Response organization in selective adaptation to speech sounds

JAMES R. SAWUSCH and DAVID B. PISONI
Indiana University, Bloomington, Indiana 47401

Previous experiments in speech perception using the selective adaptation procedure have found a shift in the locus of the category boundary for a series of speech stimuli following repeated exposure to an adapting syllable. The locus of the boundary moves toward the category of the adapting syllable. Most investigators have interpreted these findings in terms of feature detector models in which specific detectors are reduced in sensitivity through repeated adaptation. The present experiment was conducted to determine whether the adaptation results might be due to changes in response organization as a consequence of the labeling instructions presented to subjects in selective adaptation experiments. A perceptually ambiguous speech stimulus was selected from the middle of a [bi]-[di] test series and used as an adaptor under two different sets of instructions. One group of subjects was told that the adapting stimulus was the syllable [bi], while another group was told that the stimulus was the syllable [di]. The acoustically ambiguous adaptor failed to produce a shift in the locus of the category boundary in the direction predicted on the basis of the labeling instructions presented to subjects. These results indicate that the acoustic attributes and perceived quality of the adapting stimulus determine the direction and magnitude of the adaptation effects rather than the labels provided by the experimenter.

Using the selective adaptation paradigm, a number of recent studies in speech perception have reported evidence for the existence of feature detectors in speech processing. This technique, which was first employed by Eimas and Corbit (1973), has been used to investigate the perceptual changes that result from the repeated exposure to a speech stimulus. During a selective adaptation experiment, a subject listens to repeated presentations of a speech or speech-like stimulus. Typically, there are about 100 presentations of the stimulus in approximately 1 min. The subject is then required to identify one or more syllables from a speech syllable series. This procedure of adaptation and testing is repeated until a sufficient number of responses to each speech stimulus in the test series has been collected. The typical result obtained is that the locus of the adapted identification function shows a shift relative to the locus of the preadapted boundary. The shift is toward the phonetic category from which the adapting stimulus was selected. Results of this sort have been found for a number of phonetic distinctions involving place of articulation, manner, and voicing in consonants.

In most of the previous selective adaptation studies, subjects were explicitly informed about the identity of the adapting stimulus. It is possible that the specific instructions used in these selective

adaptation experiments might induce changes in response organization in subjects that produce the shifts in their identification functions following adaptation. Thus, after a sequence of adaptation trials, a subject may be more likely to favor one response category than another simply as a consequence of adopting a particular response strategy. An account of the selective adaptation results based on differences in response organization, however, would be incompatible with the feature detector models that have recently been proposed by a number of investigators (Cooper & Nager, 1975; Eimas & Corbit, 1973; Pisoni & Tash, 1975; Tartter & Eimas, 1975). A common feature of these models is that they all assume that the shifts in the phonetic boundaries found in selective adaptation are due either to fatigue of detector mechanisms sensitive to certain attributes of the adapting stimuli or to some slight returning of the relevant detectors. The effects of selective adaptation have been assumed to be primarily sensory in nature, occurring prior to the response process.

Relatively little is known about the degree to which subject control processes enter into the selective adaptation effects. Some evidence bearing on the role of labeling has been reported by Ades (1974) and Diehl (1975). Ades found a fairly high correlation (r = 0.79) between the magnitude of the boundary shift and the ratings of goodness of the adapting stimuli. In his experiment, the more B-like or D-like the adapting stimulus, the greater the adapting effect on a [bae]-[dae] continuum. In post hoc analyses, Diehl (1975) found a somewhat similar result. Diehl

used a transition cued [tɛ] syllable as an adaptor and tested subjects on a [bɛ]-[dɛ] place series. Four of the original six subjects reported that the adapting syllable [tɛ] sounded like a [pɛ] and showed a reliable shift in their identification function towards [bɛ]. Based on these post hoc findings, Diehl argued that the subject's identification of the adaptor determines the direction of the shift in the category boundary. Unfortunately, Diehl did not report whether or not subjects in his study were told the identity of the adapting syllable before exposure to it.

In a series of category judgment experiments, we have studied the possibility that response bias enters into the identification task by attempting to introduce changes in the subject's adaptation level (Sawusch & Pisoni, Note 1; Sawusch, Pisoni, & Cutting, Note 2). We found that differences in the probability of occurrence of stimuli produced category boundary shifts for nonspeech stimuli and nonlinguistic dimensions of speech stimuli but failed to produce shifts in the locus of the phonetic category boundary for several stop-vowel syllable series. In these experiments, stimuli such as tones varying in intensity or CV syllables varying along the place of articulation dimension or voicing dimension were presented to subjects for absolute identification under two conditions. In one condition, each stimulus occurred equally often, while in the other condition an end-point stimulus occurred with a greater probability than any of the other stimuli in the test continuum. The unbalanced probability paradigm has previously been used to study the judgment of visual brightness (Helson, 1964) and weight (Parducci, 1963, 1975) as well as a number of auditory dimensions (Cuddy, Pinn, & Simons, 1973; Pollack & Boynton, 1963). The results of our experiments showed that the category boundary for a nonlinguistic dimension shifted toward the category of the more frequently occurring stimulus when compared to the equal probability condition. These results are in agreement with those found by Helson (1964) and Parducci (1963, 1975) using similar paradigms. However, the place and voicing CV syllable series failed to show the analogous shifts in the category boundary. We have interpreted these findings as support for the idea that the acoustic attributes for certain phonetic features in CV syllables are identified on an absolute basis independently of the context of the experiment. The criterion for a decision about the presence or absence of an attribute appears to be based on some internally represented standard rather than on the range and spacing of stimuli presented to the subject during the course of an experiment.

If we consider the selective adaptation paradigm as an absolute identification task in which one stimulus (i.e., the adaptor) occurs more frequently than other stimuli, then our results with the un-

balanced probability paradigm would appear, at first glance, to rule out any simple response bias account of the selective adaptation effects. Indeed, some support for this position can be found in a recent paper by Cooper, Ebert & Cole (1976) who have reported differences in sensitivity to adjacent stimuli after selective adaptation with the use of a magnitude estimation paradigm.[1] However, the question of labeling per se and the possibility that some response bias might be introduced by the specific labels given to subjects has not, to our knowledge, been studied in any of the previous selective adaptation experiments. The present experiment was, therefore, carried out in order to provide a more direct experimental test of whether the labeling instructions given to subjects would have any systematic effects on selective adaptation.[2]

In previous work, the adapting syllable has typically been a good exemplar of the category being adapted and was selected from the end points of the test series. The instructions to the subjects about the identity of the adapting syllable were generally consistent with their own perception of the stimulus. In order to manipulate the instructions to the subject, it is necessary to obtain an adapting syllable that would be compatible with different sets of labeling instructions. Accordingly, in the present study, we selected the middle or boundary stimulus from a synthetic speech series, since this stimulus is identified ambiguously. Based on previous work, subjects assign the middle stimulus to either of two opposing phonetic categories about equally often. The feature detector models of selective adaptation would predict that such an adaptor would have little, if any, effect on the phonetic category boundary, since this stimulus contains sets of attributes for which both opponent detectors would be equally responsive. Moreover, these models assume that adaptation effects are primarily sensory in nature and are based on some interaction between the stimulus properties of the adaptor and test series. On the other hand, any account based on changes in response organization would predict that the instructions to the subjects about the identity of the adapting stimulus should play an important role in adaptation. Thus, telling one group of subjects that the repeated syllable is from one phonetic category or another should produce a contrast effect and therefore reduce the number of identification responses for the adapted category. This would be reflected, in turn, by a shift in the locus of the category boundary in identification.

METHOD

Subjects

Twenty-four undergraduate students at Indiana University participated as part of a course requirement. All were right-handed native speakers of English with no known history of a

hearing or speech disorder. The subjects were divided into four groups of six subjects each.

Stimuli

The stimuli were three formant synthetic CV syllables, originally prepared on the parallel resonance synthesizer at Haskins Laboratories. These stimuli were recorded on audio tape and later digitized and stored on disk memory under the control of a PDP-11 computer. The test stimuli consisted of one series of nine CV syllables that ranged perceptually on the feature of place of articulation from [bi] to [di]. These stimuli varied in the starting frequencies for the second and third formant transitions from 1,465 Hz (F2) and 2,348 Hz (F3) for the [bi] end to 2,078 Hz (F2) and 3,690 Hz (F3) for the [di] end in eight approximately equal steps. The duration of the formant transitions was 50 msec, followed by a 250-msec steady state vowel [i]. The vowel had formant center frequencies of 287, 2,307, and 3,026 Hz for the first through third formants, respectively. Formant amplitudes were predicted from the acoustic theory of speech productions.

Procedure

All experimental events were controlled by a PDP-11 computer. The digitized waveforms of the test stimuli were reconverted to analog form via a 12-bit D-A converter and presented binaurally through Telephonics (TDH-39) matched and calibrated headphones to the subjects. The stimuli were presented at a comfortable listening level (83 dB SPL for the steady state calibration vowel [i]).

The experiment consisted of two 1-h sessions conducted on consecutive days. The subjects were run in small groups. At the beginning of each day, all subjects listened to two 90-trial identification sequences. Each sequence consisted of 10 presentations of each of the nine syllables in random order. By the end of the experiment, the subjects provided 40 unadapted responses to each syllable. The subjects were instructed that they would hear synthetic speech sounds approximating the syllables [bi] and [di] and were to respond to these syllables using a 4-point rating scale. The subjects were told to use the response label 1 if they were positive that the syllable they heard was a "bi," the label 2 if it was possible that they had heard a "bi," 3 if it was possible that they had heard a "di," and the response 4 if they were positive they had heard a "di." A copy of the response scale was present in front of each subject at all times. The subjects entered their responses by pushing the appropriate button of a four-button response box.

Immediately following the identification sequences, an adaptation sequence was presented. Three different syllables were used as adaptors: the [bi] end point, the [di] end point, and the middle (fifth) syllable of the nine-syllable test series. The first group of subjects received the [bi] end point as the adaptor and were told that the repeated syllable was a [bi]. The second group was also told that their repeated syllable was a [bi], even though the middle syllable (stimulus 5) was actually used. The third group of subjects listened to the repeated [di] end-point syllable and were told that it was a [di]. The fourth group was presented with the middle stimulus repeatedly, but again they were told that it was the syllable [di].

The adapting stimulus was presented for 1 min (100 repetitions with a 300-msec interstimulus interval). After each minute of adaptation, the nine test syllables were presented in random order for identification by subjects using the rating scale. Nine of these adaptation trials were run in the adaptation sequence on each day. Thus, by the end of the experiment, each subject provided 18 adapted responses to each of the nine stimuli.

RESULTS AND DISCUSSION

The data from the first identification test sequence

on the first day were treated as a practice run for subjects to accustom themselves to the rating response scale. These data were excluded from further analysis for all subjects.

Identification functions for each subject were obtained by collapsing the 4-point rating scale into a 2-category scale. Responses 1 and 2 were treated as "bi" responses and 3 and 4 were treated as "di" responses. The two groups that received end-point syllables as adaptors showed results similar to previous adaptation studies. The data for both of these groups, before and after adaptation, are shown in the upper half of Figure 1 as percent of "bi" responses. The phonetic boundary was determined by a computer program that located the 50% point along the stimulus scale by linear interpolation. The shifts in the phonetic category boundaries for both groups were significant [t(5) = 4.72, p < .01 for the [bi] adapted group; t(5) = 11.04, p < .002 for the [di] adapted group using two-tailed, correlated t tests].

Identification functions before and after selective adaptation with the middle ambiguous syllable are shown in the lower panel of Figure 1. Neither group showed any consistent trend following adaptation [t(5) = −2.23, p > .05 for the [bi] instructions group; t(5) = −0.18, p > .8 for the [di] instructions group]. It should be noted that although there appears to be a small overall shift for the group that received the [bi] instructions (Figure 1, lower left), this shift is in the opposite direction from that anticipated solely on the basis of the instructions presented to these subjects.

The rating response data for the two groups with [bi] adaptor instructions are shown in Figure 2. The left-hand panel shows the results of [bi] instructions with the [bi] end-point stimulus (Stimulus 1) as the
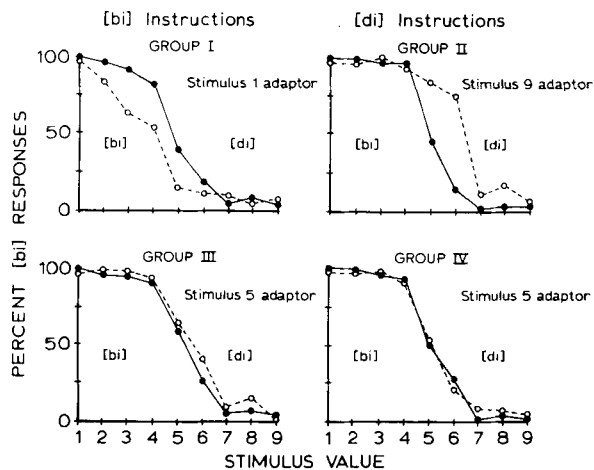

Figure 1. Unadapted (solid circles) and adapted (open circles) identification functions for the four groups on the [bi]-[di] series. The four rating responses were collapsed into two category responses.

[bi] Instructions

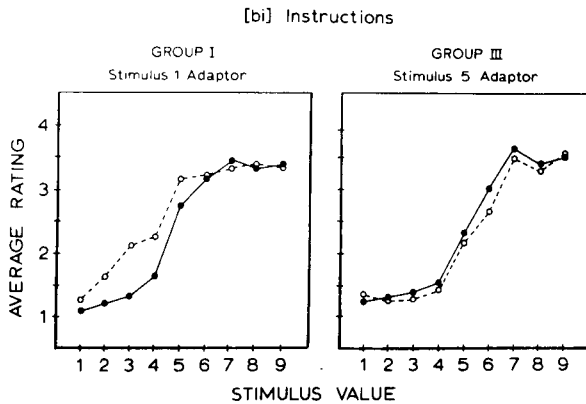GROUP I
Stimulus 1 Adaptor

GROUP III
Stimulus 5 Adaptor

Figure 2. Average unadapted (solid circles) and adapted (open circles) rating functions for the two [bi] instructions groups.

adaptor. The shifts in the rating response found within the [bi] category as well as at the category boundary are similar to those found previously by Sawusch (1976) and indicate that selective adaptation affects the whole range for which a detector is sensitive and not simply boundary values. The rating response shifts for the second through the fifth stimuli were all significant. In contrast, for the [bi] instructions group that received Stimulus 5, the post-adaptation rating responses were not significantly different from the base-line responses. The rating data for this group are shown in the right-hand panel of Figure 2.

The rating data for the two [di] instructions groups are presented in Figure 3 and are similar to the [bi] groups described previously. Again, the [di] end-point adaptor (Stimulus 9) had a significant effect within the [di] category as well as at the category boundary. These results are also consistent with those found previously by Sawusch (1976). For the Stimulus 5 adaptor with [di] instructions, no significant shift was found in the rating response for any of the stimuli in the test series.

The findings obtained in the present study support the idea that selective adaptation is not due to changes in response organization brought about by labeling instructions to subjects. Rather, adaptation seems to be primarily related to the acoustic properties of the adaptor and test series and the perceived identity of the adaptor. The absence of any overall adaptation effect with a perceptually ambiguous adaptor under different labeling conditions suggests that selective adaptation effects cannot be easily manipulated by processes typically thought to be under the control of the subject. Some additional support for our conclusions can be found in a recent paper by Blumstein and Stevens (Note 3). They reported that a stimulus containing both formant transition cues and burst cues to place of articulation produced a larger adaptation effect than a stimulus with only transitions as cues to place. This effect was obtained

even when the test continuum contained variations in only the transitions as cues to place. More importantly, however, they also produced adaptors that contained conflicting cues to place and found that although subjects identified the adaptor as a particular phonetic segment, the presence of the conflicting cue influenced the magnitude and direction of the adaptation effect. The conflicting cue stimuli sometimes produced adaptation effects in the direction opposite to that expected on the basis of the phonetic label assigned to the stimulus by the subject. Thus, the results of the present study as well as the findings of Blumstein and Stevens indicate that selective adaptation effects are primarily due to changes that occur during the relatively early stages of perceptual analysis prior to response organization.

On the basis of these results, it seems reasonable to conclude that the instructions provided to subjects regarding the identity of the adapting stimulus do not seem to play an important role in selective adaptation. However, further examination of the individual subject data for the two groups which received the middle (Stimulus 5) adaptor showed a small, but consistent, trend that is worthy of some brief discussion. As mentioned earlier, the averaged data did not show an overall shift in the direction expected on the basis of the instructions. However, most subjects failed to place their category boundary at precisely Stimulus 5. For a number of subjects, Stimulus 5 was either a marginal [bi] or a marginal [di] as defined by the probability of a B or D response during baseline testing. A truly ambiguous stimulus would receive both responses with equal probability. To explore this further, the category boundary values for the pre- and postadapted identification functions were reexamined separately for each subject. Since the instructions to subjects had no overall effect on the category shift, the two groups receiving Stimulus 5 as the adaptor were combined together for this
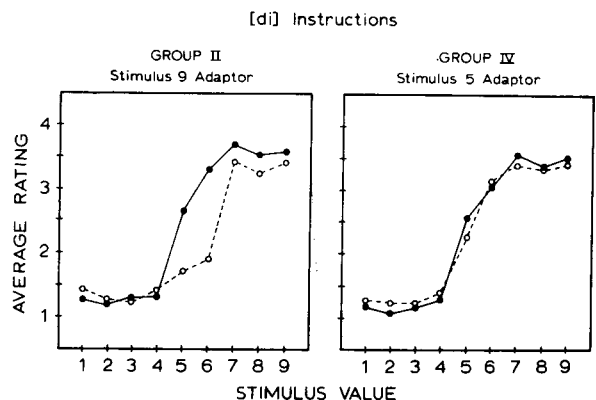
[di] Instructions

GROUP II
Stimulus 9 Adaptor

GROUP IV
Stimulus 5 Adaptor

Figure 3. Average unadapted (solid circles) and adapted (open circles) rating functions for the two [di] instructions groups.

Table 1
Predicted and Observed Direction of Shift in the [bi]-[di]
Category Boundary with Stimulus 5 as the Adaptor

Subject	Probability of [bi] Response to Stimulus 5	Predicted Direction of Shift Toward [bi]	Obtained Direction of Shift Toward [bi]
1	.933	+	+
2	.133	−	−
3	.793	+	+
4	.467	−	−
5	.90	+	−
6	.433	−	−
7	.90	+	+
8	.367	−	−
9	.467	−	+
10	.30	−	−
11	.90	+	+
12	.167	−	−

analysis. The probability of identifying Stimulus 5 as a [bi] prior to selective adaptation is given in Table 1 for each subject. If this probability is above 0.5 for a given subject, then we assume Stimulus 5 should act as a "bi" adaptor and any boundary shift found should be toward the [bi] end of the test series. Similarly, if the probability of identifying Stimulus 5 is below 0.5, the category boundary should shift toward the [di] end of the stimulus series following adaptation. The predicted and observed directions of the boundary shifts are given in Table 1. Ten of the 12 subjects showed a shift in the predicted direction, a result which is significant using a two-tailed sign test (p < .038).

Thus, the middle stimulus does have some effect on the phonetic category boundary, although the effect varied from subject to subject. For individual subjects, the Stimulus 5 adaptor shifted the category boundary toward the phonetic category to which it had originally been assigned during baseline testing. This occurred regardless of the labeling instructions to the subject as to the identity of Stimulus 5 when it was the adapting syllable. The results of this analysis are therefore in accord with those reported by Diehl (1975). However, it should be noted that although the middle stimulus did produce adaptation effects, these shifts were quantitatively much smaller than those found with end-point stimuli. Moreover, as we noted earlier, the middle adaptor produced no significant change in the rating responses within the phonetic categories, which is in marked contrast to the rating shifts found for the end-point adaptors shown in Figures 2 and 3 and those found in the earlier study by Sawusch (1976).

To summarize, an account of selective adaptation in terms of changes in response organization as a result of labeling instructions is not supported by the findings of the present experiment. The labeling instructions to the subject as to the identity of a perceptually ambiguous stimulus produced no overall systematic adaptation effect. However, two factors were identified which appear to play a role in selective adaptation. First, the relation of the acoustic properties of the adaptor to the test series appears to determine the magnitude of the shifts in the locus of the phonetic boundary. Second, the subject's own identification of the adapting stimulus contributes to the direction of the shift regardless of the experimenter's instructions to the subject about the identity of that adapting stimulus. Both findings suggest that selective adaptation to speech is a consequence of changes in perceived quality of the stimulus brought about by sensory factors at early stages of processing.

REFERENCE NOTES

1. Sawusch, J. R., & Pisoni, D. B. *Category boundaries for speech and non-speech sounds.* Paper presented at the 86th meeting of the Acoustical Society of America, November 1973, Los Angeles.
2. Sawusch, J. R., Pisoni, D. B., & Cutting, J. E. *Category boundaries for linguistic and non-linguistic dimensions of the same stimuli.* Paper presented at the 87th meeting of the Acoustical Society of America, April 1974, New York.
3. 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, April 1975, Boston.

REFERENCES

ADES, A. How phonetic is selective adaptation? Experiments on syllable position and vowel environment. *Perception & Psychophysics,* 1974, 16, 61-67.
COOPER, W. E., EBERT, R. R., & COLE, R. A. Perceptual analysis of stop consonants and glides. *Journal of Experimental Psychology: Human Perception and Performance,* 1976, 2, 92-104.
COOPER, W. E., & NAGER, R. M. Perceptuo-motor adaptation to speech: An analysis of bisyllabic utterances and a neural model. *Journal of the Acoustical Society of America,* 1975, 58, 256-265.
CUDDY, L. L., PINN, J., & SIMONS, E. Anchor effects with biased probability of occurrence in absolute judgement of pitch. *Journal of Experimental Psychology,* 1973, 100, 218-220.
DIEHL, R. L. The effect of selective adaptation on the identification of speech sounds. *Perception & Psychophysics,* 1975, 17, 48-52.
EIMAS, P. D., & CORBIT, J. D. Selective adaptation of linguistic feature detectors. *Cognitive Psychology,* 1973, 4, 99-109.
HELSON, H. *Adaptation level theory.* New York: Harper & Row, 1964.
PARDUCCI, A. Range-frequency compromise in judgment. *Psychological Monographs,* 1963, 77(2), 1-50.
PARDUCCI, A. Contextual effects: A range-frequency analysis. In E. Carterette (Ed.), *Handbook of perception* (Vol. II). New York: Academic Press, 1975.
PISONI, D. B., & TASH, J. B. Auditory property detectors and processing place features in stop consonants. *Perception & Psychophysics,* 1975, 18, 401-408.
POLLACK, I., & BOYNTON, J. Identification of elementary auditory displays: Effects of unbalanced probabilities of occurrence. *Journal of the Acoustical Society of America,* 1963, 35, 1831-1832.

SAWUSCH. J. R. Selective adaptation effects on end-point stimuli in a speech series. *Perception & Psychophysics,* 1976, in press.

TARTTER. V. C., & EIMAS, P. D. The role of auditory and phonetic feature detectors in the perception of speech. *Perception & Psychophysics,* 1975, **18,** 293-298.

NOTES

1. We feel that the conclusions drawn by Cooper, Ebert, and Cole (1976) about changes in sensitivity in selective adaptation experiments are completely unwarranted because the assumptions of the decision model that they employed were not adequately tested with these stimuli.

2. We wish to thank Professor Peter Ladefoged, UCLA, for insisting that an experiment along these lines be carried out.