

The role of auditory feature detectors in the perception of speech

VIVIEN C. TARTTER and PETER D. EIMAS

W. S. Hunter Laboratory of Psychology, Brown University, Providence, Rhode Island 02912

Two adaptation experiments were conducted to determine some of the sufficient acoustic properties for excitation of the feature detectors underlying perception of voicing and place distinctions in speech. It was found that the greater the acoustic overlap between the adapting stimulus and the test continuum, the greater the adaptation effect. In addition, evidence was obtained that various parts of the speech signal (e.g., the second formant alone) are sufficient cues for adaptation of detectors. An outline of a model in which phonetic feature information is determined at an auditory level of analysis is presented.

In recent years, a number of researchers have assumed that underlying the perception of speech sounds are feature detectors that are especially tuned to the physical properties of the speech signal (e.g., Abbs & Sussman, 1971; Eimas & Corbit, 1973; Stevens, Note 1). Abbs and Sussman (1971) and Stevens (Note 1) have assumed that these detectors are auditory property analyzers, responsive to relatively low-level acoustic information, whereas Eimas and Corbit (1973) have assumed that the detector mechanisms underlying the perception of speech are linguistic in nature. That is to say, they have assumed that these mechanisms are part of the specialized speech processing system and that they have as their input the auditory transform of the acoustic information and as their output the equivalent of distinctive phonetic features (e.g., Chomsky & Halle, 1968; Jakobson, Fant, & Halle, 1951).

In the initial experimental studies supporting a detector model, Eimas and Corbit (1973) selectively adapted listeners to one or the other of the two values of voicing found in English and numerous other languages. Adaptation produced a shift in the locus of the phonetic boundary for a series of synthetic speech stimuli that varied along the complex dimension of voice onset time (VOT), variations in which, as Lisker and Abramson (1964) have shown, are effective in signaling the perceived voicing distinctions. The direction of the shift was always toward the adapting stimulus. Thus, after adaptation with a voiceless stop consonant, stimuli near the phonetic boundary, assigned to the voiceless category in the unadapted state, were now assigned to the voiced category.

This research was supported by Grant HD 05331 from the National Institute of Child Health and Human Development. The authors wish to express their appreciation to Dr. Franklin S. Cooper for generously making the facilities of the Haskins Laboratories available to us. The studies reported herein were completed as part of the requirements for the MA degree at Brown University by the first author under supervision of the second author.

Similarly, adaptation with a voiced stop caused a greater number of responses to be assigned to the voiceless category. Furthermore, adaptation effects were not specific to the adapting phonetic segment, but rather were feature specific. Adaptation with the voiced bilabial stop [b], for example, produced a boundary shift of nearly the same magnitude as adaptation with the voiced alveolar stop [d] regardless of whether the test series signaled the bilabial or alveolar stop consonants, [b,p] or [d,t], respectively.

In their explanation of the adaptation effect, Eimas and Corbit (1973) assumed that there exist two voicing or VOT detectors, each of which is differentially sensitive to a range of partially overlapping VOT values, with peak sensitivities at different values. Either or both of the detectors may respond to an incoming speech stimulus, depending on the particular value of VOT present in the signal. It was further assumed that the voiced-voiceless decision is determined by the detector with the greater response, and that the phonetic boundary lies at that value of VOT to which both detectors are equally responsive. Selective adaptation was hypothesized to fatigue the excited detector, resulting in a diminution in the strength of the output signal for all VOT values to which the detector was sensitive. Consequently, after adaptation, the unfatigued detector makes a relatively greater response to boundary VOT values than does the fatigued detector, which results in a shift in the phonetic boundary and ultimately in perception.¹ Moreover, since Eimas, Cooper, and Corbit (1973) found that adaptation with nonspeech stimuli, the formant transitions, alone, (chirps) which contained all of the essential voicing information, produced no shift in the phonetic boundary, they argued that these voicing detectors were in fact a part of the specialized speech processing system and not a component of the more general auditory system. [Ades (personal communication), however, reports that he has obtained slight, but reliable, boundary shifts after adaptation with the voiced chirp].

Cooper (1974a) tested the effects of adaptation on a synthetically produced continuum of speech sounds

varying in place of articulation, a phonetic feature for which, unlike voicing, there is no known relatively context-invariant acoustic information. Cooper found that adaptation with [b], [d], and [g] all produced boundary shifts in accord with a model that assumed three feature detectors, one for each of the three major values of place of articulation. Of particular interest were Cooper's findings that reliable shifts in the appropriate boundary could be produced by adaptation with real speech stimuli that varied in voicing ([p^hbae]) or in vowel quality ([bi]). Inasmuch as these stimuli shared relatively little acoustic information with the test series, Cooper concluded that the detectors for the place distinction might be responsive to quite abstract phonetic information. Similar conclusions are possible from a recent study (Cooper & Blumstein, 1974) that demonstrated reliable shifts in the [b-d] boundary after adaptation with [m] and [v].

The evidence favoring feature detectors at a phonetic level is essentially based on the fact that adapting stimuli which differ in the critical acoustic information from the test series still produce significant boundary effects. However, in the studies mentioned above, where this situation existed, the adaptation effects were always smaller than the effects obtained after adaptation with a member of the test series. Inasmuch as non-test-series adapting stimuli did share some acoustic information with the test stimuli, although less than that shared by adapting and test stimuli from the same continuum, the possibility remains that a major portion, if not all, of the adaptation effect has an auditory basis. In several studies (Cooper, 1974b; Diehl, 1975; Ganong & Liberman, personal communication), reliable adaptation effects were obtained after careful attempts to eliminate the sharing of critical acoustic information between the adapting stimulus and test series, thereby providing more convincing evidence for a phonetic level of processing. For example, Cooper (1974b) used two adapting stimuli, each with a VOT value of 25 msec but differing in the amount of formant transition after the onset of voicing. The identification series varied only in the absolute VOT (cf. Stevens & Klatt, 1974), and only the initial stimulus in the series had a perceptible duration of formant transition after the onset of voicing. Cooper found that the two adapting stimuli produced differential boundary shifts, and ascribed the results to the adaptation of an integrative phonetic detector that responded to the weighted combination of the two voicing cues. Both Diehl (1975) and Ganong and Liberman (personal communication) have demonstrated that adaptation with a pattern in which place distinctions were determined by burst information can produce shifts in the boundary of a test series in which the place information was determined solely by variations in the formant transitions. The burst cues were in fact absent in the identification stimuli.

Compelling evidence for the existence of detector systems at an auditory level comes from the studies by Ades (1974a, b) and Bailey (Note 2). Ades (1974a) found that adaptation with [bae] and [dae] failed to produce a change in the phonetic boundary of a series of synthetic stimuli that varied from [aeb] to [aed], just as adaptation with [aeb] and [aed] did not alter the phonetic boundary of a [bae-dae] test series. In addition, Ades (1974b) found that adaptation with the formant transitions of a [bae] speech stimulus, a nonspeech sound containing all the relevant place information, was sufficient to alter the phonetic boundary of a [bae-dae] test series. Furthermore, Bailey (Note 2) showed that in at least one instance some acoustic overlap between the adapting stimulus and the test series was necessary for a shift in the phonetic boundary for a place continuum.

It is apparent that the available evidence is indeterminate with respect to the locus of the adaptation effect. Indeed, it is possible, and most likely quite probable, that the adaptation effect has at least two levels, an auditory and a phonetic level. The present experiments were undertaken to provide further evidence for an auditory level, as well as to determine some of the acoustic features of the speech signal that are sufficient to produce an adaptation effect.

EXPERIMENT I

If auditory detectors underlying perception of the place distinction exist, it should be possible to obtain shifts in the phonetic boundary after adaptation with nonspeech-like components of the entire speech pattern, for example, the second formant alone. Indeed, Ades (1974b) has presented some evidence supporting this contention. This experiment further investigated the sufficiency of certain acoustic components of the speech signal to act as adapting stimuli for place detectors.

Method

Subjects. Six students at Brown University served as subjects. All were native speakers of English with no known hearing loss. They were paid for their participation.

Stimuli. The identification series consisted of 13 CV syllables produced by means of a parallel resonance synthesizer at the Haskins Laboratories. The stimuli differed from one another only in the starting frequencies and directions of the second- and third-formant transitions, a sufficient cue for perception of variations in place of articulation. The formant starting frequencies and steady-state values for the stimuli are shown in Table 1. The stimuli were each 300 msec long, and consisted of 35 msec of prevoicing, 45 msec of transitions, and 220 msec of steady state.

The adapting stimuli were the first [bae] stimulus described in Table 1 and five variations of this speech pattern: a chirp consisting of the first 45 msec of the three formant transitions, the entire second and third formants together, the second formant alone, the third formant alone, and a stimulus consisting of the second and third formants of [bae] and a first formant with a transition that fell from 894 Hz to a steady-state value of 743 Hz ("bae"). The last stimulus is perceived as speech-like but without consonantal

quality. Note that the third formant of [bae] is identical to the third formant of [gae] in Stimulus 11.

Procedure. All subjects were initially presented with three different randomized sequences of the syllables in order to obtain baseline values of identification in the unadapted state. Each sequence consisted of 130 stimuli, with 10 presentations of each syllable. The interstimulus interval on all identification tapes was 3 sec. Subjects were instructed to classify the stimuli as b, d, or g, and to record their answers by writing the appropriate initial consonant on the answer blanks provided.

Following the initial identification session, subjects participated in a series of 1½-h identification sessions. Subjects were tested in two groups of three, with the order of the adaptation sessions different for each group. Each session was separated by at least 24 h. Each adaptation session began with the presentation of an identification tape as a warm-up. Subjects then listened to 2 min (120 presentations) of the adapting stimulus, with an interstimulus interval of 700 msec. There were then 26 adaptation trials, each consisting of 1 min of adaptation followed by five identification stimuli, randomly selected. This procedure took approximately 45 min. After a short break, listeners were presented with another 2 min of initial adaptation and 26 adaptation trials, resulting in a total of 20 observations for each stimulus.

The identification tapes in all sessions were played on a Sony TC-360 recorder through an Ampex AG-500 recorder, and the adaptation stimuli were presented by means of a tape loop on the Ampex recorder. The listeners heard all stimuli through Koss 600AA headsets. The adapting and identification stimuli were presented at approximately equal intensities.

Results and Discussion

Table 2 shows the mean boundary loci for each adaptation session, as well as the average unadapted boundaries. Boundary loci were computed using a least mean squares analysis on z scores derived from the percentages of identification responses assigned to each stimulus. Average unadapted boundary loci were computed by taking the mean of the boundary values found for the initial identification session and for the identification series preceding each adaptation session.

A simple analysis of variance performed on the [b-d] boundary values for the average unadapted condition and each of the adapting stimuli showed a significant overall adaptation effect $[F(5,30) = 15.95, p < .001]$. Individual comparisons revealed that adaptation with the speech syllable [bae] significantly differed from adaptation with any of the nonspeech components, that the shift magnitude obtained after adaptation with the second and third formant together differed significantly from the shift

Table 1
The Formant (F_1, F_2, F_3) Frequency Values for the [ae] Continuum

Stimulus	Starting Transition Values (Hz) ^a	
	F_2	F_3
Steady State Values (Hz): $F_1 = 743^*$ $F_2 = 1620$ $F_3 = 2862$		
1	1312	2348
2	1386	2525
3	1465	2694
4	1541	2862
5	1620	3026
6	1695	3195
7	1772	3026
8	1845	2862
9	1920	2694
10	1996	2525
11	2078	2348
12	2156	2180
13	2234	2018

^aThe starting F_1 value was 150 Hz for all stimuli.

magnitude obtained after adaptation with the third formant alone and after adaptation with "bae," and that each adaptation condition differed significantly from the unadapted condition.

A similar analysis of variance, computed on the [d-g] boundary value, also showed a significant adaptation effect $[F(5,30) = 3.08, p < .05]$. Individual comparisons showed this effect to have resulted entirely from a significant shift of the [d-g] boundary towards [g] following adaptation with the third formant alone. Since [bae] and [gae] generally have similar third formant transitions (and, in fact, [gae] Stimulus 11 has the same third formant as Stimulus 1), it is really not surprising that such a shift was obtained.

The results of the present experiment lend support to the hypothesis that there exist lower order auditory detectors which are responsive to particular acoustic features of the speech signal. That adaptation with portions of the speech signal yielded smaller boundary shifts than did adaptation with the entire speech pattern can be readily explained by assuming that the entire speech signal adapted a greater number of auditory detectors underlying the analysis of place information than did any of the component adapting stimuli.

Table 2
Mean Boundary Loci for Each Adaptation Condition (Experiment I)

Boundary		Adapting Stimuli						
		bae	F_2	F_3	F_2	F_3	[bae]-Chirp	"[bae]"
[b-d]	Mean	2.87 (4.55)*	3.43		3.68	4.06	3.76	3.89
	Difference	1.68	1.12		.87	.49	.79	.66
	p	.005	.005		.025	.05	.001	.05
	Percent of [bae] Shift	100	67		52	29	47	39
[d-g]	Mean	8.64 (8.92)*	8.97		8.70	9.34	8.86	9.00
	Difference	.32	-.05		.22	-.42	.06	-.08
	p	.50	.90		.90	.025	.20	.80

*The unadapted boundary values are shown in parentheses.

However, this explanation may not fully account for the fact that the addition of a first formant to the second and third formants (the entire [bae] stimulus) and the addition of steady-state information to the [bae] chirp resulted in a large increment in the magnitude of the boundary shift, despite the fact that neither of these acoustic parameters carried information with respect to place of articulation. We will return to this finding in the general discussion.

EXPERIMENT II

Given the results of the first experiment, it becomes more difficult to interpret the findings of Eimas et al. (1973) that nonspeech chirps containing the essential voicing information failed to alter the phonetic boundary of a speech series that varied in voice onset time. That is, while possible, it is certainly unlikely that place detectors have auditory components, whereas voicing detectors are solely phonetic in nature. In an attempt to resolve this apparent discrepancy, Experiment II investigated the effects of adaptation on voicing detectors, with a greater variety of the nonspeech, acoustic components of voicing.

Method

Subjects. An additional six students from the same population as those in the first experiment served as subjects.

Stimuli. Two identification series were used in this experiment, a place series to measure baseline adaptation effects and a voicing series. Each consisted of nine stimuli. For the place series [bae-dae], the initial 35 msec of prevoicing was removed from the first eight stimuli described above. An additional [bae] stimulus was synthesized, with a second formant starting frequency of 1,232 Hz and a third formant starting frequency of 2,180 Hz. This stimulus would have preceded No. 1 on Table 1. All stimuli had a VOT value of 5 msec. The nine stimuli for the [bae-pae] continuum were constructed by taking the first stimulus on the [bae-dae] continuum and varying the VOT in 5-msec steps from 5 to 45 msec. The length of each stimulus for both continua was 265 msec.

Adapting stimuli for the [bae-dae] series consisted of [bae] in its entirety and the second and third formants of [bae]. Adapting stimuli for the [bae-pae] series were [bae] in its entirety, the [bae] chirp, the second and third formants of [bae], and the second and third formants of [pae].

Procedure. Identification tapes for the nine-step continua contained 15 exemplars of each syllable in random order. There were 27 adaptation trials after the initial 2-min presentation of the adapting stimulus, resulting in 15 observations for each stimulus of the nine-step continua, for each adaptation condition. The testing procedure was identical to that described in Experiment I.

There were three baseline sessions to ensure that the new listeners performed in a similar manner to the listeners of Experiment I. These sessions provided estimations of the phonetic boundary for the [bae-dae] series (1) without adaptation, (2) after adaptation with [bae], and (3) after adaptation with the second and third formants of [bae]. There were five additional sessions designed to obtain estimates of the [bae-pae] boundary without adaptation and after adaptation with each of the four adapting stimuli described above.

Results and Discussion

The mean boundary data for the [bae-dae] series are displayed in Table 3. An analysis of variance of

Table 3
Mean Boundary Loci for Each Adapting
Condition: Experiment II

Stimulus	Adapting Stimuli			
	Continuum	Unadapted	[bae]	[bae] F ₂ F ₃
[bae-dae]	Mean	5.11	3.60	4.42
	Difference		1.51	.70
	p		.001	.025

Stimulus		Adapting Stimuli				
		Unadapted	[bae]	[bae]- chirp	[bae] F ₂ F ₃	[pae] F ₂ F ₃
[bae-pae]	Mean	5.45	4.84	5.28	5.16	5.94
	Difference		.60	.17	.29	.50
	p		.025	.40	.10	.05

these data revealed an overall effect of adaptation [$F(2,10) = 13.0, p < .01$]. Individual comparisons revealed that [bae] produced a significantly greater shift than did the second and third formants of [bae], but that both were effective adapting stimuli, as was found for different subjects in Experiment I.

The mean boundary loci for the [bae-pae] series are also found in Table 3. A similar analysis showed a reliable effect of adapting stimulus [$F(4,20) = 5.1, p < .01$]. Individual comparisons revealed that adaptation with [bae] shifted the boundary toward the [b] category ($p < .05$) and that adaptation with the second and third formants of [pae] shifted the boundary toward the [p] category ($p < .05$). Neither the [bae] chirp nor the second and third formants of [bae] produced reliable adaptation effects.

Voice onset time is a complex dimension, consisting of several potential cues for the voiced-voiceless distinction: (1) the duration between the onset of acoustic energy and the onset of voicing, (2) the relative onset time of the first formant, and (3) the presence or absence of perceptible formant transitions after the onset of voicing (Stevens & Klatt, 1974). The second and third formants of [bae] would have the first and third cues to signal a voiced stop, but, if the absence of a first formant was a cue for voicelessness, then this adapting stimulus would have contained conflicting cues. On the basis of this reasoning, the second and third formants of [pae] provide all three cues for a voiceless stop, and thus it is not surprising that they produced a reliable adaptation effect, whereas the second and third formants of [bae] produced a smaller, nonsignificant adaptation effect.

The ineffectiveness of the chirp as an adapting stimulus for the voiced-voiceless distinction suggests that the analysis of timing relations necessary for voicing decisions requires relatively longer stimulus durations than does analysis of the cues for place distinctions. Nevertheless, given the reliable adaptation effect obtained with the second and third formants of [pae], Eimas et al. (1973) were incorrect in concluding that the analysis of voicing was solely a function of speech-specific mechanisms.

GENERAL DISCUSSION

The results of the present experiments clearly indicate that the detector systems that underlie the perception of phonetic contrasts based on voicing and place of articulation have an auditory component. That is to say, at least some, if not all, of the analyzing channels are part of the general auditory processing system, and, as such, do not require phonetic information for their activation. Consider first the major findings with respect to place of articulation: (1) a speech pattern, [bae] in this instance, was a more effective adapting stimulus than any of the several isolated components of the entire pattern. (2) Adaptation with the second and third formants of [bae] together produced a larger adaptation effect than either the second or third formant alone. (3) There are multiple cues for bilabial feature information, and presumably this is true for the remaining values of the place feature. (4) The initial formant transitions (the [b] chirp) produced a small, but highly reliable, alteration of the phonetic boundary (cf. Ades, 1974b). And (5) two different values of place of articulation may have effective cues in common, as evidenced by the reliable shifts in the [b-d] and [d-g] boundaries after adaptation with a common third formant.

If a model of auditory detectors similar to that proposed by Abbs and Sussman (1971) is assumed and is combined with the assumptions of the model proffered by Eimas and Corbit (1973) for phonetic detectors, most of the present results are readily explicable. Such a model must first assume that there are multiple auditory analyzers for place of articulation information, each of which is sensitive to a restricted range of acoustic information. Thus, there might exist auditory analyzers for the detection of burst information, the absolute frequencies and directions of the second- and third-formant transitions, and perhaps, the relation between the directions of the formant transitions (cf. Stevens, Note 1). If these analyzers are arranged in groups of opposing channels, as suggested by Eimas and Corbit (1973) for phonetic analyzers, then adaptation of one channel will result in a greater relative contribution of the opposing channel and a reliable boundary shift. For example, if there is a detector pair that is maximally sensitive to a sharply rising and a slightly falling second-formant transition (respectively, the approximately modal stimuli for [b] and [d] in the context of the vowel [ae]), then a transition of intermediate value would be expected to represent a boundary value in the unadapted state. However, after adaptation with a rising transition, the same intermediate value would be expected to produce a relatively weaker response by the fatigued rising transition detector, and hence would be more likely to be assigned to the category mediated by the opposing

detector (i.e., the detector for slightly falling second-formant transitions).

This model would predict that the greater the number of auditory detectors that are adapted, the greater the shift in the locus of the phonetic boundary. And, indeed, the second and third formant together produce a larger effect than either formant alone. Moreover, the second formant alone appears to be a more effective cue for place than the third formant alone, a conclusion reached by Liberman, Cooper, Shankweiler, and Studdert-Kennedy (1967) from quite different data. What is difficult for an auditory detector model to accommodate, as mentioned above, is the much larger adaptation effect obtained by adapting with the entire speech pattern. The addition of a first formant, or steady-state information, to adapting stimuli that contained all the relevant place of articulation information produced a large increment in the magnitude of the boundary shift in both experiments. Inasmuch as neither of these additions contained any information regarding place of articulation, it is difficult, but not impossible, to explain how they could have had an effect at the auditory level.³ If, however, we were to assume a second level of detectors, phonetic in nature, and if these detectors were activated only when place information was presented in conjunction with manner and vowel information, then the greater adaptation effect with the entire [bae] pattern is explicable (cf. Darwin, in press).

One other finding of Experiment I deserves comment. The fact that an isolated third formant shifted both the [b-d] and [d-g] boundaries indicates that the detector for this information must be responding not only to the direction of the third-formant transition, but also to its absolute frequency. Since the third-formant information was presented in isolation, there was no way of tagging the incoming energy as belonging to a particular formant. Furthermore, the result cannot be attributed to the adaptation of a rising formant detector, since adaptation with the second formant alone, which also had a rising transition, did not shift the [d-g] boundary. The effectiveness of the third formant as an adapting stimulus for velar stops also precludes the possibility that a "spreading detector" (Stevens, Note 1) is responsible for the perception of velar stops, at least at the lowest auditory level. Presumably, such a detector would require the simultaneous presentation of *both* a falling second-formant transition and a rising third-formant transition in order to be activated. That a detector which integrates transition information may exist at higher levels is certainly possible and cannot be excluded by the present data.

The second experiment revealed that detectors underlying the perception of voicing can likewise be activated by nonspeech components of the speech

pattern. It is of particular interest that the detectors for voicing information at the auditory level require more sustained information than do comparable detectors for place. In addition, the data of Experiment II confirm the findings of others (e.g., Cooper, 1974b; Stevens & Klatt, 1974) that the VOT dimension is complex and consists of a number of dimensions, which individually may be sufficient for the activation of voicing detectors (cf. Cooper, 1974b).

In summary, these experiments clearly demonstrate an auditory level of analysis for the information that determines place and voicing decisions. Moreover, the mechanisms that analyze these forms of acoustic information are, in all likelihood, feature detectors that serve both the perception of speech and other complex sounds, that is, they are a part of our general auditory processing system. That there exists another, higher level of detectors that serves to integrate the lower order auditory information, having as its output phonetic feature information, is strongly suggested by a number of studies (Cooper, 1974b; Cooper & Blumstein, 1974; Diehl, 1975; Ganong & Liberman, personal communication). However, inasmuch as alternative explanations are often possible, any final statement concerning the existence of phonetic feature detectors must await additional and more precise research.

REFERENCE NOTES

1. Stevens, K. N. *The potential role of property detectors in the perception of consonants*. Paper presented at Symposium on Auditory Analysis and Perception of Speech, Leningrad, USSR, 1973.
2. Bailey, P. Perceptual adaptation for acoustical features in speech. *Speech perception: Report on Research in Progress in the Department of Psychology*. Belfast: The Queen's University of Belfast, 1973, 2.2, 29-34.
3. Sawusch, J. R., & Pisoni, D. B. *Category boundaries in speech and nonspeech sounds*. Paper presented at the 86th meeting of the Acoustical Society of America, Los Angeles, 1973.
4. Sawusch, J. R., Pisoni, D. B., & Cutting, J. C. *Category boundaries for linguistic and nonlinguistic dimensions of the same stimuli*. Paper presented at the 87th meeting of the Acoustical Society of America, New York, 1974.

REFERENCES

- ABBS, J. H., & SUSSMAN, H. M. Neurophysiological feature detectors and speech perception: A discussion of theoretical implications. *Journal of Speech and Hearing Research*, 1971, 14, 23-36.
- ADES, A. E. How phonetic is selective adaptation? Experiments on syllable position and vowel environment. *Perception & Psychophysics*, 1974, 16, 61-67. (a)
- ADES, A. E. Bilateral component in speech perception? *Journal of the Acoustical Society of America*, 1974, 56, 610-617. (b)

- 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, 617-628. (a)
- COOPER, W. E. Selective adaptation for acoustic cues of voicing in initial stops. *Journal of Phonetics*, 1974, 2, 303-313. (b)
- COOPER, W. E. Selective adaptation to speech. In F. Restle, R. M. Shiffrin, N. J. Castellan, H. Landman, and D. B. Pisoni (Eds.), *Cognitive theory*. Potomac, Md: Erlbaum (in press).
- COOPER, W. E., & BLUMSTEIN, S. E. A labial feature analyzer in speech perception. *Perception & Psychophysics*, 1974, 15, 591-600.
- DARWIN, C. J. The perception of speech. In M. Freedman and E. C. Carterette (Eds.), *Handbook of perception*. New York: Academic Press (in press).
- DIEHL, R. H. 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.
- JAKOBSON, R., FANT, G. M., & HALLE, M. *Preliminaries to speech analysis*. Cambridge, Mass: M.I.T. Press, 1951.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-461.
- LISKER, L., & ABRAMSON, A. S. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 1964, 20, 384-422.
- STEVENS, K. N., & KLATT, D. H. Role of formant transitions in the voiced-voiceless distinction for stops. *Journal of the Acoustical Society of America*, 1974, 55, 653-659.

NOTES

1. For arguments and evidence that the adaptation effect occurs at a sensory level and not at the level of response decisions, the reader is referred to Cooper (in press), Eimas and Corbit (1972), Sawusch and Pisoni (Note 3), and Sawusch, Pisoni, and Cutting (Note 4).
2. The second and third formant frequencies are quite close for some of the last stimuli, and for the last stimulus, in fact, overlap. Spectrograms of these stimuli suggest that the overlap may have resulted in a burst-like segment. Subjects had no difficulty identifying these stimuli as [gae]. Indeed, for all subjects, the percent identified as [gae] for Stimuli 11, 12, and 13 was very near 100% in the unadapted state, as well as after adaptation with each of the adapting stimuli.
3. One possible explanation at an auditory level of the large boundary shift produced by adaptation with the speech stimulus is that there are a number of auditory detectors sensitive to combinations of acoustic information that involve first formant input, e.g., a detector sensitive to the direction of first- and second-formant transitions or sensitive to transition information in combination with steady-state information. With the first formants or steady-state portion absent, these analyzers would not be fatigued, resulting in a smaller shift. It is important to note, however, that the greater adaptation effect with the entire speech stimulus cannot be attributed to a first formant detector per se or to detectors for steady-state information in isolation inasmuch as all of the stimuli shared the same first formant and steady-state information. As a consequence, adaptation of these detectors could not produce a shift in the locus of the phonetic boundary.

(Received for publication June 30, 1975;
revision received August 1, 1975.)