# Selective anchoring and adaptation of phonetic and nonphonetic continua

Helen J. Simon

*Arizona State University, Department of Speech and Hearing Science, Tempe, Arizona 85281*

Michael Studdert-Kennedy

*Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511. Queens College and the Graduate Center, City University of New York*

(Received 30 March 1978; revised 31 July 1978)

A series of four experiments compared the effects of unequal probability anchoring and selective adaptation on phonetic and nonphonetic judgments. The basic stimulus series was a synthetic stop consonant continuum ranging from /b/ to /d/. On this continuum were superimposed covariations in fundamental frequency, intensity or vowel. In each experiment subjects listened to identical test tapes under two judgment conditions: place of articulation, and pitch or loudness or vowel judgments. The two types of judgment were significantly dissociated under both anchoring and adaptation paradigms, thus demonstrating that the former may be no less selective than the latter. From this and other evidence, it was concluded that the two paradigms are, in principle, equivalent, and that the main factors in speech adaptation effects are peripheral fatigue and central auditory contrast. If the selective processes of fatigue and contrast are taken to reflect functional channels of analysis rather than the operation of feature detectors, the same broad processes can be seen at work in both speech and nonspeech adaptation.

PACS numbers: 43.70.Dn, 43.70.Ve

## INTRODUCTION

Over the past five years, several dozen papers have reported studies of the selective adaptation of speech sounds. The series began with a paper by Eimas and Corbit (1973). They asked listeners to categorize members of a synthetic voice onset time (VOT) continuum (Lisker and Abramson, 1964), and demonstrated that the perceptual boundary between voiced and voiceless categories along that continuum was shifted by repeated exposure to (that is, adaptation with) either of the endpoint stimuli: there was a decrease in the frequency with which stimuli close to the original boundary were assigned to the adapted category and a consequent shift of the boundary toward the adapting stimulus. Since the effect could be obtained on a labial VOT contiuum after adaptation with a syllable drawn from an alveolar VOT continuum, and vice versa, adaptation was clearly neither of the syllable as a whole nor of the unanalyzed phoneme, but of a feature within the phonemic segment. Eimas and Corbit therefore termed the adaptation "selective" and attributed their results to the fatigue of specialized linguistic feature detectors and the relative "sensitization" of opponent detectors. Subsequent studies have replicated the results for VOT and have extended them to other phonetic feature oppositions, such as place and manner of articulation. For reviews, see Ades (1976), Cooper (1975), and Eimas and Miller (1978).

These later studies have generally continued to favor a detector fatigue model, but have tended to modify the hypothesized level of adaptation by attributing the effects to acoustic rather than linguistic feature detectors. Thus, in a recent version of the model, Eimas and Miller (1978) propose "... multiply-tuned feature detectors, each... sensitive to a range of complex acoustic information... sufficient to signal a single phonetic feature value...," and each having its "...greatest sensitivity... at some set of acoustic values that correspond to the modal acoustic consequences of articulating that feature value in a particular set of conditions...."

Attractive as such an account may be for those concerned with the biology of language in general, and with the ontogeny of speech perception in particular, there are both broad and narrow grounds for scepticism. Broadly, the hypothesis that selectively tuned feature detectors segment the speech signal, at an early stage of its processing, into sets of invariant properties is not easy to mesh with the accumulated evidence that cues to phonetic structure are highly variable and are often distributed dynamically across an entire syllable (for example, Cooper *et al.*, 1952; Liberman *et al.*, 1967; Shankweiler, Strange, and Verbrugge, 1976; Studdert-Kennedy, 1976a; Dorman, Studdert-Kennedy, and Raphael, 1977; and for recent discussions bearing directly on the question of feature detectors, see Repp, 1977, and Repp *et al.*, 1978). Moreover, evidence to support the several assumptions of the model comes entirely from speech adaptation studies. While there is no question that these studies have persuasively demonstrated channels of feature analysis in speech perception, they scarcely warrant literal interpretation of a physiological metaphor. For although complementary (or opponent) detectors may be plausibly invoked to account for the perception of, say, color or movement, the function of such detectors in the perception of, say, laryngeal periodicity or formant transition is less easily imagined.

Among the narrower grounds for doubt are proliferating reports of contingent effects in adaptation studies (Ades, 1976): adaptation of initial stop consonants on both voicing (Cooper, 1974) and place of articulation (Bailey, 1973; Pisoni, Sawusch, and Adams, 1975; Miller and Eimas, 1976) is now known to be contingent on the following vowel, while adaptation of place of articulation has been found to depend also on syllable position (Ades, 1974; Miller and Eimas, 1976) and funda-

 0001-4966/78/6405-1338$00.80  

mental frequency (Ades, 1977). As Eimas and Miller (1978) explicitly recognize, the theoretical utility of selectively tuned feature detectors goes down as the number of contexts to which they must be tuned goes up.

A second reason for doubt is that the detector fatigue, supposedly induced by adaptation, has usually been inferred from shifts in response frequency rather than measured directly by threshold determination in the manner of, for example, Kay and Matthews (1972) who traced the tuning curves of human auditory channels sensitive to specific ranges of frequency modulation. Recent work by Miller (1977), Miller, Eimas, and Root (1977), and Sawusch (1976, 1977) has begun to fill this gap, and we will take account of their results in the final discussion, but the only other attempt to estimate sensitivity changes directly, in a speech adaptation study, produced results precisely the opposite of those predicted by a fatigue model. Cooper, Ebert and Cole (1976) used a magnitude estimation procedure (Durlach and Braida, 1969) to assess $d'$ values, before and after adaptation, for pairs of stimuli along a labial stop-to-glide continuum. Although they found a general decrease in sensitivity after adaptation, by far the largest decrease was at the phoneme boundary rather than at the adapted stimulus, as a fatigue model would predict. In fact, for the two pairs of stimuli that included the endpoint adaptors, Cooper *et al.* (1976) reported an *increase* in sensitivity, which, though not tested for significance, led them to countenance the "...strong possibility...that the increased sensitivity following adaptation for a pair of stimuli including or neighboring the adapting stimulus reflects the listener's heightened ability to utilize the adapting stimulus as an anchor, or reference stimulus..." (Cooper *et al.*, 1976, p. 103).

Such a possibility is exactly what the present experiments were designed to explore. Our starting point was a study by Sawusch and Pisoni (1973). They collected identification functions for a series of synthetic speech sounds varying in VOT from /b/ to /p/ and for a series of tones varying in intensity from 60 to 80 dB SPL ("soft" to "loud"). In one condition (control), all stimuli occurred equally often. In the other (anchor), a particular endpoint stimulus occurred twice as often as other stimuli in the series (2:1 ratio).[1] The result was that, in the anchor condition, the category boundary for the tones shifted toward the anchoring stimulus, while the phonetic boundary remained stable. In a second experiment, Sawusch, Pisoni, and Cutting (1974) used the same anchoring procedure (at a 4:1 ratio) on a synthetic continuum in which the stimuli varied simultaneously in both place of articulation (/b/ to /d/) and fundamental frequency ($F_0$) ("high" to "low"). Once again the nonphonetic boundary shifted, while the phonetic boundary remained stable. Since the authors believed that the effect of anchoring is to bias a subject's responses rather than his percepts, they inferred that stop consonants are largely immune to response bias. They attributed this immunity to the listener's having an internal standard for the stop consonant and concluded that the effects of selective adaptation to which stop consonants are not immune were therefore sensory, or perceptual.

However, this argument reaches the right conclusion for the wrong reason. Not all anchoring effects reflect response bias (see, for example, Helson and Kozaki, 1968; Helson, 1971), and it seems unlikely that two procedures, differing as little as do anchoring and adaptation, would engage qualitatively different mechanisms in judgments of the same continuum. The main difference between the procedures is simply in the number and distribution of anchor or adaptor repetitions. In adaptation, the adaptor typically occurs many times at short interstimulus intervals (ISI) in a single block before the test stimuli. In anchoring, the anchor occurs less often and with a longer ISI, its repetitions being scattered randomly among the test stimuli. The two procedures lie at opposite ends of a continuum of anchor/adaptor energy concentration, so that differences between the consonantal results in the two types of experiment may well be of degree rather than of kind.

The following experiments, therefore, address three questions: (1) Is there a reliable dissociation in the effects of anchoring between judgments of phonetic and judgments of nonphonetic continua? (2) Is there a similar dissociation in the effects of adaptation? (3) Can a single, unified account be developed for the effects of both experimental paradigms on both types of continua? Experiments I–III approach the first question by comparing the effects of anchoring on the identification of a synthetic stop consonant continuum with those on the identification of fundamental frequency, intensity and synthetic vowel continua. The two types of stimulus variation in each experiment covary, that is to say, are perfectly correlated and are carried simultaneously on the same series of syllables. Experiment IV approaches the second question by comparing the effects of both anchoring and adaptation on the identification of covarying synthetic stop consonant and fundamental frequency continua. We reserve the third question for our concluding discussion.

## I. EXPERIMENT I

The first experiment replicated the study of Sawusch *et al.* (1974) with one difference. These authors used a seven-step synthetic stop place of production continuum (/b/ to /d/), and paired each stimulus value along the continuum with each of seven variations in $F_0$. They thus produced 49 stimuli in which formant structure and $F_0$ varied independently. In the present study, each stimulus along a seven-step synthetic stop place of production continuum was assigned a different $F_0$ contour. This yielded only seven stimuli in which formant structure and $F_0$ were perfectly correlated. If under these conditions a shift in the fundamental frequency boundary was observed without a shift in the phoneme boundary, an even stronger case might be made for the dissociation of phonetic and nonphonetic judgments in an anchoring paradigm.

### A. Method

#### 1. Stimuli

A series of three-formant consonant vowel (CV) syllables was generated on the Haskins Laboratories paral-

TABLE I. Fundamental frequency and the starting frequencies (in hertz) of the second- and third-formant transitions for the synthetic CV syllables of experiment I.

| Stimulus Number | $F_0$ | $F_2$ | $F_3$ |
|---|---|---|---|
| 1 | 114 | 1385 | 2525 |
| 2 | 120 | 1468 | 2694 |
| 3 | 126 | 1541 | 2862 |
| 4 | 132 | 1620 | 3026 |
| 5 | 138 | 1695 | 3195 |
| 6 | 144 | 1772 | 3363 |
| 7 | 150 | 1845 | 3530 |

Note: The fixed steady state formants were centered at 666 Hz ($F_1$), 1620 ($F_2$) and 3026 ($F_3$).

lel resonance synthesizer. The series consisted of seven syllables, each with a duration of 300 ms. The final 260 ms was a steady-state portion appropriate to the American English vowel /æ/ and was identical in all syllables: 660 Hz for $F_1$, 1620 Hz for $F_2$ and 3026 Hz for $F_3$. The seven stimuli ranged perceptually from /bæ/ to /dæ/ in approximately equal steps in second- and third-formant transition starting frequencies (Table I). For all seven stimuli, $F_1$ rose over the first 40 ms from a starting frequency of 234 Hz to its steady-state frequency of 660 Hz.

Fundamental frequency for the first 225 ms of each syllable varied from 114 Hz to 150 Hz in 6 Hz steps over the series from /b/ to /d/, and fell linearly from its initial value to 80 Hz during the last 75 ms of the syllable. Therefore, a particular fundamental frequency contour characterized each syllable along the phoneme continuum. Overall amplitude was attenuated linearly by 28 dB in the last 75 ms of each syllable.

The syllables were recorded on magnetic tape. They were then digitized, edited and stored on the Haskins pulse code modulation (PCM) system for subsequent tape preparation. Two identification tapes were prepared. On the first, the control tape, the seven syllables were recorded equally often: ten random permutations of the series to produce 70 stimuli. On the second, the anchoring tape, the anchor stimulus was recorded four times as often as the other six stimuli: ten random permutations of ten stimuli (four anchors plus six others) to yield a total of 100 stimuli. The anchor stimulus was the first syllable (/bæ/) described in Table I, with an $F_0$ of 114 Hz. Since a differential effect of the two end points was not of interest in this experiment, only one end point was used as an anchor.

The stimuli on both tapes were recorded on a Crown 800 tape recorder connected directly to the output of the PCM system, with a 2-s interval between stimuli and a 10-s interval after every tenth stimulus. A 1000-Hz calibration tone, set at the peak vowel amplitude on the VU meter of the tape recorder, was recorded at the beginning of each tape to permit uniform playback levels.

### 2. *Procedure*

The experiment called for identification of the two stimulus series, control and anchor, under two task conditions, pitch identification and phoneme identification. Each subject heard the same control (70 item) and the same anchor (100 item) tape twice. Only the instructions varied as the task changed. For the pitch task, subjects were told that they would hear a series of sounds presented on two tapes similar in every way save that the second had 100 stimuli, while the first had only 70, and that while other aspects of the stimuli would vary, their task was simply to write "high" or "low" to identify the pitch. For the phoneme identification task, the subjects were given similar instructions, but were asked to identify each syllable as beginning with either /b/ or /d/. Each experimental tape was preceded by display and practice tapes.

The order of testing was the same for all subjects: pitch control, pitch anchor, phoneme control, and phoneme anchor. The pitch task was deliberately given first, since the expected absence of an anchoring effect in the phoneme task would then have occurred in spite of the expected anchoring effect in the preceding pitch task.

All experimental tapes were reproduced binaurally from the output of an Ampex AG 500 tape recorder over calibrated Telephonics (TDH-39) matched headphones with a circumaural seal. The gain of the tape recorder playback for the 1000-Hz calibration tone was adjusted to give a voltage across the earphones equivalent to approximately 75 dB SPL *re* 0.0002 dyn/cm$^2$.

### 3. *Subjects*

Nine normal hearing (screened at 20 dB HTL *re* ANSI 1969 standard) undergraduate students at Yale University participated in this experiment. All were native speakers of English, had no past history of speech or hearing problems, and were paid at a rate of $2.00 per hour. They were tested in a quiet room, either alone or in pairs, in a single session with a short break between the pitch and phoneme judgment conditions.

## B. Results

Figure 1 displays the pitch and consonant identification functions for control and anchor conditions. Apart from the marked discontinuity in the pitch control function (discussed below), the most obvious features of the figure are the shallowness of the pitch functions—as compared with the steepness of the consonantal functions—and the clear leftward shift of the pitch anchor function—as compared with the trivial shift of the consonantal anchor function.

As a simple measure of the anchoring effect, we can compare the distribution of the two response categories over the stimulus continuum in the control and anchor conditions. If we assume an effect of contrast, rather than assimilation, we can then predict a decrease in the total number of responses in the class to which the anchor is typically assigned. Since, in the anchoring condition, the anchor stimulus was presented more often than the other stimuli, and was almost always assigned

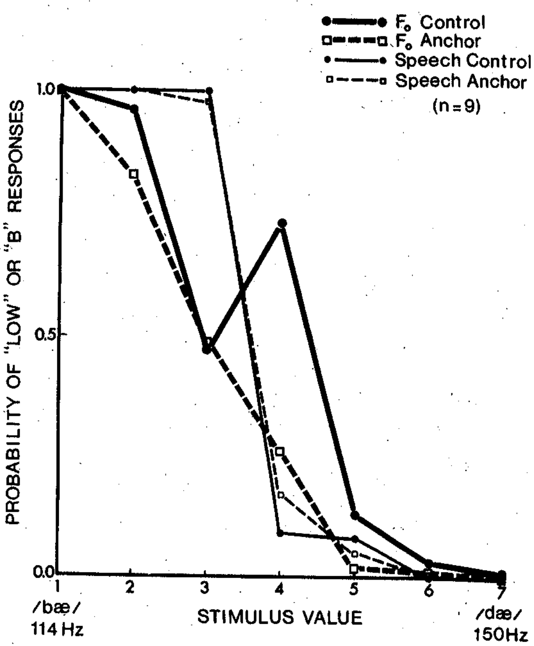


FIG. 1. The probability of "low" or "b" responses as a function of stimulus value, under control and anchor conditions, for the pitch and speech tasks in experiment I.

to the same response class, responses to this stimulus were omitted from the totals. Table II, therefore, lists individual and mean numbers of "low" and "b" identification responses to the sixty presentations of the remaining stimuli under the two task conditions. For the pitch task there is a mean decrease of 7.2 "low" responses in the anchor condition; the decrease is significant by a one-tailed, matched pairs $t$-test ($t=4.07$, $p<0.005$). For the speech task there is an insignificant mean increase of 0.2 "b" responses in the anchoring condition.

A second, derived measure of an anchoring effect is the shift in the estimated boundary between response categories: here, a contrast effect would appear as a shift in the boundary toward the anchor stimulus. Normal ogives were fitted to the individual data by the method of least squares[2] (Woodworth, 1938), and the results are listed in Table III. For the pitch task, the mean control and anchor boundaries (that is, means of the fitted ogives) are 3.63 and 3.24, respectively, a significant shift in the anchor condition of 0.34 continuum steps toward the anchor stimulus (matched pairs $t=2.03$, $p<0.05$, one-tailed). For the speech task, the mean control and anchor boundaries are 3.74 and 3.83, respectively, an insignificant shift in the anchor condition of 0.07 steps away from the anchor.

TABLE III. Individual and mean boundaries (means of the normal ogives for the $F_0$ and speech tasks in experiment I.

| | $F_0$ | | Speech | |
|---|---|---|---|---|
| Subject | Control | Anchor | Control | Anchor |
| 1 | 3.28 | 3.87 | 3.67 | 3.67 |
| 2 | 3.72 | 3.42 | 3.92 | 3.96 |
| 3 | 3.15 | 2.93 | 3.67 | 3.80 |
| 4 | 3.68 | 3.23 | 3.67 | 3.50 |
| 5 | 3.60 | 2.09 | 3.87 | 3.97 |
| 6 | 3.79 | 3.39 | 3.72 | 3.92 |
| 7 | 3.63 | 3.48 | 3.67 | 3.88 |
| 8 | 3.83 | 2.85 | 3.67 | 3.92 |
| 9 | 4.04 | 3.91 | 3.81 | 3.90 |
| Mean | 3.63 | 3.24 | 3.74 | 3.83 |

The individual and mean standard deviations (reciprocally related to the slopes) of the ogives, fitted to the control data from the two task conditions, are listed in Table IV. The speech standard deviations are remarkably consistent across individuals, and every subject, except subject 8, gives a larger standard deviation on the pitch task than on the speech task. The mean pitch standard deviation of 1.16 is significantly larger than the mean speech standard deviation of 1.02 on a two-tailed matched pairs $t$-test ($t=2.40$, $p<0.05$). There is no significant Spearman coefficient of rank order correlation between control standard deviation and boundary shift for either pitch ($\rho=-0.14$) or speech ($\rho=0.09$).

The relatively large pitch standard deviations clearly result, in part, from the sharp discontinuity in the pitch control function. This discontinuity itself seems to have resulted from a simple contrast effect, precipitated by an accident of the randomized test order. By a chance not noticed until after the data had been gathered, stimulus 3 was always preceded by either stimulus 1 or stim-

TABLE II. Individual and mean number "low" and "b" responses to the sixty presentations of stimuli 2–7 in experiment I.

| | $F_0$ | | Speech | |
|---|---|---|---|---|
| Subject | Control | Anchor | Control | Anchor |
| 1 | 28 | 22 | 20 | 20 |
| 2 | 22 | 18 | 24 | 24 |
| 3 | 15 | 14 | 20 | 21 |
| 4 | 22 | 15 | 20 | 19 |
| 5 | 20 | 16 | 22 | 20 |
| 6 | 21 | 18 | 28 | 24 |
| 7 | 25 | 17 | 20 | 22 |
| 8 | 21 | 11 | 20 | 23 |
| 9 | 26 | 24 | 21 | 24 |
| Mean | 23.3 | 16.1 | 21.7 | 21.9 |

TABLE IV. Individual and mean standard deviations of the normal ogives fitted to the control identification data for the $F_0$ and speech tasks in experiment I.

| Subject | $F_0$ | Speech |
|---|---|---|
| 1 | 1.13 | 1.00 |
| 2 | 1.08 | 1.04 |
| 3 | 1.27 | 1.00 |
| 4 | 1.09 | 1.00 |
| 5 | 1.20 | 1.06 |
| 6 | 1.15 | 1.07 |
| 7 | 1.54 | 1.00 |
| 8 | 0.93 | 1.00 |
| 9 | 1.08 | 1.04 |
| Mean | 1.16 | 1.02 |

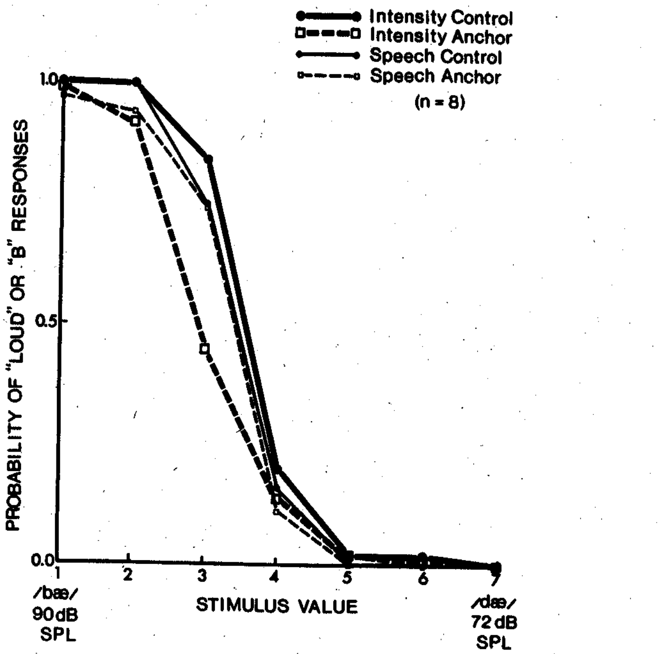


FIG. 2. The probability of "loud" or "b" responses as a function of stimulus value, under control and anchor conditions, for the loudness and speech tasks in experiment II.

ulus 2, both of which were, in this position, always identified as "low." Stimulus 4, on the other hand, was always preceded by stimulus 5, 6, or 7 which, on 96% of their occurrences in this position, were identified as "high." The result was a disproportionate number of "high" judgments for stimulus 3 and a disproportionate number of "low" judgments for stimulus 4. Note that, although precisely the same stimuli and test order were used for the consonant task, no discontinuity appears in the consonant function.

We defer discussion of these results until we have reported on the next experiment.

## II. EXPERIMENT II

The purpose of this experiment was to explore further the effects of anchoring in correlated phonetic and nonphonetic continua, this time using intensity as the nonphonetic dimension varied over the stimulus continuum.

### A. Method

#### *1. Stimuli*

The stimuli were the series of consonant–vowel syllables used in experiment I (see Table I), but with three differences: (1) fundamental frequency for the first 175 ms was held constant at 114 Hz for all seven stimuli in the continuum, and fell linearly to 80 Hz during the last 75 ms of each syllable; (2) in order to conform to most previous adaptation experiments, the stimuli were shortened to a duration of 250 ms by dropping 50 ms of the steady-state portion; (3) amplitude was attenuated by 18 dB in 3 dB steps from /bæ/ to /dæ/. Therefore, in this experiment, the correlate of position on the phoneme continuum was overall amplitude. Tapes were prepared in the same way as for experiment I.

#### *2. Procedure*

Order of testing followed the same pattern as in experiment I: loudness control, loudness anchor, phoneme control, and phoneme anchor. In the loudness condition, subjects were asked to listen for loudness differences in the syllables and to ignore other differences observed. Instructions were to write "L" for a loud sound or "S" for a soft sound for the loudness judgments, and "b" or "d" for the phoneme judgments. Each experimental session was preceded by a display and a practice tape. The gain of the tape recorder playback for the 1000 Hz calibration tone was adjusted to give a voltage reading across the earphones equivalent to approximately 90 dB *re* 0.0002 dyn/cm$^2$. Therefore, the test stimuli varied from 90 dB SPL for stimulus 1 (the anchor stimulus) to 72 dB SPL for stimulus 7.

Nine paid listeners served as subjects. All met the criteria established in the previous experiment. They were tested either singly or in pairs, and testing was completed in one session with a short break between the loudness and phoneme judgment conditions.

### B. Results

One subject was dropped from the data analysis because he did not categorize the stimuli consistently in the phoneme identification task. Figure 2 displays the loudness and consonant group identification functions for the remaining eight subjects in the control and anchor conditions. The slopes of the control functions seem equally steep for the two tasks, but the anchor function is appreciably lowered for loudness, very little for the consonants.

Table V presents the total number of "loud" and "b" responses (excluding those to the anchor stimulus) for each of the eight subjects under both conditions. For the loudness task, there is a significant mean decrease of 5.4 "loud" responses in the anchoring condition (matched pairs $t = 4.80$, $p < 0.005$, one-tailed). For the speech task there is an insignificant decrease of 1.2 "b" responses in the anchoring condition (matched pairs $t = 1.2$, $p > 0.10$, one-tailed).

Normal ogives were again fitted to the individual data, and the results are listed in Table VI. For the loudness

TABLE V. Individual and mean number of "loud" and "b" responses to the sixty presentations of stimuli 2–7 in experiment II.

| | Intensity | | Speech | |
|---|---|---|---|---|
| Subject | Control | Anchor | Control | Anchor |
| 1 | 17 | 10 | 20 | 19 |
| 2 | 19 | 12 | 24 | 23 |
| 3 | 21 | 9 | 13 | 8 |
| 4 | 21 | 18 | 23 | 17 |
| 5 | 17 | 13 | 20 | 18 |
| 6 | 32 | 28 | 19 | 26 |
| 7 | 19 | 16 | 14 | 15 |
| 8 | 21 | 18 | 20 | 17 |
| Mean | 20.9 | 15.5 | 19.1 | 17.9 |

TABLE VI. Individual and mean boundaries for the intensity and speech tasks in experiment II.

| | Intensity | | Speech | |
|---|---|---|---|---|
| Subject | Control | Anchor | Control | Anchor |
| 1 | 3.36 | 2.72 | 3.40 | 3.84 |
| 2 | 3.58 | 3.12 | 3.96 | 3.92 |
| 3 | 3.80 | 3.10 | 3.54 | 2.22 |
| 4 | 3.82 | 3.68 | 3.78 | 3.36 |
| 5 | 3.48 | 3.54 | 3.66 | 3.38 |
| 6 | 4.64 | 4.26 | 3.50 | 4.04 |
| 7 | 3.58 | 3.32 | 3.22 | 3.28 |
| 8 | 4.00 | 3.56 | 3.64 | 3.58 |
| Mean | 3.78 | 3.41 | 3.59 | 3.45 |

task, the mean control and anchor boundaries are 3.78 and 3.41, respectively, a significant shift in the anchor condition of 0.37 continuum steps toward the anchor stimulus (matched pairs $t=4.17$, $p<0.005$, one-tailed). For the speech task, the mean control and anchor boundaries are 3.59 and 3.45, respectively, an insignificant shift in the anchor condition of 0.14 continuum steps toward the anchor stimulus (matched pairs $t=0.09$, $p>0.45$, one-tailed).

Table VII lists individual and mean standard deviations of the fitted ogives. The mean loudness standard deviation of 1.07 is insignificantly larger than the mean speech standard deviation of 1.05, on a two-tailed matched pairs $t$-test ($t=0.82$, $p>0.40$). There is no significant rank order correlation between control standard deviation and boundary shift for either loudness ($\rho=0.30$) or speech ($\rho=0.22$).

## III. DISCUSSION OF EXPERIMENTS I AND II

The overall conclusion to be drawn from these experiments is essentially similar to that of Sawusch and Pisoni (1973) and of Sawusch *et al.* (1974): A synthetic stop consonant continuum is less susceptible to psychophysical anchoring effects than an arbitrary continuum of pitch or loudness. However, the conclusion is even stronger in the present experiments, since each value of place was always paired with a given $F_0$ or intensity value, so that one dimension was fully predictable from the other. Despite this deliberately imposed association of stimulus dimensions, phonetic and nonphonetic response patterns displayed definite dissociation.

One possible account of this dissociation is suggested by the fact that the stimuli most susceptible to anchoring effects are usually those for which the variance of judgment is highest (Volkmann, 1951; Parducci, 1965).[3] Certainly, the slopes of the pitch control functions in experiment I were significantly greater than those of the speech control functions. Furthermore, all anchoring effects in both experiments (apart from the small effect on stimulus 2 in the loudness task) were confined to "boundary" stimuli that were not consistently identified in the control conditions. However, this cannot be the complete explanation for the dissociation, since there was no significant rank order correlation for either task in either experiment between the standard deviation of the control function and the degree of boundary shift; nor was there any significant difference in the slopes of the loudness and speech control functions of experiment II. This last fact demonstrates that, even if members of a nonphonetic continuum are classified consistently as members of a stop consonant continuum, they still may be more susceptible to anchoring. In other words, while some degree of stimulus ambiguity may be a necessary condition of anchoring effects, ambiguity alone cannot account for all of the variance. We return to this matter in the general discussion.

## IV. EXPERIMENT III

The results of the first two experiments may be taken as instances of a general susceptibility to context effects, typical of continuously perceived dimensions, as compared with a general resistance to context effects, typical of categorically perceived dimensions. Since strong context effects have been reported for synthetic vowels (Fry *et al.*, 1962; Stevens *et al.*, 1969), but less for synthetic stop consonants (Liberman *et al.*, 1961; Eimas, 1963), the next experiment extends the anchoring paradigm, to compare its effects on correlated stop consonant and vowel continua.

### A. Method

#### 1. Stimuli

The stimuli in this experiment were a series of two-formant consonant–vowel syllables ranging perceptually from /bæ/ to /dɛ/ in approximately equal steps of second-formant transition onsets and first- and second-formant steady-state frequencies. The third-formant circuit on the Haskins parallel resonance synthesizer was turned off during synthesis, because of the difficulty encountered in generating a perceptually acceptable three-formant correlated vowel and consonant continuum.

The stimulus values, including transition durations and extents, are displayed in Table VIII. All $F_1$ transitions were 40 ms in duration and rose linearly from 234 Hz to a steady-state frequency, ranging from 718 Hz (stimulus 1) to 562 (stimulus 7), so that the transition extents ranged from 484 Hz to 328 Hz. The $F_2$ transitions rose linearly from onset frequencies, ranging be-

TABLE VII. Individual and mean standard deviations of the normal ogives fitted to the control identification data for the intensity and speech tasks in experiment II.

| Subject | Intensity | Speech |
|---|---|---|
| 1 | 1.08 | 1.06 |
| 2 | 1.06 | 1.00 |
| 3 | 1.10 | 1.12 |
| 4 | 1.00 | 1.06 |
| 5 | 1.08 | 1.00 |
| 6 | 1.18 | 1.04 |
| 7 | 1.06 | 1.10 |
| 8 | 1.00 | 1.06 |
| Mean | 1.07 | 1.05 |

TABLE VIII. Onset and steady-state frequencies, extent of transition rises and transition durations for the two formants of the correlated vowel and consonant continuum (experiment III). All syllables were 250 ms in duration.

| | $F_1$ | | | | $F_2$ | | | |
|---|---|---|---|---|---|---|---|---|
| Stimulus number | Onset frequency in Hz | Steady-state frequency in Hz | Transition extent in Hz | Transition duration in ms | Onset frequency in Hz | Steady-state frequency in Hz | Transition extent in Hz | Transition duration in ms |
| 1 | 234 | 718 | 484 | 40 | 1385 | 1541 | 156 | 40 |
| 2 | 234 | 692 | 458 | 40 | 1468 | 1620 | 152 | 40 |
| 3 | 234 | 666 | 432 | 40 | 1541 | 1695 | 154 | 35 |
| 4 | 234 | 640 | 406 | 40 | 1620 | 1772 | 152 | 30 |
| 5 | 234 | 614 | 380 | 40 | 1695 | 1845 | 150 | 25 |
| 6 | 234 | 588 | 354 | 40 | 1772 | 1920 | 148 | 25 |
| 7 | 234 | 562 | 328 | 40 | 1845 | 1996 | 151 | 20 |

tween 1385 Hz (stimulus 1) and 1845 Hz (stimulus 7), to steady-state frequencies, ranging between 1541 Hz and 1996 Hz, so that transition extents remained roughly constant around 152 Hz. The rising $F_2$ transitions were necessary to maintain roughly equal frequency steps along the continuum. To counter the resulting labial percepts, $F_2$ transitions were reduced in duration from 40 ms (stimulus 1) to 20 ms (stimulus 7). The effect of this maneuver was presumably to reduce the perceptual salience of the transitions and, thus, "flatten" the second formants, granting the more abbreviated transitions (stimuli 5–7) effective values closer to the more or less level transitions appropriate for /d/ in these vowel contexts. A pilot experiment established the acceptability of the stimulus series and estimated the phoneme boundaries. For each stimulus, fundamental frequency was constant at 114 Hz for the first 175 ms and dropped linearly to 80 Hz, while overall amplitude dropped by 28 dB over the last 75 ms.

All experimental tapes were produced as previously described. Two test tapes were again prepared: a 70-item control tape with all stimuli presented equally often, and a 100-item anchor tape with stimulus 1 (/bæ/) presented four times as often as stimuli 2–7.

### 2. *Procedure*

As in previous experiments, the study called for identification of items on the two test tapes under two different sets of instructions. In the first condition, the task was to categorize the vowels as /æ/ or /ɛ/; in the second condition the task was to identify the consonants as /b/ or /d/. Since we feared that the obvious correlation between vowel and consonant might tempt listeners to base their responses in both conditions on the same dimension, we decided to enhance the correlation by drawing subjects' attention to it, while asking them to disregard it in making their judgments. The order of testing was the same for all subjects: vowel control, vowel anchor, stop–consonant control, and stop–consonant anchor. All other aspects of the procedure were identical to those of the previous experiments. There were nine paid listeners, all meeting the earlier criteria.

The experimental tapes were reproduced by an Ampex AG 500 tape recorder and presented binaurally through Telephonics (TDH-39) matched and calibrated headphones. The voltage across the headphones of the 1000-Hz calibration tone was set to the equivalent of approximately 80 dB *re* 0.0002 dyn/cm$^2$.

## B. Results

One subject was dropped from the data analysis because she did not categorize the stimuli consistently in the consonant identification task. Figure 3 displays the vowel and consonant group identification functions for the remaining eight subjects in the control and anchor conditions. The slopes of the consonant functions are notably less steep over the lower end of the continuum than are those of the vowel functions. The anchor functions are appreciably lower than the control functions for both consonants and vowels.

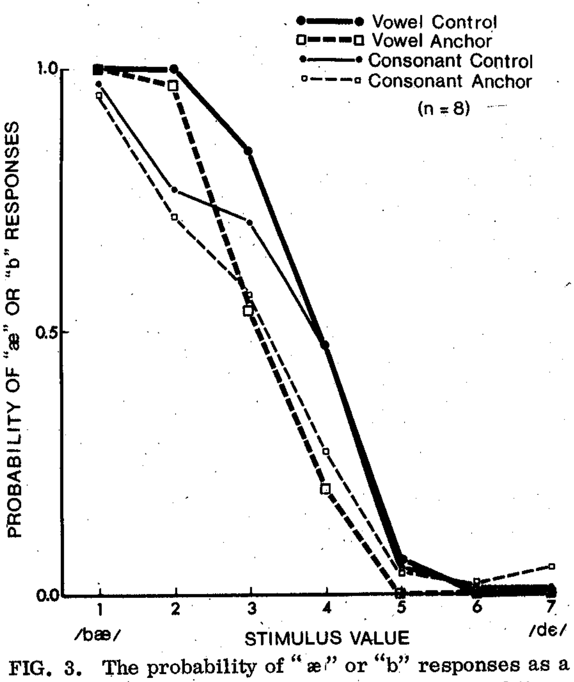


FIG. 3. The probability of "æ" or "b" responses as a function of stimulus value, under control and anchor conditions, for the vowel and consonant task in experiment III.

TABLE IX. Individual and mean number of "æ" and "b" responses to the sixty presentations of stimuli 2–7 in the control and anchor conditions of experiment III.

| | Vowel | | Consonant | |
|---|---|---|---|---|
| Subject | Control | Anchor | Control | Anchor |
| 1 | 26 | 25 | 24 | 25 |
| 2 | 21 | 14 | 27 | 24 |
| 3 | 18 | 16 | 27 | 25 |
| 4 | 28 | 16 | 8 | 3 |
| 5 | 27 | 19 | 30 | 24 |
| 6 | 19 | 15 | 29 | 18 |
| 7 | 24 | 17 | 6 | 8 |
| 8 | 27 | 15 | 12 | 8 |
| Mean | 23.8 | 17.1 | 20.4 | 16.9 |

TABLE XI. Individual and mean standard deviations of the normal ogives fitted to the control identification data for the vowel and consonant tasks in experiment III.

| Subject | Vowel | Consonant |
|---|---|---|
| 1 | 1.06 | 1.38 |
| 2 | 1.08 | 1.04 |
| 3 | 1.08 | 1.00 |
| 4 | 1.00 | 1.40 |
| 5 | 1.04 | 1.04 |
| 6 | 1.08 | 1.04 |
| 7 | 1.00 | 1.56 |
| 8 | 1.12 | 1.34 |
| Mean | 1.06 | 1.23 |

Table IX presents the total number of "æ" and "b" responses (excluding those to the anchor stimulus) for each of the eight subjects under both conditions. For the vowel task, there is a significant mean decrease of 6.7 "æ" responses in the anchor condition (matched pairs $t = 4.60$, $p < 0.005$, one-tailed). For the consonant task, there is a significant mean decrease of 3.5 "b" responses in the anchor condition (matched pairs $t = 2.4$, $p < 0.025$, one-tailed).

Normal ogives were fitted to the individual data, and the results are listed in Table X. For the vowel task, the mean control and anchor boundaries are 3.87 and 3.46, respectively, a significant shift in the anchor condition of 0.41 continuum steps toward the anchor stimulus (matched pairs $t = 4.7$, $p < 0.005$, one-tailed). For the consonant task, the mean control and anchor boundaries are 3.62 and 3.32, respectively, a marginally significant shift in the anchor condition of 0.30 continuum steps toward the anchor stimulus (matched pairs $t = 1.51$, $p < 0.10$, one-tailed). Based on the results of experiments I and II and of previous work with consonants and vowels, we may predict that any shift in the consonant boundary due to contextual anchoring effects will be significantly less than the shift in the vowel boundary on a correlated continuum, and this proves to be the case (matched pairs $t = 2.02$, $p < 0.05$, one-tailed).

Finally, the greater ambiguity of the consonants than of the vowels at the /bæ/ end of the continuum (Fig. 3) was not evidenced by every subject. As may be seen from Table XI, which lists individual and mean standard deviations of the fitted ogives, only subjects 1, 4, 7, and 8 show a larger consonant than vowel standard deviation. For the other four subjects, consonant values are equal to or less than those for the vowels. The difference between the mean vowel standard deviation of 1.06 and the mean consonant standard deviation of 1.23 is marginally significant ($t = 1.99$, $p < 0.10$, two-tailed). There is no significant rank order correlation between control standard deviation and boundary shift for either vowels ($\rho = -0.15$) or consonants ($\rho = 0.02$).

TABLE X. Individual and mean boundaries for the vowel and consonant tasks in experiment III.

| | Vowel | | Consonant | |
|---|---|---|---|---|
| Subject | Control | Anchor | Control | Anchor |
| 1 | 4.18 | 3.88 | 4.22 | 4.18 |
| 2 | 3.68 | 3.34 | 4.12 | 3.96 |
| 3 | 3.54 | 3.52 | 4.08 | 4.36 |
| 4 | 3.88 | 3.32 | 2.60 | 1.26 |
| 5 | 3.96 | 3.60 | 4.34 | 4.06 |
| 6 | 3.60 | 3.28 | 4.28 | 3.42 |
| 7 | 3.96 | 3.48 | 2.36 | 2.68 |
| 8 | 4.12 | 3.24 | 2.96 | 2.62 |
| Mean | 3.87 | 3.46 | 3.62 | 3.32 |

## V. DISCUSSION

Members of a synthetic stop consonant continuum are not always immune to the psychophysical effects of anchoring. The present effects are not strong, though clearly significant ($p < 0.025$) when measured by the number of responses in the anchor response class, they are only marginally significant ($p < 0.10$) when measured by the boundary shift estimated from fitted normal ogives. Nonetheless, even on the latter measure, six of the eight subjects display the effect, and the mean boundary shift of 0.30 continuum steps is more than double that observed on the (differently constructed) continuum of experiment II.

If we take these results to be reliable, two questions arise. First, why were synthetic stop consonants varying in the acoustic cues to place of articulation susceptible to anchoring along a correlated vowel continuum, but not along correlated continua of fundamental frequency or intensity? Second, why were such consonants less susceptible to anchoring than their correlated vowels?

We can immediately dismiss one possible answer to the first question. This is the suggestion that, having had their attention drawn to the consonant–vowel correlation, subjects tended to judge the vowel rather than the consonant, so that consonant judgments "followed" vowel judgments. If this were the case, we would expect some correlation both between consonant and vowel boundaries and between their boundary shifts. However, a scan of Table X suggests no such relations, and the Spearman coefficients of rank order correlation are not significant either for the control boundaries ($\rho = -0.05$) or for the boundary shifts ($\rho = 0.43$).

A second possible answer is that the synthetic stops on the correlated vowel continuum were, at least toward the /bæ/ end of the continuum, poor exemplars of a phonetic class and, therefore, susceptible to contrast effects induced by anchoring. To the extent that ambiguous stimuli are more susceptible to anchoring than unambiguous stimuli, this interpretation would seem to be correct. We must then ask why these syllables were more ambiguous than those of experiments I and II.

Here, we may recall that, since the correlated consonant–vowel continuum lacked a third formant, variations in place of articulation were cued entirely by second formant transitions. Furthermore, the onsets of these transitions were not adjusted to their following steady-state frequencies. While $F_1$ and $F_2$ steady-state frequencies varied to yield a range of vowels from /æ/ to /ε/, $F_1$ and $F_2$ onset frequencies were identical to those of experiments I and II, where steady-state formant frequencies were constant at values for /æ/. The onset frequencies (and transitions) were therefore somewhat inappropriate to the following vowels. If we add to this the fact that the vowel itself was unpredictable from trial to trial, so that the listener lacked the invariant spectral reference against which the variations of a synthetic stop consonant continuum are usually judged, it is not surprising that listeners found the continuum of experiment III somewhat ambiguous.

To this uncertainty the /bæ/ anchor perhaps brought a note of stability. The recurrence of its relatively low $F_2$ transitional and steady-state frequencies may have established a standard against which the transitions and steady states of the boundary stimuli (2–4) were sometimes heard as raised and "flattened," the vowel more like /ε/, the consonant more like /d/. This explanation, in terms of psychophysical contrast, must, of course, be distinguished from the usual account in terms of detector "fatigue."

The answer to our first question follows naturally from this account. Synthetic stop consonants were susceptible to anchoring on the correlated vowel continuum because simultaneous variations in transitional and steady-state formant frequencies produced ambiguous stimuli that contrasted with the anchor stimulus over their entire length. The syllables of the correlated $F_0$ and intensity continua, on the other hand, were relatively unambiguous and contrasted with the anchor stimulus only in their first 40 ms. As we will suggest below in the general discussion, the total relevant energy in the anchoring stimulus and its accumulation over the test may largely determine its effect. It is precisely this which may also answer our second question and explain why the consonants of experiment III were, despite their ambiguity, less susceptible to anchoring than their correlated vowels. This again is a matter to which we return in the general discussion.

## VI. EXPERIMENT IV

The preceding experiment demonstrated that, under certain conditions, an anchoring effect, similar to that of the standard adaptation paradigm, can be induced on a synthetic stop consonant continuum. In our final experiment, we undertook a direct comparison of the two paradigms on the covarying consonant and $F_0$ continua of experiment I.

The defining distinction between the paradigms is in the distribution of anchor/adaptor energy over the test: the adaptation paradigm rapidly repeats the adaptors in a number of blocks immediately before the test stimuli, while the anchoring paradigm distributes its repetitions of the anchor stimulus randomly among the test syllables. A second distinction is in the ratio of anchor/adaptor stimuli to test stimuli: typically, the adaptor ratio is much the larger, often by a factor of several hundreds, if not thousands. (If the anchor ratio were appreciably increased, the effective differences between the two paradigms would presumably disappear.) A third difference between the paradigms is in the interstimulus intervals (ISI): while repetitions of the anchor stimulus, even if in immediate succession, are separated by the two or more seconds needed for a subject to respond, repetitions of the adaptor stimulus typically follow at a rate of one or two per second. A final, less important, difference is that subjects are usually asked to respond to every stimulus presentation in the anchoring paradigm, but only to test syllables in the adaptation paradigm.

Taken together, these distinctions add up to a simple difference between the paradigms in the accumulated anchor/adaptor energy over the course of a test, a difference of degree rather than of kind. Since, further, adaptation effects have already been shown to depend on the number (Hillenbrand, 1975; Simon, 1977), intensity (Hillenbrand, 1975; Sawusch, 1977), and ISI (Simon, 1977) of the adaptors, we hoped to demonstrate by judicious juggling of these variables, the essential equivalence of the two paradigms.

As a preliminary, it was established in two other experiments (Simon, 1977) that significant adaptation effects could be induced for synthetic stop consonants at a 16 : 1 adaptor-to-test stimulus ratio and at an ISI of 1750 ms. We took the 16 : 1 ratio to be close to the maximum at which the anchor and adaptation paradigms could still be effectively distinguished. We took the 1750 ms ISI to be close to the minimum within which subjects could be expected to respond. With these values of ISI and anchor/adaptor ratio, we were in a position to construct two tests, identical in every respect except for the placement of the anchor/adaptor syllables.

### A. Method

#### 1. Stimuli

The stimuli were the series of correlated $F_0$-consonant–vowel syllables used in experiment I, reduced to 250 ms in duration by dropping the last 50 ms of the steady-state portion and making appropriate adjustments in the $F_0$ and overall amplitude contours.

Tests were prepared in the same way as for previous experiments. The control identification tests were the same for both anchoring and adaptation: seven stimuli recorded ten times each in a random order with 1750 ms

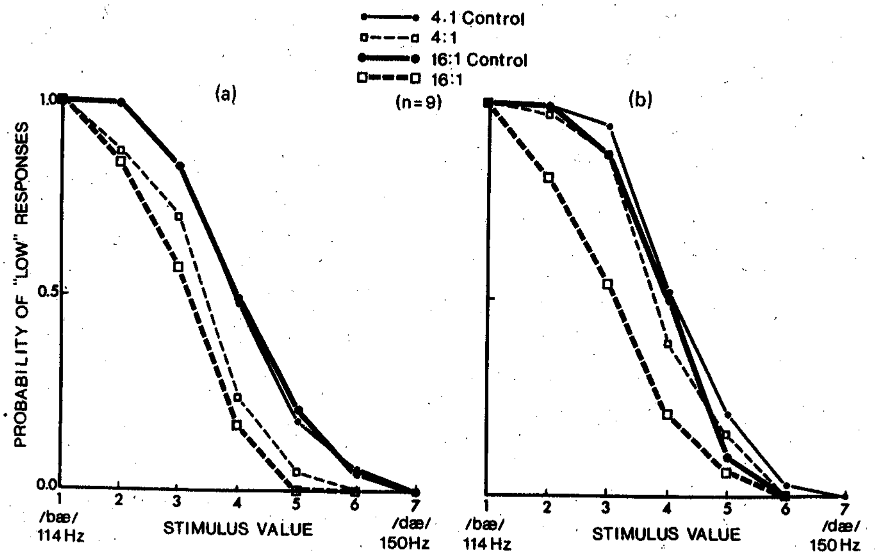


FIG. 4. The probability of "low" responses as a function of stimulus value, in the two control and two ratio conditions of experiment IV, for the pitch task, under (a) anchoring and (b) adaptation.

between stimuli and a 5-s pause after every 20th stimulus.

Two adaptation tests were prepared, with adaptor-to-test-stimulus ratios of 4 : 1 and 16 : 1. Each test consisted of ten blocks of the specified number of adaptors (either 4 or 16), followed by one presentation of each of the remaining six syllables on the continuum, to make a total of either 100 (4 : 1) or 220 (16 : 1) stimuli, separated by 1750 ms within a block and by 5 s between blocks.

Two anchor tests were prepared, identical in every respect to the adaptation tests, except that the repetitions of the anchor stimulus were distributed randomly among the six test syllables of each block.

## 2. *Procedure*

The procedure was identical in both the adaptation and anchor experiments. Subjects were required to identify all stimuli: both anchors/adaptors and test syllables, under both control and experimental conditions for both tasks (pitch and phoneme) at each ratio (4 : 1 and 16 : 1). Each subject heard the same tapes for each of the two task conditions. Only the instructions to the subjects varied as the task changed; they were identical with those of experiment I.

The adaptation tests were administered to all subjects before the anchoring tests, and the entire set of pitch tests was presented first, as in experiment I. Each experimental test (4 : 1 or 16 : 1) was preceded by its own control; the order of the tapes was counterbalanced within each paradigm and each task condition.

All experimental tapes were reproduced binaurally from the output of an Ampex AG 500 tape recorder over calibrated Telephonics (TDH-39) matched headphones with a circumaural seal. The 1000-Hz calibration tone was set to deliver a voltage across the phones equivalent to approximately 75 dB SPL *re* 0.0002 dyn/cm$^2$.

Nine listeners meeting the previously stated criteria participated in the experiments. They were tested singly or in pairs, in a quiet room, with a minimum of 24 hours and a maximum of a week between conditions.

## B. Results

Figure 4 displays the pitch group identification functions in the control and experimental conditions for anchoring (left) and adaptation (right). The experimental functions are appreciably lower than the control functions under both paradigms.

Table XII presents the total number of "low" responses

TABLE XII. Individual and mean numbers of "low" responses to the sixty presentations of stimuli 2–7 for the two ratio conditions and their associated controls under the anchoring and adaptation procedures of experiment IV.

| | Anchoring | | | | Adaptation | | | |
|---|---|---|---|---|---|---|---|---|
| Subject | Control | 4:1 | Control | 16:1 | Control | 4:1 | Control | 16:1 |
| 1 | 22 | 17 | 23 | 9 | 25 | 27 | 25 | 17 |
| 2 | 16 | 11 | 19 | 9 | 26 | 26 | 23 | 5 |
| 3 | 23 | 19 | 23 | 10 | 23 | 19 | 25 | 17 |
| 4 | 17 | 12 | 27 | 16 | 28 | 20 | 19 | 5 |
| 5 | 25 | 19 | 21 | 16 | 24 | 20 | 20 | 16 |
| 6 | 26 | 21 | 29 | 22 | 23 | 22 | 25 | 27 |
| 7 | 32 | 18 | 27 | 14 | 32 | 29 | 27 | 13 |
| 8 | 38 | 32 | 36 | 27 | 34 | 29 | 35 | 30 |
| 9 | 28 | 19 | 26 | 20 | 26 | 18 | 22 | 16 |
| Mean | 25.2 | 19.7 | 25.7 | 15.9 | 26.8 | 23.3 | 24.6 | 16.2 |

TABLE XIII. Individual and mean boundaries on the $F_0$ functions for the two ratio conditions and their associated controls under the anchoring and adaptation procedures of experiment IV.

| | Anchoring | | | | Adaptation | | | |
|---|---|---|---|---|---|---|---|---|
| Subject | Control | 4:1 | Control | 16:1 | Control | 4:1 | Control | 16:1 |
| 1 | 3.74 | 3.28 | 3.92 | 2.62 | 4.04 | 4.19 | 3.85 | 3.50 |
| 2 | 3.36 | 2.77 | 3.71 | 2.74 | 3.95 | 3.84 | 3.87 | 2.22 |
| 3 | 3.81 | 3.49 | 3.92 | 2.70 | 3.30 | 3.58 | 3.81 | 3.31 |
| 4 | 3.50 | 2.96 | 4.11 | 3.20 | 4.31 | 3.58 | 3.45 | 2.22 |
| 5 | 4.00 | 3.58 | 3.72 | 3.32 | 3.94 | 3.65 | 3.65 | 3.14 |
| 6 | 3.93 | 3.80 | 4.20 | 3.77 | 3.36 | 3.77 | 4.00 | 4.12 |
| 7 | 4.62 | 3.21 | 4.11 | 3.03 | 4.45 | 4.14 | 4.22 | 3.00 |
| 8 | 5.04 | 4.45 | 5.03 | 4.08 | 4.74 | 4.18 | 4.73 | 3.99 |
| 9 | 4.20 | 3.43 | 4.19 | 3.64 | 4.07 | 3.55 | 3.77 | 3.43 |
| Mean | 4.03 | 3.44 | 4.10 | 3.23 | 4.02 | 3.83 | 3.93 | 3.21 |

(excluding those to the anchor/adaptor stimulus) under the two control and the two ratio conditions for both paradigms. In the anchor condition, there is a significant mean decrease in "low" responses of 6.5 at the 4:1 ratio (matched pairs $t=6.24$, $p<0.0005$, one-tailed), and of 9.8 at the 16:1 ratio (matched pairs $t=4.23$, $p<0.005$, one-tailed). In the adaptation condition, there is a significant mean decrease in "low" responses of 3.5 at the 4:1 ratio (matched pairs $t=3.09$, $p<0.01$, one-tailed), and of 8.4 at the 16:1 ratio (matched pairs $t=4.12$, $p<0.005$, one-tailed).

Normal ogives were fitted to the pitch data, and the results are listed in Table XIII. In the anchor condition, the mean 4:1 control and experimental boundaries are 4.03 and 3.44, respectively, a significant shift of 0.59 continuum steps toward the anchor stimulus (matched pairs $t=4.84$, $p<0.005$, one-tailed). The mean 16:1 control and experimental boundaries are 4.10 and 3.23, respectively, a significant shift of 0.87 continuum steps toward the anchor stimulus (matched pairs $t=7.84$, $p<0.005$, one-tailed). In the adaptation condition, the mean 4:1 control and experimental boundaries are 4.02 and 3.83, respectively, a marginally significant shift of 0.19 steps (matched pairs $t=1.46$, $p<0.10$, one-tailed); the mean 16:1 control and experimental boundaries are 3.93 and 3.21, respectively, a significant shift of 0.71 steps (matched pairs $t=3.85$, $p<0.005$, one-tailed).

Figure 5 displays the group speech identification functions in the control and experimental conditions for anchoring (left) and adaptation (right). The experimental functions are clearly lower than the control functions at the 16:1, but scarcely at the 4:1 ratio, under both paradigms.

Table XIV presents the total number of "b" responses (excluding those to the anchor/adaptor stimulus) under the two control and the two experimental conditions for both paradigms. In the anchor condition, there is an insignificant mean decrease in "b" responses of 1.1 at the 4:1 ratio (matched pairs $t=1.26$, $p>0.10$, one-tailed), and a marginally significant mean decrease of 2.9 responses at the 16:1 ratio (matched pairs $t=1.69$, $p<0.10$, one-tailed). In the adaptation condition, there is a marginally significant mean decrease in "b" responses of 0.7 at the 4:1 ratio (matched pairs $t=1.40$, $p<0.10$, one-tailed), and a significant decrease of 4.5 responses at the 16:1 ratio (matched pairs $t=3.54$, $p<0.005$, one-tailed).

Normal ogives were fitted to the speech data, and the

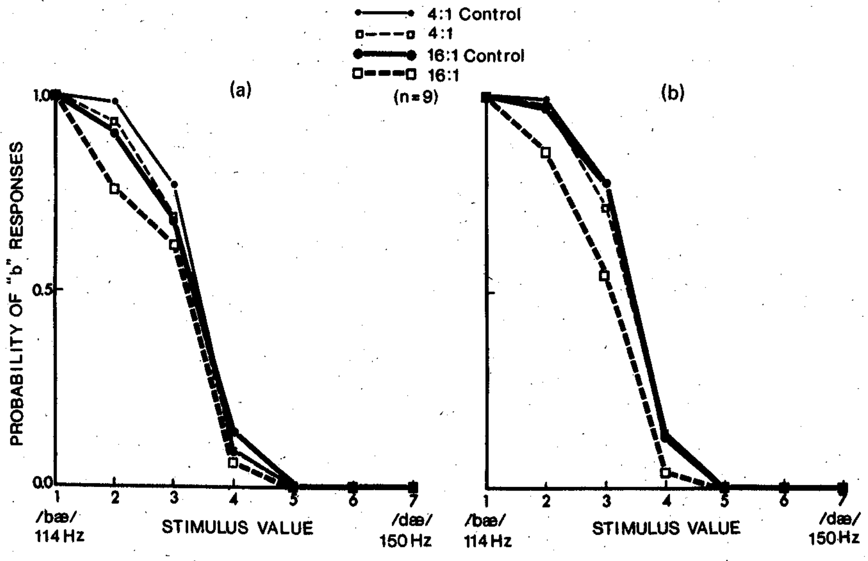


FIG. 5. The probability of "b" responses as a function of stimulus value, in the two control and two ratio conditions of experiment IV, for the speech task, under (a) anchoring and (b) adaptation.

TABLE XIV. Individual and mean number of "b" responses to the sixty presentations of stimuli 2–7 for the two ratio conditions and their associated controls under the anchoring and adaptation procedures of experiment IV.

| | Anchoring | | | | Adaptation | | | |
|---|---|---|---|---|---|---|---|---|
| Subject | Control | 4:1 | Control | 16:1 | Control | 4:1 | Control | 16:1 |
| 1 | 13 | 8 | 10 | 3 | 12 | 11 | 14 | 9 |
| 2 | 21 | 20 | 18 | 18 | 21 | 23 | 20 | 20 |
| 3 | 20 | 20 | 21 | 16 | 21 | 20 | 21 | 17 |
| 4 | 17 | 14 | 19 | 20 | 19 | 18 | 17 | 14 |
| 5 | 25 | 26 | 27 | 22 | 23 | 23 | 24 | 24 |
| 6 | 19 | 20 | 14 | 20 | 14 | 13 | 20 | 11 |
| 7 | 20 | 15 | 18 | 6 | 22 | 19 | 20 | 11 |
| 8 | 9 | 11 | 4 | 2 | 16 | 14 | 9 | 8 |
| 9 | 21 | 21 | 24 | 22 | 24 | 25 | 24 | 15 |
| Mean | 18.3 | 17.2 | 17.2 | 14.3 | 19.1 | 18.4 | 18.8 | 14.3 |

results are listed in Table XV. In the anchor condition the mean 4:1 control and experimental boundaries are 3.49 and 3.36, respectively, an insignificant shift of 0.13 continuum steps toward the anchor stimulus (matched pairs $t=1.23$, $p>0.10$, one-tailed). The mean 16:1 control and experimental boundaries are 3.33 and 3.07, respectively, a marginally significant shift of 0.26 continuum steps toward the anchor stimulus (matched pairs $t=1.57$, $p<0.10$, one-tailed). In the adaptation condition, the mean 4:1 control and experimental boundaries are 3.62 and 3.56, respectively, a significant shift of 0.06 steps (matched pairs $t=1.90$, $p<0.05$, one-tailed). The mean 16:1 control and experimental boundaries are 3.51 and 3.16 steps, respectively, a significant shift of 0.35 steps (matched pairs $t=2.52$, $p<0.025$, one-tailed).

Two within-subjects analyses of variance were carried out. The first examined the differences between decreases in the number of "low" and "b" responses as a function of ratio, paradigm, and task. The mean decrease in response due to the 16:1 ratio (6.3 responses, averaged across paradigms and tasks) was significantly greater than that due to the 4:1 ratio (3.0 responses) ($F=35.28$; df$=1, 8$; $p<0.01$). The mean decrease in responses on the pitch task (7.0 responses) was significantly larger than that on the speech task (2.3 responses) ($F=19.42$; df$=1, 8$; $p<0.01$). There was no significant effect of paradigm ($F=1.36$; df$=1, 8$; $p>0.25$) and no significant interactions.

The second analysis of variance examined the differences between boundary shifts as a function of ratio, paradigm, and task. The mean boundary shift due to the 16:1 ratio (0.55 continuum steps, averaged across paradigms and tasks) was significantly greater than that due to the 4:1 ratio (0.24 steps) ($F=12.11$; df$=1, 8$; $p<0.01$). The mean boundary shift on the pitch task (0.59 steps) was significantly greater than that on the speech task (0.20 steps) ($F=19.61$; df$=1, 8$; $p<0.01$). There was no significant effect of paradigm ($F=2.05$; df$=1, 8$; $p>0.10$), and there were no significant interactions.

In order to illustrate how the effects of anchoring and adaptation are distributed over the stimulus series, Fig. 6 replots the data of Figs. 4 and 5 in a format borrowed from Bailey (1973, 1975). The mean decrease in the probability of "low" or "b" responses, due to anchoring or adaptation, is plotted as a function of stimulus value, with task and ratio as parameters of the curves. Since there was no significant paradigm effect, anchoring and adaptation are combined. The arrows over the curves mark the positions of the mean boundaries estimated for the control conditions. The response shifts are obviously very much larger for pitch than for the phoneme, but

TABLE XV. Individual and mean boundaries on the speech functions for the two ratio conditions and their associated controls under the anchoring and adaptation procedures of experiment IV.

| | Anchoring | | | | Adaptation | | | |
|---|---|---|---|---|---|---|---|---|
| Subject | Control | 4:1 | Control | 16:1 | Control | 4:1 | Control | 16:1 |
| 1 | 3.17 | 2.57 | 2.72 | 1.97 | 2.92 | 2.79 | 3.03 | 2.62 |
| 2 | 3.82 | 3.67 | 3.42 | 3.52 | 3.82 | 3.92 | 3.67 | 3.67 |
| 3 | 3.67 | 3.67 | 3.82 | 3.14 | 3.82 | 3.67 | 3.72 | 3.37 |
| 4 | 3.17 | 3.06 | 3.50 | 3.67 | 3.58 | 3.53 | 3.37 | 3.03 |
| 5 | 4.00 | 3.77 | 4.07 | 3.87 | 3.92 | 3.92 | 3.96 | 3.96 |
| 6 | 3.50 | 3.67 | 3.27 | 3.67 | 3.22 | 3.17 | 3.65 | 2.77 |
| 7 | 3.65 | 3.11 | 3.42 | 2.32 | 3.88 | 3.82 | 3.65 | 2.77 |
| 8 | 2.61 | 3.02 | 1.89 | 1.63 | 3.44 | 3.23 | 2.61 | 2.95 |
| 9 | 3.82 | 3.72 | 3.84 | 3.88 | 3.96 | 4.00 | 3.96 | 3.28 |
| Mean | 3.49 | 3.36 | 3.33 | 3.07 | 3.62 | 3.56 | 3.51 | 3.16 |

TABLE XVI. Observed decreases percentaged against possible decreases in probability of "low" or "b" responses for combined anchoring and adaptation conditions of experiment IV at two ratios of anchor (or adaptor) to test stimulus.

| | Fundamental frequency ("low") | | Speech ("b") | |
|---|---|---|---|---|
| Stimulus | 4:1 | 16:1 | 4:1 | 16:1 |
| 2 | 6 | 16 | 3 | 13 |
| 3 | 9 | 35 | 9 | 20 |
| 4 | 36 | 62 | - | 57 |
| 5 | 45 | 81 | - | - |
| 6 | 100 | 100 | - | - |

for both tasks the shifts are confined to stimuli that were not identified with total consistency in the control conditions (that is, stimuli 2–4 for the speech task and 2–6 for the pitch task). Notice that, on every function, the absolute response shift increases with distance from the adaptor to a peak near the category boundary and thereafter declines. The decline is a floor effect entailed by the dichotomous scale: If the observed decreases in response probability are computed as percentages of possible decreases, monotonic increasing functions over the entire region of ambiguity result (see Table XVI).

Finally, Table XVII lists individual and mean standard deviations of the normal ogives fitted to the control identification data for the pitch and speech tasks under both paradigms. The grand mean standard deviation for pitch of 1.12 is slightly larger than the grand mean standard deviation for speech of 1.08, but one-way analysis of variance shows no significant effect of task ($F<1$). The Spearman coefficients of rank order correlation between control standard deviation and boundary shift were significant for the pitch task under anchoring at 4:1 ($\rho=0.63$, $p<0.05$) and under adaptation at 16:1 ($\rho=0.59$, $p<0.05$), but not under anchoring at 16:1 ($\rho=0.17$) or under adaptation at 4:1 ($\rho=0.33$). None of the coefficients was significant for the speech task under either anchoring (4:1, $\rho=-0.17$; 16:1, $\rho=0.10$) or adaptation (4:1, $\rho=0.48$; 16:1, $\rho=-0.37$).

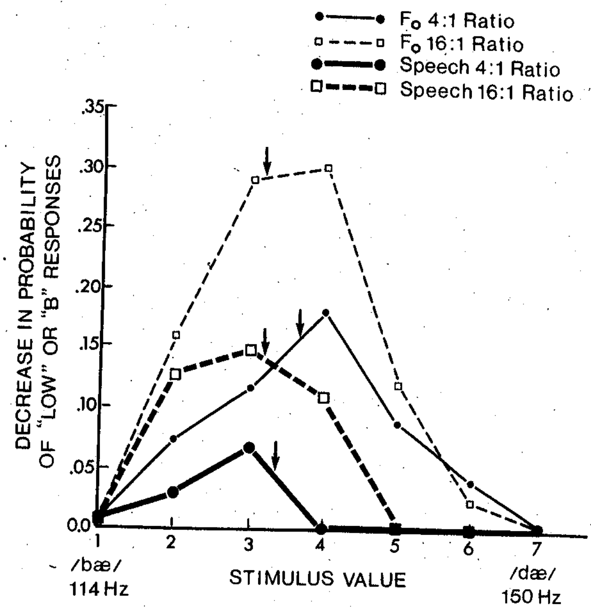


FIG. 6. Mean decrease in the probability of "low" or "b" responses as a function of stimulus value, combined over anchoring and adaptation conditions, for the pitch and speech tasks in experiment IV.

## VII. DISCUSSION

The effect of paradigm was not significant by analysis of variance and did not interact significantly with either task or ratio. Since, further, the effect of ratio was significant, without interactions, we may conclude that selective adaptation and psychophysical anchoring are equivalent procedures having similar effects on pitch and stop consonant identification, at least under the conditions of this experiment. The conclusion is weak, since the effect of anchoring on the consonant task was certainly marginal. Furthermore, the effect of adaptation on both tasks was probably reduced by the use of low ratios and a relatively long interval between adaptor presentations. Yet, taken with the small, but significant effect of anchoring on the consonants of experiment III, the results do suggest that the two paradigms differ in degree rather than in kind.

TABLE XVII. Individual and mean standard deviations of the normal ogives fitted to the control identification data for the $F_0$ and phoneme tasks of experiment IV.

| | $F_0$ | | | | Phoneme | | | |
|---|---|---|---|---|---|---|---|---|
| | Anchor | | Adaptation | | Anchor | | Adaptation | |
| Subject | 4:1 | 16:1 | 4:1 | 16:1 | 4:1 | 16:1 | 4:1 | 16:1 |
| 1 | 1.06 | 1.08 | 1.12 | 1.06 | 1.12 | 1.29 | 1.22 | 1.19 |
| 2 | 1.25 | 1.15 | 1.22 | 1.13 | 1.00 | 1.06 | 1.00 | 1.00 |
| 3 | 1.04 | 1.00 | 1.00 | 1.04 | 1.00 | 1.00 | 1.00 | 1.04 |
| 4 | 1.04 | 1.41 | 1.22 | 1.15 | 1.12 | 1.04 | 1.06 | 1.07 |
| 5 | 1.00 | 1.04 | 1.00 | 1.04 | 1.00 | 1.00 | 1.00 | 1.00 |
| 6 | 1.10 | 1.08 | 1.15 | 1.08 | 1.04 | 1.09 | 1.11 | 1.04 |
| 7 | 1.19 | 1.10 | 1.08 | 1.06 | 1.04 | 1.06 | 1.00 | 1.04 |
| 8 | 1.29 | 1.22 | 1.18 | 1.09 | 1.31 | 1.55 | 1.09 | 1.31 |
| 9 | 1.22 | 1.25 | 1.10 | 1.04 | 1.00 | 1.04 | 1.00 | 1.00 |
| Mean | 1.13 | 1.15 | 1.12 | 1.08 | 1.07 | 1.13 | 1.05 | 1.08 |
| Grand mean | | 1.12 | | | | 1.08 | | |

The effect of task was clearly significant, again without interactions. This result confirms for anchoring, and extends to adaptation, the finding of experiment I and of Sawusch, Pisoni, and Cutting (1974) that the category boundary on an arbitrary $F_0$ continuum is significantly more susceptible to movement than is the category boundary on a stop consonant continuum. We will consider the possible origins of this difference more fully below. Here we note merely that, although the difference evidently reflects a differential influence of test context, it cannot be predicted simply from the ambiguity of the stimulus series, as measured by the standard deviations of the fitted control ogives, since these were neither significantly different between tasks, nor reliably correlated with degree of boundary shift among subjects.

Finally, the lack of a paradigm effect, taken with the significant dissociation in anchoring effects between consonant and vowel tasks in experiment III, and between consonant and pitch tasks in the present experiment, show that anchoring may be no less "selective" than adaptation: its effects are on perceptually distinct dimensions of the syllable rather than on the syllable as a whole.

## VIII. GENERAL DISCUSSION

### A. Task differences

One unequivocal conclusion from these experiments is that a continuum of fundamental frequency is significantly more susceptible to the contextual effects of both anchoring and adaptation than a synthetic stop consonant place continuum, even when the two types of variation are carried, in perfect correlation, on the same syllables. If we may take the anchoring results of experiments II and III, and extrapolate from the lack of a paradigm effect in experiment IV, both intensity and vowels are also more susceptible to anchoring and adaptation than stop consonants. The addition of vowels to the list forbids us to attribute the different susceptibilities to differences between phonetic and nonphonetic dimensions. However, we defer further discussion of the task differences until we have considered possible processes underlying anchoring and adaptation.

### B. Psychophysical processes

If we reframe our unequivocal conclusion to make the point that a continuous, suprasegmental dimension of a syllable has proved even more susceptible to selective adaptation than a categorical, segmental dimension, the hypothesis that speech adaptation reflects fatigue of specialized feature detectors is not strengthened. For it was the facts of categorical perception that, at least in part, initially spurred the search for specialized detectors. Nor does the hypothesis gather strength from the finding that the selective effects of adaptation, which lent plausibility to the postulated fatigue of selectively tuned feature detectors, not only may result from the far weaker contextual manipulation of simple psychophysical anchoring, but may "select" fundamental frequency or intensity no less readily than consonantal place of articulation.

What, in fact, seems to be needed is a unified account that will handle both selective anchoring and adaptation of both phonetic and nonphonetic dimensions. In the following discussion we consider several standard psychophysical approaches, paying particular attention to how well they handle the speech adaptation data, for which the broader account seems most needed.

#### 1. *Adaptation level theory and frequency analysis*

According to adaptation level theory (Helson, 1964, 1971), a subject, invited to classify members of a stimulus series into two or more categories and lacking any explicit standard with which to compare them, takes the psychological midpoint of the series as his standard. This midpoint, or point of subjective equality (usually termed the "category boundary" in speech adaptation studies), proves to be some appropriately weighted measure of the central tendency of the stimulus series. If the series is skewed by addition of an extreme stimulus at some distance from an endpoint or by presentation of a particular stimulus more often than others, the weighted mean of the series is shifted and the entire distribution of the subject's responses shifts also to match the new mean. Thus, in the anchor/adaptor conditions of the present experiments, the increased presentations of the lower endpoint of the series gave a negative skew to the stimulus distribution, reducing both the mean of the stimuli, and, in theory, the matching mean of the subject's responses.

An alternative, closely related formulation is the frequency analysis of Parducci (1965, 1974) who argues that the general tendency of subjects, faced with an absolute judgment task for which they lack an explicit standard, is, "...to place the same number of stimuli in each of the available [response] categories" (Parducci, 1974, p. 134). If stimulus frequencies are unequal, subjects will distribute a frequent stimulus over several response categories and/or combine several rare stimuli into a single category. Thus, in the control conditions of the present experiments, subjects would tend to call half the stimuli "low" or "*b*" and the other half "high" or "*d*". In the anchor/adaptor conditions the increased presentations of the lower endpoint of the series would lead to a corresponding decrease in the number of times that the remaining stimuli were assigned to the "low" or "*b*" category.

While both adaptation level theory and frequency analysis correctly predict the direction of the changes in the subjects' responses, neither of them predicts their extent with any precision. In experiment IV, for example, the boundary shifts induced by the 4:1 and 16:1 ratios (Tables XIII and XV) are in the direction predicted by adaptation level theory, but clearly do not match the shifts in the mean of the stimulus series, whatever the weighting procedure. Similarly, the drops in the number of "low" and "*b*" responses induced by the two ratios (Tables XII and XIV) are predicted by a frequency analysis, but obviously fall far short of placing equal numbers of stimuli in each category.

Furthermore, since both theories, "...treat the context or frame of reference as a frequency distribution of stimulus values" (Parducci, 1974, p. 128), they both predict that changes in that distribution will be reflected by response shifts across its entire range. However, most speech adaptation studies, including the present one, have found that response shifts are confined to a few boundary stimuli toward the adaptor end of the continuum. Certainly, none has shown a shift of the entire response distribution.

In short, neither adaptation level theory nor frequency analysis can provide a satisfactory account either of the effects induced by anchoring and adaptation in the present series of experiments or of the effects reported in the speech adaptation literature.

### 2. *Response organization*

As we saw in the introduction, Sawusch and Pisoni (1973) and Sawusch *et al.* (1974), took the immunity of consonant continua to anchoring effects in their experiments as evidence that consonants are not subject to response bias and inferred that speech adaptation effects are therefore perceptual. However, reliable range effects on consonant continua, probably attributable to response bias, have been reported by Studdert-Kennedy (1976b). Furthermore, if we accept either the contention of Helson (1971) that anchoring effects are, in fact, sensory, or the conclusion of the present study that anchoring and selective adaptation are essentially the same process, the arguments of Sawusch and his colleagues lose their force.

Nonetheless, there is other, more direct evidence against a response bias interpretation of selective adaptation [see Cooper (1975), Ades (1976), and Eimas and Miller (1978) for reviews]. First are the cross-series effects, the repeated finding that boundary shifts may be induced by an adapting stimulus not drawn from the test series. These effects rule out response bias at the level of the syllable or phoneme. Second are all those studies in which the degree of adaptation declines as spectral overlap between adaptor and test series declines. Third is the evidence that the degree of adaptation varies with the extent to which the adapting stimulus is a good exemplar of the adapted category (Sawusch and Pisoni, 1976; Miller, 1977a; Cole and Cooper, 1977). Finally, asymmetries in the effects of adapting stimuli drawn from opposite ends of a continuum have been repeatedly reported, since the initial observation of Eimas and Corbit (1973) that their voiceless adaptors were more effective than their voiced adaptors. The combined weight of the evidence does not preclude response bias as a factor but certainly rules it out as a major determinant of adaptation effects.

### 3. *Perceptual sharpening*

A third hypothesis, briefly considered by Cole, Cooper, Singer, and Allard (1975), and the focus of experiments by Cole and Cooper (1977), and by Ainsworth (1977), would attribute selective adaptation to "retuning" or sharpening of the perceptual mechanism responsible for assigning stimuli to their categories: "...repeated listening to a single speech sound narrows the *range* of stimuli which are now acceptable as belonging to the same phonetic category as the adapting syllable" (Cole *et al.*, 1975, p. 243). Cole and Cooper (1977) falsified the hypothesis for speech by showing that a single sound was not required: they induced significant adaptation effects with a randomly repeated series of six syllables, spanning an entire phonetic category. Ainsworth (1977) reached a similar conclusion, and further grounds for rejecting this hypothesis come from any study demonstrating an effect of adaptation on the adaptor itself. Instances of this are cited below.

### 4. *Anchor contrast*

A fourth hypothesis is that repeated presentations of the anchor/adaptor stimulus establish an auditory ground with which ambiguous test stimuli contrast. For example, adaptation with a stimulus having a short VOT makes the VOT of a midrange stimulus seem longer than it is. Or, as was suggested in our discussion of experiment III, repeated presentation of a syllable with an initially rising $F_2$ may make the flat transition of a midrange stimulus seem to fall (cf. Blumstein, Stevens, and Nigro, 1977, Fig. 13, p. 1312).

There is no doubt that contrast can be a potent force in shaping absolute judgments of vowel continua (Fry *et al.*, 1962; Eimas, 1963) and of arbitrarily labeled nonphonetic continua (Sherif, Taub, and Hovland, 1958; Eimas, 1963; Helson, 1964). Recall, for example, the large contrast effect exerted on pitch judgments by immediately preceding stimuli in experiment I. Of course, we may also recall, from the same experiment, the complete absence of an effect on the stop consonants. But this may be an instance of the general resistance of consonants to context effects (which we will discuss below) rather than of their total immunity. In fact, contrast effects on judgments of both voicing and place of articulation in stop consonants were reported by Eimas (1963). More recently, Brady and Darwin (1978) manipulated the range of synthetic VOT stimuli within an identification test block and found that, "...the perceived voicing of a sound depends quite markedly on the range of other sounds presented before it.... The more voiced are the previous sounds, the more voiceless it will appear" (Brady and Darwin, 1978, p. 1557). Finally, our own place of articulation effects with a 4:1 anchoring ratio in experiment III can hardly be explained by any mechanism other than contrast.

If we are to extend the hypothesis to an account of speech adaptation, the results summarized in Fig. 6 and Table XVI are particularly important, because their form is typical of both speech adaptation and psychophysical contrast experiments. For example, Sherif *et al.*, (1958) studied contrast effects in the rating of weights against a fixed standard on a six-point scale. For stimuli close to the standard they found a modest shift in ratings toward the anchor value (assimilation), but beyond a certain point, ratings shifted away from the anchor (contrast) and the magnitude of the shift increased with the distance between the anchor and the stimulus being rated. Eimas (1963) reports the same increasing function in his analysis of contrast effects in

ABX triads drawn from a variety of phonetic and nonphonetic, visual and auditory continua. In the typical speech adaptation study, with its endpoint adaptor and two-point scale, assimilation of stimuli close to the adaptor could hardly be detected, even if it were occurring, since, already in the control condition, these stimuli are assigned to the same class as the adaptor. But for stimuli in the region of ambiguity between categories, where most adaptation effects occur, speech adaptation studies have invariably found that response shifts increase with distance from the adaptor (see Fig. 6 and Table XVI). Thus, a function that characterizes the results of experiments in simple psychophysical contrast turns out to characterize the results of speech adaptation studies as well.

A corollary of this increasing function is that the degree of adaptation should decrease as the stimulus selected for use as an adaptor is moved inward along the continuum (cf. Foreit, 1977, p. 347). Precisely this outcome (though usually taken to reflect the "tuning curve" of a detector) has been reported by Sawusch and Pisoni (1976), Miller (1977a), and Cole and Cooper (1977).

For the moment, then, we may reasonably entertain the hypothesis that speech adaptation effects on ambiguous stimuli in the middle of the stimulus range, the region to which most adaptation effects are confined, result from auditory contrast. The hypothesis is attractive, since it might comfortably gather speech and nonspeech under a single rubric. However, this cannot be the whole story because within-category adaptation effects also occur: as we shall see shortly, adaptation may even affect response to the adapting stimulus itself. This outcome could only result from contrast, if the contrast were between the adaptor and some property of the stimulus distribution, a possibility that we have already rejected in our discussion of adaptation level and frequency analysis.

### 5. *Fatigue*

A final hypothesis, and the one usually preferred in the speech adaptation literature, is that adaptation response shifts reflect fatigue, or desensitization, of a tuned detector. An obvious prediction from this model is that adaptation will shift the response probabilities for the adapting stimulus itself. Such an effect was, in fact, reported by Warren and Gregory (1958), using a quite different experimental paradigm ("verbal transformation"), later systematically elaborated by Goldstein and Lackner (1973) and by Lackner and Goldstein (1975). However, the standard procedure of most adaptation studies conceals the effect, if it is present, by using a closed response set and eschewing any estimate of sensitivity changes. This fact has not troubled feature detector theorists because the adaptor is typically close to the modal value of its category, a value to which the opponent detector is supposedly insensitive, so that, no matter how great the fatigue, the adapted detector will always have an output greater than that of its supposed opponent. Yet a direct test of the fatigue hypothesis is clearly crucial to feature detector accounts and has recently begun to draw experimental attention.

Miller (1975, 1977) has attacked the problem with an ingenious dichotic procedure. She has shown, for both labial/alveolar and voiced/voiceless oppositions, that adaptation with a good exemplar of a phonetic category reduces the effectiveness of that exemplar in dichotic competition with a good exemplar of an opponent phonetic category. Unfortunately, while this is precisely the outcome that a fatigue account predicts, it is also the outcome that the contrast hypothesis predicts: adaptation will increase the relative salience of the unadapted feature and, therefore, its probability of being correctly identified against the background of (that is, in dichotic competition with) the adapted feature.

However, Miller *et al.* (1977) have recently reported a more direct test of the hypothesis. The details of the experiment are complicated, but, briefly, they find that, in a three-choice /b, d, g/ identification test, using a good exemplar of each phoneme, adaptation with either /bæ/ or /gæ/ raises the intensity required to maintain a given probability of correctly identifying the adapting syllable. To our knowledge, this is the only direct demonstration of the drop in sensitivity predicted by the fatigue hypothesis.

Ample indirect support comes from the work of Sawusch (1976, 1977). He increased response sensitivity, by using a rating procedure rather than straight identification on a /b–d/ continuum, and repeatedly demonstrated significant drops, after adaptation, in the rated quality of an endpoint adaptor itself and of its near neighbors. Interestingly, these within category drops in rated quality occurred if the adaptor was a test series endpoint, but not if the three formants of the adapting syllable, though identical in structure to those of a test series endpoint, were raised above them by $1-1\frac{1}{2}$ critical bandwidths. In other words, a standard endpoint adaptor produced drops in the rated quality of stimuli both within and between categories, while a spectrally displaced adaptor produced rated quality drops only of stimuli between categories (Sawusch, 1977, experiment II). Furthermore, in this same experiment, Sawusch found only 50% interaural transfer of the adaptation effect for the standard adaptors, but 100% transfer for the spectrally displaced adaptors. From these and other results he inferred two levels of processing in selective adaptation: "...a frequency specific, peripheral auditory level and a more abstract, central, integrative level..." (p. 748). He suggested further that within-category response shifts may reflect fatigue of peripheral feature detectors, while between-category shifts may reflect central processes, either fatigue or, "...other mechanisms...possibly involving decision rules..." (p. 749).

Broadly, this account goes well with our earlier evidence and arguments. Apart from adaptation-level theory, frequency analysis or response organization, each of which we have already rejected for other reasons, there seems no alternative to "fatigue" as an account of the adapting effect of a stimulus on itself, and the data of Sawusch (1977) suggest that the account should be extended to adapting effects on close spectral neighbors. We need not concern ourselves at this point with just what is fatigued beyond remarking that, if the fatigue is indeed peripheral, "feature detectors" would not seem to be likely candidates. As for the second level, reflected by the between-category response shifts, our earlier arguments suggest that auditory contrast—a somewhat "abstract," presumably "central" and perhaps even "integrative" process—may be an important factor. However, other factors must also be involved. For example, Diehl (1975) and Sawusch and Pisoni (1976) have shown that the adapting effect of a syllable may be determined less by its acoustic structure or by experimental labeling instructions than by the phoneme category to which a listener assigns it. In other words, whatever role auditory contrast may play in speech adaptation, it plays within linguistic constraints.

## C. Phonemic anchoring and auditory contrast

The central assumption of all psychophysical accounts of anchoring or adaptation is that the frame of reference is established by the experiment itself. It is assumed that the subject has no prior standard with which to compare items presented for judgment and therefore derives a standard from the conditions of the experiment. While this assumption may be valid for judgments of arbitrarily labeled dimensions, such as brightness, weight or pitch, it clearly does not hold for judgments of consonantal status.

Many studies have demonstrated that a synthetic consonant continuum has a relatively firm perceptual structure, brought to it by the listener from his native language (for a review see Strange and Jenkins, 1978). A two-category continuum, for example, ranges from one or more acceptable tokens of a particular phoneme to one or more acceptable tokens of an opponent phoneme; these extremes sufficiently resemble naturally occurring types for them to be assigned to their categories with high consistency. Between the extremes occur one or more ambiguous stimuli of types that perhaps never occur naturally and may even be articulatorily impossible. These are the stimuli that succumb to varying degrees of contrast with other stimuli within the series.

Whether the surrounding fixed frame either limits or facilitates contrast between these ambiguous central stimuli and other portions of the continuum is not known. However, some suggestions come from a study by Donald (1976), independently replicated in all its essentials by Foreit (1977). As a test continuum Donald used a labial VOT series from −80 to +70 ms, a range that spans three phonological categories (with boundaries at roughly −20 and +25 ms) for speakers of Thai, but only two categories (with a boundary at roughly +15 ms) for speakers of English. Among Donald's adaptors were syllables with VOT values of −80 and +5 ms. The latter produced a significant shift in the +15/+25 ms boundary for both groups. A simple auditory contrast hypothesis would predict even greater shifts after adaptation with the more distant adaptor (−80 ms). In the event, the distant adaptor (−80 ms) produced the same shift as the near adaptor (+5 ms) for the English speakers who judged them to be tokens of the same phoneme, but for Thai speakers, who judged the distant adaptor to be a different phoneme than the near, and to be separated from the +25 ms boundary by an entire phonological category, the −80 ms adaptor produced no effect at all.

Are we forced to conclude, as does Foreit (1977) after reporting similar results, that, "...the effects of acoustic manipulations on selective adaptation are strongly limited by their linguistic implications..." (Foreit, 1977, p. 351), and that auditory contrast plays no role at all in producing boundary shifts? Perhaps; but consider a curious finding from Cole and Cooper (1977, experiment II). They constructed a four-item test series from /ǰa/ to /da/ by trimming the duration of the initial friction on a naturally spoken syllable (37, 29, 21, and 14 ms of friction). In a similar manner they constructed two adaptors, /či/ and /ǰi/, with friction durations of 120 and 60 ms. Since the 60 ms adaptor /ǰi/ produced a significant shift in the /ǰa–da/ boundary, we might reasonably predict either, by simple contrast theory, that the 120 ms adaptor (/či/) would produce an even greater shift or, by analogy with the Thai results reported above, that it would produce none at all. As it happens, both predictions are wrong: the two adaptors produced identical boundary shifts.

Although the cross-language differences remain to be explained, we may now be less confident that an intervening phonological category protects a phoneme boundary from adaptation effects or that different adaptors have the same effect because they are judged to be the same phoneme. In fact, phonological status may be irrelevant. Sarris (1967) has shown that contrast effects do not increase indefinitely with distance between standard and comparison: there is a critical distance beyond which they begin to decline. The equivalence of the near and distant adaptors for English speakers in these three studies may therefore, be the equivalence of symmetrical points on a parabola, with the peak of auditory contrast lying between them.

In short, while the preceived structure of a speech series probably influences the acoustic range over which auditory contrast occurs, we know essentially nothing about how it does so. Before we dismiss auditory contrast as a factor in speech adaptation, we need systematic parametric studies of contrast over acoustic ranges within which representative speech continua lie.

## D. The role of stimulus energy

We come finally to the question with which we began this discussion: Why were judgments of place of articulation in stop consonants less susceptible to anchoring or adaptation than judgments of pitch, loudness and vowels? We saw repeatedly in the reports of results that this fact could not be attributed to systematic differences in stimulus ambiguity as measured by judgment variability. Nor have we come upon any factor in our review of possible psychophysical processes that might be implicated.

Perhaps, in fact, the most likely source of the difference is accumulated anchor or adaptor energy. We already know that the degree of adaptation varies with the number of adaptors (experiment IV; see also Hillenbrand, 1975; Simon, 1977) and with the interval between them (Simon, 1977). Since we know further that the degree of adaptation increases with the intensity of the adaptor (Hillenbrand, 1975; Sawusch, 1977), it is reasonable to suppose that variations in duration, the other determinant of stimulus energy, should have an equivalent effect: the accumulation of anchor/adaptor energy over a test and the resulting degree of effect should be greater for long stimuli than for short. If this is so, we can explain the differences between pitch, loudness, and vowels, on the one hand, and stop consonants, on the other, by appealing to the relatively brief duration of the acoustic events that carried the consonantal information and the relatively long duration of the events that carried the pitch, loudness, and vowel information.

A virtue of this account is that it meshes neatly with acoustic memory explanations of consonant–vowel differences in other experimental paradigms. Stimulus duration has been shown to be a factor in the variance of discrimination (Fujisaki and Kawashima, 1969; Pisoni, 1973), precategorical acoustic storage (Crowder, 1973; Hall and Blumstein, 1977) and ear advantages in dichotic listening (Godfrey, 1974).

## IX. CONCLUSIONS

A fully unified account of the processes underlying the effects of anchoring and adaptation on phonetic and nonphonetic continua is hardly possible. For even if we rule out various psychophysical processes as factors in phonetic tasks, there is no reason why they should not contribute to effects on nonphonetic tasks. There seem, in fact, to be several processes, any or all of which may contribute in different degrees, depending on the paradigm and the type of task or continuum. As far as the two paradigms are concerned, they do not differ in principle, and, in practice, they are equally "selective." They do, however, differ in anchor/adaptor energy concentration, making "fatigue" more likely to occur under adaptation than under anchoring. Otherwise, there seem to be no grounds for distinguishing the paradigms in terms of the psychophysical processes that they activate.

The cardinal difference between speech and nonspeech continua is that the speech continuum has a fixed perceptual structure: it comes to the listener (or perhaps the listener to it) with its endpoints already anchored. Thus, only the ambiguous center of a speech continuum is liable to the contextual influences that may shape an entire nonspeech continuum. This means that speech is protected from those modes of contextual influence by which some parameter of the stimulus distribution, such as its mean or form, comes to serve as the reference for judgments. In short, speech is not subject to the contextual processes modeled by adaptation-level theory or range-frequency analysis.

The two portions of the speech continuum—its rigid ends and loose center—are subject to different main influences: fatigue and contrast. Since fatigue, or desensitization, is not peculiar to speech (see Elliot and Fraser, 1970, for a review), and since the auditory dimensions and values subject to fatigue may be no fewer than all discriminable properties of all sounds, it is misleading to characterize the fatigued entities as "feature detectors." A more neutral, descriptive term, such as "channels of analysis," is to be preferred (Kay and Matthews, 1972; cf. Eimas and Miller, 1978), particularly if the fatigue is indeed peripheral. By thus turning from a structural claim to a functional description, we resist the temptation to press a physiological metaphor in the absence of converging evidence. The two processes, peripheral fatigue and central contrast, may then be seen as contributing to adaptation effects in nonspeech no less than in speech.

## ACKNOWLEDGMENT

The experiments reported here are drawn from a thesis by the first author submitted to The Graduate School of The City University of New York in partial fulfillment of the requirements for the Ph.D. degree. The work was supported in part by NICHD Grant HD-01994 to Haskins Laboratories. (Portions of the paper were completed while the second author was on sabbatical leave as a research fellow at the Center for Interdisciplinary Research, University of Bielefeld, Germany.) For advice, discussion, and comments we thank Peter Bailey, Joanne Miller, David Pisoni, Bruno Repp, and Quentin Summerfield.

[1]The term "anchor" has several usages (see Parducci, 1974, p. 132). Most commonly it is used for a reference or standard stimulus outside the series being judged. However, we here follow Sawusch and Pisoni (1973) and use the term for a stimulus within the series to be judged, but assigned a higher frequency of occurrence.

[2]Although this procedure usually leads to the same conclusion as does a direct count of the decrease in the number of responses assigned to the anchor class, it also permits systematic measurement of response variability. Since the degree of response variability (that is, of stimulus ambiguity) may bear importantly on the interpretation of the results, and since boundary shifts are both easily read on the figures and have become standard measures in the adaptation literature, we report both direct and indirect measures throughout this paper.

[3]We are indebted to Katherine S. Harris for emphasizing this fact and for drawing our attention to the work of Volkmann.

Ades, A. E. (1974). "How phonetic is selective adaptation? Experiments on syllable position and vowel environment," Percept. Psychophys. 16, 61–66.

Ades, A. E. (1976). Adapting the property detectors for speech perception. In *New Approaches to Language Mechanisms*, edited by R. J. Wales and E. Walker (North Holland, Amsterdam).

Ades, A. E. (1977). "Source assignment and feature extraction in speech," J. Exp. Psychol. 3, 673–685.

Ainsworth, W. A. (1977). "Mechanisms of selective feature adaptation," Percept. Psychophys. 21, 365–370.

Bailey, P. (1973). "Perceptual adaptation for acoustical features in speech," Speech Perception (Dept. Psychol., The Queen's U., Belfast), 2, 29–34.

Bailey, P. (1975). "Perceptual adaptation in speech: Some properties of detectors for acoustical cues to phonetic distinctions," Ph.D. dissertation (Univ. of Cambridge), (unpublished).

Blumstein, S. E., Stevens, K. N., and Nigro, G. N. (1977). "Property detectors for bursts and transitions in speech perception," J. Acoust. Soc. Am. 61, 1301–1313.

Brady, S. A., and Darwin, C. J. (1978). "Range effect in the perception of voicing," J. Acoust. Soc. Am. 63, 1556–1558.

Cole, R. A., Cooper, W. E., Singer, J., and Allard, F. (1975). "Selective adaptation of English consonants using real speech," Percept. Psychophys. 18, 227–244.

Cole, R. A., and Cooper, W. E. (1977). "Properties of friction analyzers for [j]," J. Acoust. Soc. Am. 62, 177–182.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., and Gerstman, L. J. (1952). "Some experiments on the perception of synthetic speech sounds," J. Acoust. Soc. Am. 24, 597–606.

Cooper, W. E. (1974). "Contingent feature analysis in speech perception," Percept. Psychophys. 16, 201–204.

Cooper, W. E. (1975). "Selective adaptation to speech, in *Cognitive Theory I*, edited by F. Restle, R. M. Shiffrin, N. J. Castellan, H. Lindman, and D. B. Pisoni (Lawrence Erlbaum Associates, Hillsdale, NJ).

Cooper, W. E., Ebert, R. R., and Cole, R. A. (1976). "Perceptual analysis of stop consonants and glides," J. Exp. Psychol. 2, 92–104.

Crowder, R. G. (1973). "Precategorical acoustic storage for vowels of short and long duration," Percept. Psychophys. 13, 502–506.

Deihl, R. (1975). "The effect of selective adaptation on the identification of speech sounds," Percept. Psychophys. 17, 48–52.

Donald, S. L. (1976). "The effects of selective adaptation on voicing in Thai and English," Haskins Lab. Stat. Rep. Speech Res. 47, 129–135.

Dorman, M. F., Studdert-Kennedy, M., and Raphael, L. J. (1977). "Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues," Percept. Psychophys. 22, 109–122.

Durlach, N. I., and Braida, L. D. (1969). "Intensity perception, I. Preliminary theory of intensity resolution," J. Acoust. Soc. Am. 46, 372–383.

Eimas, P. D. (1963). "The relation between identification and discrimination along speech and non-speech continua," Lang. Speech 6, 206–217.

Eimas, P. D., and Corbit, J. D. (1973). "Selective adaptation of linguistic feature detectors," Cogn. Psychol. 4, 99–109.

Eimas, P. D., and Miller, J. L. (1978). "Effects of selective adaptation on the perception of speech and visual patterns: Evidence for feature detectors," in *Perception and Experience*, edited by R. D. Walk and H. L. Pick, Jr. (Plenum, New York).

Elliot, D. N., and Fraser, W. R. (1970). Fatigue and adaptation. In *Foundations of Modern Auditory Theory*, Vol. 1, edited by J. V. Tobias. (Academic, New York), 117–155.

Foreit, K. G. (1977). "Linguistic relativism and selective adaptation for speech: A comparative study of English and Thai," Percept. Psychophys. 21, 347–351.

Fry, D. B., Abramson, A. S., Eimas, P. D., and Liberman, A. M. (1962). "The identification and discrimination of synthetic vowels," Lang. Speech 5, 171–189.

Fujisaki, H., and Kawashima, T. (1969). "On the modes and mechanisms of speech perception," Annu. Rep. Eng. Res. Inst. (Univ. Tokyo), 28, 67–73.

Godfrey, J. J. (1974). "Perceptual difficulty and the right-ear advantage for vowels," Brain Lang. 4, 323–336.

Goldstein, L. M., and Lackner, J. R. (1973). "Alterations of the phonetic coding of speech sounds during repetition," Cognition 2, 279–297.

Hall, L. L., and Blumstein, S. E. (1977). "The effect of vowel similarity and syllable length on acoustic memory," Percept. Psychophys. 22, 95–99.

Helson, H. (1964). *Adaptation-Level Theory: An Experimental and Systematic Approach to Behavior* (Harper, New York).

Helson, H. (1971). "Adaptation-level theory: 1970 and after," in *Adaptation Level Theory*, edited by M. H. Appley (Academic, New York), 5–17.

Helson, H., and Kozaki, A. (1968). "Anchor effects using numerical estimates of simple dot patterns," Percept. Psychophys. 4, 163–164.

Hillenbrand, J. M. (1975). "Intensity and repetition effects on selective adaptation to speech," Res. Speech Percept. (Indiana University, Bloomington) 2, 56–137.

Kay, R. H. and Matthews, D. R. (1972). "On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency-modulated tones," J. Physiol. 225, 657–677.

Lackner, J. R. and Goldstein, L. M. (1975). "The psychological representation of speech sounds, Q. J. Exp. Psychol. 27, 173–185.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). "Perception of the speech code," Psychol. Rev. 74, 431–461.

Liberman, A. M., Harris, K. S., Kinney, J., and Lane, H. (1961). "The discrimination of relative onset time of the components of certain speech and nonspeech patterns," J. Exp. Psychol. 61, 379–388.

Lisker, L. L., and Abramson, A. S. (1964). "A cross-language study of voicing in initial stops: acoustical measurements," Word 20, 384–422.

Miller, J. L. (1975). "Properties of feature detectors for speech: Evidence from the effects of selective adaptation on dichotic listening," Percept. Psychophys. 18, 389–397.

Miller, J. L. (1977). "Properties of feature detectors for VOT: The voiceless channel of analysis," J. Acoust. Soc. Am. 62, 641–648.

Miller, J. L., Eimas, P. D., and Root, J. (1977). "Properties of feature detectors for place of articulation," J. Acoust. Soc. Am. 61, S48 (A).

Miller, J. L. and Eimas, P. D. (1976). "Studies on the selective tuning of feature detectors for speech," J. Phonetics 4, 119–127.

Parducci, A. (1965). "Category judgement: A range-frequency model," Psychol. Rev. 72, 407–418.

Parducci, A. (1974). "Contextual effects: A range-frequency analysis," in *Handbook of Perception* 2, edited by E. C. Carterette and M. P. Friedman (Academic, New York), 127–141.

Pisoni, D. B. (1973). "Auditory and phonetic memory codes in the discrimination of consonants and vowels," Percept. Psychophys. 13, 253–260.

Pisoni, D. B., Sawusch, J. R., and Adams, F. T. (1975). "Simple and contingent adaptation effects in speech perception," Res. Speech Percept. (Dept. Psychol., Indiana U., Bloomington), 2, 22–55.

Repp, B. H. (1977). "Dichotic competition of speech sounds: The role of acoustic stimulus structure," J. Exp. Psychol. 3, 37–50.

Repp, B. H., Liberman, A. M., Eccardt, T., and Pesetsky, D. (1978). "Perceptual integration of cues for stop, fricative and affricate manner," Haskins Lab. Stat. Rep. Speech Res. 53, 2, 61–83.

Sarris, V. (1967). "Adaptation-level theory: Two critical experiments on the Helson's weighted average model," American J. Psychol. 80, 331–355.

Sawusch, J. R. (1976). "Selective adaptation effects on end-point stimuli in a speech series," Percept. Psychophys. 20, 61–65.

Sawusch, J. R. (1977). "Peripheral and central processing in speech perception," J. Acoust. Soc. Am. 62, 738–750.

Sawusch, J. R., and Pisoni, D. B. (1973). "Category boundaries for speech and nonspeech sounds," J. Acoust. Soc. Am. 54, 76 (A).

Sawusch, J. R., and Pisoni, D. B. (1976). "Response organization and selective adaptation to speech sounds," Percept. Psychophys. **20**, 413–418.

Sawusch, J. R., Pisoni, D. B., and Cutting, J. E. (1974). "Category boundaries for linguistic and non-linguistic dimensions of the same stimuli," J. Acoust. Soc. Am. **55**, S55 (A).

Shankweiler, D. P., Strange, W., and Verbrugge, R. (1976). Speech and the problem of perceptual constancy. In *Perceiving, Acting and Comprehending: Toward an Ecological Psychology*, edited by R. Shaw and J. Bransford (Erlbaum, Potomac, Maryland).

Sherif, M., Taub, D., and Hovland, C. I. (1958). "Assimilation and contrast effects of anchoring stimuli on judgments," J. Exp. Psychol. **55**, 150–156.

Simon, H. J. (1977). Anchoring and selective adaptation of phonetic and nonphonetic categories in speech perception. Ph.D. dissertation (City University of New York), (unpublished).

Stevens, K. N., Liberman, A. M., Studdert-Kennedy, M., and Öhman, S. E. G. (1969). "Cross-language study of vowel perception," Lang. Speech **12**, 1–23.

Strange, W., and Jenkins, J. J. (1978). "The role of linguistic experience in the perception of speech," in *Perception and Experience*, edited by R. D. Walk and H. L. Pick, Jr. (Plenum, New York).

Studdert-Kennedy, M. (1976a). "Speech Perception," in *Contemporary Issues in Experimental Phonetics*, edited by N. J. Lass (Academic, New York), 243–293.

Studdert-Kennedy, M. (1976b). "Stimulus range as a determinant of phoneme boundaries along synthetic continua," J. Acoust. Soc. Am. **60**, S92(A).

Volkmann, J. (1951). "Scales of judgment and their implications for social psychology," in *Social Psychology at the Crossroads*, edited by J. H. Rohrer and M. Sherif (Harper, New York).

Warren, R. M., and Gregory, R. L. (1958). "An auditory analogue of the visual reversible figure," Am. J. Psychol. **71**, 612–613.

Woodworth, R. S. (1938). *Experimental Psychology* (Holt, New York).