CUES FOR THE DISCRIMINATION OF AMERICAN ENGLISH FRICATIVES IN SPOKEN SYLLABLES

KATHERINE SAFFORD HARRIS

Haskins Laboratories, New York

Reprinted from

LANGUAGE AND SPEECH

Vol. 1, Part 1, Jan.-March 1958, pp. 1-7.

CUES FOR THE DISCRIMINATION OF AMERICAN ENGLISH FRICATIVES IN SPOKEN SYLLABLES*

KATHERINE SAFFORD HARRIS
Haskins Laboratories, New York

It has sometimes been assumed that the identification of the fricatives of American English in CV syllables depends primarily on the characteristics of the noise (i.e., nonvocalic) portion of the speech sound. A second possibility is that characteristics of the vocalic portion—previously shown to be cues for the perception of other consonants—are important for the fricatives. These alternatives were tested by combining the noise from one spoken fricative-vowel syllable with the voiced portion of another. Results indicate that the important cues for the fricatives /s/ and /ʃ/ are given by the noise but that the differentiation of /f/ and /θ/ is accomplished primarily on the basis of cues contained in the vocalic part of the syllable. Similar results were obtained for the voiced counterparts of these sounds.

This paper will be concerned with the first steps in an investigation of the cues which listeners use in discriminating among the members of the class of unvoiced fricatives /f/, /θ/, /s/ and /ʃ/, and among their voiced counterparts, /v/, /ð/, /z/, and /ʒ/.

If one examines spectrograms of these sounds in consonant-vowel syllables, he finds them to be made up of two successive segments—a period of noise, which we shall call the friction, succeeded by a segment with well-marked formant structure, which we shall call the vocalic portion.[1] Cues for identifying the phonemes might well be in either or both of the two portions. Indeed, one might infer from research with other groups of consonants that both parts are important. For example, the friction of the fricatives is much like the burst of the stop consonants ; it has previously been found that the frequency position of this burst is significant in determining which stop will be perceived (Liberman, 1952). Transitions of the second and third formant in the vocalic part of the syllable have been shown to be important for distinguishing among the liquids and semi-vowels (O'Connor, 1957), among the nasals, and among the stops (Liberman, 1954).

As a step towards isolating the acoustic cues for the fricatives, then, it has seemed reasonable in this experiment to assess the overall relative contribution of the friction and vocalic portions of fricative-vowel syllables. Detailed examination of particular cues within each of the two parts will be left for later study.

**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 DA 49-170-sc-2159. Some of the results were reported at a meeting of the Acoustical Society of America in New York in June 1954.*

[1] *These characteristics of spectrograms of fricatives have been noted in two previous studies. Potter, Kopp and Green (1947) have called the two parts of fricatives described above the 'fill' and the 'consonant influence on the vowel', while Joos (1948) has called them the 'noise patch' and the 'glide' or 'transition'.*

Procedure

A simple means of assessing the relative importance of cues in the friction and vocalic portions would be to split the friction segment away from the vocalic segment of the syllable, and then recombine friction and vocalic portions from different syllables. Presumably, the more important part of the sound would determine which phoneme a listener would hear.

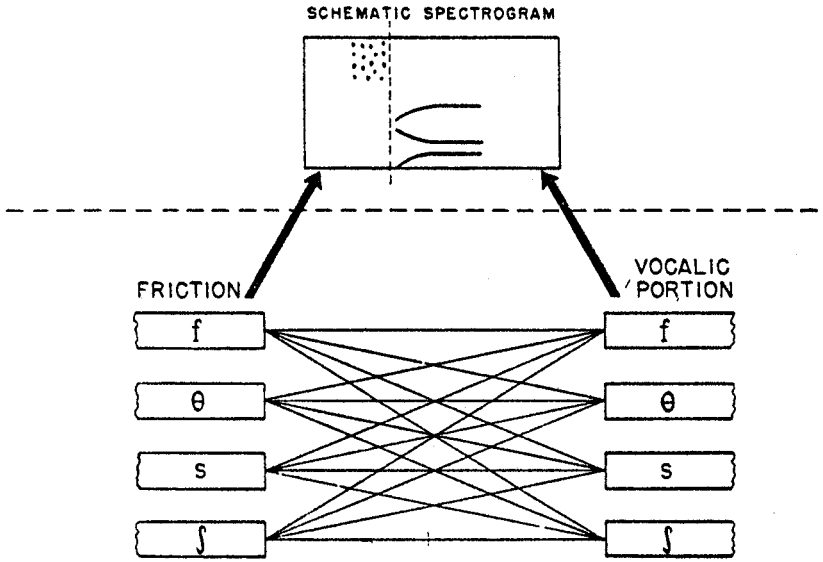

Fig. 1. Test stimuli generated from the spoken syllables /fi/, /θi/, /si/, and /ʃi/. The dotted line on the schematic spectrogram at the top of the figure indicates the point in the recorded sound at which the magnetic tapes were cut. Each of the resulting four types of friction was combined with each of the four types of vocalic portion to make 16 new combination stimuli, as indicated in the lower half of the figure.

A similar technique of recombination has been used by Schatz (1954) in a study of changes in stop consonants from one vowel to another. In her study, magnetic tape recordings of the stop consonants were made, and then the burst of noise at the beginning of the stop was split away from the vocalic portion, and interchanged with the burst from another stop consonant-vowel syllable. In the present experiment, we have used the same technique to interchange friction and vocalic parts of different fricative-vowel syllables.

The first step was to make tape recordings of a number of repetitions of each of the four syllables /fi/, /θi/, /si/, and /ʃi/, spoken by a single male speaker. Since we wanted to split the syllables between friction and vocalic portions, we needed to put a marker on the magnetic tape at the join of the two. To do this, the tape was run back and forth over the playback head of a tape recorder by hand, and the output monitored by listening and by watching the face of an oscilloscope ; the join of the

friction and vocalic parts of the syllable could be seen and heard by the change from low-intensity, high-frequency noise to high-intensity, low-frequency periodic sound waves Through the use of this method, all the tapes were marked and cut at the marked points A friction portion from each of the four syllable types was combined with a vocalic portion from each, to make 16 combination stimuli. The recombinations produced are schematized in Fig. 1.

A similar set of 4 repetitions of each of the four fricatives before each of the vowels /e/, /o/, and /u/ was recorded, and each set was recombined to make 16 stimuli for each vowel, analogous to the combinations for the vowel /i/ designated in the preceding paragraph. All 64 stimuli, 16 for each of the four vowels, were then re-recorded.

The naturalness of the recombined syllables was dependent on the accuracy of the location of the original cut between the friction and the vocalic parts of the syllable. This location was checked after the rejoining operation and the re-recording by turning the original tape to the oxide side and running a magnetized knife blade along the splices between friction and vocalic parts of each syllable, so that there was a sharp click placed at the join of the two. When spectrograms were made, the click appeared as a black line at the point of join. One could then tell by visual examination whether or not the cut had been made at the point intended. (Of course, it was necessary to use the re-recording made before the insertion of the test clicks for the final test tape.)

After all the rejoined syllables had been checked in the manner described above, the re-recording was made into a test tape by rearranging all the stimuli in random order, and spacing them in such a way that each syllable was repeated once after an interval of 0·9 sec., and successive pairs of syllables appeared 6 sec. apart. This recording of the 64 stimuli will be referred to below as the unvoiced fricatives test.

A similar test was made for the four voiced fricatives. The male speaker who recorded the unvoiced fricatives recorded /v/, /ð/, /z/ and /ʒ/, before the same set of four vowels, /i/, /e/, /o/ and /u/. The syllables were recombined as before and checked for splice position. We should note, however, that the spectrographic checking technique used is somewhat less accurate in the case of a voiced fricative, since the boundary between friction and vocalic portions is harder to define from a spectrogram than the boundary for an unvoiced fricative. The 64 stimuli were made into a test in the manner described above ; this recording will be referred to below as the voiced fricatives test.

The voiced and unvoiced fricatives tests were presented to 22 listeners, volunteers from undergraduate and graduate courses at the University of Connecticut. The subjects were present for two sessions, each of which contained one presentation of each test. Within the session, half the subjects heard the unvoiced fricatives test first, while half heard the voiced fricatives test first. All subjects were instructed to judge the stimuli as /f/, /θ/, /s/ or /ʃ/, for the former test, or as /v/, /ð/, /z/ or /ʒ/, for the latter.

We thought that the judgments made might be affected by the intensity level of the tests, since the friction portions of the different phonemes vary considerably in intensity; to control for this possibility we presented the stimuli at two different intensity levels in the two experimental periods. These intensity levels cannot be specified meaningfully in the usual units, because the listeners were seated in an experimental room listening to a loudspeaker, rather than listening on earphones. The two levels selected were, first, the lowest speaker output level at which the experimenter could hear the friction when seated in the subject's chair farthest from the loudspeaker, and, second, the highest level at which the sound could be presented without distortion. Analysis of the data has shown that this variable had no effect on the listeners' judgments over the range selected; therefore, results are not presented separately for the two intensities.

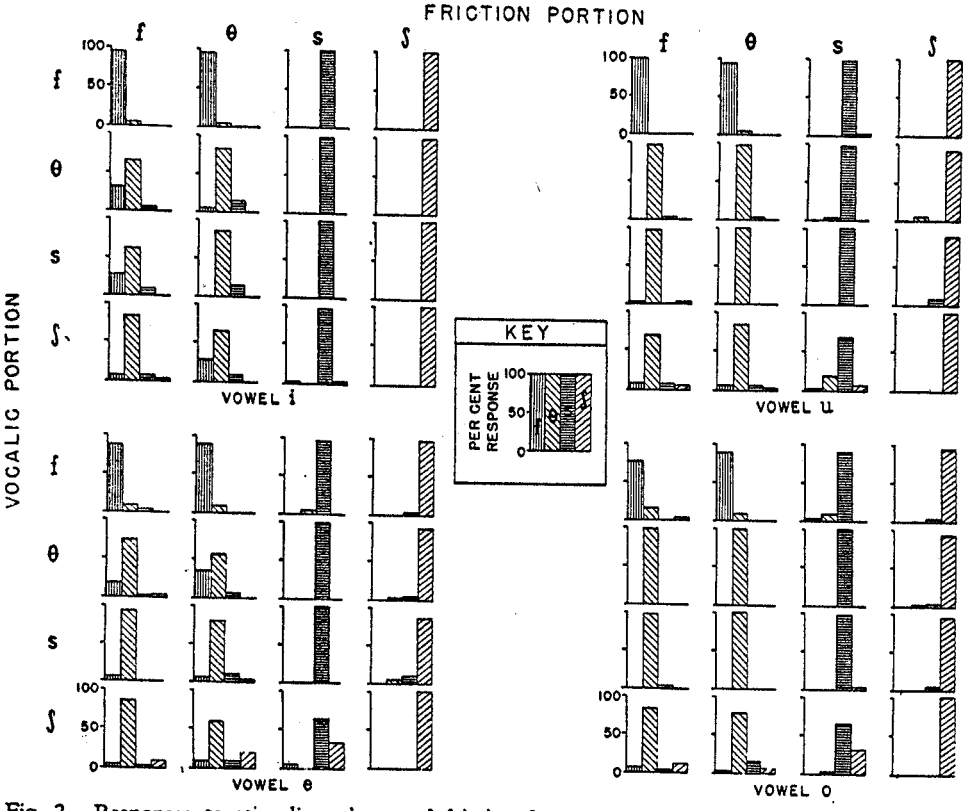

Fig. 2. Responses to stimuli made up of friction from one unvoiced fricative-vowel syllable paired with the vocalic portion of another. Within each small rectangle, the height of each of the four bars of the histogram indicates the per cent judgments of each of the four phonemes /f/, /θ/, /s/, and /ʃ/. The data were obtained from 22 subjects, each of whom heard each syllable twice

Results

The results for the unvoiced fricatives test are shown in Fig. 2. Each histogram represents the responses of the subjects to one of the 64 stimuli of the experiment. For example, the upper left-hand histogram in the upper left-hand quadrant indicates the responses of the subjects to a stimulus made up of /f/ friction and /f/ vocalic portion, with the vowel /i/. Approximately 95% of the subjects identified the resulting syllable as /f/, while 5% heard the syllable as /θ/.

Before describing the results in detail, we should note that the results are the same for the four vowels. This can be seen by examining the corresponding histograms in each of the four quadrants, which represent the four vowels. We will therefore describe the data hereafter without reference to specific vowels.

As can be seen in Fig. 2, the results of the test were quite different for /s/ and /ʃ/, on the one hand, and /f/ and /θ/, on the other. When /s/ friction was paired with any vocalic portion, the resulting stimulus was judged as /s/ ; similarly, when /ʃ/ friction was paired with any vocalic portion, the resulting stimulus was judged as /ʃ/. None of the other eight stimuli was judged as /s/ and /ʃ/ with any great frequency. Apparently, then, the friction of /s/ and /ʃ/ provide the necessary and sufficient cues for their identification, and override whatever cues may be provided by the vocalic portions.

The situation was somewhat more complicated for /f/ and /θ/ judgments. In general, only stimuli with /f/ or /θ/ friction were judged as either /f/ or /θ/, but which of the two judgments was made depended largely on the vocalic part of the syllable. Most of the listeners tended to judge a syllable with /f-θ/ friction as /f/ when it had /f/ vocalic portion, and as /θ/ when it had any other vocalic portion.[2]

The results of the voiced fricatives test, shown in Fig. 3, were similar to those just described, though not as clear. The phonemes /z/ and /ʒ/ behaved like their unvoiced counterparts /s/ and /ʃ/ for all vowels, in that both /z/ and /ʒ/ were identified almost entirely by their friction portions. The results for /v/ and /ð/, on the other hand, were variable from vowel to vowel. When the vowel was /o/ or /u/, syllables with /v/ friction and /ð/ friction were identified as /v/ or /ð/ depending on the vocalic portion ; in other words, with these vowels the sounds behaved in the same way as /f/ and /θ/. For /i/ and /e/, however, there

[2] *Hughes and Halle (1956) have recently reported some experimental results which are in general agreement with the results of the unvoiced fricatives test. In their study, several speakers produced syllables containing /f/, /s/ and /ʃ/ with various vowels. The isolated friction segments of the syllables were then presented to listeners who were asked which of the three phonemes had been spoken. All three were identified quite well. The result is not surprising for /s/ and /ʃ/, since we had concluded that friction provided the necessary and sufficient cues for their identification. Furthermore, we would expect that /f/ friction would be discriminable in the set of alternatives presented, since in our experiment, /f/ friction was not confused with any friction except /θ/, and /θ/ was not a possible response in Hughes' and Halle's experiment.*

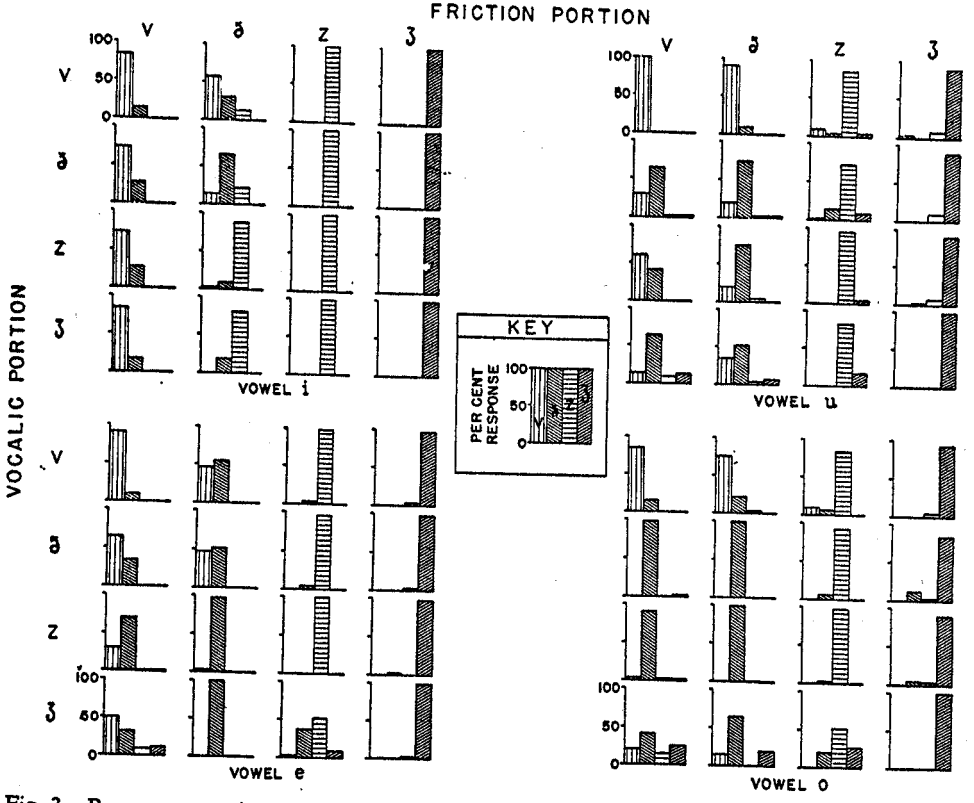

Fig. 3. Responses to stimuli made up of friction from one voiced fricative-vowel syllable paired with the vocalic portion of another. The data are displayed as in Fig. 2.

appeared to be some contribution from both friction and vocalic portions. In this connection, it should be remembered that, as we noted above, the join of the friction and vocalic portions is less clear for the voiced fricatives ; consequently, the original splicing may not have been made as well and we might expect somewhat more variability in the results.

By way of summary, then, we may say that the results suggest a general way of describing the perception of the four fricatives (with their voiced counterparts). The listener may be said to behave as if he first decided on the basis of friction, whether the syllable belonged to the /s-ʃ/ class or to the /f-θ/ class. If /s/ or /ʃ/ he uses the friction again to decide which of these alternatives it was. If, on the other hand, the first decision had been that the sound belonged in the /f-θ/ class, then the listener uses the vocalic portion to decide which of the two sounds, /f/ or /θ/, he had heard.

References

HUGHES, G. W. and HALLE, M. (1956). Spectral properties of fricative consonants. *J. Acoust. Soc. Amer.*, 28, 303-10.

JOOS, M. (1948). Acoustic phonetics. *Language Suppl.*, 27, 93.

LIBERMAN, A. M., DELATTRE, P. C. and COOPER, F. S. (1952). The rôle of selected stimulus-variables in the perception of the unvoiced stop consonants. *Amer. J. Psychol.*, 65, 497-516.

LIBERMAN, A. M., DELATTRE, P. C., COOPER, F. S. and GERSTMAN, L. J. (1954). The rôle of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychol. Monogr.*, 68, no. 8, 1-13.

O'CONNOR, J. D , GERSTMAN, L. J., LIBERMAN, A. M., DELATTRE, P. C. and COOPER, F. S. (1957). Acoustic cues for the perception of initial /w j r l/ in English. *Word*, 13, 24-43.

POTTER, R. K., KOPP, G. A. and GREEN, H. C. (1947). Visible Speech (New York).

SCHATZ, C. D. (1954). The rôle of context in the perception of stops. *Language*, 30, 47-56.