

Feature analyzers for the phonetic dimension *stop vs. continuant*

RANDY DIEHL

University of Texas at Austin, Austin, Texas 78712

Perception of sounds along the phonetic dimension *stop vs. continuant* was studied by means of a selective adaptation procedure. Subjects first identified a series of synthetic consonant-vowel syllables whose formant transitions varied in duration, slope, and amplitude characteristics. They were perceived as either [ba] or [wa]. After the initial identification test, an adapting stimulus was presented repeatedly, and then the subjects again identified the original test series. Adapting with a stop (either [ba] or [da]) led to a decrease in the number of test stimuli identified as [ba], whereas adapting with the continuant sound [wa] led to an increase in the number of [ba] identification responses. Removing the vowel portion of an adapting stimulus greatly reduced the identification shift only when the resulting stimulus was no longer perceived as speech-like. A reduction in the number of [ba] identifications occurred even when a nonspeech "stop" (the sound of a plucked string) was used as the adapting stimulus, suggesting that phonetic processing is not a necessary condition for an adaptation effect.

A growing body of experimental evidence suggests that linguistic feature analyzers mediate the process of speech perception. Much of this evidence derives from studies using a selective adaptation procedure of the following format: First, a series of speech sounds which vary along some phonetic dimension is presented to the subjects for identification. Next, an adapting stimulus is presented repeatedly, after which the subjects again identify the original test series. Eimas and Corbit (1973) hypothesized that if feature analyzers mediate the perception of voiced and voiceless stops, then selective adaptation with a stimulus having, e.g., the feature *voiced*, should cause the detector of that feature to fatigue and become less sensitive. In support of this hypothesis, they found that a voiced adapting stimulus led to a reduction in the number of test stimuli identified as voiced, whereas a voiceless adapting stimulus produced an increase in voiced responses.

Similar findings were obtained for the phonetic dimension *place-of-articulation* (Ades, 1974; Cooper, 1974; Cooper & Blumstein, 1974; Diehl, 1975). Adaptation with a stimulus of a particular place

value, e.g., bilabial, typically led to a decrease in the number of test stimuli identified as having that place value. Shifts in stimulus identification along the place dimension were obtained even when the adapting stimulus was voiceless and the test stimuli were voiced, suggesting that selective adaptation affects feature-specific, rather than exclusively phoneme-specific, analyzers (Cooper, 1974; Diehl, 1975). In addition, Diehl (1975) found reliable identification shifts along the place dimension when the acoustic information specifying place value for the adapting stimulus (burst-cued) was different from the information specifying place value for the test stimuli (transition cued). This suggests that at least part of the adaptation effect occurs at a site of phonetic, not merely acoustic, feature analysis. Further evidence of a phonetic level of adaptation is the fact that identification shifts were reduced or eliminated when the vowel portion of the adapting stimulus was removed, producing a stimulus which typically was not perceived as speech-like (Diehl, 1975; Eimas, Cooper, & Corbit, 1973).

One might argue that the observed identification shift following adaptation results not from decreased detector sensitivity but rather from a change in response criterion. Perhaps the adapting stimulus serves as a new standard against which judgments are made. To test this interpretation, Eimas and Corbit (1973) had subjects perform an ABX discrimination task on the test stimuli before and after selective adaptation.¹ In previous studies (e.g., Abramson & Lisker, 1970), discrimination performance has been shown to peak rather sharply at the boundary between two stop consonant categories. If, in the adaptation experiments, the identification shift were purely the result of a change in response criterion, we would not

This research was conducted while the author was at the University of Minnesota and was supported in part by a National Science Foundation graduate fellowship to the author and by grants to the University of Minnesota, Center for Research in Human Learning, from the National Science Foundation (GB-35703X), the National Institute of Child Health and Human Development (HD-01136), and from the Graduate School of the University of Minnesota. An earlier version of this paper was presented at the 89th meeting of the Acoustical Society of America, Austin, Texas, April 1975. The author is indebted to Neal Viemeister for his technical assistance and to James Jenkins for his helpful comments and criticisms. Reprint requests should be sent to Randy L. Diehl, Department of Psychology, 330 Mezes Hall, University of Texas, Austin, Texas 78712.

expect the discrimination peak to have a corresponding shift, since the ABX procedure should be free of any criterion effect. But, in fact, Eimas and Corbit did observe shifts in the identification boundary. This appears to rule out an account of the adaptation effect which is based purely on response factors.

As we have seen, previous studies of selective adaptation have examined either the voicing dimension or the place-of-articulation dimension. The present experiment sought to investigate whether adaptation effects could be obtained on yet a different phonetic dimension, that of *stop vs. continuant*. In English, the class of stops includes the phonetic segments [b d g p^h t^h k^h], whereas the class of continuants includes vowels, liquids, e.g., [r], fricatives, e.g., [s], and glides [w j] (the latter is pronounced "y"). When stops are produced, the air flow through the mouth is effectively blocked; in the production of continuants, the primary constriction of the vocal tract is not narrowed to the point of completely blocking the air flow (Chomsky & Halle, 1968). In word-initial position, stops are perceptually characterized by an abrupt or explosive onset, continuants by a more gradual onset.

The stop-continuant dimension was selected for study for two principal reasons. First, it would allow one to test the generality of the selective adaptation effects that have been observed on the voicing and place dimensions. If those effects are truly general, then one would expect the following outcomes: (1) Adapting with stops or continuants should produce identification shifts along the stop-continuant dimension whether or not the adapting stimulus is phonemically equivalent with some of the test stimuli. (2) The identification shifts should be reduced or eliminated when the vowel portion of the adapting stimulus is removed, to the extent that such a stimulus is no longer perceived as speech-like.

Second, the stop-continuant contrast appears to be unique among phonetic feature dimensions in that it also characterizes a wide range of nonphonetic acoustic stimuli, e.g., the sound of a plucked violin string vs. that of a bowed violin string. (Here I am obviously referring to the perceptual contrast between stops and continuants rather than to the articulatory contrast between them. Perceptually, the sound of a plucked string has a much more abrupt quality at onset than the sound of a bowed string.) An interesting theoretical question was whether adapting with a nonphonetic stop or continuant would produce shifts in identification of a series of phonetic stimuli which vary along the stop-continuant dimension.

METHOD

Stimuli

The identification test stimuli were a series of 11 three-formant synthetic consonant-vowel syllables prepared by means of a

computer-controlled parallel resonance synthesizer (Glace-Holmes) at the University of Minnesota. They varied in duration of the formant transitions and were perceived as either the stop sound [b] or the continuant sound [w],² plus the vowel [a]. [That the distinction between [b] and [w] and between other stops and glides can be signaled by varying transition duration was first demonstrated in a set of experiments by Liberman, Delattre, Gerstman, and Cooper (1956).] Each stimulus consisted of an initial 20 msec of prevoicing, followed by formant transitions which in turn were followed by 290 msec of steady state formants corresponding to the vowel. The nominal starting frequencies of the transitions were: first formant (F_1), 230 Hz; second formant (F_2), 1,100 Hz; third formant (F_3), 1,900 Hz. The nominal terminal frequencies were: F_1 , 740 Hz; F_2 , 1,400 Hz; F_3 , 2,460 Hz. Transition durations in milliseconds for the 11 stimuli were: 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66. Because the starting and terminal frequencies of the transition were invariant for all the stimuli, an increase in transition duration was accompanied by a decrease in transition slope. After the period of prevoicing (which was obtained by setting F_1 at 230 Hz while zeroing the amplitudes of F_2 and F_3), the formant amplitudes were set at a constant nominal level during the entire transition period and most of the steady state period. Despite these nominal amplitude settings, there was considerable amplitude fluctuation or ringing during roughly the first 100 msec of each stimulus. The actual amplitude characteristics of the stimuli of the test series varied systematically as a function of transition duration, with shorter transition durations being accompanied by higher onset amplitudes. Figure 1 displays the amplitude envelopes of the first and last stimuli of the test series.

Eight different adapting stimuli were prepared. Two were copies of the test stimuli having transition durations of 12 and 60 msec and were perceived as [ba] and [wa], respectively. A [da] and [ja] (pronounced "ya") were synthesized in a manner similar to the [ba] and [wa], except that the nominal starting frequency of the F_2

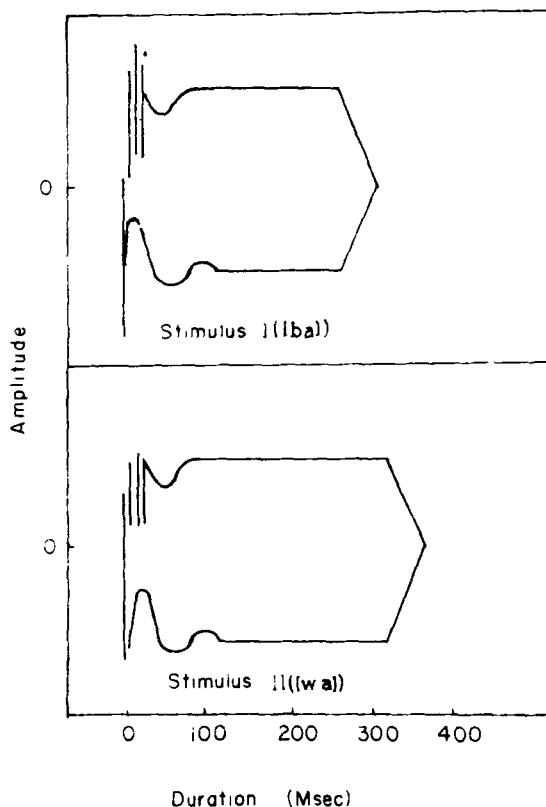


Figure 1. Amplitude envelopes of the first and last stimuli of the test series.

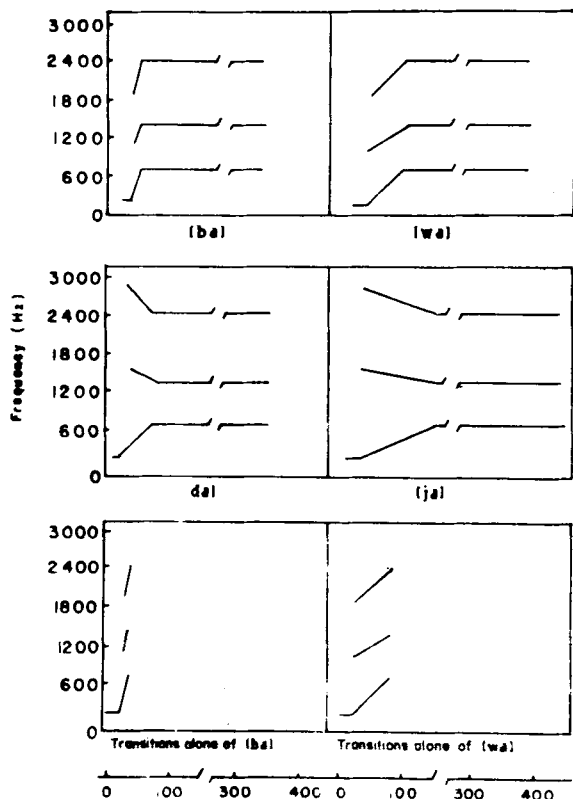


Figure 2. Schematized spectrograms of six of the adapting stimuli.

transition was 1,580 Hz and that of the F_3 transition was 2,900 Hz. (The terminal frequencies remained the same.) In addition, the transition durations of [da] and [ja] were 36 and 120 msec, respectively. It should be noted that [d] shares the feature *stop* with [b], whereas [j] shares the feature *continuant* with [w]. Two other stimuli were acoustically identical to the [ba] and [wa], except that the steady state vowel portions were removed. These will be referred to as "[b] transitions alone" and "[w] transitions alone." Schematized spectrograms of these first six adapting stimuli appear in Figure 2.

The two final adapting stimuli were nonphonetic. They consisted of a sawtooth waveform at 440 Hz. Each had an exponential rise and decay. They differed only in rise time: one stimulus had a 10 ± 5 msec rise time and was similar to the "stop" sound produced by plucking a stringed instrument. The other stimulus had a rise time of 80 ± 5 msec and was similar to the "continuant" sound produced by bowing a stringed instrument. The duration of the two stimuli was 620 ± 10 msec, approximately 150 msec of which occurred prior to the onset of decay. The amplitude envelopes of the "pluck" and "bow" stimuli are shown in Figure 3. (For a description of a similar set of stimuli, see Cutting & Rosner, 1974.)

Procedure

Following an initial stimulus familiarization session, nine experimental sessions, each lasting about 25 min, were run on consecutive weekdays. In the first experimental session, the subjects identified the test stimuli and then, after a 15-min interval with no stimulus presentation, identified the test series a second time. In each of the remaining sessions, the subjects first identified the test stimuli and then listened to one of the adapting stimuli presented repeatedly, after which they again identified the test stimuli. A Crown CX822 tape recorder was used to present the test stimuli, and the adapting stimuli were presented using a Revox A77 tape

recorder. Both sets of stimuli were routed through a mixer and heard through sets of Koss PRO 600AA earphones at a comfortable listening level.

The identification functions for the subjects in the unadapted state were obtained as follows. The test stimuli were presented in 10 blocks, each consisting of a different randomization of the 11 stimuli. The interval separating the stimuli was 2 sec within a block and 4 sec between blocks. The subjects were instructed to identify each stimulus by writing "b" or "w" on an answer sheet provided.

Next, identification functions were obtained after selective adaptation. The subjects listened to an adapting stimulus for periods of 1 min and, after each such period, identified one block of the 11 test stimuli. This continued until all of the original 10 blocks were identified. Two different orders of stimulus presentation were alternately used across the nine experimental sessions, but in a given session subjects were always given the same order of presentation for both identification tests. During the adaptation sessions, the onset-to-onset interval for the pluck and bow stimuli was 900 msec, while for the rest of the adapting stimuli it was 655 msec. There was a 4-sec interval between the end of an adaptation period and the onset of an initial test stimulus. At the end of each session, each subject was asked to describe on paper his perception of the adapting stimulus.

The adapting stimuli were assigned to the last eight experimental sessions in the following order: [w] transitions alone, [ba], [da], [b] transitions alone, [wa], pluck, bow, [ja].

Subjects

Eight students at the University of Minnesota served as subjects. All were native speakers of English and had no known hearing deficits. They received \$2 per session and were not informed of the experimental hypothesis beforehand.

RESULTS

Table 1 shows the net change in [b] identification responses between the first and second identification tests for the eight adaptation conditions and for the control condition in which no adapting stimulus was presented. For this control condition, the subjects showed no significant shift in the number of stimuli

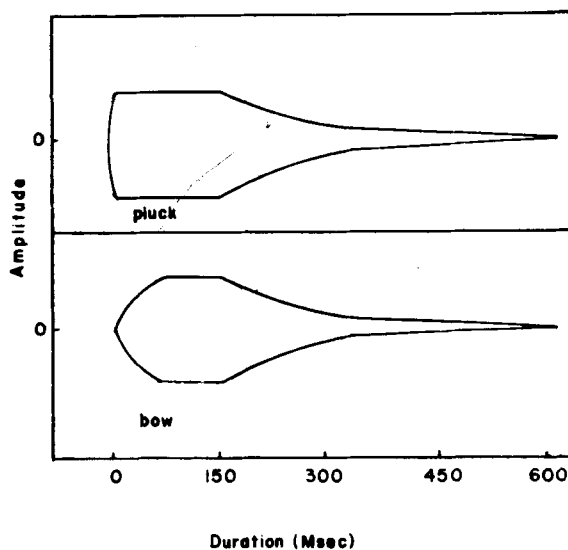


Figure 3. Amplitude envelopes of the "pluck" and "bow" adapting stimuli.

Table 1
Net Change in Number of Stimuli Identified as [b] Between First and Second Identification Tests

Adapting Stimulus	Subjects									Mean	p <
	L. B.	P. B.	B. B.	L. P.	E. B.	M. L.	O. N.	J. M.			
No adaptation	0	+3	-2	0	0	+2	+2	+1	+7.5	n.s.	
[ba]	-34	-16	-5	-4	-7	-14	-13	-12	-13.12	.005	
[wa]	+17	+11	+14	+16	+14	+13	+3	+18	+13.25	.0005	
[da]	-18	-5	-3	-1	-2	-11	-3	-3	-5.75	.025	
[ya]	0*	0*	+3	+6	-2*	+15	0*	+3	+3.12	n.s.	
[b] transitions alone	-7	+3	-2	-7	-2	0	0	-7	-2.75	.05	
[w] transitions alone	-7†	+6	+10	+14	+11	+10	+10	+14	+8.50	.005	
Pluck	-4	-2	-4	-3	-8	-6	-2	-7	-4.50	.0005	
Bow	-8	-8	+1	+2	+1	0	-3	+5	-1.25	n.s.	

Note—A positive number indicates an increase, from the first test to the second, in the number of stimuli identified as [b]; a negative number indicates a decrease.
*Did not consistently hear a continuant
†Heard [bla]

identified as [b]. In contrast, adapting with [ba] led to a large decrease in the number of [b] identification responses ($p < .005$),³ while adapting with [wa] produced an equally large increase in [b] responses ($p < .0005$).

The [da] was also an effective adapting stimulus, producing a decrease in the number of test stimuli identified as [b] ($p < .025$). On the other hand, no reliable identification shift was obtained using the stimulus [ja]. The four subjects (P.B., K.P., M.L., and J.M.) who reported hearing this consistently as a continuant did show a nonsignificant increase in [b] responses, but the other four subjects reported hearing the initial part of this stimulus variously as a stop and as a continuant (e.g., [dia], [dra], and [ja]),

and these subjects individually showed no identification shift.

Adapting with "[b] transitions alone produced a small decrease in [b] responses ($p < .05$), and an increase in [b] responses was obtained using "[w] transitions alone ($p < .005$). None of the subjects reported perceiving "[b] transitions alone as speech, but all of the subjects heard "[w] transitions alone" as speech (L.B. heard [bla], all others heard [w] or [wə]).

The two nonphonetic adapting stimuli produced mixed results. Adapting with the pluck sound led to a significant reduction in the number of [b] identifications ($p < .0005$); however, no reliable identification shift was obtained with the bow

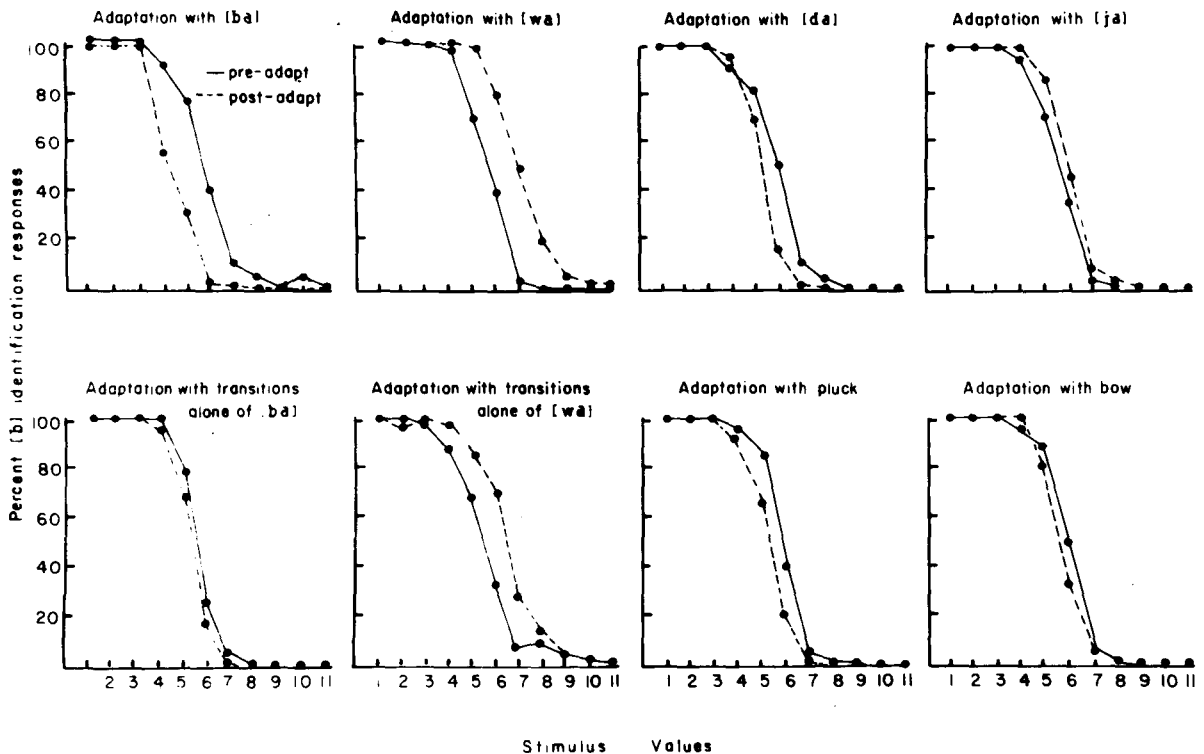


Figure 4. Group identification functions for each of the eight adaptation conditions.

adapting stimulus. The group identification functions for the eight adaptation conditions are shown in Figure 4.

DISCUSSION

The present results generally support the claim that feature detectors mediate the perception of stop and continuant sounds. In most instances, if the subjects perceived the adapting stimulus as a stop, they later identified fewer of the test stimuli as stops in comparison to the preadaptation test. On the other hand, if they perceived the adapting stimulus as a continuant, they typically produced a greater number of stop identifications. As we have seen, such a pattern of results may be explained by the hypothesis of Eimas and Corbit (1973) that there exist linguistic feature detectors which are reduced in sensitivity through selective adaptation.

A close correspondence exists between the results of this study and those of previous selective adaptation experiments that have dealt with the voicing and place dimensions. For example, identification shifts were obtained on the [ba-wa] test series using the adapting stimulus [da], indicating that some part of the adaptation effect occurs at a feature-specific, rather than a purely phoneme-specific, level of analysis. In addition, removal of the vowel portion of an adapting stimulus greatly reduced the amount of identification shift when the resulting stimulus was not perceived as speech-like (as in the case of "[b] transitions alone"). When the resulting stimulus was still heard as speech-like (as in the case of "[w] transitions alone"), the amount of identification shift remained relatively large. These findings are in accord with those of Diehl (1975) in suggesting that some component of the adaptation effect occurs at a level of phonetic feature detection. However, the fact that "[b] transitions alone" produced a small identification shift indicates that phonetic processing is probably not a necessary condition for an adaptation effect. Some part of the effect must occur at a level of acoustic analysis. [Additional evidence that some adaptation takes place at a site of acoustic feature detection has recently been found by Blumstein and Stevens (Note 1) and by Bailey and Haggard (Note 2).]

The results obtained using the pluck sound as the adapting stimulus are perhaps the strongest evidence available that phonetic processing is not required to produce reliable adaptation effects. Recall that this stimulus, a nonphonetic stop, led to a significant decrease in the number of stop identifications along the [ba-wa] test series. An acoustic explanation of this result is possible. As was mentioned, the stimuli of the [ba-wa] test series varied not only in the slope and duration of the formant transitions but also in their initial amplitude characteristics (see Figure 3). Since the pluck and bow sounds also differed in initial

amplitude characteristics (i.e., rise time), it is reasonable to assume that such acoustic characteristics were the basis of the adaptation effect. But it is also possible that the effect occurred at a more abstract level of feature detection. Perhaps the "stop detector" would be adapted regardless of how the stop quality of the adapting stimulus was acoustically realized. A test of this hypothesis could be made by using a [ba-wa] test series which did *not* vary in amplitude characteristics as a function of transition duration. If the pluck stimulus still produced a decrease in [b] identifications, then one could argue that the stop detector operates at a more abstract level than that of a simple acoustic analysis. Unfortunately, such an experiment has not yet been performed. Given the identification shift which the pluck sound produced, it is somewhat surprising that the bow sound failed to produce a similar shift in the opposite direction. The reason for this asymmetry in the results is not presently clear.

In summary, the overall findings of the present study point to the existence of feature detector mechanisms for the stop-continuant dimension, mechanisms whose sensitivity is reduced through selective adaptation.

REFERENCE NOTES

1. Blumstein, S. E., & Stevens, K. N. *Property detectors for bursts and transitions in speech perception*. Paper presented at the 89th meeting of the Acoustical Society of America, Austin, Texas, April 1975.
2. Bailey, P. J., & Haggard, M. P. *Cue-specific adaptation to speech sounds*. Paper presented at the 89th meeting of the Acoustical Society of America, Austin, Texas, April 1975.

REFERENCES

- ABRAMSON, A. S., & LISKER, L. Discriminability along the voicing continuum: Cross-language tests. In *Proceedings of the 6th International Congress of Phonetic Sciences, Prague, 1967*. Prague: Academia, 1970. Pp. 569-573.
- ADES, A. E. How phonetic is selective adaptation? Experiments on syllable position and vowel environment. *Perception & Psychophysics*, 1974, **16**, 61-66.
- CHOMSKY, N., & HALLE, M. *The sound pattern of English*. New York: Harper and Row, 1968.
- COOPER, W. E. Adaptation of phonetic feature analyzers for place of articulation. *Journal of the Acoustical Society of America*, 1974, **56**, 617-627.
- COOPER, W. E., & BLUMSTEIN, S. E. A "labial" feature analyzer in speech perception. *Perception & Psychophysics*, 1974, **15**, 591-600.
- CUTTING, J. E., & ROSNER, B. S. Categories and boundaries in speech and music. *Perception & Psychophysics*, 1974, **16**, 564-570.
- DIEHL, R. L. The effect of selective adaptation on the identification of speech sounds. *Perception & Psychophysics*, 1975, **17**, 48-52.
- EIMAS, P. D., COOPER, W. E., & CORBIT, J. D. Some properties of linguistic feature detectors. *Perception & Psychophysics*, 1973, **13**, 247-252.
- EIMAS, P. D., & CORBIT, J. D. Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, 1973, **4**, 99-109.

LIBERMAN, A. M., DELATRE, P. C., GERSTMAN, L. J., & COOPER, F. S. Tempo of frequency change as a cue for distinguishing classes of speech sounds. *Journal of Experimental Psychology*, 1956, 52, 127-137.

NOTES

1. In the ABX procedure, three stimuli are presented in succession. The first two are different from each other, while the third is identical to either the first or the second. The subject must decide which of the first two stimuli is most like the third.

2. Some of the subjects perceived the stimulus as having [r] color. Since [r] and [w] are both continuants, this did not affect the interpretation of the results in the present study.

3. In all instances, the statistical significance of identification shifts was determined by one-tailed t tests for correlated observations.

(Received for publication February 7, 1975;
revision received November 3, 1975.)