

<sup>5</sup> It is not appropriate in this paper to attempt a detailed explanation of the first-formant intensity change in terms of articulatory-acoustic considerations, and, indeed, we are not at this point fully prepared to do so. In general, however, such an explanation is possible on the basis of several factors which singly or in combination might produce the effect. One consideration, suggested to us by Dr. K. N. Stevens, is that the noise source of affrication may be weak in the low frequencies. A second suggestion, for which we are indebted to Dr. G. Fant, is that in English speech the vocal cords are open during the articulation of the aspirated portion of the voiceless stop, and that, in consequence, the lower resonance may be effectively "lost" in a large back cavity that extends beyond the open cords into the bronchial tubes. Also, there is at least a possibility that the noise source is not in such a position as to excite the entire tract.

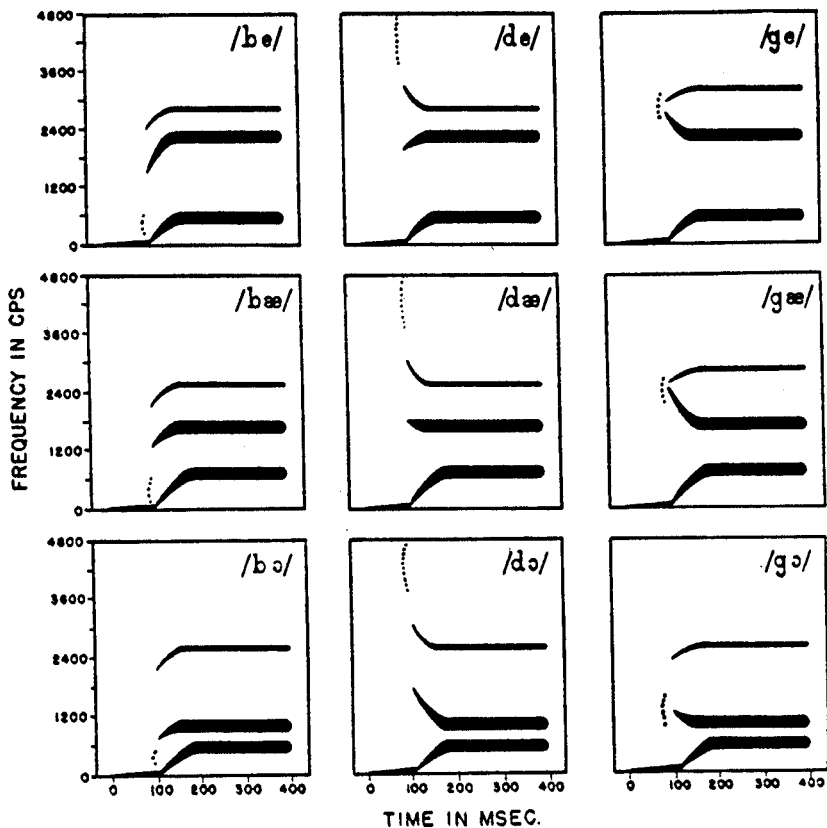


Fig. 1. Hand-painted spectrographic patterns used to produce the most extremely voiced stops (plus vowels).

#### FIRST-FORMANT CUTBACK

In this experiment we studied the effects of progressively eliminating the transition of the first formant. For that purpose we first prepared hand-painted spectrographic patterns like those shown in Fig. 1. These patterns were made in accordance with the results of several experiments (Liberman, Delattre, and Cooper, 1952 ;

\* Mlle. Durand (1956) has found that increasing the rate of first formant-transition contributes some voicelessness to the stop consonants. We have not ourselves investigated this cue systematically, nor have we carefully compared its effects with those produced by reducing the intensity of the first-formant transition. It is our impression, however, that in the case of English the latter change has by far the larger and more realistic effect.

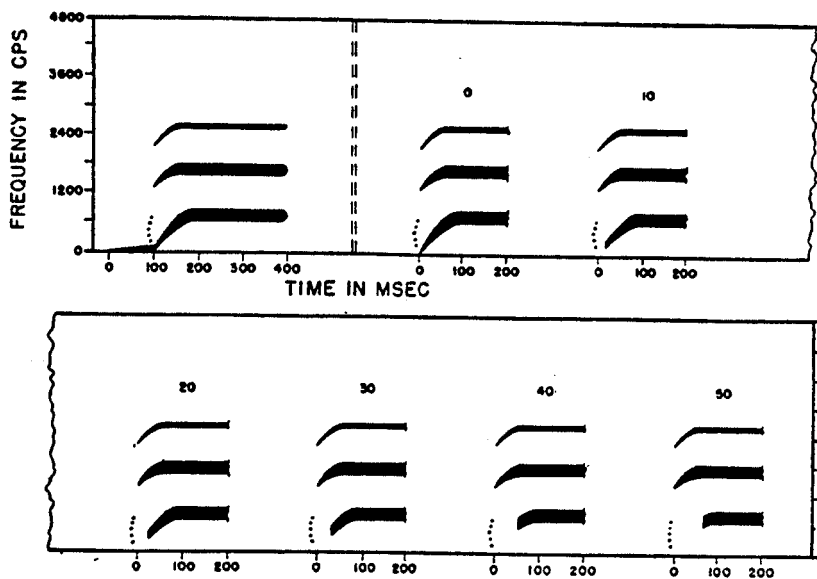


Fig. 2. Spectrographic patterns illustrating the way in which the voice bar was removed and the first formant was cut back to produce one stimulus series. The patterns shown are appropriate for /b,p/ with /æ/. The numbers above the patterns show the amount of first-formant cutback in msec.

Delattre, Liberman, and Cooper, 1955; Harris, Hoffman, Liberman, Delattre, and Cooper, 1958) on the acoustic cues for the stop consonants and were intended to synthesize stop consonant-plus-vowel syllables, the bursts, transitions, and steady states being appropriate for /b/, /d/, and /g/ before /e/, /æ/, and /ɔ/. Since the intensity of the burst is itself a cue for the voiced-voiceless distinction, great care had to be exercised to make this feature essentially neutral. This was accomplished, separately for each pattern, on a trial-and-error basis.

The patterns of Fig. 1 have the fully rising first formant which is presumed to produce the strongest voiced effect. Each of these patterns was then altered in the manner illustrated (for /b,p/ with /æ/) in Fig. 2. The first step was to remove the voice bar and so produce the pattern labelled "O". Successive cutbacks of the first formant in steps of 10 msec. produced the stimuli labelled "10" to "50".

It should be emphasized here that the duration of cutback in the first formant corresponds to the period of aspiration, during which the vocal cords are open. Any energy in the second and third formants during this period must, therefore, consist of noise rather than harmonics. In the first set of experiments, the second and third formants were nevertheless filled with harmonics. The effects of replacing the harmonics with noise will be dealt with in the third part of this paper.

Cutbacks of the first formant analogous to those described above were made on the other synthetic stop-plus-vowel patterns. In this way nine series (three stops times three vowels), each consisting of seven stimulus patterns (the six cutback conditions plus the zero cutback with the voice bar) were produced. These patterns were converted into sound by the playback, recorded on magnetic tape, and spliced into a random order. The resulting tape was played for 28 listeners who were instructed to judge each sound as /b/, /d/, /g/, /p/, /t/, or /k/ and to guess if necessary. Of the 28 listeners, 20 were phonetically naive students at the University of Connecticut, five were workers at the Haskins Laboratories who had no special linguistic training, and three were trained linguists.

Before considering the major results of the experiment, we should note that in making their responses the listeners had to choose from among six alternatives. In effect, they had not only to determine whether the sound was voiced or voiceless, but also to identify its place of production (*i.e.*, to decide whether it was a labial, alveolar, or velar stop). Out of a total of 1,764 judgments made, only 15 represented "place" errors—that is, less than one per cent of the sounds were identified by the listeners as having a place of production other than that intended by the experimenters. This aspect of the data has, therefore, been ignored and in Fig. 3 the judgments for each "place" category have been plotted separately in such a way as to show how the identification as voiced or voiceless varied as a function of the changes in the stimulus variable.

It is seen from Fig. 3 that the first-formant cutback effectively converted voiced stops into voiceless. The alveolar stop /d,t/ required the greatest amount of cutback to be heard as voiceless, and the velars /g,k/ only slightly less, while the labials /b,p/ changed to voiceless with relatively small amounts of the first formant removed. Indeed, in the case of the labials many listeners judged them as voiceless when the voice bar had been taken away and before any part of the first-formant transition had been eliminated. This is the only case in which the effect of the voice bar can be seen, since all the alveolar and velar stops were heard as voiced by all the listeners whether the voice bar was in or out.

The effect of the first-formant cutback appears to be largely independent of the vowel for /b,p/ and /g,k/; in the case of /d,t/, however, there is a tendency for the shift to the voiceless member to occur at progressively larger cutbacks through the series /e, æ, o/.

Although each stimulus was judged only once by each listener, it is nevertheless possible to assess the consistency of the individual subject's responses. For this

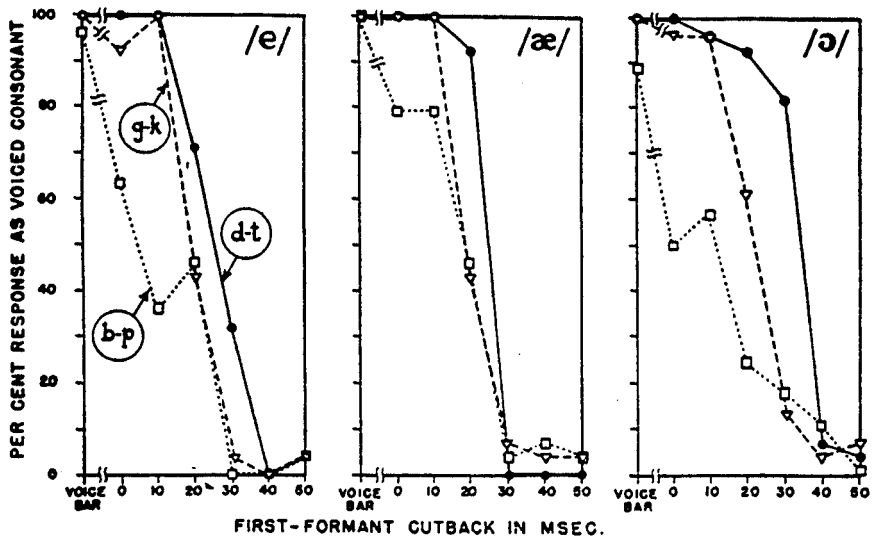


Fig. 3. Responses of 28 phonetically naive listeners as a function of amount of cutback of the first formant. The data are shown as the percentage of listeners who identified each pattern as a voiced stop. For the stimulus condition in which the voice bar preceded the patterns, the responses are shown on the separate ordinate at the left of each graph. The graphs are arranged in groups according to the vowels /e, æ, ɔ/ with which the consonant transitions were paired.

purpose we tabulated the responses of each subject separately against the various values of the stimulus variable. Since the stimuli were presented in random order, we can take it as an indication that the subject was self-consistent if it is possible to locate a dividing line on the stimulus series such that all responses on one side of the line are voiced and on the other side unvoiced—that is, if the subject has sorted perfectly across one step of the stimulus scale. The results of this kind of analysis are presented in Table I, where we see, for each stop paired with each vowel, the percentage of listeners who sorted the stimuli perfectly. (For the purposes of this analysis the voice-bar condition has been treated as if it were the first step on the stimulus continuum.)

It can be seen that individual self-consistency is highest for /d,t/, slightly lower for /g,k/, and, relatively, quite a bit lower for /b,p/. Although it is difficult to define standards for a situation like this, it would appear in general that the self-

TABLE I

## STOP

		STOP		
		/b,p/	/d,t/	/g,k/
VOWEL	/e/	64	86	82
	/æ/	86	100	96
	/ɔ/	71	89	86

Percentages of 28 listeners who sorted the stimuli perfectly.

consistency is remarkably high. In the case of /d,t/ with /æ/, all subjects sorted perfectly.

In addition to questions about self-consistency or variability of the individual, we can ask about variability between individuals. It is clear from the plots of group data in Fig. 3 that the inter-subject agreement must be quite high for /d,t/ and /g,k/, especially with /æ/. Indeed, for /d,t/ with /æ/, one can see from the group data that 26 of the 28 listeners changed their judgments from /d/ to /t/ at the same point on the abscissa (30 msec.); the remaining two listeners changed at a point one step removed (20 msec.). In the case of /b,p/, there is, as we have seen, considerably more intra-subject variability, and it becomes more difficult to compare individuals. It is possible, however, to find enough listeners who were sufficiently consistent with their own judgments to permit conclusions concerning differences between individuals (as distinct from random variability). By examining the responses of these listeners, we find that individual differences are indeed relatively greater for /b,p/ than for /d,t/ or /g,k/.



To provide a further test of these conclusions about intra- and inter-subject variability, we arranged several additional random orders of the stimuli and presented each order a number of times to each of 11 phonetically naive subjects. In this way we built up a sufficient number of responses to enable us to plot, separately for each subject, functions similar to those of Fig. 3. These plots, which are not shown here, fully support the conclusions reached on the basis of the group data. Both intra- and inter-individual differences are very small for /d,t/ and /g,k/ and relatively larger for /b,p/.

These facts about intra- and inter-individual differences may mean that the first-formant cutback is less important for /b,p/ than for /d,t/ or /g,k/. On the other hand it is possible that these results will prove to be peculiar to the types of patterns used or to the general methods of synthesis. Thus, it may be that the synthetic labial stops are less realistic than the others and that this produces the relatively greater variability we observed. Similar considerations must, of course, be taken into account in interpreting the fact that the labial, alveolar, and velar stops changed from voiced to voiceless at different amounts of first-formant cutback. Certain constant aspects of the patterns (such as the relative strength of the burst) may have contributed to voicing or voicelessness. If this were the case, the shift from voiced to voiceless might well occur at different first-formant cutbacks in different experiments, depending on the particular balance of cues. In exploratory work with quite a wide variety of patterns, however, we have obtained results that are, in general, much like those reported for the particular patterns of this experiment, and we think, therefore, that those aspects of the results being discussed here do have considerable generality.

Comparisons among the stops in regard to the relative importance of the first-formant cutback or the amount of cutback required do not in any event affect the major conclusion of this part of the study: namely, that the presence or absence of a "cutback" in the first formant is a sufficient cue, and very likely an important one, for distinguishing voiced and voiceless stops in initial position.

#### EXPERIMENTAL SEPARATION OF TIME AND FREQUENCY CHANGES INVOLVED IN THE FIRST-FORMANT CUTBACK

In the patterns of the first series of experiments, the cutback of the first-formant produced two correlated changes in the stimulus: (1) a progressively greater elevation in the starting frequency, and (2) a progressively greater delay in the time at which the first formant began. In the second series, we carried out a number of studies in an attempt to separate these two variables and to determine the role of each in the effects just described. In general we have found that both factors are important: a rising first formant is a cue for the class of voiced stops, and a time delay in the first formant, without the rising transition, is a cue for the voiceless stops. Thus, a continuum of patterns from voiced to voiceless can, in general, be produced only

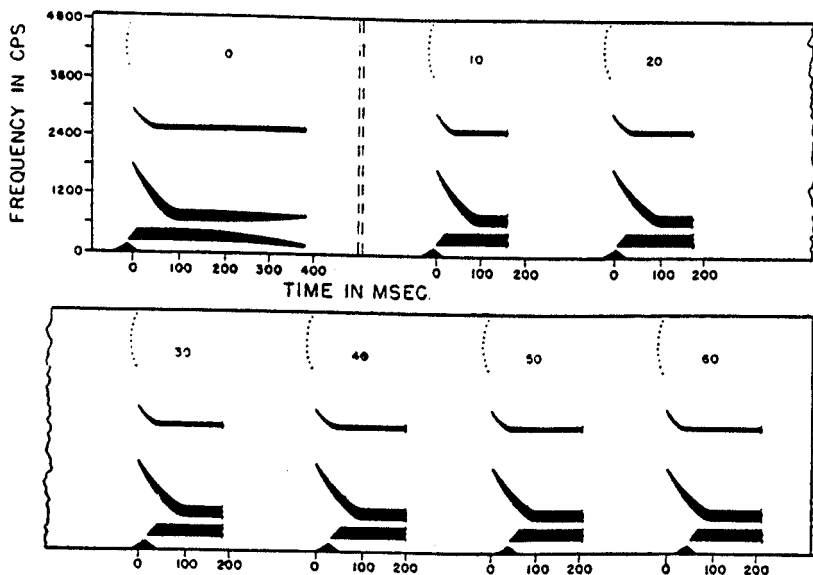


Fig. 4. Illustrations of the patterns used to study the effects of the first-formant time of onset on the perceived distinction between /d/ and /t/. The numbers above the patterns show the extent of cutback in msec.

by raising the first-formant starting frequency and also delaying its onset. This is what we did in the first experiments, and this is presumably what occurs in nature.

In a few special cases, however, it is possible to go from voiced to voiceless by making only one change. We were motivated to try to find such a set of stimuli by the desire to know whether the high degree of accuracy with which listeners identified the stimuli in the first experiment could be maintained when the physical difference was reduced to a single acoustic dimension—in this case, time of onset of the first formant.

One of the special cases in which time of onset of the first formant is the only variable is shown in Fig. 4. When the stop is in front of /o/, as in this case, the first formant does not have far to rise, and, given a large second-formant transition, a reasonably compelling voiced stop can be produced with a first formant like that shown.<sup>7</sup> The various patterns in the figure show successive delays of 10 msec. in the onset of the first formant. These patterns were recorded on magnetic tape and assembled into three random orders, each stimulus pattern being represented twice

<sup>7</sup> The triangular patch just below and to the left of the first formant was added because it was found, in this special case, to increase the realism of the sound. It may function as a voice bar or possibly, in combination with the first formant, as a transition. In either event its position is constant in relation to the first formant, and the only difference among the patterns is in the time of onset of the first formant-plus-triangular-patch.

in each order. All of these stimuli were then presented to a group of 27 phonetically naive listeners for judgment as /d/ or /t/. The judgments of these listeners, plotted in Fig. 5, make it clear that the delay in first-formant onset is sufficient to cue the distinction between /d/ and /t/. More interesting, perhaps, is the fact that quite a few listeners sorted these randomly presented stimuli perfectly across a stimulus difference equal to only 10 msec. of delay in the first formant. Out of the total group of 27 subjects, 7 sorted perfectly on the first random order, 12 on the second, and 15 on the third. (It should be remembered in this connection that going through a single random order means going through the stimulus set twice.) In effect, these listeners were applying phoneme labels with perfect consistency to stimuli which differed by 10 msec. of first-formant delay. Given the appropriate non-speech controls, it should be possible in future research to determine whether this represents, as we think it does, an exceptionally great sensitivity brought about by long experience with the language.

#### NOISE IN PLACE OF HARMONICS IN THE FORMANT TRANSITIONS

As was pointed out earlier in this paper, it is to be presumed that the first-formant cutback corresponds to the period of aspiration. If so, the second and third formants should, of course, contain noise rather than harmonics for the duration of the first-formant cutback. This noise might be expected to be a cue for voicelessness, either by itself or in combination with the cutback. In this part of the study we have undertaken to measure the effect of combining noise in the second and third formants with first-formant cutback, and to compare this effect with that produced by first-formant cutback alone (harmonics rather than noise in the transitions of the second and third formants, as in the first experiment) and by noise alone (noise in all three transitions without cutback of the first formant). For this purpose we used the Voback, a relatively new synthesizer (Borst and Cooper, 1957). Like the Pattern Playback used in the first two parts of this study, the Voback converts hand-painted spectrograms into sound; indeed, the same spectrogram can be used with either instrument. There are many differences between these synthesizers, but for the present study the most relevant one is that with the Pattern Playback the experimenter can use only an harmonic series of tones as a source, while with the Voback he can use either an harmonic series or noise (though not both at the same time), as he wishes. It may prove to be of some consequence for this experiment that the Voback is significantly superior to the Pattern Playback in signal-to-noise ratio.

Taking advantage, then, of the possibility of putting noise rather than harmonics into the formant transitions, we experimented with the types of stimulus conditions illustrated in Fig. 6. The pattern at the left, which is very similar to the zero cutback pattern of Fig. 2, is a starting point from which the changes illustrated in the other patterns are made. It has the fully rising first formant previously found to be a voicing cue combined, in this case, with place cues (burst and transitions of first and second formants) appropriate for /d/ before /ɔ/. The next pattern to the right,

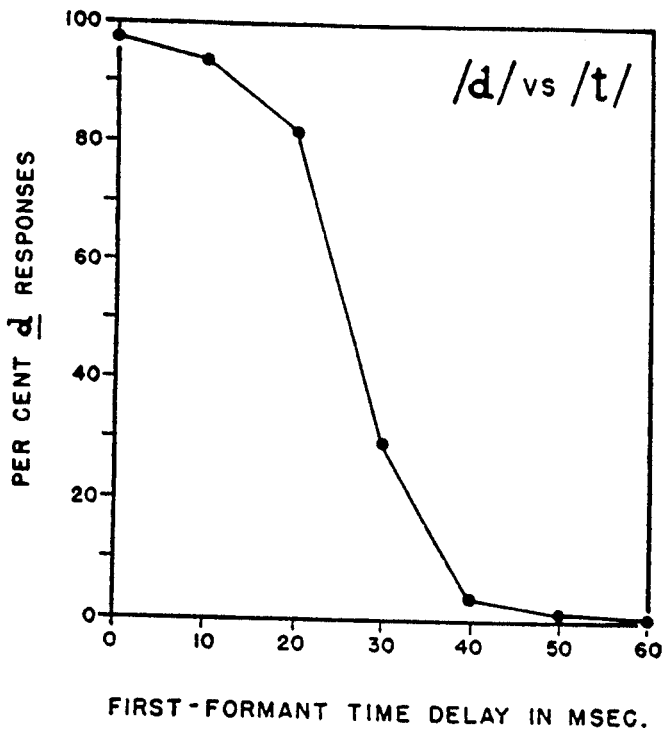


Fig. 5. Responses of 27 phonetically naive listeners to the patterns in which time of onset of the first formant was varied (see Fig. 4). Each listener judged each stimulus six times, making a total of 162 judgments per stimulus. The data are shown as the percentage of times each stimulus was heard as /d/.

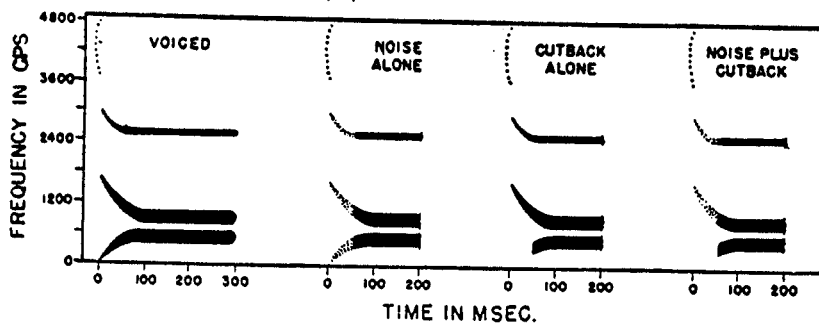


Fig. 6. Patterns illustrating the stimulus conditions designed to evaluate the effects of noise in the formants.

labelled "noise alone," differs only in that noise has been substituted for harmonics in all three formants.<sup>3</sup> Third from the left is the "cutback alone" condition, which

<sup>3</sup> In this case, and in all others in which noise was substituted for harmonics, the noise was 2 db down from the harmonics. These measurements were made with a VU meter, the noise and harmonics being in essentially steady state during the measurement.

essentially duplicates the cutback patterns of the first experiment. The "noise plus cutback" at the extreme right differs from "cutback alone" only in that noise has replaced harmonics in the first and second formants for the duration of the cutback of the first formant. In all cases the stimuli were changed in steps of 10 msec.; the patterns shown in Fig. 6 illustrate the extreme, or 50 msec., condition. The stimulus conditions described here were arranged for /b,p/, /d,t/, and /g,k/ paired in all combinations with /e/, /æ/, and /ɔ/, a total of 54 stimuli. (As was pointed out above, the patterns shown in Fig. 6 are for /d,t/ before /ɔ/.)

Exploratory work with these various patterns made it clear that the use of noise alone (*i.e.*, substitution of noise for harmonics in all three formants) produces essentially no voicelessness. This conclusion was based on the judgments of phonetically naive listeners. We proceeded, then, to look into the possibility that the rising first formant in the "noise alone" pattern (see Fig. 6) was such a potent voicing cue as to override any impression of voicelessness contributed by the noise in the formants. For these experiments we straightened the first formant in patterns that were otherwise like the one labelled "noise alone." Listening to these patterns made it clear that while this manoeuvre may have reduced the effect of voicing, it did not produce voiceless consonants.

The conclusion that voicelessness in stops is not produced by filling all the transitions with noise (without cutting back the first formant) is strengthened by suggestions that have come from some as yet incomplete studies on /h/. Here we find that steady-state ("straight") formants which are filled with noise for the first 50 msec. and then with harmonics for the remainder of the sound produce an impression of a brief whispered vowel followed by a longer voiced vowel. If, with this same pattern, we remove the first formant for the duration of the initial noise section, we hear something more closely approximating /h/ followed by a voiced vowel. To see these results in relation to those obtained in the attempt to produce voiceless aspirated stops, one need only take account of the common assumption that aspiration is to be equated, at least roughly, with /h/.

The exploratory work with the patterns of Fig. 6 confirmed that, while putting noise into the formants does not convert voiced stops into voiceless ones, cutting back the first formant is quite effective for this purpose, as it was in the first experiments. Further exploratory work indicated that substituting noise in the second and third formants seemed to produce a greater impression of voicelessness than can be obtained with first-formant cutback alone. To permit a better comparison of the effects of "cutback alone" and "cutback plus noise," we obtained judgments from 32 phonetically naive listeners. The stimuli they judged were those illustrated by the two patterns at the right in Fig. 6 and described earlier in this section of the paper; the "noise alone" condition was omitted. Thus, for each stop-vowel combination there were six patterns of first-formant cutback alone (from 0 cutback through 50 msec. in steps of 10 msec.) and six patterns of noise plus cutback (again from 0 cutback through 50 msec. in steps of 10 msec.). These conditions

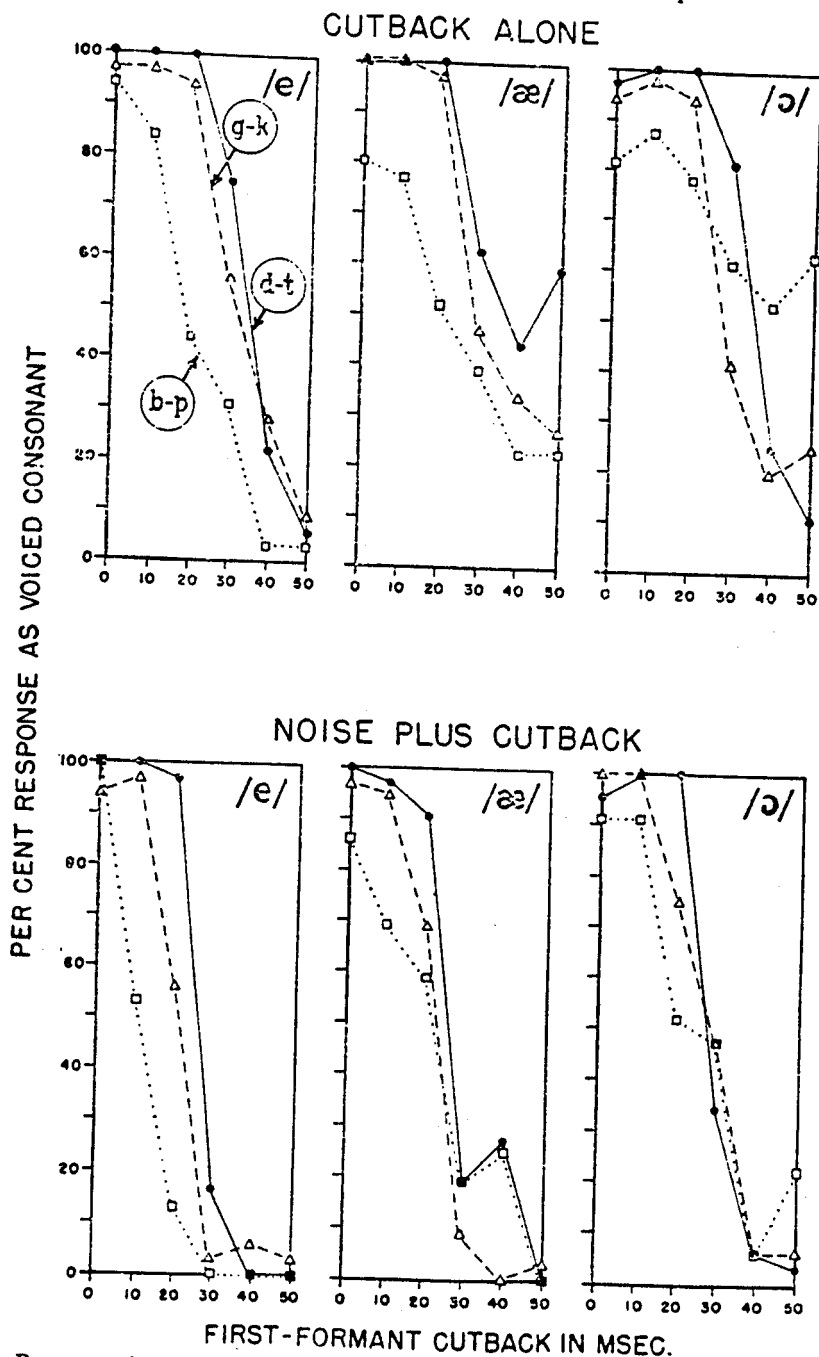


Fig. 7. Responses by 32 phonetically naive listeners to patterns designed to compare the effects of cutback alone and noise plus cutback (see Fig. 5). The data are shown as the percentage of listeners who identified each of the patterns as a voiced stop. The graphs are arranged in the rows according to the experimental conditions and in the columns according to the vowel with which the consonant transitions were paired.

were arranged for each of the three stops paired with each of three vowels, /e/, /æ/, and /ɔ/.

As in the earlier experiment, the listener had to identify each sound according to place and voicing. Out of a total of 3,456 judgments made, only 87 (or 2.5%) were in error as to place. We have, therefore, ignored this aspect of the subjects' judgments, and, as in the first part of the paper, have dealt only with the judgments of the sounds as voiced or voiceless. In Fig. 7 the judgments are plotted against the experimental variables: for the data plotted in the top row the variable was first-formant cutback alone; in the bottom row the variable was first-formant cutback with noise in the second and third formants for the duration of the cutback.

Clearly, cutting back the first formant on Voback produced a significant shift toward voicelessness. In the graphs of the bottom row it is seen that a further contribution to voicelessness was made by substituting noise in the second and third formants for the duration of the cutback in the first formant.

The condition of cutback alone essentially duplicates with Voback the experiment previously done on the Pattern Playback and reported in the first part of the paper. A comparison of the results obtained in the two studies shows that the first-formant cutback was somewhat less effective in the present experiment than in the earlier one. There is, however, a very close correspondence in all respects between the results obtained in the earlier experiment and those obtained with the Voback in the condition of first-formant cutback plus noise. Bearing in mind that there is considerably less background noise in the Voback than in the Pattern Playback, we can assume that the absence of noise in the transitions is more noticeable in sounds produced by the former machine than in sounds produced by the latter. Given the noise background in the sounds produced by the Pattern Playback the listener might well have "filled in" with noise wherever appropriate, and it would, then, have made less difference whether we had put noise into the transitions or not. On this basis we should suppose that the data obtained with the Voback show a clearer separation of effects, and we should conclude that, while cutting back the first formant is a most effective cue to voicelessness, it is perhaps, somewhat less effective than we might have been led to suppose from looking at the data obtained with the Pattern Playback. We should conclude, further, that the addition of noise to the second and third formants for the duration of cutback in the first formant increases voicelessness, even though the addition of noise alone (without cutback) is not at all effective for this purpose.

## REFERENCES

- BORST, J. M. and COOPER, F. S. (1957). Speech research devices based on a channel vocoder. *J. acoust. Soc. Amer.*, 29, 777 (Abstract).
- DELATTRE, P. C., LIBERMAN, A. M., and COOPER, F. S. (1955). Acoustic loci and transitional cues for consonants. *J. acoust. Soc. Amer.*, 27, 769.
- DURAND, M. (1956). De la perception des consonnes occlusives ; questions de sonorité. *Word*, 12, 12.
- FISCHER-JØRGENSEN, E. (1954). Acoustic analysis of stop consonants. *Miscellanea Phonetica*, 2, 42.
- HARRIS, K. S., HOFFMAN, H. S., LIBERMAN, A. M., DELATTRE, P. C. and COOPER, F. S. (1958). Effect of third-formant transitions in the perception of the voiced stop consonants. *J. acoust. Soc. Amer.*, 30, 122.
- LIBERMAN, A. M., DELATTRE, P. C., and COOPER, F. S. (1952). The role of selected stimulus variables in the perception of the unvoiced stop consonants. *Amer. J. Psychol.*, 65, 497.