INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR

77-20,522

SIMON, Helen J., 1940-
ANCHORING AND SELECTIVE ADAPTATION
OF PHONETIC AND NONPHONETIC CATEGORIES
IN SPEECH PERCEPTION.

City University of New York, Ph.D., 1977
Psychology, experimental

Xerox University Microfilms, Ann Arbor, Michigan 48106

ANCHORING AND SELECTIVE ADAPTATION OF PHONETIC AND NONPHONETIC CATEGORIES IN SPEECH PERCEPTION

by

Helen J. Simon

A dissertation submitted to the Graduate
Faculty in Speech and Hearing Sciences in
partial fulfillment of the requirements
for the degree of Doctor of Philosophy
The City University of New York.

1977

This manuscript has been read and accepted for the Graduate Faculty in Speech and Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

May 10th 1977
date

Michael Studdert-Kennedy
Chairman of Examining Committee

May 10, 1977
date

Irving Hochberg
Executive Officer

Katherine S. Harris
Harry Levitt
David B. Pisoni
Supervisory Committee

The City University of New York

ABSTRACT

Recent experiments in speech perception have shown that category boundaries along a synthetic stop consonant continuum can be shifted in a selective adaptation paradigm, but seem immune to the psychophysical effects of anchoring. With nonspeech stimuli, boundary shifts can be obtained in both paradigms. The finding in anchoring studies, that consonantal category boundaries are not subject to the same context effects as boundaries between fundamental frequency categories, has been used to dispute accounts of selective adaptation in terms of response bias or response organization. The selective adaptation effects for speech have then been interpreted as evidence for the fatiguing of feature detectors, tuned to acoustic and/or phonetic segments.

In this series of experiments, the two paradigms, anchoring and selective adaptation, were compared, using phonetic and nonphonetic categories. Experiments I, II, and III compared anchoring effects on the identification of a synthetic stop consonant continuum with those on the identification of correlated fundamental frequency, intensity and vowel continua carried on the same synthetic syllables. The results of Experiments I and II demonstrated category boundary shifts on the fundamental frequency and intensity continua as a function of variations in the probability of occurrence of the stimulus anchor, but no such changes on the consonant continuum. Experiment III, however, demonstrated category boundary shifts for both vowel and consonant continua.

Experiments IV and V were designed to examine the effects of

varying total adaptor energy in selective adaptation along a synthetic stop consonant continuum. Experiment IV considered the effect of the number of adaptors presented in the adaptation sequence preceding each test sequence, with repetition rate held constant: over the range from 8 to 32 adaptors, the magnitude of the phonetic boundary shift was found to be a linear function of the logarithm of the number of adaptors. Experiment V explored the effect of variations in the repetition rate (or density) of adaptors with the number held constant. A non-monotonic relation was found between the phoneme boundary shift and the inter-adaptor-interval: a greater shift was observed for a 750 msec interval than for either a 250 msec interval or a 1750 msec interval. These low-level stimulus energy variables (adaptor number and repetition rate) affect the magnitude of the phonetic boundary shift in what may be a trading relationship.

Experiment VI compared the effects of selective adaptation and anchoring on the correlated fundamental frequency and stop consonant continuum of Experiment I. The independent variable was the placement of the adapting and anchoring stimuli: in a single block immediately before a test sequence (adaptation) or randomly distributed among the test syllables (anchoring). Both procedures yielded significant boundary shifts on both continua, although the adaptation effects were systematically greater than the anchor effects. This outcome suggests that the difference between the two procedures lies simply in the distribution of adaptor/anchor stimulus energy and is therefore one of degree rather than kind.

Finally, although the effects of both procedures were systematically greater on the fundamental frequency than on the stop consonant continuum, there were significant effects on both continua. This outcome suggests that speech/nonspeech differences are also of degree rather than of kind. Since it is implausible to attribute pitch perception to the operation of feature detectors, and since it is unparsimonious to posit a different mechanism of selective adaptation for nonspeech than for speech, the results of these experiments call into question the claim that selective adaptation effects on speech continua reflect the operation of specialized feature analyzing devices.

ACKNOWLEDGEMENTS

I would first like to thank my major advisor and chairman, Michael Studdert-Kennedy, for his dynamic intellect and teaching ability, patience and insight. His guidance and critical comments in designing the study and the preparation of the manuscript were invaluable. He introduced me to speech perception and made me appreciate that it is "special." Throughout my studies, the other members of my committee, Katherine S. Harris and Harry Levitt, were instrumental in promoting my intellectual growth and development, and gave freely and generously of their time and help.

I am grateful to Alvin M. Liberman, Director and President of Haskins Laboratories and Franklin S. Cooper, Past President and currently Associate Director of the Laboratories, for making the facilities available to me. Special thanks must go to the staff at Haskins Laboratories: to Alice Dadourian for her support, friendship, and East 60's address; to Nancy O'Brien for her unparalleled support, instructions and supervision in the ways of RUNOFF, and marvelous organizational abilities; to Agnes McKeon for her suggestions and help in the preparation of the figures; and to Sabina Koroluk for her suggestions and aid in the preparation of the tables and the earlier drafts of this manuscript, and above all, to my typist, the PDP 11/45, who only failed me once. Special thanks must go to Donald Hailey, for maintaining calibration of the equipment.

The years at Haskins Laboratories have been a rewarding and enriching experience. There is insufficient space to name all those staff members and fellow students who have added to my experience at the Laboratories. Peter Bailey and Quentin Summerfield supported, suggested, and designed computer programs guaranteed to lessen the "fatigue" in adaptation experiments. My thanks must also extend to Fredericka Bell-Berti, Terry Halwes and Harriet Greisser, who were always receptive to questions and discussions, at all levels.

I owe special thanks to Douglas Kahn, who always had faith in my abilities and sustained me through the most difficult of times, and to my family and friends, especialy John Goodlet, who encouraged my preoccupation with my work.

This research was partially supported by a grant to Haskins Laboratories from the National Institute of Child Health and Human Development (HD-01994).

LIST OF TABLES

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (cont'd)

TABLE OF CONTENTS

PAGE

CHAPTER

TABLE OF CONTENTS (cont'd)

TABLE OF CONTENTS (cont'd)

PAGE

CHAPTER

IV. EXPERIMENT III. ANCHORING OF CORRELATED VOWEL AND STOP CONSONANT CONTINUUM

V. EXPERIMENT IV. ADAPTATION AND THE NUMBER OF ADAPTOR REPETITIONS

VI. EXPERIMENT V. ADAPTATION AND INTER-ADAPTOR-INTERVAL

TABLE OF CONTENTS (cont'd)

TABLE OF CONTENTS (cont'd)

CHAPTER I: INTRODUCTION

Analyzing the Acoustic Speech Signal

Current theories in speech perception have sought to explain the process by which the acoustic speech signal is analyzed into segmental phonetic percepts at the level of the linguistic message (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967; Stevens and House, 1972). Although these phonetic segments can be transcribed as alphabetic symbols, are perceptually distinct, and are described in terms of their articulation, they are abstract in that they are not inherent in the speech signal. Research in speech perception has sought to define the relation of these phonetic segments to the physical signal. Theorists have recently used an information processing model to account for the extraction of these segments from the sound wave.

These models assume that the perceptual processing of the speech signal can be divided into a hierarchy of levels (Studdert-Kennedy, 1974, 1976; Pisoni, 1975; Wood, 1975; Cutting, 1976). These levels are generally taken to be auditory, phonetic, phonological, syntactic, and semantic. The concern of this discussion will be with the processing of the signal at the auditory and phonetic levels.

The first level of processing is generally assumed to process all auditory information, whether speech or non-speech, in the same way. This preliminary auditory analysis stage automatically transforms the acoustic waveform into a spectral display of frequency, intensity, time, and durational information. The second stage of auditory analysis is the first stage of the sensory analysis and the recognition process (Pisoni, 1975) and is assumed to extract spectral and temporal auditory features from the Preliminary Auditory Analysis (Stevens, 1972; Pisoni, 1975). The type of information extracted from the first stage by these specialized detectors is the presence, absence, bandwidth, intensity, duration of noise, transition and fundamental frequency information, and spectral change information (Pisoni, 1975). These attributes of the sound wave are termed acoustic features. When these features of the sound wave have perceptual or linguistic significance, they are called acoustic cues.

The third stage is assumed to map many auditory features from the previous stage into single invariant abstract features. In addition, more than one abstract feature may be cued by any one auditory feature (Fant, 1962; Lisker and Abramson, 1967; Liberman, 1970). The question of whether the output of this abstract level should be treated as a phonetic feature or an auditory feature is moot. Pisoni (1975) describes this level as based on decision rules, the output of which is a set of abstract phonetic features, while Sawusch (1976) contends that this level is auditory only, abstracting particular auditory patterns from the secondary auditory stage. At this least abstract level (phoneme-sized phonetic segment level being the most abstract), according to Stevens and House (1972), the phonetic features are related to the overt attributes of speech by a set of rules describing the behavior of the vocal mechanism. Linguists have proposed 20-30 phonetic features to describe the speech segments (Stevens and House, 1972). The features describe the basic dichotomy between consonant and vowel, the place of articulation (e.g. labial, apical, velar), the presence or absence of secondary vocal tract constrictions, the manner of articulation (e.g. stop, nasal), and the type of vocal tract excitation (e.g.ivoiced, voiceless).

The next process is assumed to transform the features (or complex of cues) from the previous stage into an abstract phonetic feature matrix (Chomsky and Halle, 1968). Whether this process occurs in parallel with the abstract auditory combination above is still not clear. However, this information is processed at the higher levels of linguistic analysis (phonological, syntactic, semantic).

Experimental Issues in Speech Perception

Two basic problems have emerged in the study of speech perception over the past thirty years: the problems of invariance and segmentation. The first refers to the lack of invariance between the acoustic signal and the perceived message (Liberman, Delattre, and Cooper, 1952; see Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967, for a review). In general, both consonant and vowel features depend upon context; different acoustic cues may then signal the same phonetic percept and the same acoustic cues may signal different percepts.

Two classic Haskins' Laboratory experiments illustrate these findings. First, Liberman, Delattre, and Cooper (1952) showed that a single acoustic cue, a burst at 1440Hz, gave rise to different percepts depending on the vowel. The 1440Hz burst was perceived as /p/ if the following vowel was /i/ or /u/, but was perceived as /k/ if the vowel was /a/. Second, Liberman, Delattre, Cooper, and Gerstman (1954) showed that two different acoustic cues, a rising versus a falling second formant transition gave rise to the same percept. A rising second formant transition was perceived as /d/ before the vowel /i/, while a falling second formant transition, was perceived as /d/ before the vowel /u/. Therefore, different acoustic cues may give rise to the same phonetic percept.

The second basic problem in speech perception, that of segmentation, is closely related to the invariance problem. Studies of speech spectrograms revealed that few segments of the acoustic message correspond to perceived segments of the message and that single segments of the acoustic signal were found to carry information about several successive segments of the message (Fant, 1962). These departures from a one-to-one relationship between the acoustic and phonetic segments are assumed to be the result of co-articulation (Fant, 1968; Liberman, 1970). Variables such as phonetic context, stress, speaking rate, juncture, idiosyncracies of the speaker and language, have been shown to influence the acoustic structure of a given phone (Lindblom, 1963; Kozhevnikov and Chistovich, 1965; Öhman, 1966; Ladefoged, 1967).

The resulting acoustic signal, therefore, cannot be described as a sequence of discrete components. This lack of a mapping system of discrete acoustic events onto discretely perceived phonetic properties led to the description of speech as a complex acoustic code, continuously changing in pattern (Liberman, Cooper, Shankweiler, and Studdert-

Kennedy, 1967). The speech message is therefore encoded in the sound stream and decoded in perceptual analysis. The process by which these contextually varying cues are perceptually decoded into an invariant linguistic unit is one of the central issues in acoustic phonetics.

Cue and Feature Combination

Since there is no one-to-one relationship between the acoustic features and the phonetic segments, cue and feature combinations are fundamental to any information processing model of speech perception. Early work on the exploration of the acoustic cues raised the notion of cue combination in the perception of stop consonants.

That a complex of acoustic cues underlies phonetic feature distinctions for place of articulation was first suggested by Cooper, Delattre, Liberman, Borst, and Gerstman (1952) where the second-formant transition information was needed in addition to the frequency of the preceding burst for the identification of /p/ and /k/. It was also found that both friction and vocalic information were needed to distinguish between /f/ and /θ/ (Harris, 1958). While the second-formant transition cue was found to be nearly sufficient for the place distinction among the voiced stops (Cooper, Delattre, Liberman, Borst, and Gerstman 1952), Harris, Hoffman, Liberman, Delattre, and Cooper (1959) revealed that the third-formant transition cues are independent of the first- and second-formant transitions to which they are added, and combine with the other cues in an additive fashion to enhance perception.

Hoffman (1958) continued the investigation by varying the burst frequency as well as the second- and third-formant transitions. Although he found that the cues appear to be independent, he suggested that the addition of cues was vectorially summated. For example, if a cue for /d/ is added to a cue for /g/, perception is equivocal. However, if the same cue for /d/ is added to a sound already equivocally perceived between /d/ and /g/, a /d/ is perceived as the result of the summation of the two vectors.

Trading relations such as those among the acoustic cues for place of articulation have also been shown for the voiced-voiceless distinction. In a series of experiments that manipulated the acoustic parameters of synthetic stimuli, Liberman, Delattre, and Cooper (1958) first observed that a cue to voicelessness was the delay in the onset of the first formant relative to the higher formants. A further cue was the aspiration of these higher formants during this first formant "lag." Measurements on spoken words (Lisker and Abramson, 1964) found that the voice-onset time (VOT) (the time between consonant release and the beginning of glottal pulsing) for the consonants in the pre-stressed initial position were from 0-20 msec for the voiced consonant and 50 msec or more for the voiceless consonant depending on the place of articulation.

Voice onset time is considered to be the underlying articulatory variable and a primary acoustic cue for voicing distinctions in the initial position. However, several additional acoustic cues have been shown to contribute to the percept, including, for example, the duration of the first-formant transition after the onset of voicing. As VOT is increased from 0 msec in a voiced stop, the onset of the voicing is delayed. At the same time, the transition of the first-formant is shortened until, in voiceless stops, it is no longer present. Experiments independently manipulating these two cues have found a trading relationship between them (Stevens and Klatt, 1974; Summerfield, 1974; Summerfield and Haggard, 1974; Lisker, 1975). They have found that short duration of the first formant transition cued a voiceless stop when paired with a 30 msec cutback (VOT), but a longer transition needed a longer cutback to cue voicelessness.

Independently, Lisker (1975) and Summerfield (1975) suggested that a low onset frequency of the first formant rather than a voiced F_1 transition, may be a requirement for the voiced stops. Summerfield and Haggard (In press), manipulated F_1 onset frequency and F_1 transition duration orthogonally. They confirmed that the onset frequency of F_1 is a major cue to voicing, not the transition itself nor the overall distribution of stimulus energy. The trade-off relation is that the lower the frequency of the onset of F_1, the longer the VOT required to cue voicelessness and vice versa.

Stevens (1973, 1975) attempted to conceptualize the problem of cue combination by proposing a set of "auditory property detectors" that would classify the place of articulation for most consonants in the pre-stressed position by integrating burst and transition cues. Stevens (1973, 1975) proposed a rising spectral frequency detector for the labials, a falling frequency detector for the alveolars, and either a diverging frequency detector or a combination of the first two for the velars. These proposed property detectors are possible examples of the kind of mechanism by which multiple acoustic features such as bursts and transitions would signal one abstract phonetic (or auditory) feature. One auditory feature also might cue more than one abstract feature or property. These detectors would also be context independent for the most part, and, therefore, relatively invariant.

The Possible Role of Feature Detectors in Speech Perception

The interest in the possible role of "feature detectors" in speech perception as devices for extracting the relevant information from the acoustic signal grew out of several findings and theoretical positions. Liberman, Cooper, Shankweiler, and Studdert-Kennedy (1967), in a discussion of the contextually-dependent acoustic cues specifying a particular phonetic feature, drew attention to the feature analyzing devices discovered in animals (see below) and suggested that such devices might form part of a complex auditory processor or decoder to deal with the invariance problem. Abbs and Sussman (1971) specified the acoustic parameters to which such feature detectors might be sensitive: frequency, intensity, and rates and durational characteristics of frequency and intensity change. In a further refinement, Stevens (1973) as previously discussed, suggested a form of global "property detector," sensitive to a combination of cues such as bursts and transitions, as a mediator between the acoustic waveform and the phonetic message.

Several authors have attempted to relate broadly tuned feature analyzers to the results of selective adaptation studies. For example, as a result of their first adaptation experiment in speech perception (discussed in detail below), Eimas and Corbit (1973) suggested that the existence of linguistic feature detectors would: 1. be consistent with the theoretical notion of distinctive features in speech; 2. provide an example of a complex linguistic analysis analogous to that described in the visual system; 3. provide an explanation for the phenomenon of categorical perception; and, 4. suggest the innateness of feature detecting systems.

With respect to the supposed innate detectors, it has been theorized that infants must have some innate mechanism to extract the acoustic cues that are important in making phonetic distinctions from the speech signal. Stevens and House (1972) note that the same features occur repeatedly from one language to another and that children learn rules for manipulating these features after exposure to a relatively small number. This constitutes evidence that the auditory system and central processes are capable of encoding and decoding sounds as segments and features. This implies both that the predisposition of the auditory system towards the gestures used in speech processing may be innate and that certain analytic processes must be learned (Stevens and House, 1972; Stevens, 1975). Stevens (1975), further suggests that the child first uses simple context-independent property detectors for categorizing consonants and later uses these detectors as a base for the acquisition of a set of context-dependent cues to perceive phonetic features. He thereby resolves both the problems of infant speech perception and the problem of invariance at a single blow. Studies of infant perception, measured by non-nutritive sucking (Eimas, Siqueland, Juscyk, and Vigorito, 1971; Morse, 1972; see Eimas, 1974 for a review) and heart rate deceleration (Moffit, 1971) have, in fact, found that as early as one month, infants are able to make distinctions between voiced and voiceless sounds and among places of articulation, paralleling those observed in adults.

In short, there is evidence that some form of feature-detecting system mediates between the acoustic signal and the phonetic feature, and research in speech perception is now attempting to define the nature of these features more exactly.

"Feature Detecting" Systems in Animals

The introduction of microelectrode study of individual cells at various levels in the visual and auditory system in animals prompted speculation on feature-detecting mechanisms (cells selectively sensitive to particular features of complex stimuli) in both sensory modalities. Recording from the visual ganglion cells of the frog, Lettvin, Maturana, McCulloch, and Pitts (1959) and Maturana, Lettvin, McCulloch, and Pitts (1959) were the first to demonstrate selectively responsive stimulus units. Three classes of stimulus-specific units were sensitive to general characterisitcs of the visual world: edge detectors responsive to the border between light and dark regions, moving contrast detectors responsive to moving edges, and dimming detectors responsive to the lowering of overall illumination. The fourth class of detector was the convex edge or "bug" detector, responsive only to small, dark objects moving into the visual field.

Hubel and Wiesel (1962, 1965), recording from the visual cortex of the cat, described single units with simple, complex and hypercomplex receptive fields (the region of the retina over which one can influence the firing of that cell). The simple cells were most responsive to an edge or bar (edge and line detectors). The complex cortical receptive field cells were responsive to a stimulus moving over the retina while the hypercomplex cortical cells appeared to be the output of combinations of complex cells that coded combinations of stimulus features.

Although auditory analogues to the visual feature detector systems have been demonstrated, they have not produced such relatively direct evidence as their visual counterparts. This is largely due to the problems involved in using complex acoustic stimuli and to the differences in the anatomical organizations of the visual and auditory systems (see Evans, 1974 for a review). Evans and Whitfield (1964) and Whitfield and Evans (1965) studied single units in the primary auditory cortex of unanaesthetized, unrestrained cat. They found units responsive to amplitude and frequency modulation (rate and direction of change). Nelson, Erulkar, and Bryan (1966) and Suga (1965) have shown related findings in the activity patterns of units in the inferior colliculus of cat and bat respectively. Møller (1969) found units responsive to the direction of frequency change as peripheral as the cochlear nucleus. Whitfield (1967) equates such directionally specific (rather than frequency specific) units to the previously discussed units in the visual cortex responsive to direction of movement and orientation of lines. Concerning the auditory system, Whitfield states that units sensitive to more complex patterns but not to single pure tones whether amplitude or frequency modulated, probably exist because of the large number of units in the auditory cortex responsive to complex noises. Whitfield (1965) also indicates the possible relevance of such units for formant tracking in speech.

One of the first studies to investigate communicatively significant sounds was that of Frishkopf and Goldstein (1963). They found two main spectral components of the mature male bullfrog's mating call at 300 and 1500 Hz that corresponded with the characteristic frequencies of the two groups of neurons in the frog's auditory nerve, but not in the cortex. However, related studies in the cricket frog (Capranica, 1965) found characteristic frequencies of auditory neurons in the medulla that corresponded to the frequency spectrum of the call of male cricket frogs from the same geographical region.

Recent studies by Funkenstein, Nelson, Winter, Wollberg, and Newman (1971) reported results that suggest cortical cell specialization for detection of certain features relevant to the recognition of squirrel monkey calls. The highly developed repertoire of meaningful calls has been spectrographically and behaviorally divided into five distinct groups. Many neurons in the primary auditory cortex of the monkey were responsive to certain of these vocalizations consisting of patterns of formants with characteristic frequencies and transitions.

Also interesting was the finding that units that did not respond maximally to the monkey calls responded to acoustically similar calls (frequency modulated or noise). Electronic elimination of portions of the call resulted in response patterns with correspondingly absent portions. The different responses of various cells to the same vocalization suggests that more than one feature may be recognized in each cell. However, all cells responding to calls in the absence of a response to noise and tone, and most of the cells responding more often to the vocalizations than the noise and tone, were noted to respond either to a single call type or a single class of calls.

"Feature Detecting" Systems in Humans

Since microelectrode techniques are not possible with human subjects, a psychophysical procedure was needed to test for "detectors" in the visual and auditory system of humans. One of the earliest methods to obtain information about the operation of feature extraction systems in vision was to prolong the stimulation to the eyes that changes subsequent perceptions (Kohler and Wallach, 1944). The effects of these after-images provided cues and suggestions to the properties and operations of the feature extraction system in vision. The prolonged exposure to a stimulus is the basis for a great deal of recent experimentation in vision, audition, and now, speech perception.

The effects of the adaptation paradigm as first used in vision have been found for orientation and spatial frequency (Blakemore and Campbell, 1969), size (Blakemore and Sutton, 1969), and color (McCullough, 1965), to name a few. A threshold is determined before and after prolonged exposure to a stimulus consisting of a particular visual feature and the typical finding is that sensitivity to that feature is reduced by the exposure. For example, Blakemore and Sutton, (1969) found that after prolonged exposure to a pattern with stripes of a certain width, narrow bars appeared thinner and broad bars appeared wider than they had appeared before prolonged exposure. They suggested that the prolonged exposure to one of two complementary systems, "fatigues" or "adapts" the responsiveness of the neural mechanism. The opposing system or "channel" becomes more active in relation to its "fatigued" antagonist, thereby causing a threshold change.

In audition, Kay and Matthews (1972) adapted human subjects with frequency modulated (FM) and amplitude modulated (AM) tones. Adapting with an FM tone raised the threshold at the adapted modulation frequency and at neighboring frequencies. However, while the FM adaptors raised the threshold of the AM tone, the AM detectors had no effect on the FM threshold. Monaural and contra-aural adaptation revealed an interaural transfer of about 60-80%. The results indicated "channels" in the auditory pathways sensitive to frequency modulation, more readily explained by place than by periodicity-coding theory. The interaural findings suggest that the effects are predominantly central phenomena.

Selective Adaptation in Speech Perception

Following modification of the adaptation paradigm for speech by Eimas and Corbit (1973), a number of studies in speech perception have used an adaptation paradigm to support the notion that feature detectors exist to analyze the complex speech signal into component acoustic features at an early stage in the listening process. In the adaptation paradigm used in speech perception, a series of synthetic speech sounds is generated that varies in one phonetic dimension along an acoustic continuum.

Before discussing the adaptation work, we must digress to describe synthetic speech continua. Machine synthesis permits a series of acoustic stimuli to be prepared by varying, for example, the onset frequencies and direction of the second- and third- formant transitions in small steps while holding all other acoustic features constant. This acoustic dimension has been found sufficient to cue place of articulation for /b/, /d/, and /g/ (Liberman, Delattre, Cooper, and Gerstman, 1954). Similarly, a vowel continuum (/i/ to /ɪ/ to /ɛ/ to /æ/ can be prepared by systematically varying the center frequencies of the first three formants (Fry, Abramson, Eimas, and Liberman, 1962), while a voiced-voiceless continuum could be prepared by varying the Voice-Onset-Time (VOT) (the time between the release burst and the onset of periodic pulsing) of syllable pairs (/p,b/, /d,t/, /k,g/) (Lisker and Abramson, 1970).

If now, a seven-item continuum of sounds ranging from, for example, /ba/-/pa/ is randomly ordered and presented a number of times for identification, subjects do not perceive the continuum as gradually changing, but divide the continuum into two discrete perceptual categories, /ba/ and /pa/. This identification function differs considerably from the results obtained in psychophysical studies (Liberman, Harris, Hoffman, and Griffith, 1957). The phoneme boundaries (50% cross-over point of the identification function) are relatively sharp due to the almost perfect identification within a category.

When the stimuli are presented for pair-wise discrimination, stimuli within the same phoneme category tend to be discriminated at near-chance levels while stimuli taken from different phoneme categories are nearly perfectly discriminated, even though all stimuli are separated by the same acoustic difference. Therefore, there is a peak in the discrimination function that corresponds to the phoneme boundary in the identification function. These discrimination functions suggest that listeners can only hear these sounds categorically and cannot discriminate other differences. This responsiveness to stimuli in absolute terms has been labeled "categorical perception" (Liberman, et.al., 1957).

While the original claim had been that categorical perception occurs only for speech sounds (Liberman, Cooper, Harris, and MacNeilage, 1962), recent experimentation has demonstrated categorical perception for non-speech sounds as well (Locke and Kellar, 1973; Cutting and Rosner, 1974; Miller, Pastore, Wier, Kelly, and Dooling, 1974; Cutting, Rosner, and Foard, 1975; Pisoni, In press). However, the original work on categorical perception led Eimas and Corbit (1973) to explore its basis by means of an adaptation paradigm.

In the adaptation procedure as originally modified for speech by Eimas and Corbit (1973), listeners were presented with two series of synthetic consonant-vowel syllables varying in their initial consonant, one series from /ba/ to /pa/ and the second from /da/ to /ta/. They, therefore, used a synthetic acoustic continuum to simulate the articulatory dimension of VOT or the time between the onset of the release burst and the onset of the periodic pulsing (Lisker and Abramson, 1964).

Variations in VOT have been found sufficent to cue the distinctions between the voiced and voiceless stop consonants in English in syllable initial position (Lisker and Abramson, 1970), and the perception of these distinctions has been found to be nearly categorical (Abramson and Lisker, 1970).

Listeners were required to identify the randomized stimuli from the two continua singly from a forced-choice set of categories /ba/ or /pa/, /da/ or /ta/. The phoneme boundary was then determined. This identification or pre-adaptation series was followed by an adaptation sequence consisting of repetitive presentations of the adapting stimulus (one of the extreme end points of the continuum) and the randomized test syllables. The adaptors were presented for 2 minutes at the beginning of each session (150 adaptors) and for 1 minute (75 adaptors, a 75:1 ratio) before each single stimulus to be identified. Identification of the presented stimuli was compared in the pre-adaptation and the adaptation conditions by determing the 50% VOT cross-over point.

After adaptation with the voiceless stop, /pa/, subjects gave more /ba/ responses to the /ba/-/pa/ series, especially for the stimuli near the phoneme boundary. Therefore, there was a shift in the phoneme boundary in the adaptation condition relative to the phoneme boundary in the pre-adaptation condition. This shift was towards the /pa/ or the phonemic category from which the adapting stimulus was selected. In other words, there were fewer stimuli identified in the adaptor category subsequent to adaptation. Similar effects were found for the /da/-/ta/ series using the voiceless adaptor and for both series using the voiced adaptors. Also, the peak in the ABX discrimination task presented to the subjects subsequent to adaptation with /pa/ shifted to coincide with the adapted phonetic boundary.

However, the most interesting result of this study was that the shift in the phoneme boundary occurred whether or not the adaptor and test stimuli belonged to the same place of articulation (bilabial or alveolar). Crossed-series effects were weaker than those obtained on the same-series tests. Nonetheless, the importance of this result, according to Eimas and Corbit (1973), was that it ruled out the notion of adaptation of the phoneme-sized unit and pointed to an effect on the extraction of information underlying the differences in voicing that was independent of information underlying differences in place of articulation. In other words, the adaptation effect was selective.

Their original interpretation of these results was used to explain the high discrimination between the phoneme categories in categorical perception: the assumption of two linguistic feature detectors, each differentially sensitive to a range of VOT values within a phonetic category and equally sensitive to the VOT value that lies at the phonetic boundary. Adaptation with an extreme acoustic token on the continuum reduced the output of the detector responsive to that token, and at the same time, increased the sensitivity of the opposing detector. The result was then a shift in the category boundary, in the region of the continuum in which the supposed detectors are equally sensitive to both VOT extremes.

Although there has been some disagreement with the linguistic detector interpretation, the importance of this study lies in the fact that an adaptation effect was obtained and that the crossed-series effect was 80% of the magnitude of the same-series effect. That the "effect has proved to be on a feature within the phonemic segment, has provided the strongest evidence to date of a physiologically grounded feature system" (Studdert-Kennedy, 1976, p. 34) and has provided evidence that adaptation for speech is selective.

Studies employing the Eimas and Corbit paradigm have proliferated. Experiments have obtained boundary shifts for the phonetic distinctions of place of articulation, manner, and voicing in consonants. Most of the recent experimentation attempts to answer two fundamental questions concerning the nature of the adaptation effect. The first question concerns precisely what is adapted. Is it a phonetic feature detector or an acoustic feature detector underlying the phonetic percept? The second question concerns the site of the adaptation effect in the perceptual system: Does the adaptation of these features take place central to or peripheral to binaural interaction?

Other Explanations for the Boundary Shifts in Selective Adaptation

Before reviewing the literature dealing with these questions, another question has emerged concerning the extent to which selective adaptation is perceptual rather than the result of response biases or response organizational effects. Response biases, or the tendency of the listener to respond on the basis of mental sets (the expectancy of the stimulus presentation, familiarity of the stimulus) rather than on the perception of the stimulus, has long been recognized as a problem in psychophysics (Guilford, 1954; Engen, 1971). It is known that judgments are more often made in favor of certain categories than others (Johnson, 1971). The effects of these biases make data interpretation difficult. Response biases will be most evident when the stimuli are ambiguous (Engen, 1971).

Context effects in judgments of a stimulus continuum are well-established effects of an external anchor. The end categories of the response continuum, normally used for the extreme stimuli, occur less frequently when a more external anchor is used, so that there is a shift in the average judgments. If the boundary shifts towards the anchor, the effect is one of contrast (as opposed to assimilation). A psychological relativity theory such as Adaptation Level Theory (Helson, 1964) or the range-frequency theory (Parducci, 1974) would predict this boundary movement towards the more frequently occurring or designated anchor stimulus.

Evidence has accumulated against the position that the selective adaptation effect is solely due to response bias in the above sense or response organization effects (cf. Eimas and Miller, In press). Initially, the shifts in the ABX discrimination functions subsequent to adaptation (Eimas and Corbit, 1973) were used as evidence against the response bias or contrast interpretation since this process is presumed to operate on a decision rule involving only discrimination. Since discrimination is determined by identification (Liberman, Harris, Hoffman, and Griffith, 1957; Studdert-Kennedy, Liberman, Harris, and Cooper, 1970), it has been argued (Cooper, 1975) that such discrimination peaks are influenced by the identification of the stimuli. Therefore, discrimination effects as well as identification effects could reflect response biases.

The cross-series effect (a shift in the phonetic boundary along one stimulus series after adaptation with an end-point stimulus from another series) indicates that adaptation does not occur at the level of the syllable or even of the phoneme-sized phonetic segment (Eimas and Corbit, 1973; Cooper, 1974a; Cooper and Blumstein, 1974) and this has also been interpreted as evidence against a response bias effect at the syllabic or segmental level (Eimas and Miller, In press). However, as Sawusch (1976) points out, the cross-series identification results would not rule out a response organization explanation at a feature level, since theories of judgment would predict the category boundary shift as a result of a repeated feature as well as of a repeated syllable.

Other relevant evidence against response bias effects comes from the differences in the magnitude of category boundary shifts as a function of the end-point stimulus used as an adaptor. For example, Cooper (1974a), using a three-category synthetic CV syllable series that ranged from /bæ/ to /dæ/ to /gæ/, measured category boundary shifts in identification and in ABX discrimination using three syllables from the series as adaptors. Significant boundary shifts in the direction of the adapting syllable were found in the identification task for the /bæ/ adaptor (Stimulus 1) on the /bæ/-/dæ/ boundary, the /gæ/ adaptor (Stimulus 13) on the /dæ/-/gæ/ boundary, and the /dæ/ adaptor (Stimulus 7) on both boundaries. Although there was no shift for the /bæ/ adaptor on the /dæ/-/gæ/ boundary, there was a small shift for the /gæ/ adaptor on the /bæ/-/dæ/ boundary (significant on a one-tailed test but not on a two-tailed test).

The same pattern of results was found for the peaks in the discrimination functions that bordered the categories of the adapting stimulus. However, a discrepancy was found for the phonetic boundaries and the discrimination peaks that did not border the adapting stimulus categories. There was no shift in the discrimination function in the /bæ/-/dæ/ boundary with the /gæ/ adaptor that corresponded to the shift in the identification function. Cooper (1975) suggested that adaptation with one extreme end of the continuum should produce significant shifts of equal magnitude on both boundaries if adaptation was operating on relative feature values. A response bias interpretation would also predict a shift in all the boundaries towards the adapting stimulus.

Another set of experiments that might rule out response bias effects are those where significant boundary shifts are obtained in a phonetic continuum using only portions of the speech signal as the adaptor (Ades, 1974a; Tartter and Eimas, 1975). Since these portions are not perceived as speech, shifts in the phonetic boundary are difficult to explain on the basis of a response bias model operating on the segmental or syllabic level. However, Eimas and Miller (In press) argue that they might be explained by response bias factors operating at a phonetic feature level since the portions of the speech signal contained some feature information.

Finally, the anchoring paradigm first used in speech perception by Sawusch and Pisoni (1973) and Sawusch, Pisoni, and Cutting (1974) sought a direct experimental test of the response bias account.

Anchoring effects have a long history in studies of psychological and psychophysical judgments. Judgments of randomly presented stimuli differing from each other along some dimension shift systematically with certain changes in contextual conditions. Although the first stimulus to be presented may be judged almost at random, subsequent stimuli are judged on the basis of the range of stimuli in the experiment, the other stimuli in the series, intraserial effects (the effects of the preceding stimulus), the relative frequency with which the stimuli are presented, and the past experience of the subject (Rogers, 1941; Woodworth and Schlosberg, 1954; Engen, 1971; Parducci, 1974). The alignment of the scale of judgment with the series of stimuli is influenced by the stimulus context as well as by special anchor stimuli, and therefore, the scale can be shifted by manipulation of these stimuli. Such shifts have been called context or anchoring effects.

Sawusch and Pisoni (1973) and Sawusch, Pisoni, and Cutting (1974) took an unbalanced probability continuum, previously used in the study of judgment of visual brightness (Helson, 1964) and weight (Parducci, 1965, 1974) and adapted it for speech perception. In the anchoring paradigm, as in the selective adaptation paradigm, an end-point stimulus is presented more frequently than any of the other stimuli in the continuum. However, these stimuli as well as the test syllables are presented for judgment. This procedure differs from the adaptation paradigm where the adaptors are presented before the test syllables. The category boundary in the anchoring condition is then compared with the boundary in the pre-anchor condition where all stimuli are presented equally often, in a random order.

Specifically, Sawusch, Pisoni, and Cutting (1974) used a seven-step synthetic stop consonant continuum where each value of place was paired with each of seven variations in fundamental frequency ranging from 114 Hz to 150 Hz in 6 dB steps. Forty-nine stimuli were, therefore, produced making one dimension totally independent of the other. The fact that the experimenters obtained a boundary shift in the nonlinguistic or fundamental frequency condition, but no shift in the linguistic condition suggested that subjects' identifications along a synthetic stop consonant continuum are not subject to simple psychological response bias effects. From this the experimenters inferred that the boundary shifts typically observed in adaptation studies of stop consonants are probably perceptual in nature rather than the result of response bias or response organization effects. This interpretation was based on the assumption that the selective adaptation paradigm is an absolute identification task where one stimulus occurs more frequently than other stimuli (Sawusch and Pisoni, 1976).

While it can be questioned whether the failure to show category boundary shifts for the linguistic stimuli in anchoring is evidence that the shifts in consonant boundaries in adaptation are sensory in nature, the results do suggest that the phonetic features in the CV syllables are identified in an absolute, non-arbitrary manner, independent of the context of the experimental series. The decision concerning the presence or absence of an acoustic attribute underlying the phonetic features in the CV syllables appears to be based on an internally represented standard rather than on the range and frequency of the stimulus presented to the subject, during the course of an experimental session (Sawusch and Pisoni, 1976).

Recent studies (Ades, 1974b; Diehl, 1975; Sawusch and Pisoni, 1976) have provided more direct experimental tests of the extent to which instructions given to, or labels used by, the subjects have systematic effects on selective adaptation. Ades (1974b) found a high correlation ($r = 0.79$) between subjects' rating of goodness of the adapting stimuli as phonetic entities and the magnitude of the boundary shifts. He found a greater adapting effect the more /b/-like or /d/-like were the adapting stimuli. Diehl (1975) found a similar effect using a /te/ syllable as the adaptor on a /be/-/de/ place continuum. A reliable boundary shift towards the /be/ stimulus was found for 4 of the 6 subjects who reported that the adapting stimulus /te/ sounded like /pe/.

Sawusch and Pisoni (1976), however, found that labelling instructions to the subjects had no systematic effect in selective adaptation. Instead of using a good example of the category being adapted as the adaptor (usually the end-points of the series), Sawusch and Pisoni used the middle stimulus of the series or the boundary stimulus. Subjects usually assign this stimulus to either of the two opposing phonetic categories equally often. Two groups of subjects were used, one group told that the repeating syllable was a /di/, the other that it was /bi/. Since neither group showed a shift in their identification function, Sawusch and Pisoni (1976) concluded that stimulus variables and not response organization effects were the determining factors in selective adaptation.

Cooper, Ebert, and Cole (1976) applied a signal detection analysis to data from a selective adaptation experiment using a /ba/-/wa/ continuum with end-point adaptors in order to assess the contribution of response criterion shifts. They based their analysis on the absolute identification of magnitude estimation used by Durlach and Braida (1972). The d' values were obtained on syllable pairs in the pre-adaptation and adaptation conditions and then converted into just noticeable differences (JNDs). Substantial changes were found following adaptation with /ba/ and /wa/. Cooper, Ebert, and Cole (1976) concluded that sensory changes accompanied the selective adaptation procedure. There was a decline in the d' following adaptation with both adaptors, indicating a decline in sensitivity. However, Sawusch (1976) points out that there is a question of the applicability of the Durlach and Braida (1972) model in this situation, based on the larger magnitude estimations of the JND in the Cooper, Ebert, and Cole (1976) analysis. Sawusch (1976) suggests that further analysis of the model is necessary before it can be accepted for use in selective adaptation and before any firm conclusions can be drawn about the supposed sensory basis of their results.

While it has been previously suggested that the phonetic features in a CV syllable are identified on an absolute basis, independent of context, Helson (1964) argues that there are no absolute quantities in psychology. Adapatation Level Theory (ALT) attempts to quantify perceptual reality by proposing that the effect of a stimulus on the organism is related to the adaptaton level (AL) at the time of the stimulus input. With the organism in this neutral state, adaptation level is defined as the weighted mean of all stimuli (focal, background and residual) presented to the subject. The various values of the stimuli presented successively to the subject are neurologically pooled to form a certain adaptation level for the perceiver. Therefore, judgment of a stimulus is made in comparison to the prevailing adaptation level.

Before discussing selective adaptation in speech perception, it is necessary to review the concept of absolute judgments, the definition of an "anchor" and theories of "anchoring."

The Relativity of Absolute Judgments

The method of "single stimuli" (Wever and Zener, 1928), essentially the same as the method of "absolute judgment," is a psychophysical procedure in which a series of stimuli differing from each other along some dimension are randomly presented for judgment. Subjects are instructed to judge the stimuli in categories appropriate to the attribute being judged (e.g. "loud," "medium," "soft") rather than relative to the other stimuli in the continuum (Woodworth and Schlosberg, 1954; Bieri, Atkins, Briar, Leaman, Miller, and Tripodi, 1966; Engen, 1971). Measures of central tendency or thresholds between categories are then established.

Despite the requirement that category judgments be absolute, they tend to shift systematically with changes in experimental conditions: stimuli are judged relative to context and anchoring effects are observed (Bieri, et al., 1966). Thus, although the first stimulus may be judged independently, subsequent stimuli are judged on the basis of the range of preceding stimuli, the relative frequency with which they have been presented, and the past experience of the subject (Rogers, 1941; Woodworth and Schlosberg, 1954; Engen, 1971; Parducci, 1974). Judgments may also be influenced by the use of comparison or anchor stimuli, and the alignment of the physical and psychological scales can be shifted by manipulation of their values. These context and anchoring effects have been the starting points for various theories of relativity in psychological scaling experiments.

Classical psychophysics undertook to formulate the relation between particular stimuli and particular responses (Engen, 1971). Later theories (see Johnson, 1958; Helson, 1964; Parducci, 1965, 1974) attempted to predict responses not only from individual stimuli but also from the range and frequency of the stimulus series and the experience of the listener. These theories gave rise to numerous experiments investigating "anchor," contextual, or frame of reference effects. Such experiments have a double purpose: first, to tie down a scale so that the meaning of a particular judgment can be relatively fixed, and, second, to determine the principles that account for the shifting of the scale as stimulus conditons are varied (Guilford, 1954).

The Definition of "Anchor"

Contradictory statements regarding the status of an anchor stimulus and the concept of anchoring evidently prompted Guilford's statement, "The concept of an anchor stimulus itself has needed anchoring..." (1954, p.312). The earliest use of the term "anchor" was to indicate a new stimulus added to the series being presented, with the subject being informed both of the stimulus and of its response value. This definition has evolved to include stimuli within or without the stimulus series originally presented, stimuli unlabeled as well as labeled, and stimuli introduced to the subject but not judged as part of the response series (Bieri, et al., 1966; Johnson, 1972). For example, one view differentiates an anchor from other stimuli in the test series by assuming the anchor to be a stable reference point. Anchoring is then the effect of the presumed fixed position of this anchor upon the judged position of another stimulus (Bieri, et al., 1966). A second, not incompatible view, regards the end stimuli as ever-present anchors. Since these extreme stimuli are judged more reliably than stimuli in the center of the series and manifest the effect of introduced anchors upon the judgment of the other stimuli in the series, they are termed natural anchors. Thus, anchoring is never truly absent, and judgments of a particular stimulus are always a function of its position with respect to the end of the series (Bieri, et al., 1966).

A third view regards <u>every</u> stimulus in the experiment as an anchor stimulus, whether intended or not, since there is an effect of each stimulus upon the total judgment pattern (Guilford, 1954).

Looked at in this light, an anchor may lie anywhere within or without the stimulus continuum (Johnson, 1972), and may be presented with the same relative frequency or with different frequency than other stimuli in the series.

Thus, it is evident that contextual effects or frame of reference effects might be observed even where an introduced or more frequently presented anchor is not used.

Definition of "Anchoring Effects"

According to Johnson (1972), the most "unambiguous evidence" for anchoring is that the anchor stimulus is judged more consistently than the other stimuli on the continuum. Variability and errors decrease as the anchor is approached: the anchor stimulus itself is judged with less error than other stimuli. The reduction in variability for the anchor stimulus is likely to be enhanced when the anchor stimulus is an end stimulus (natural anchor), since end stimuli are judged with fewer errors and less variability (Volkman, 1951). If the anchor is a midrange stimulus, a similar increase in reliability is noted and the reduction in variability is even more striking since initial variability in this range is higher than in the vicinity of the natural end-anchor (Bieri, et al., 1966). The effects of end-anchoring spread inwards, leading to the assumption that the scale is organized from each end towards the center (Johnson, 1972).

The effects of the anchor stimuli on the scale appear to depend upon the position of the anchor. Both assimilation effects (judgments shifted towards the value of the anchor stimulus) and contrast effects (judgments shifted away from the anchor stimulus) have been observed, but contrast effects predominate. Assimilation effects are reported to be smaller, less stable, and not as predictable as contrast effects, and the conditions under which assimilation effects occur are not clear

(Bieri, et al., 1966).

Theories of Anchoring

Theories of anchoring are grouped into two categories, centering theories and distance theories. In general, a stimulus of midrange value in the series of stimuli presented will be judged to be in the middle of the response scale (whether calculated by an arithmetic or geometric mean). A stimulus scale that is extended in range by the introduction of a more extreme anchor value at one end will shift the judged mean on the response scale in a direction away from the anchor position (contrast effect). In the simplest sense, the difference in theories is in the rationale for this shift. In the distance theories, this contrast effect is the result of the increase in the absolute distance between the anchor stimulus and the judged stimulus; while in the centering theories, it is either the result of a change in the relative distance of the critical stimulus from the two extreme values in the series (Bieri, et al., 1966), or the effect of the range of the series.

Although not the first centering theory nor the first theory of relativity used in psychophysical judgments, Helson's Adaptation-Level Theory is so well known that the term "adaptation level" is used generally as a synonym for context or frame of reference, without any necessarily quantitative implication (Johnson, 1971). The original work of Helson (1938) attempted to predict the perception of the color of an object from measurable aspects of the visual stimulation rather than from cognitive aspects of contrast (Hochberg, 1971). He attempted to show that the subject uses the general level of the visual field as a neutral point from which judgments of color are made according to their divergence from this level. Subsequently, he attempted to make this approach general, objective, and quantitative (Hochberg, 1971).

In Helson's first visual adaptation experiment (Helson, 1938), the subject was placed in a light-tight booth, flooded with nearly homogeneous light (all one wavelength). In one session, the booth was lined with gray paper and flooded with red light. When the subject first entered the booth, everything looked red, or at least red-illuminated. After five minutes of adaptation to the illumination, the subject judged a set of 19 gray nonselective reflectors ranging in equal-interval steps from white to black under normal illumination. Those samples with the same reflectance as the gray walls and background were judged a gray of medium brightness. Brighter samples were judged red and, most surprisingly, samples darker than the walls were judged green or blue-green, the afterimage complement of the red illuminant. In another session, where the walls were covered with white paper instead of the gray so that the eyes received more red stimulation, a higher adaptation level was established. With this higher level, only two of the lightest grays appeared red. In contrast, when the walls were black all except the darkest samples appeared red (Woodworth and Schlosberg, 1954; Hochberg, 1971). Helson attributed the results to the adaptation level established by the prevailing reflectance. Grays near that level appeared to be the same hue as the illuminant, while the darker grays took on the after-image complement of that hue. Subsequent research (see Hochberg,

1971) established that it was the conditions of stimulation rather than the subject's knowledge of the surfaces and sources, that brought about the adaptation.

The concept of adaptation was originally defined in sensory physiology as a decrement in sensitivity following intense or prolonged exposure to a stimulus and was regarded as a functional characteristic of all modalities (Small, 1963; Elliot and Fraser, 1970); in general biology adaptation was viewed as an adjustment of an organism to its environment. Starting from the purely physiological concept, Helson (1947, 1964) moved toward the general biological view by working out a formula for computing the adaptation level so as to explain contextual and frame of reference effects in psychological judgments. He applied this formula not only to the effects observed in the 1938 experiment, but to all manner of judgments in sensory processing, perception, learning, cognition, affective behavior, social psychology, personality, and motivation (Helson, 1964).

According to Hochberg (1971), this extended concept should be viewed based on judgment scales rather than identifiable physiological processes. When Helson attempts to apply the same concept to constancy and contrast, lifted weights and social phenomena, there is, according to Woodworth and Schlosberg, "little left of physiological value" (1954, p. 449). However, the present study is concerned with Helson's basic physiological account of adaptation level: that there is an increased sensitization to complementary stimulation and a decreased response to predominant stimulation. Adaptation is effected by reaction of the organism to the stimulation as well as by the action of the stimulation upon the organism. Recent theorists in speech perception (Eimas and Corbit, 1973; Eimas and Miller, In press) have used Helson's sensory approach as a basis for defining the characteristics of feature detectors in selective adaptation.

One criticism of ALT relevant to this study has been refuted: the notion that the locus of the contextual effects found in AL is in response organization. Both Helson and other theorists (Helson, 1971) maintain that it is the changes in stimulation that affect the perceived quality, magnitude and other dimensions of stimuli. By contrast, critics of AL theory state that the changes are merely semantic, verbal, or judgmental, and therefore, reflect response change rather than perceptual change.

Helson (1971) attributes the belief that the changes are judgmental rather than perceptual to the use of verbal category rating scales. He cites a 1938 study (Helson and Kozaki, 1938, cited in Helson, 1971) where numbers instead of verbal categories were used as the response mode. Random patterns of 10, 12, 14, 16, or 18 dots were exposed for different anchor, 4 dots, 13 dots, and 32 dots and one group served as a control with no anchor. The anchors were exposed before each of the stimuli but were not judged. The results were as anticipated: the presence of the small anchor (4 dots) increased the number of dots perceived while the large anchor (32 dots) decreased the number. Both the control and the group with the 13 dot anchor (close to the geometric mean of the series) maintained their judgments close to the objective number. Since anchor effects could be reliably obtained with numbers instead of verbal categories that might influence response judgment, Helson concludes that the locus of adaptation effects is in perception rather than response organization.

If we accept this conclusion, as seems reasonable, then the distinction between adaptation and anchoring paradigms made by Sawusch and Pisoni (1973) and Sawusch, Pisoni, and Cutting (1974) breaks down. This is a point to which we will return.

Before reviewing the now extensive literature in selective adaptation in speech perception, it might be informative to examine auditory adaptation and fatigue in order to clarify terminology differences between the two adaptation procedures, to contrast the procedures and effects obtained, and to present evidence on stimulus energy effects that are relevant to both areas of adaptation research.

Studies of Auditory Adaptation and Fatigue

Studies of auditory adaptation and fatigue at the peripheral level usually distinguish between adaptation and fatigue (Small, 1963, 1973; Ward, 1963; Elliot and Fraser, 1970) because of differences in measurement, stimuli involved, and physiological changes. Studies of adaptation and fatigue define fatigue as a "reversible increase in threshold" (Small, 1973, p. 400) and adaptation as "a reversible decline in loudness" (Small, 1973, p. 400) with the definitions based on the measurement procedures.

Studies of auditory fatigue routinely establish threshold before and after exposure to a moderate-to-intense level of stimulation (pure tones and wide-band noise). The difference between the two measurements is defined as the amount of fatigue or temporary threshold shift (TTS) (Small, 1973; Elliot and Fraser, 1970). The intensity of the stimulus, the time interval between the stimulus exposure and the post-exposure threshold measurement, the duration and frequency of the fatiguing stimulus, and the frequency of the thresold test signal(s) have an effect on TTS (Elliot and Fraser, 1970). Generally, the obtained threshold shifts increase as the duration and intensity of the exposing stimulus increases (Hood, 1950). There also seems to be an interrelationship between the intensity and frequencies affected. At low to moderate stimulus intensities, TTS is equally distributed about the frequency of the stimulus. However, as intensity increases there is a spread of fatigue to adjacent frequencies, usually with the maximum fatigue at frequencies one-half octave above the stimulating frequency (Hood, 1950). Recovery from TTS is usually measured after two minutes, because the initial portion is irregular. However, recovery generally increases with longer post-exposure intervals, and recovery time is proportional to the amount of TTS (Elliot and Fraser, 1970).

The development of fatigue and the recovery process are complicated and confounded by measurement errors and the fact that TTS changes constantly during the recovery time. Generally, however, the results of studies suggest that the amount of threshold shift is related to the total stimulus energy reaching the receptor, this being a function of the intensity, duration, and presentation frequency of the stimulus.

Studies of auditory adaptation generally measure the effect indirectly (Small, 1973) during stimulation by a technique known as simultaneous dichotic loudness balance (SDLB). An adapting stimulus is presented to one ear and a comparison stimulus is presented to the control ear at three different time periods: pre-stimulatory (before adaptation), per-stimulatory (during adaptation), and post-stimulatory (after adaptation). The subjects adjust the comparison stimulus to the loudness of the adapting stimulus.

Long-term and short-term adaptation stimuli have been used. In long-term adaptation stimuli are usually presented for 20 seconds, with a 90 second interval between stimuli. In short-term adaptation the test stimulus is presented for approximately 20 msec and the adapting stimulus for 400 msec (Small, 1973). No information was available comparing the effects of long-term and short-term adaptation.

Generally, the course of development of adaptation is negatively accelerated, with the greatest rate of adaptation during the first one to two minutes (Elliot and Fraser, 1970) and the maximum amount of adaptation being reached after three to six minutes of stimulation, depending on the intensity of the adapting stimulus (Small, 1973). Adaptation can be obtained with pure tones at only weak to moderate levels. Hood (1950) and Jerger (1957) found a linear relationship between the degree of adaptation and the intensity of the stimulus tone from 10dB to 60dB SPL, followed by a flattening of the function. Other studies have found that the linear relations continue above 60dB SPL, and that there is little adaptation below 50dB SPL (Small, 1973). However, as in measurement of TTS, methodological differences, as well as stimulus cross-over between ears that might cause adaptation in the control ear, complicate the results.

While most of the experimentation in adaptation of intensity has been done with the test intensity the same as the adapting intensity, Hood (1950) adapted at 100dB SPL (with a pure tone) and tested at sensation levels from 20dB through 100dB. He found that adaptation, at least for test intensities lower than the adapting intensity, was dependent upon the test intensity and not the adaptor intensity.

In contrast to TTS, the spread of adaptation away from the adapting frequency is symmetrical, and if frequency is increased, an increase in adaptation is found only up to about 2000 Hz. Above this frequency, there is little or no relation to the frequency of the adapting stimulus (Elliot and Fraser, 1970; Small, 1973).

Carterette (1955), using noise interrupted at various rates with a 50% duty cycle, found that adaptation increased with the faster or increased adaptation rate. He suggested that the period of adaptation is shortened as is the recovery time due to a cumulative energy effect. Elliot and Fraser (1970) suggest that the recovery processes suffer more damage from the increase in the interruption rate than do the adapting processes.

Voicing in Selective Adaptation

Adaptation experiments have considered the nature of the adaptation effects, what they imply about the processing of speech, and the suggestion that speech processing involves detectors. What aspect of the stimulus determines the adaptation effect? Are phonetic features of a segment adapted, or the underlying acoustic structures corresponding to these phonetic units? Does adaptation occur at an auditory level, or at a more central integrative level? Is it specifically sensitive to individual acoustic cues, or to combinations of cues?

Eimas, Corbit, and Cooper (1973) continued their investigation of the perception of voiced and voiceless stop consonants using the selective adaptation paradigm in a second set of experiments. To support the interpretation that adaptation operates in terms of linguistic feature detectors differentially sensitive to variations in VOT, an experiment was designed to determine if this analysis of VOT is operating on general auditory information or on specialized speech information.

A comparison was made of the adaptation effect with the first 50 msec of the voiced syllable /da/ (d-chirp), as a nonspeech adaptor, and the syllable /da/, as a speech adaptor. Since the d-chirp contained the same VOT information as did the /da/ adaptor, it was assumed that the absence of a boundary shift with the non-speech adaptor would support the notion that the VOT detectors are part of the speech processing mechanism rather than the general auditory mechanism. As predicted, adaptation with the speech adaptor produced a category boundary shift while adaptation with the non-speech adaptor produced no significant shift. Taken with their cross-series evidence (Eimas and Corbit, 1973) that the /ba/ adaptor had an effect on both the /ba/-/pa/ and /da/-/ta/ test series, Eimas, Corbit, and Cooper (1973) concluded that the effect is due to adaptation of the phonetic segment: the VOT detectors are part of the speech processing system, and in order for this mechanism to be activated by auditory information, it must be embedded in the whole speech signal.

A number of problems exist in the interpretation of the Eimas and Corbit (1973) results as stated. First, VOT is not the only cue for voicing distinctions in pre-stressed stop consonants. The results of the Eimas and Corbit (1973) study do not clearly differentiate between results based on analyzers of voicing or on analyzers for the VOT cue alone (Cooper, 1975). Second, the cross-series effect could be due to the fatiguing of the first-formant detector that remained constant in both series. While the feature being tested in the experiment was VOT, the differences in the series were in the acoustic cues to place of articulation (the first- and second-formant transitions) (Studdert-Kennedy, 1976). Third, the lack of an adaptation effect with the d-chirp is not an indictment against an auditory level of adaptation. As Studdert-Kennedy (1976) points out, control and test items that are acoustically identical but phonetically distinct may yield ambiguous results, if the control item does not shift the phonetic boundary. It may mean either that the detector being fatigued is phonetic or that there may not be enough information in the incomplete signal (d-chirp) to cause fatigue of the appropriate detector.

Other studies using non-speech adaptors have yielded contradictory results. Ades (1973), using the first 38 msec of the adaptor, obtained statistically significant shifts in the phoneme boundary using a place continuum (/bæ/-/dæ/), as did Tartter and Eimas (1975), who also used a number of different non-speech adaptors on a voicing continuum. Tartter and Eimas (1975) used a p -F2F3 adaptor consisting of the second and third formants of a voiceless /pæ/ syllable, a bæ-F2F3 adaptor consisting of the second and third formants of a voiced /bæ/ syllable, a /bæ/ chirp (45 msec of the first three formants) and the whole /bæ/ syllable. The entire syllable adaptor /bæ/ and pæ-F2F3 produced significant boundary shifts. However, the pæ-F2F3 adaptor also had a significant effect on the boundary, moving it towards the /pæ/ end of the continuum. Neither of the other adaptors produced significant shifts.

They concluded that reliable adaptation effects obtained with the pæ-F2F3 adaptor are contrary to an interpretation of adaptation at a phonetic level. The difference betweeen the two non-speech adaptors was explained by the fact that the pæ-F2F3 adaptor contained all three cues of voicelessness (see Cues and Feature Combination) (Stevens and Klatt, 1974) and the bæ-F2F3 adaptor contained conflicting cues to voicelessness and voicedness (formant transitions and acoustic energy cues available but omitting the first formant onset cue).

A study by Cooper (1974b) was designed to test the assumption that the adaptation effects observed by Eimas and Corbit (1973) and Eimas et al., (1973) are due to absolute VOT, independent of the formant transition cues. Cooper used two adaptors from the Stevens and Klatt (1974) series, both stimuli with VOT's of +25 msec, but differing in the amount of transition after voice onset. The /da/-long adaptor contained a 40 msec transition after voice onset and the /da/-short adaptor contained a 10 msec transition. The test series consisted of 11 synthetic CV syllables from /ba/ to /pa/ ranging in VOT from +5 to +55 msec in 5 msec steps, all with formant transitions of 20 msec. The /da/-long adaptor shifted the phonetic boundary significantly toward the voiced (/ba/) end of the series while the /da/-short had a nonsignificant effect in the opposite direction.

Cooper (1974b) concluded, first, that the results cannot be attributed to VOT per se, since the two adaptors had identical VOT values, yet produced different effects on the phonetic boundary. Second, the adaptation effect could not be due to fatigue of detectors sensitive to the formant transitions, since the test stimuli near the phonetic boundary contained little or no formant transitions after voicing onset. The /da/-long adaptor, with a 40 msec transition, therefore, had no effect on the boundary stimuli. Third, Cooper concluded that the adaptation effects were due to fatiguing of an integrative analyzer sensitive to a number of auditory cues. Ades (1976) argued, however, that the integrative level as the site of adaptation is tenuous. Cooper (1974b) based his conclusions on the assumption that the boundary stimuli had no first-formant transition, and that there were therefore no first-formant cues to voicelessness available. However, Ades (1976) stated that as long as some first-formant cues to voicedness exist, adaptation can be attributed to this information. A cue that is still available to be fatigued by the first-formant detectors and that might explain the adapting effect of the /da/-long adaptor is the low first-formant onset frequency (Lisker, 1975). Ades (1976) also stated that since there were 80% /ba/ responses to the stimuli with no first-formant transitions in Cooper's study, some cue to voicedness must be present in the stimuli.

A second study by Cooper (1974c) has been interpreted as evidence that non-phonetic mechanisms are involved in the voiced/voiceless distinction in selective adaptation. Cooper (1974c) used a contingent adaptation paradigm where both the voicing of the consonant and the vowel were varied systematically. Two test continua, /ba/-/pa/ and /bi/-/pi/ were used with the adaptors /da/-/ti/ and /di/-/ta/ alternatively repeated.

It was assumed that if adaptation operated on phonetic features independent of vowel environment, there would be no shift in the boundary since the adapting stimuli were of opposing values of the phonetic feature of voicing. Therefore, the two detectors, voiced and voiceless, would be equally fatigued. On the other hand, a shift in the phonetic boundary would show a vowel-dependent relationship.

The results showed that adaptation with the /da/-/ti/ adaptors effected a vowel-dependent shift since the /ba/-pa/ boundary moved towards the /ba/ or voiced end of the series, while the /bi/-/pi/ boundary moved towards the /pi/ or voiceless end. The shifts for the /di/-/ta/ adaptors were in the opposite direction, also towards the vowel of the test series. This effect has since been replicated by Pisoni, Sawusch, and Adams (1975) and Miller and Eimas (1976). All three studies attribute the adaptation results not to the fatiguing of phonetic features, but rather to vowel-dependent, and therefore, according to Sawusch (1976), frequency-specific features. However, Pisoni, Sawusch, and Adams (1975) found no effect with either adaptor sequence in the /bi/-/pi/ series, a fact that clearly complicates all the results.

Place of Articulation and Selective Adaptation

Concurrent with the original selective adaptation studies on VOT, Bailey (1973, cf. Eimas and Corbit, 1973; Eimas et al., 1973) was investigating place of articulation. The original Eimas et al. (1973) interpretation of adaptation was in terms of linguistic feature detector systems for whole linguistic features, in the Chomsky-Halle (1968) sense. Bailey envisioned feature detector systems for individual acoustic cues. The place feature was of special interest to Bailey, as well as Cooper (1974a), because of the acoustically non-invariant cues, the direction and extent of the second- and third-formant transitions. This is opposed to the relatively invariant acoustic information in VOT (the timing information) and the complexity of the numerous cues underlying voicing. Cooper (1974a) reasoned that adaptation should not occur for place of articulation if only relatively invariant acoustic information, such as VOT, is adapted. On the other hand, category boundary shifts would be expected if adaptation operates on analyzers tuned to phonetic features.

Cooper (1974a) used a three-category series (/bæ/-/dæ/-/gæ/). Significant boundary shifts were produced on the /bæ/-/dæ/ boundary with the /bæ/ adaptor, but not on the /dæ/-/gæ/ boundary. The reverse occurred with the /gæ/ adaptor: significant effects were produced on the /dæ/-/gæ/ boundary, but not on the /bæ/-/dæ/ boundary. The /dæ/ adaptor affected both boundaries. In a second experiment, the peaks in the ABX discrimination functions shifted to coincide with the shift in the identification boundary. Cooper (1974a) attributed these results to acoustically non-invariant or phonetic adaptation, because the adaptor /dæ/ consisted of slightly falling second- and third-formant transitions, but still affected the /bæ/-/dæ/ boundary consisting of rising transitions as well as the /dæ/-/gæ/ boundary containing falling transitions.

In another experiment, Cooper (1974a) used five different adaptors on the same /bæ/-/dæ/-/gæ/ continuum. He interpreted a shift produced by the /pæ/ on the /bæ/-/dæ/ boundary to a feature-specific unit of adaptation rather than to the consonant as a whole. A shift on the same boundary by a natural /bi/ was interpreted to show that adaptation was operating on a feature-specific component, mainly at the phonetic level of processing. The fact that less shift was produced by a real speech /bæ/ than by a synthetic /bæ/ suggested that some of the acoustic information provided by the real syllable is unnecessary for the perception of a phoneme. A larger shift for the real /bæ/ as opposed to the real /pæ/ was accounted for as adaptation of a "phonetic unit" operating on the consonant sound as a whole. The reduced shift produced by the /bi/ is suggestive of a "syllable unit." Therefore, Cooper (1974a) proposed that adaptation may occur at more than one level of analysis: feature, phoneme, and syllable, corresponding to the multileveled processing of speech.

Cooper and Blumstein (1974) extended the previous study to include five more labial adaptors differing in manner (/mæ/, /væ/, and /wæ/; /pæ/ and /bæ/) to determine if adaptation in the /b/-/d/-/g/ series operates in a feature-specific manner. All the real speech CV syllables, except /wæ/, produced significant shifts in the /bæ/-/dæ/

phonetic boundary towards the /bæ/ end of the series. Cooper and Blumstein (1974) attributed their results to a labial feature analyzer where all cues to place are integrated. However, they noted that the data could also be explained on the basis of the acoustic similarity between the adaptors and test syllables (rising second- and third-formant transitions). Similar experiments using the place dimension for cross-series adaptation have been reported (Bailey, 1973, 1975; Ades, 1974b). Ades (1974b) used a /bæ/-/dæ/ test series with rising F_2 and F_3 transitions for the /bæ/ syllable and falling transitions for the /dæ/. Five adaptors were used, all with the vowel /e/. The /be/+F3 contained a rising F2 and F3, the /de/+F3 contained a slightly rising F2 and a falling F3. The other three adaptors, /be/-F3, /de/-F3, and /de/-F3 had flat transitions with different degrees of rising F2 transitions. It was reasoned that if the detectors were phoneme or feature specific, all three /de/-adaptors should effectively shift the /bæ/-/dæ/ series. However, if the detectors were tuned to the rising or falling frequency transitions, only the /de/-F3 would be an effective adaptor on this series. Although the /de/+F3 showed a greater magnitude of shift, all three /de/ adaptors shifted the phonetic boundary in the predicted direction, towards the /dæ/ syllable. Although Ades (1976b) concluded that there is a phonetic component to adaptation, this interpretation is tempered by three results. First, a previous experiment (Ades, 1974b) found no cross-series adaptation of CV and VC syllables. The adapting stimuli /bæ/ and /dæ/ in the initial position were unable to effect a shift on the /æb/ -/æd/ continuum and vice versa. Linguistic theory would not predict this positionally dependent effect. Second, the effect of the /de/ adaptors was only 30% as great as the /dæ/ adaptor.

If only a phonetic level was involved, there should be complete transfer. Third, the starting frequencies of the formants for /de/ were closer to /dæ/ than to /bæ/. If spectral similarity rather than the direction of the formant transitions is the critical cue, then adaptation could be taking place at an auditory rather than a phonetic level.

Bailey (1973, 1975) found adaptation for the adaptors /bɛ/ and /dɛ/ on a /ba/-/da/ test series and /ba/ and /da/ on the /bɛ/-/dɛ/ test series. Since the adaptation effects were largest when the adaptors and the test stimuli were from the same series, Bailey (1975) concluded that adaptation effects occur at a level of "simple acoustic variables" (p.53) in this case, the formant transition cues to place of articulation.

In a second experiment, Bailey used two /ba/-/da/ test series, one where the transition cue was in the third-formant with a fixed F_2 and one where the information was carried in the second formant with no F_3. Both the /ba/-fixed F_2 and the /ba/-no F_3 had an effect on the fixed F_2 series. The /ba/-fixed F_2 had no effect on the no-F_3 series. This lack of a cross-series adaptation would be predicted because of the lack of stimulus energy in the F_3 region, and the ambiguous F_2 information supplied by the slightly falling F_2 transition in the middle range of the second-formant transitions in a /ba/ syllable. Yet, there was a small but significant effect of the /da/-fixed F_2 adaptor on the no-F_3 series. As a result of this experiment, Bailey (1975) suggested that spectral similarity between the test syllables and the adaptor are important in adaptation and Ades (1976) ruled out any single-level explanation of adaptation in terms of syllabic, phonetic, or linguisitc features where /ba/-fixed and /ba/-no F_3 would be considered the same.

Further support for adaptation at a frequency-specific auditory level is another cross-series experiment using a /bi/-/di/ and /bu/-/du/ place series (Bailey, 1975). There was no adaptation of the /bi/ adaptor on the /bu/-/du/ series and vice versa although both series shared the labial and phoneme feature /b/.

That phonetic boundary shifts only occur when there is spectral overlap between the adaptor and test series was supported by Tartter and Eimas (1975) who used a /bæ/-/dæ/-/gæ/ continuum with five non-speech adaptors from the /bæ/ syllable. Signficant boundary shifts were found on the /bæ/-/dæ/ boundary, however, none of the non-speech adaptors produced effects of the same magnitude as did the entire syllable, and, the fewer acoustic cues shared between adaptor and test syllable, the smaller the effect.

Tartter and Eimas (1975) draw three conclusions from these experiments. First, there is some auditory component in selective adaptation not requiring phonetic information. However, if adaptation was solely dependent upon spectral information, the F_2F_3 patterns and the isolated F_2 and F_3 adaptors should have produced an effect equal to the whole syllable, since all relevant place information was available. Second, the greater effect obtained with the larger number of acoustic cues shared is consistent with a model of multiple auditory analyzers where each analyzer is responsive to a restricted range of acoustic information. This model would explain the larger effect of the second- and third-formants together than with either one alone. Third, the addition of the first formant or the steady-state information that produced the largest shift was explained on the basis of a second level of detectors, activated only when the adapting stimuli are speech signals.

Another study using non-speech adaptors (Pisoni and Tash, 1975) also supports the notion that some part of the adaptation effect takes place at an auditory level. It was found that the speech embedded chirps (SEC's)(300 msec patterns consisting of a 250 msec initial steady-state vowel, followed by 50 msec transitions in the final position) containing all the acoustic information underlying the place feature, without the accompanying phonetic information, shifted the /ba/-/da/ boundary. However, these SEC's were only about 30% as effective as the speech adaptors.

Further support for an auditory component in selective adaptation is also evident in contingent adaptation effects (Miller and Eimas, 1976). In one condition, /bæ/ and /di/ adaptors were alternated on a /bæ/-/dæ/ and /bi/-/di/ test series. Adaptation with /bæ/-/di/ shifted the phonetic boundary on the /bæ/-/dæ/ series towards the /b/ and on the /bi/-/di/ series towards the /d/. As with the voicing series, Miller and Eimas (1976) interpreted these results to indicate that channels for the analysis of place information are selectively tuned to respond to contextual information on the basis of the vowel. Pisoni, Sawusch, and Adams (1975) reported similar contingent adaptation results using a /bi/-/di/ and /ba/-da/ test series. These results were interpreted as evidence for selective adaptation at a frequency specific auditory level.

Several recent experiments have attempted to obtain shifts in the phonetic boundary using adaptors and test stimuli that differ in the kind of information that specifies the place of articulation. Diehl (1975) used a test series /bɛ/ to /dɛ/ cued by formant transitions with an adaptor, /tɛ/, that contained no transitions but was cued by a high frequency burst. Small, but reliable shifts in the phoneme boundary towards the /dɛ/ end of the series were found.

Following Diehl (1975), Ganong (1975) used a /bæ/-/dæ/ test series with four adaptors: the end-point /dæ/, /sæ/ without formant transitions (SWT) synthesized by joining a 100 msec noise burst characteristic of /s/ to the steady-state vowel /æ/, /sæ/ with a noise burst followed by an acoustic pattern similar to /dæ/, and /tæ/ without formant transitions (TWT) consisting of a 15 msec noise burst followed by the steady-state vowel /æ/. Both syllable adaptors showed large phonetic boundry shifts while the SWT and the TWT adaptors showed reliable but smaller effects, towards the /dæ/ end of the series.

The shift in the phonetic boundary of a transition-cued test series by burst-cued adaptors has been interpreted as evidence for adaptation partly occuring at a phonetic level (Diehl, 1975), both at an acoustic and phonetic level (Ganong, 1975), at an auditory integrative level (Sawusch, 1976) and at one integrative level (Ades, 1976).

Blumstein and Stevens (1975) performed a similar experiment where a /ba/-/da/-/ga/ test series was used with transition-cued and transition- and appropriate burst-cued /da/ and /ga/ syllables. Significant boundary shifts were obtaind with both sets of adaptors. However, a greater effect was found with the adaptor containing both cues. No significant effect was obtained with conflicting transition and burst cues, for example with a /d/ burst and /g/ transition. Since the test syllables contained only transitions, it was argued that the adaptation effect occurs at an integrative level only, where burst and transition cues are combined (Blumstein and Stevens, 1975).

Ades (1976), however, argues that the same objections made against phonetic and linguistic feature levels can be made against the concept of integrative property detectors. Two of Bailey's experiments previously discussed, the /bi/ and /bu/ experiment and the Fixed F_2 and No F_3, were cited. Although both the stimuli in the Bailey (1975) experiments contained the same spectral patterns appropriate for activiation of an integrative dectector, there was no cross-series adaptation. If the integrative level exists, it must be spectrally specific (Ades, 1976). However, Ades (1976) further states that these experimental results can be explained by transition detectors, that are independently sensitive to frequency changes at each formant, and, therefore would be responsive to the burst cue. This is consistent with Stevens (1973, 1975) who argues that place distinctions could be achieved with three different kinds of property detectors; each sensitive to a different set of acoustic characteristics of transition and burst. If the spectral energy of the burst is in the same frequency area as the major spectral peak for the vowel, the burst cue is not regarded as a separate parallel cue, but as one component of the formant transition cue (Blumstein and Stevens, 1975; Dorman, Studdert-Kennedy, and Raphael, In press).

In the first of a series of experiments exploring the levels of perceptual processing, Sawusch (1976) used a /bæ/-/dæ/ test series with eight different adapting stimuli. Four full syllable adaptors were used: the end-points of the test series, /bæ/-low (frequency) and /dæ/-low (frequency) and /bæ/-high (frequency) and /dæ/-high (frequency) constructed from the original test series but with the center frequencies of the formants logarithmically scaled one and one-half critical bandwidths higher. The last four adaptors ("chirps") were the first 45 msec of each of the two "high" and "low" adaptors previously described, containing only the transitional information of the syllables.

The results indicated that all four of the full syllable adaptors produced significant boundary shifts towards the adapting syllables. All the chirp adaptors except the /bæ/-high chip, also produced significant boundary shifts toward the adapting stimulus. The high adaptors were approximately 33-40% as effective as their low counterparts. Sawusch (1976) cited these results, as well as those of Ganong (1975) using burst-cued adaptors and transition cued adaptors, as evidence for two levels of processing: an integrative level where frequency-specific transition information is irrelevant, and a frequency-specific auditory level based on the larger adaptation effects for the low syllables and chirps. Also, the fact that the chirp adaptors generally produced adaptation effects as large as their full syllable counterparts, was evidence for a mechanism tuned to formant transitions and not affected by steady-state vowels.

The question concerning the location of the adaptation effect has been investigated for place of articulation. Ades (1974a), using a

/bæ/-/dæ/ test series, investigated dichotic and interaural transfer of adaptation. In the first experiment, the unadpated boundaries were found monaurally, while the adaptors /bæ/ and /dæ/ were presened simultaneously to each ear. There was a significant dichotic adaptation effect in that the boundary could be shifted towards the /bæ/ in one ear and the /dæ/ in the other.

In a second experiment, Ades (1974a) presented the adaptors to one ear and the test syllables to the other ear. Adaptation in the contralateral ear yielded 55% as great an effect as did adaptation in the ipsilateral ear, with no greater ipsilateral effect in either ear. Ades (1974a) cited the transfer data as evidence for a central site of adaptation, although a peripheral mechanism could also be operating, indicating a bilateral component to the effect. A peripheral effect would be ruled out by 100% transfer, but since only a 55% effect was obtained, Ades (1974a) sought to establish the credibility of a central site of adaptation by investigating adaptation of binaurally fused stimuli.

In a third experiment, Ades (1974a) presented the first 38 msec of the syllable /bæ/ as an adaptor in a spectrally fused condition (the first formant was presented to one ear and the second- and third-formants to the other, simultaneously) and an unfused condition (the first formant was presented to one ear and the second- and third-formants to the other ear, not simultaneously). Although significant adaptation effects were found in both the fused and unfused conditions, greater effects were obtained in the fused condition. Ades (1974a) concluded that some component of adaptation takes place subsequent to binaural fusion, and therefore, centrally.

Ades (1974a) proposed two models to account for these findings, a one-level bilaterally represented (central) model and a two-level monaurally represented model with a more central second tiered component.

Sawusch (1976) sought to distinguish between these two models by finding an adapting stimulus that would show 100% interaural transfer. Sawusch assumed that a two-level model would account for 100% transfer due to an integrative dector, however, a one-level bilateral model could not account for 100% transfer since it has no central component. A /bæ/-/dæ/ test series with the eight adaptors previously discussed was used. The test series was always presented in the right ear while the adaptor was presented to either the right or left ears. For the /bæ/-low syllable adaptor (the same bandwidth as the test syllables), significant boundary shifts were found for both the monaurally and dichoticaly presented adaptors with the monaural condition producing significantly greater shifts than the dichotic. However, the /bæ/-high adaptor produced no significant differences between the two presented conditions although significant shifts were found for both.

Similar shifts were found with the /dæ/-low and chirp adaptors. The /bæ/-low and /dæ/-low chirp adaptors were similar to the full syllables with greater boundary shifts for the monaural presentation. All the high-chirp conditions were significant except for the /dæ/-high chirp presented monaurally. As predicted, the high syllables and chirps yielded an average transfer of 100% (monaural and dichotic conditions yielded the same boundary shifts) while the low syllables and chirps yielded only an average transfer of 47%. Sawusch (1976) suggested that these results, as well as the results of the previously discussed experiment from this series, are strong evidence of two levels of adaptation, ruling out the one level bilateral model proposed by Ades (1974a).

Other consonant features studied include frication, manner, and continuants. Cole, Cooper, Singer, and Allard (1975) studied manner of articulation using natural speech stimuli. These stimuli were constructed by splicing successively longer segments from the initial syllables of the following pairs of test stimuli: /m/-/b/, /j/-/d/, /v/-/b/ and /f/-/b/. Significant boundary effects were found for all the end-point adaptors except the /ba/ in the /fa/-/ba/ test series. Adaptation was also found with the alveolar nasal /na/ and with a 150 msec segment of nasal resonance alone. Cross-series adaptation was found for the /di/ and /ji/ syllables on the /ja/-/da/ test series.

Cole and Cooper (1975) investigated selective adaptation of a fricative continuum from /sa/ to /za/ to /ða/. In a previous experiment, they found that the duration of the friction prior to a vowel was sufficient to distinguish between a voiced and voiceless fricative or affricate in the syllable-inital position. The stimuli used in this experiment were a series of CV syllables generated from a naturally spoken /sa/ by removing successively larger segments of friction prior to the vowel. Significant boundary shifts were found for the /sa/ and /za/ adaptors in the predicted direction and nonsignificant shifts with the /ða/ adaptor. Cole and Cooper (1975) conclude on the basis of the /s/ and /ð / data, that the magnitude of selective adaptation is primarily dependent upon the duration of frication although the series of stimuli differ by other acoustic features. However, the ambiguity of the /z/ stimulus and the differences in the waveform envelope for these stimuli make this conclusion tentative at best.

Two recent experiments have studied the selective adaptation of stop consonants and glides. Cooper, Ebert, and Cole (1976) used two synthetic continua /ba/ to /wa/ and /ga/ to /ja/. The two end-point adaptors in each series differed from each other in the duration and rate of the first three formant transitions. The stimuli from the two series differed from each other in the starting frequencies and directions of the second-formant transitions. Systematic shifts in the phonetic boundaries were found for both series.

In a second experiment, two /ga/ adaptors, each of 35 msec duration were used with the formant transitions of the /ba/-/wa/ series. One /ga/ adaptor contained the first- and second-formant starting frequencies equal to those of the /ga/-/ja/ series, while the second /ga/ contained second- and third-formant starting frequencies further apart from each other in order to eliminate the simulated burst in the first set. After adapatation with the /ga/ adaptors, significant boundary shifts were found with fewer responses assigned to the /b/ category. These results, according to Cooper et al., (1976) suggest the operation of an analyzer that processes rate and duration information differently for both labial and velar consonants. The fact that the two /ga/ adaptors contained the same formant transition durations as the /wa/ syllables, yet moved the category boundary toward /ba/, suggests that rate of transition is analyzed in a manner not independent of place.

Diehl (1976) used a /ba/-/wa/ test series with six different phonetic adapting stimul: the /ba/ and /wa/ end-point stimuli, a /da/ and /ja/ that contained different F_2 and F_3 starting frequencies and transition durations than the series stimuli, and the /ba/ and /wa/ transitions in isolation with no steady-state porition. Significant boundary shifts were found with all the adaptors except the /ja/, although the full syllable /ba/ and /wa/ adaptors produced the greatest shifts. These results support two assumptions according to Diehl (1976): that selective adaptation takes place at an auditory level of analysis and that phonetic analysis is probably not a necessary condition of adaptation.

Similar phoneme boundary shifts were found by Bailey (1975) who used four test series: /bʒ/-/d3/, /w3/ to /j3/,/b3/-/w3/ and /d3/-/j3/. The first two series above differed in the onset frequency of the second- and third-formant transitions, the transition durations (30 msec and 126 msec respectively) and the relative transition rate, but were similar in the rate and direction of the transitions. The last two series differed in the direction of the second- and third-formant transitions (rising in the /b-w/ series and falling in the /d-j/ series.) Within a series, transition onset frequencies were constant but transition duration and rate differed. The adaptors were the four end-points of the series.

Significant boundary shifts were found for all the within series adaptors except the /bʒ/ adaptor in the /b-w/ series. Surpisingly, the /bʒ/ adaptor shifted the boundary in the /w-j/ series, a cross-series effect. Bailey attributed these cross-series results to the spectral similarity of the F_2 and F_3 transitions of the stop adaptor on the /w/-/j/. The failure of the /w/ and /j/ adaptors to produce significant boundary shifts on the /b/-/d/ boundaries was attributed to adaptation at a temporal representative stage of processing.

Summary of Selective Adaptation Studies

Although there is early experimental evidence for both voicing and place that supports selective adaptation at a phonetic feature level (Ades, 1974b; Cooper, 1974a; Diehl, 1975; Ganong, 1975; Tartter and Eimas, 1975), most theorists have assumed that a single acoustic site is more valid and parsimonious (Ades, 1976; Eimas and Miller, In press). The evidence for selective adaptation at the phonetic feature level was based on the generalization of adaptation for place of articulation across vowels. However, this generalization can be explained, for the most part, by detectors at an auditory level (Darwin, 1975; Ades, 1976; Eimas and Miller, 1976; Sawusch, 1976).

Experiments where the spectral information in the adapting and test stimuli have differed (Cooper, 1974c; Bailey, 1975) with the resultant failure or reduction in the adaptation effect, would rule out a phonetic level of adaptation, as would shifts in the phonetic boundary after adaptation with selected components or truncated portions of the speech signal (Ades, 1974b; Pisoni and Tash, 1975; Tartter and Eimas, 1975). Finally, position sensitive adaptation (Ades, 1974b; Pisoni, 1975b) would rule out adaptation based on an abstract invariant phonetic feature.

Although there is evidence for adaptation at the acoustic feature level (Ades, 1973, 1974b; Bailey, 1975; Tartter and Eimas, 1975), there is also evidence for adaptation at an integrative level, where input from more than one acoustic feature is received and the response is to a particular pattern of auditory cues (Cooper, 1974b; Blumstein and Stevens, 1975; Cooper, Ebert and Cole, 1975; Ganong, 1975), much the same as the "property detectors" proposed by Stevens, (1973, 1975).

Adaptation has also been attributed to two levels of processing (Bailey, 1975; Cooper, 1975; Darwin, 1975). Tartter and Eimas (1975) suggest that syllables adapt at both an auditory and phonetic level while non-speech sounds adapt at an auditory level only. Sawusch (1976) attributes adaptation to a frequency specific peripheral level and a second integrative auditory level, not frequency specific.

However, Ades (1976) points out that two strengths of adaptation, as indicated by the Tartter and Eimas (1975) data, is not evidence for two levels of adaptation, but could mean that there is just one level, engaged more fully by the whole syllable than by parts of it. Darwin (1975) also states that if adaptation is occuring at a higher level than the acoustic feature, then the syllable would explain the position specific effect and the greater effect of the entire speech sound than of any of the acoustic features.

The other question that has emerged in adaptation concerns the extent to which adaptation is the result of sensory fatigue rather than the result of response bias or response organization effects. Although other experimental evidence has accumulated against the position that selective adaptation is solely due to response bias or response oranization effects (see: Other Explanations for Boundary Shifts in Selective Adaptation), the results of Sawusch and Pisoni (1973) and Sawusch, Pisoni, and Cutting (1974) are of major interest in this series of experiments. As previously discussed, in an anchoring experiment, Sawusch and Pisoni (1973) and Sawusch, Pisoni, and Cutting (1974) assumed that anchoring effects due to response biases might be evident in a category boundary shift for a non-linguistic continuum, but that linguistic stimuli would be immune to such effects.

In the pre-anchoring condition, both the speech and the non-speech dimensions of the stimuli were presented equally often while on the second condition, the end-point stimuli occured two or four times as often as the other stimuli in the series. Subjects were to respond using the categories of the non-linguistic dimension (loud or soft for intensity, for example) in one task, and the categories of the linguistic dimension (/ba/ or /pa/ for the voicing), for the second task. The results of these experiments demonstrated perceptual changes as a result of the differences in the probability of occurrence of the non-speechh stimulus dimension (fundamental frequency or tones varying in intensity), although these changes were not demonstrated for the speech stimulus. They concluded that the category boundaries for the consonants are not subject to the same context effects as boundaries between fundamental frequency categories, and therefore, ruled out response bias explanations of category boundary shifts in selective adaptation.

However, the failure to show category boundary shifts for the phonetic stimuli with the anchoring paradigm, is at best, indirect evidence that these shifts in the consonant boundaries with the adaptation paradigm are sensory in nature, for it rests on the possibly erroneous assumption that the two paradigms differ in kind. Since the main difference between the two paradigms is essentially in the placement of the adaptors or anchor stimulus, it is possible that the two paradigms are simply different versions of the same procedure.

The results of these studies have raised the following questions: 1) Does the difference between the phonetic and non-phonetic boundary shifts obtained by the Sawusch, Pisoni, and Cutting (1974) study reliably occur with other non-speech tasks? 2) Is there a fundamental difference betweeen the selective adaptation paradigm frequently used in speech perception experiments and the "anchoring" paradigm used by Sawusch, Pisoni, and Cutting (1974)? The purpose of this series of studies is to explore the proposed feature detecting systems by comparing the two paradigms, selective adaptation and anchoring in speech and non-speech tasks.

CHAPTER II: EXPERIMENT I

ANCHORING OF CORRELATED FUNDAMENTAL FREQUENCY AND PHONEME CONTINUUM

The purpose of this experiment was to compare anchoring effects on the identification of a synthetic stop consonant continuum with those on the identification of a fundamental frequency continuum, using the same stimulus continuum for both tasks. The aim was to replicate a study of Sawusch, Pisoni, and Cutting (1974) as a baseline for a series of parametric studies, involving speech versus nonspeech tasks. Sawusch, Pisoni, and Cutting (1974) used a 7-step synthetic stop, place-of-production continuum and paired each stimulus value along the continuum with each of 7 variations in fundamental frequency ranging from 114 Hz to 150 Hz in 6 Hz steps. Forty-nine stimuli were therefore produced, making one dimension totally independent of the other. The fact that the experimenters obtained a boundary shift in the nonlinguistic or fundamental frequency condition, but no shift in the place, or linguistic condition, suggested that subjects' identification along a synthetic stop consonant continuum is not subject to simple psychophysical bias effects. From this the experimenters inferred that the boundary shifts typically observed in adaptation studies of stop consonant continua are probably perceptual in nature rather than the result of response bias.

While the evidence of this study and of an earlier study using stop syllables and tones (Sawusch and Pisoni, 1973) disputes a response bias account of speech adaptation, an interesting question arises. Why has the selective adaptation paradigm first employed in speech by Eimas and Corbit (1973) and used in numerous studies since, obtained shifts in the phoneme boundary, while the Sawusch, Pisoni, and Cutting "anchoring" paradigm obtained no shift in the boundary for speech but obtained a significant shift in the boundary for non-speech?

The following experiments attempt, therefore, to answer two questions: (1) Does the difference between the linguistic and non-linguistic boundary shifts obtained by the Sawusch, Pisoni, and Cutting study reliably occur with other non-speech tasks? (2) Is there a fundamental difference between the selective adaptation paradigm frequently used in speech perception experiments and the "anchoring" paradigm used by Sawusch, Pisoni, and Cutting?

Experiments I, II, and III attempted to answer the first question by comparing the anchoring effects in the identification of synthetic stop consonant continua with those in the identification of fundamental frequency, intensity, and vowel continua. It was expected that the non-linguistic dimensions as well as the vowel continuum could be manipulated by this paradigm but that the stop-consonant continuum would be immune to psychophysical anchoring effects.

The first experiment replicated the above study with one exception. Perfect correlation between the speech and non-speech continuum was introduced by constructing only 7 stimuli instead of the 49 stimuli used above. This meant that a given CV syllable was paired consistently with only one F_0 value. If a shift in the fundamental frequency boundary occurs without a shift in the phoneme boundary, an even stronger case can then be made for the lack of a response bias effect in adaptation.

Method

Description of Stimuli

A series of three-formant consonant-vowel syllables was generated on the Haskins Laboratories parallel resonance synthesizer. The set of stimuli consisted of seven syllables each with a duration of 300 msec. The final 260 msec was a steady-state portion appropriate to the American English vowel /æ/ and was identical in all stimuli: 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. The starting frequencies and directions sufficient to cue changes in place of articulation between the stop consonants /b/ and /d/ are presented in Table 1. F_1 remained constant throughout the seven stimuli.

Fundamental frequency at syllable onset varied from 114 Hz to 150 Hz in 6 Hz steps from /bæ/ to /dæ/ and fell linearly from the initial value of each syllable to 80 Hz during the last 75 msec of the syllable. Therefore, a particular fundamental frequency contour characterized each stimulus along the phoneme continuum. Overall amplitude was attenuated linearly in the last 75 msec of each stimulus. Formant frequency amplitude rose and declined linearly. The stimuli, therefore, differed from one another only in the starting frequencies and the direction of

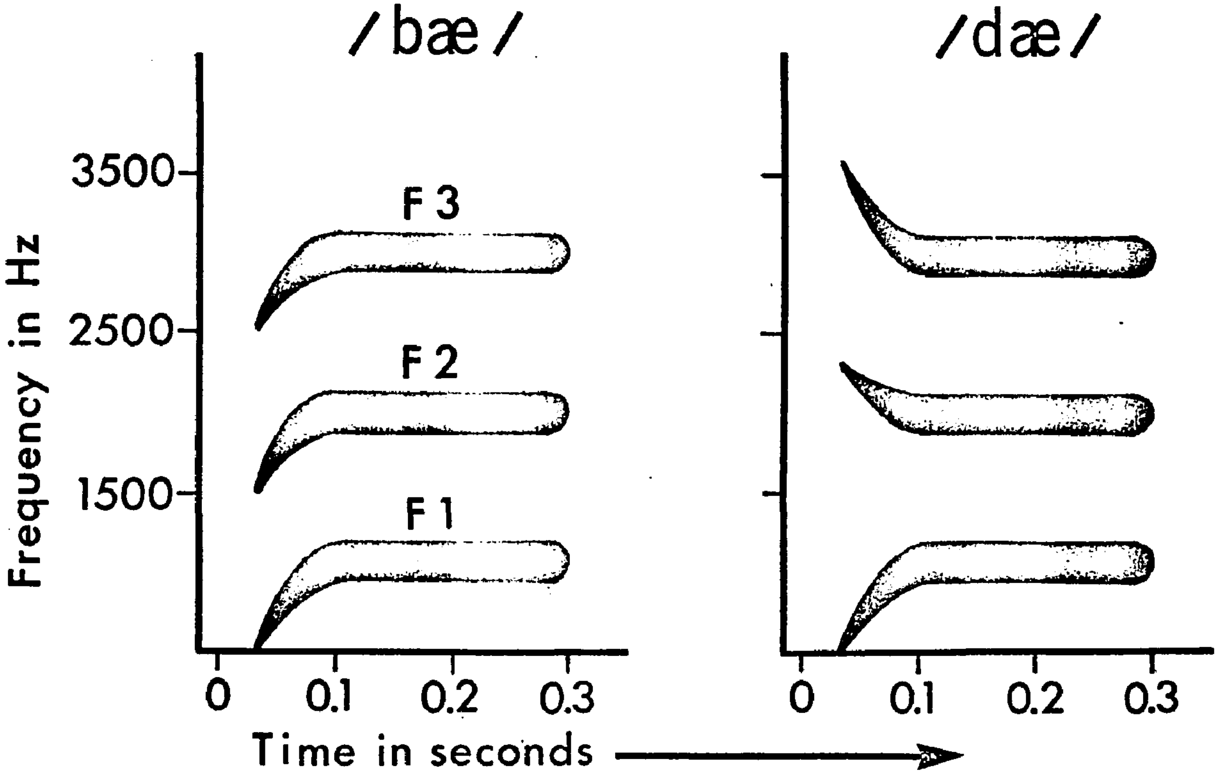
Table 1. Starting frequencies of the second- and third-formant transitions for the synthetic CV stimuli and correlated fundamental frequency.

Starting frequencies of the second- and third-formant transitions for the synthetic CV stimuli and correlated fundamental frequency.

Stimulus Number	F_0	F_2	F_3
	(Frequency in Hz)		
1	114	1386	2525
2	120	1465	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).

Figure 1. Schematized spectrographic patterns appropriate for the syllables /bæ/ and /dæ/.



STIMULUS 1 STIMULUS 7

the second- and third-formant frequency transitions and in the fundamental frequency of each syllable. Schematized spectrographic patterns of the two end-point stimuli, /bæ/ and /dæ/ are illustrated in Figure 1.

The anchoring stimulus was the first /bæ/ stimulus at 114 Hz F_0 described in Table 1. Since a differential effect of the two end points (/bæ/ and /dæ/) was not of interest in this experiment, only one end-point, /bæ/, was used as the anchoring stimulus. This anchoring stimulus (/bae/ -low F_0) will hereafter be referred to as the anchor or stimulus 1. The six other syllables (stimuli 2-7 on the continuum) will be referred to as the test syllables. The stimuli were recorded on magnetic tape, digitized, edited, and stored on the Haskins pulse code modulation (PCM) system (Cooper and Mattingly, 1969). Stimuli were reconverted to analog form and experimental tapes were prepared from these stimuli.

Experimental Tapes

Two different experimental tapes were prepared for this experiment. In the pre-anchoring or identification tape, the seven stimuuli were recorded equally often: 10 times in a randomized test order producing 70 stimuli. In the /bæ/-low F_0 anchor tape, the anchor stimulus was recorded 4 times, the other six stimuli once each; ten randomly ordered groups of these 10 stimuli (4 anchors + 6 others) were then produced, making a total of 100 stimuli for the anchor condition. The stimuli on both tests were recorded singly with a two second interval between stimuli and a 10 second interval after every tenth stimulus.

The stimuli were recorded on a Crown 800 tape recorder, connected directly to the output of the PCM system. A 1000 Hz calibration tone, recorded at the maximum vowel amplitude on the VU meter of the tape recorder, was placed at the beginning of each tape to insure uniform record and playback levels.

Subjects

Nine normal hearing (screened at 20dBHTL 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.

Procedure

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 75dB SPL re 0.0002 dyne/cm^2. Subjects were tested in a small isolated room at Haskins Laboratories.

The experimental design for this study required identification of the members of the stimulus series under two conditions: unanchored (pre-anchor) and anchored, for each of the two task conditions, pitch identification and phoneme identification. Each subject heard the same pre-anchor (70 item) tape and the same anchor (100 item identification) tape for each of the two task conditions. Only the directions to the subjects varied as the task changed. The order of testing was as follows for all subjects: fundamental frequency pre-anchor, fundamental frequency anchor, phoneme pre-anchor, 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 prior anchoring effect in the pitch task, and in spite of the correlation between fundamental frequency and phoneme category.

A standard set of instructions was read aloud to the subjects. They were told they would hear a series of sounds differing in pitch presented on two tapes identical in every way save that the second had 100 stimuli while the first had only 70. Other aspects of the stimuli would vary but their task was to write down "high" or "low" to identify the pitch. Similarly, for the phoneme identification task, the subjects listened to the identical tapes, but were required to identify each syllable as beginning with /b/ or /d/. Each experimental tape was preceded by listening and practice tapes.

Results

The average pre-anchor and anchor functions plotted for the subjects as a group are shown in Figures 2 and 3. Figure 2 shows the probability of high pitch judgments as a function of the stimulus continuum in the fundamental frequency task, while Figure 3 shows the probability of /d/ judgments in the phoneme task.

Figure 2. Average identification function for fundamental frequency task.

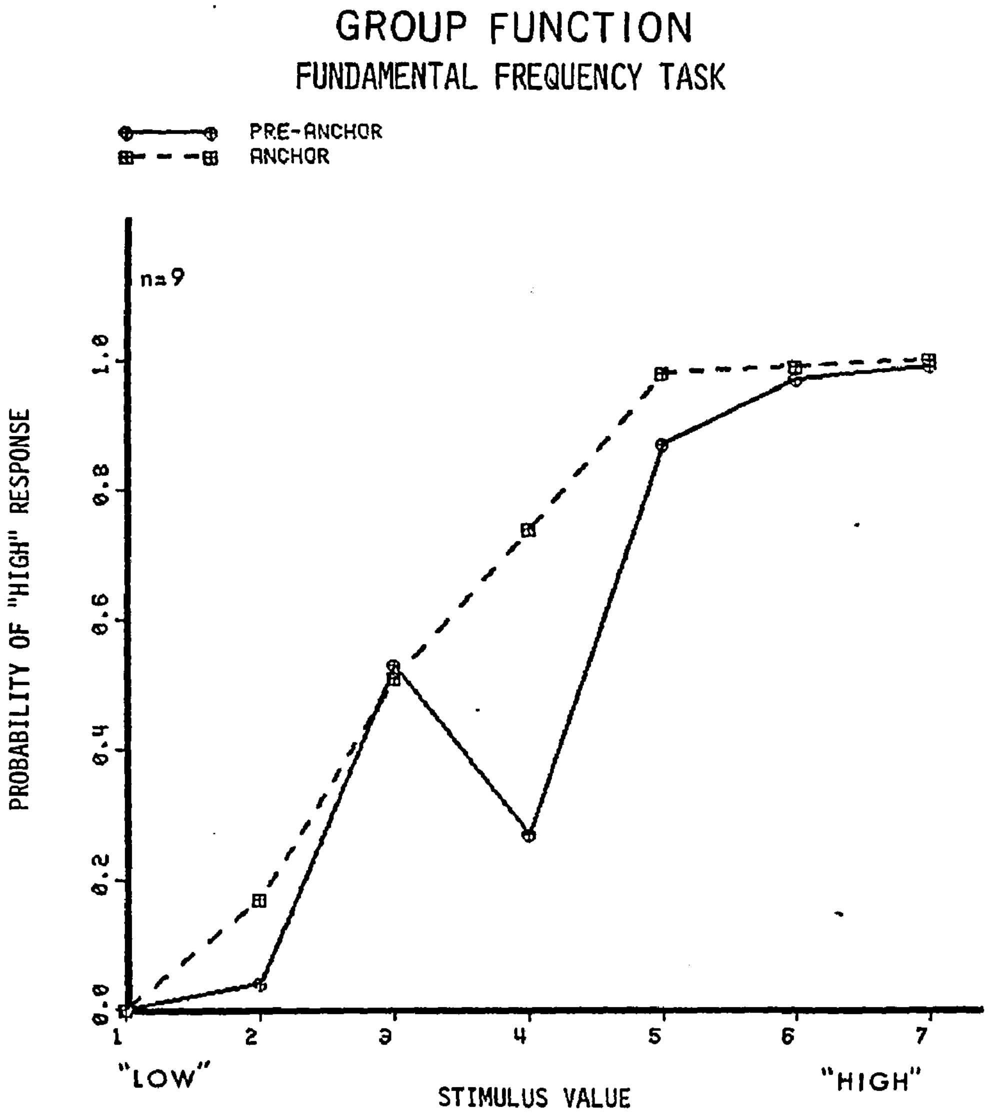
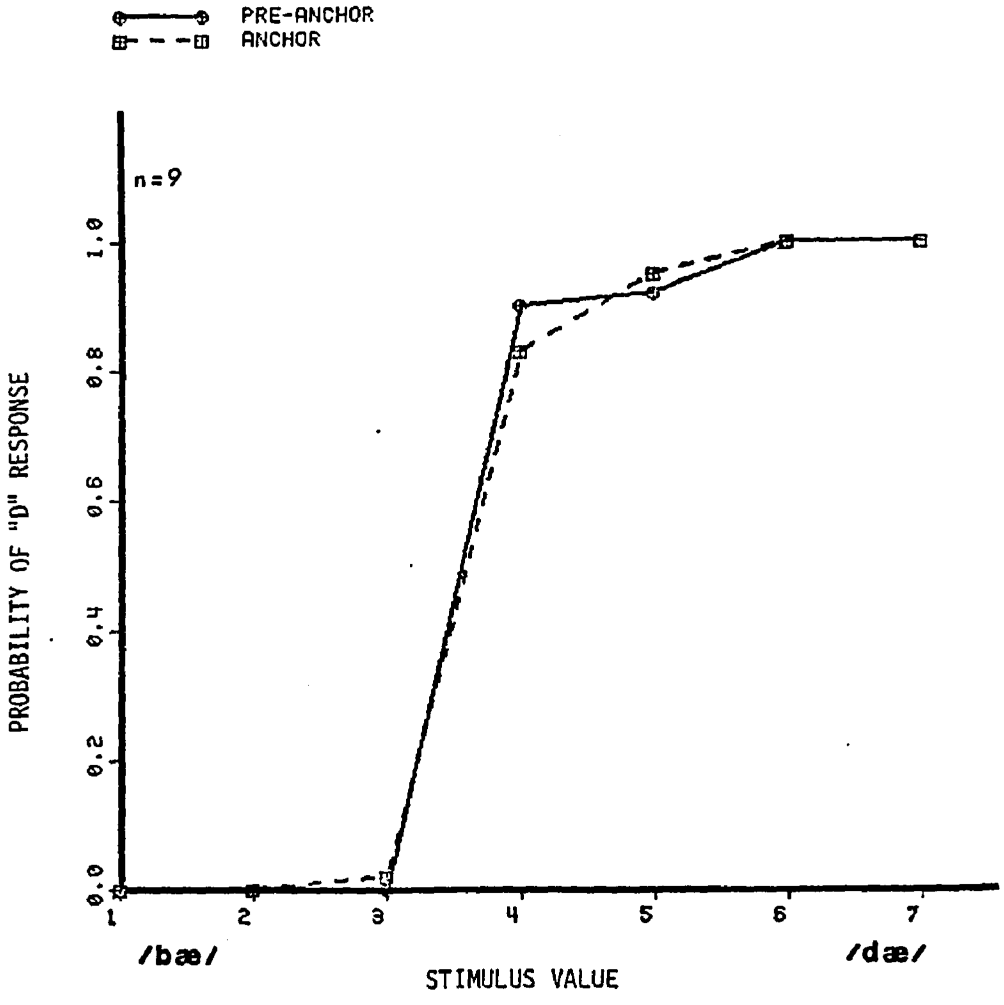

GROUP FUNCTION
FUNDAMENTAL FREQUENCY TASK
PRE-ANCHOR
ANCHOR
n=9
PROBABILITY OF "HIGH" RESPONSE
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
"LOW"
STIMULUS VALUE
"HIGH"

Figure 3. Average identification function for phoneme task.

GROUP FUNCTION
SPEECH TASK



The pre-anchor phoneme function indicates that the subjects partitioned the stimulus continuum into two distinct categories: "b" versus "d". Comparison of the phoneme function (Figure 3) with the fundamental frequency function (Figure 2) reveals that the shift from one response category to another is sharper for the speech judgment than for the pitch judgment, especially in the pre-anchor condition.

The fundamental frequency anchoring task function shows a shift toward the anchor or stimulus 1 relative to the pre-anchor function, in other words, subjects made more "high" pitch responses on the anchor tape than on the pre-anchor tape (Figure 2). Figure 3 illustrates that the phoneme task function shows almost no shift of the anchor task judgments relative to the pre-anchor judgments.

The slopes of the fitted functions were of some concern, since it might be argued that the degree of boundary shift is simply a function of the number of ambiguous stimuli susceptible to anchoring. Since, further, the choice of response steps on the fundamental frequency continuum was arbitrary, and since that choice determines both the number of ambiguous stimuli and the slope of the function, it might be supposed that the difference in boundary shift between stop consonants and fundamental frequency was simply the result of an arbitrary choice of fundamental frequency values.

Boundary loci for individual subjects were computed by a least squares analysis on z scores (Probit Analysis) (Woodworth, 1938), derived from the proportion of "high" or /d/ responses assigned to each stimulus. The individual and group boundaries, together with pre-anchor-to-anchor boundary shifts for each test condition, are presented in Table 2. Individual and mean standard deviations are presented in Table 3. The boundary locations for each of the continua were 3.63 and 3.24 for the pre-anchor and anchor conditions in the fundamental frequency task and 3.74 and 3.83 for the pre-anchor and anchor conditions in the speech task. Inspection of the mean boundary locations for the two task conditions (Table 2) indicates that for eight out of the nine subjects, the boundary in the fundamental frequency task shifted towards the anchor stimulus as expected, while in the phoneme task, slight boundary shifts were measured in the opposite direction for 7 out of 9 subjects. The mean boundary displacement for the fundamental frequency condition was 0.39 continuum steps toward the anchor, while for the phoneme condition, it was 0.09 steps away from the anchor. However, although the standard deviations (i.e., the reciprocal of the slopes) of the ogives were with one exception (ST) greater for the fundamental frequency tasks (1.16 pre-anchor steps and 1.32 anchor steps) than for the phoneme tasks (1.02 pre-anchor and 1.03 anchor), these differences were not significant ($T_8 = \underline{p} < 0.05$).

Table 4, which shows the individual and mean number of "high" and /d/ responses and shifts for each condition, also indicates a difference between the two tasks. In the fundamental frequency task, there was a mean shift of -7.22 responses from the pre-anchor condition to the anchor condition while in the speech task there was a shift of .22 responses in the opposite direction. Stimulus 1 was the anchor stimulus so that a negative value reflects a shift towards the anchor stimulus or

Table 2. Individual and mean boundaries and boundary shifts for each test condition.

Individual and mean boundaries and boundary shifts for each test condition.

	F_0			SPEECH		
Subject	Control	Anchor	Control-Anchor	Control	Anchor	Control-Anchor
SG	3.28	3.87	-0.59	3.67	3.67	0
CL	3.72	3.42	0.30	3.92	3.96	-0.04
LS	3.15	2.93	0.22	3.67	3.80	-0.13
SR	3.68	3.23	0.45	3.67	3.50	0.17
GC	3.60	2.09	1.51	3.87	3.97	-0.10
HL	3.79	3.39	0.40	3.72	3.92	-0.20
MW	3.63	3.48	0.15	3.67	3.88	-0.21
ST	3.83	2.85	0.98	3.67	3.92	-0.25
DR	4.04	3.91	0.13	3.81	3.90	-0.09
MEAN	3.63	3.24	0.39	3.74	3.83	-0.09

Table 3. Individual and mean standard deviations for each test condition.

Individual and mean standard deviations

for each test condition.

	F_0		SPEECH	
Subject	Control	Anchor	Control	Anchor
SG	1.13	1.10	1.00	1.00
CL	1.08	1.06	1.04	1.00
LS	1.27	1.52	1.00	1.08
SR	1.09	1.20	1.00	1.04
GC	1.20	1.61	1.06	1.00
HL	1.15	1.18	1.07	1.04
MW	1.54	1.08	1.00	1.00
ST	.93	1.35	1.00	1.00
DR	1.08	1.78	1.04	1.06
MEAN	1.16	1.32	1.02	1.03

Table 4. Individual and mean number of "high" and /d/ responses and shifts for each test condition.

Individual and mean number of "high"

and /d/ responses and shifts for each test condition.

	F_0			SPEECH		
Subject	Control	Anchor	Control-Anchor	Control	Anchor	Control-Anchor
SG	22	38	-16	40	40	0
CL	38	42	- 4	36	36	0
LS	45	46	- 1	40	39	1
SR	38	45	- 7	40	41	- 1
GC	40	54	-14	38	40	- 2
HL	39	42	- 3	32	36	- 4
MW	35	43	- 8	40	38	2
ST	39	49	-10	40	37	3
DR	34	36	- 2	39	36	3
MEAN	36.7	43.9	-7.2	38.3	38.1	.2

an increase in the number of "high" or /d/ responses while a positive value reflects a shift away from the anchor stimulus.

From Tables 2 and 3, it is evident that individual means in the speech-place task show little variability as compared with those in the fundamental frequency task: on the speech task, 5 subjects had identical means on the pre-anchor tape, two on the anchor tape, and no two subjects had identical means on either of the pitch tests.

A further difference between the results of the two tasks lies in the contextual effects present in the fundamental frequency task but absent in the phoneme task, as suggested by the differences in the standard deviations (Table 3), the identification functions in the pre-anchor conditions (Figures 2 and 3), and in the raw data for the two boundary stimuli, 3 and 4 (Table 4). In order to assess the contextual effects, the subjects' <u>response</u> to the stimulus immediately preceding stimuli 3 and 4 was evaluated. Table 5 gives the stimulus responses preceding stimuli 3 and 4 as well as the responses to stimuli 3 and 4. The fact that there were fewer "high" judgments for stimulus number 4 (28%) than for stimulus number 3 judgments (54%) in the pre-anchor condition suggests a contextually dependent response mode. In fact, when stimulus number 4 on the continuum followed stimulus number 5, it was called "low" by the same subjects who had called stimulus number 3 "high". When stimulus number 4 did not follow stimulus number 5, it was labeled "high." No such context effects, however, were found for the speech task, although the same stimulus ordering was followed.

Table 5. Responses to stimuli 3 and 4 in the pre-anchor fundamental frequency condition by preceding "low"/"high" stimulus response.

Responses to stimuli 3 and 4

in the pre-anchor fundmanental frequency condition

by preceding "low" "high" stimulus response.

		Preceding Stimulus Response		
Stimulus Number	Response	"Low"	"High"	Total
* 3	"Low"	46%	0%	46%
	"High"	54%	0%	54%
	TOTAL	100%	0%	
**4	"Low"	4%	68%	72%
	"High"	0%	28%	28%
	TOTAL	4%	96%	

* In the pre-anchor and anchor randomizations, stimulus number 3 was always preceded by either stimulus number 1 or stimulus number 2. Therefore, there were no immediately preceding higher-pitched stimuli or more /d/-like stimuli.

**In the pre-anchor randomization, stimulus number 4 was always preceded by a higher-numbered stimulus, therefore, there were no immediately preceding lower-pitched stimuli or more /b/-like stimuli.

In the pre-anchor randomization, stimulus number 4 was always preceded by a higher-numbered stimulus, therefore, there were no immediately preceding lower pitched or more /b/-like stimuli. Ninety-six percent of these preceding stimuli were judged to be "high" by the subjects. Of these "high" judgments, Table 5 shows that 25 (28%) of the following stimulus number 4 responses were judged to be "high" while 61 (68%) were judged to be "low." For the 4 preceding stimuli judged to be "low", 100% of the 4 following stimulus number 4 responses were also judged "low." These results alone, however, are not proof of contextual effects affecting the subjects' responses. However, inspection of the data for stimulus number 3 may be more illuminating.

In contrast to stimulus number 4, stimulus number 3 was always preceded by a lower-numbered stimulus in the pre-anchor condition. If the responses to this stimulus were based on the stimulus continuum alone, a "low" response would generally be expected since this stimulus lies below the midpoint of the continuum. On the other hand, if the responses to this stimulus were contextually based on the preceding stimulus (always judged "low"), a large number of stimulus number 3 responses would be expected to be "high," indicating a response bias in the opposite direction. However, since stimulus number 3 responses were only a little more than 50% "high," the result is ambiguous. It is possible that this level of "high" responsiveness is due to the fact that stimulus number 3 is the boundary stimulus and not due to the effect of the preceding stimulus responses. In fact, inspection of the data for the anchor condition, where there were always lower-numbered stimuli preceding stimulus number 3, shows the same pattern of results as does the pre-anchor condition, with 51% "high" responses. But it is possible that stimulus number 3 in both the pre-anchor and anchor conditions is influenced by the preceding stimulus and that the true boundary is stimulus number 4. In fact, inspection of the pre-anchor functions for fundamental frequency in Experiment VI (Chapter VII) shows that there was only 17% and 18% "high" responses for stimulus number 3. However, inspection of the stimulus ordering for this experiment shows that 8 out of the 10 stimuli preceding stimulus number 3 were higher-numbered stimuli. Therefore, just the opposite condition exists for Experiment VI, so that the lower percentage of "high" responses for stimulus number 3 might be the effect of more higher-numbered preceding stimuli.

In the anchor randomization for stimulus number 4, the distribution of immediately preceding higher-numbered and lower-numbered stimuli was equal. However, the effects of the preceding stimulus alone cannot account for the differences between the pre-anchor and the anchor results, for obviously, the effects of the anchoring stimulus must be considered. No such context effects, however, were found in the speech task, although the same stimulus ordering was followed. For stimulus number 4, for example, all 90 stimuli in the pre-anchor condition were preceded by higher-numbered and /d/ judged responses, (See Table 6), but only 9 (10%) of the occurrences of stimulus number 4 were judged /b/, in comparison to 61 (68%) judged "low" for the fundamental frequency task. Therefore, it seems that the phoneme judgment is independent of the immediately preceding stimulus item. Similar contextual effects have been found in judgments of vowel continua (Fry et al., 1962; Stevens et

Table 6. Response to stimuli 3 and 4 in the pre-anchor speech condition by preceding /b/, /d/ stimulus response.

Responses to stimuli 3 and 4

in the pre-anchor speech condition

by preceding /b/, /d/ stimulus response.

Stimulus Number	Response	Preceding Stimulus Response /b/	/d/	Total
* 3	/b/	100%		100%
	/d/	0%		0%
	TOTAL	100%		
**4	/b/	0%	10%	10%
	/d/	0%	90%	90%
	TOTAL	100%	100%	

* In the pre-anchor and anchor randomizations, stimulus number 3 was always preceded by either stimulus number 1 or stimulus number 2. Therefore, there were no immediately preceding higher pitched or more /d/-like stimuli.

**In the pre-anchor randomization, stimulus number 4 was always preceded by a higher-numbered stimulus, therefore, there were no immediately preceding lower-pitched stimuli or more /b/-like stimuli.

al., 1969), but not found in judgments of stop consonant continua (Liberman, et al., 1961; Studdert-Kennedy, et al., 1970).

A three-way analysis of variance was carried out on the data (Table 7). The three variables were Subjects (9), ratio of each of the test stimuli (#'s 2-7) to the anchor (1:1 pre-anchor or 4:1 anchor) and continuum (fundamental frequency, phoneme). The absolute numbers of "high" and /d/ responses were used for the analysis. The main effect of ratio was significant ($F_{1,8} = 12.3356$; $p < 0.01$) indicating that a significant anchoring effect was obtained (Table 7). Although the identification functions representing each of the two continua differ considerably, the effect of continuum was not significant. This is presumably due to the large individual differences among subjects evidenced by the large error term of the CS value. However, the ratio by continuum interaction was significant ($F_{1,8} = 15.820$; $p < 0.01$), presumably due to the greater anchor effect on the pitch task than on the phoneme task.

A Wilcoxon Signed-Rank Test was used to assess the effect of the anchoring condition for each continuum. For the fundamental frequency condition, there was a significant shift in the category boundary in the direction of the /b /-low F_0 anchor ($T = 0$; $p < 0.005$) while there was no significant shift in the category boundary for the phoneme condition. The difference between the two anchor conditions was also significant ($T = 1$; $p < 0.05$).

Table 7. Summary table of analysis of variance for the total number of "high" and /d/ responses.

Summary table of analysis of variance for the total number of "high" and /d/ responses. Ratio (R): 1:1, 4:1; Continua (C): F_0, Phoneme; Subjects (S).

Source	df	MS	F	p
Subjects (S)	8	32.188		
R	1	110.250	12.335	< 0.01*
RxS	8	8.938		
C	1	38.027	1.2	NS
CxS	8	32.215		
RxC	1	124.694	15.820	< 0.01*
RxCxS	8	7.881		
TOTAL	27			

Summary and Discussion

The following general observations, concerning anchoring effects on the category boundaries of a stop consonant judgment and a fundamental frequency judgment, using the same stimulus continuum for both tasks can be made.

1) In the nonlinguistic task, a shift in the category boundary towards the anchor stimulus was observed when the probabilities of occurrence of the anchor stimulus in relation to the other stimuli were unequal. There was no shift in the predicted direction with the linguistic judgment.

2) The identification functions for stop consonants were sharper and more consistent for most subjects than the corresponding functions for fundamental frequency, and also did not display the contextual effects found in the fundamental frequency function.

The overall conclusion to be drawn from this experiment is essentially the same as that of Sawusch, Pisoni, and Cutting (1974): synthetic stop consonant continua are less susceptible to psychophysical anchoring (or bias) effects than are fundamental frequency continua. However, the conclusion is even stronger in the present study, since each value of place was always paired with a given fundamental frequency, so that one dimension was fully predictable from the other. Despite this deliberately imposed association of stimulus dimensions, linguistic and nonlinguistic response patterns displayed definite dissociation.

There appear, therefore, to be differences in the processing of the two dimensions used in this study. The next study will attempt to replicate these results using intensity instead of fundamental frequency as the nonlinguistic dimension.

Footnote

[1] All Wilcoxon Matched-Pairs Signed-Ranks Tests, unless otherwise stated, were one-tailed tests. The one-tailed test was employed because of the expected directional changes in the category boundaries.

CHAPTER III: EXPERIMENT II

ANCHORING OF CORRELATED AMPLITUDE AND PHONEME CONTINUUM

The purpose of this experiment was to explore further the anchoring paradigm in correlated linguistic and nonlinguistic conditions, this time using intensity as the nonlinguistic dimension varied over the stimulus continuum.

Method

Description of Stimuli

The stimuli were again the series of consonant-vowel syllables, shown schematically in Figure 1 and used in Experiment 1, but with three differences: (1) fundamental frequency was held constant at 114 Hz for all seven stimuli in the continuum and fell linearly to 80 Hz during the last 75 msec of the syllable, while (2) initial overall amplitude was attenuated from 0 dB to 18 dB in 3 dB steps from /bæ/ to /dæ/ and fell linearly over each syllable to -28 dB during the last 75 msec of the syllable; (3) in order to conform to most adaptation experiments, the stimuli were shortened to a duration of 250 msec by dropping 50 msec of the steady-state portion. Therefore, in this experiment, a particular overall amplitude was correlated with position on the phoneme continuum.

Again in this experiment, only one end of the continuum was used as the anchor. Parameters of the anchor, /bæ/ -high amplitude and the other test stimuli are presented in Table 8.

Table 8. Starting frequencies for the second- and third-formant transitions for the synthetic CV stimuli and correlated amplitude.

Starting frequencies for the second- and third-formant transitions for the synthetic CV stimuli and correlated amplitude.

(Frequency in Hz)

Stimulus Number	F_2	F_3	Relative Overall Amplitude
1	1386	2525	0 dB
2	1465	2694	- 3 dB
3	1541	2862	- 6 dB
4	1620	3026	- 9 dB
5	1695	3195	-12 dB
6	1772	3363	-15 dB
7	1845	3530	-18 dB

Experimental Tapes

Tapes were prepared in the same way for this experiment as for Experiment I.

Procedure

The stimuli of this experiment were different from, but the procedures were the same as, those of Experiment I. 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 90 dB SPL re 0.0002 dyne/cm^2. Therefore the test stimuli varied from 90 dB SPL for stimulus #1 to 72 dB SPL for stimulus #7.

Order of testing followed the same pattern as in the previous experiment: pre-anchor intensity judgment, anchored intensity judgment, pre-anchor phoneme judgment, anchored phoneme judgment. Initially subjects were told to listen for intensity differences in the syllables and to ignore other differences observed. Instructions were to write "L" for a loud sound, "S" for a soft sound on the intensity judgment, and 'b' or 'd' for the phoneme judgment. Each experimental session was preceded by a listening and a practice tape. Subjects were tested either singly or in pairs, and testing was completed in one session with a short break between the intensity and phoneme judgment conditions.

Nine paid listeners served as subjects in this experiment. All met the criteria established in the previous experiment.

Results

Figure 4 illustrates the group probability of "soft" intensity judgments for the seven test stimuli in both the pre-anchor and the anchor conditions. The figure indicates a shift in the perceived boundary of the loudness (50% crossover point) from the midpoint between stimulus 3 and 4 in the pre-anchor condition to stimulus 3 in the anchor condition. This half-step shift towards the anchor stimulus or stimulus 1, indicates that the subjects made more "soft" responses in the anchor condition (or less "loud" responses) than in the pre-anchor condition.

Group phoneme boundary functions in both the pre-anchoring and anchoring conditions are shown in Figure 5. No shift of the anchor task judgments relative to the pre-anchor judgments is seen except for small shifts for stimuli 1, 2, and 4.

One subject was dropped from the data analysis due to his inability to categorize consistently the phoneme in the identification task. The contextual effects evident in the fundamental frequency judgment are not apparent in the intensity judgment. Listeners were more consistent in identifying the intensity than in identifying the fundamental frequency.

Table 8 presents the total number of "soft" and /d/ responses for all subjects under all conditions together with the shift in numbers of pre-anchor to anchor conditions. This table parallels the one for Experiment I using boundary loci computed by a least squares analysis on z scores (Table 2). Since no additional information was provided by the least mean squares analysis, it was not used in the present analysis, and the total number of responses in a category was substituted. Inspection of the mean shift in numbers in Table 9 indicates that there

Figure 4. Average identification functions for intensity task.

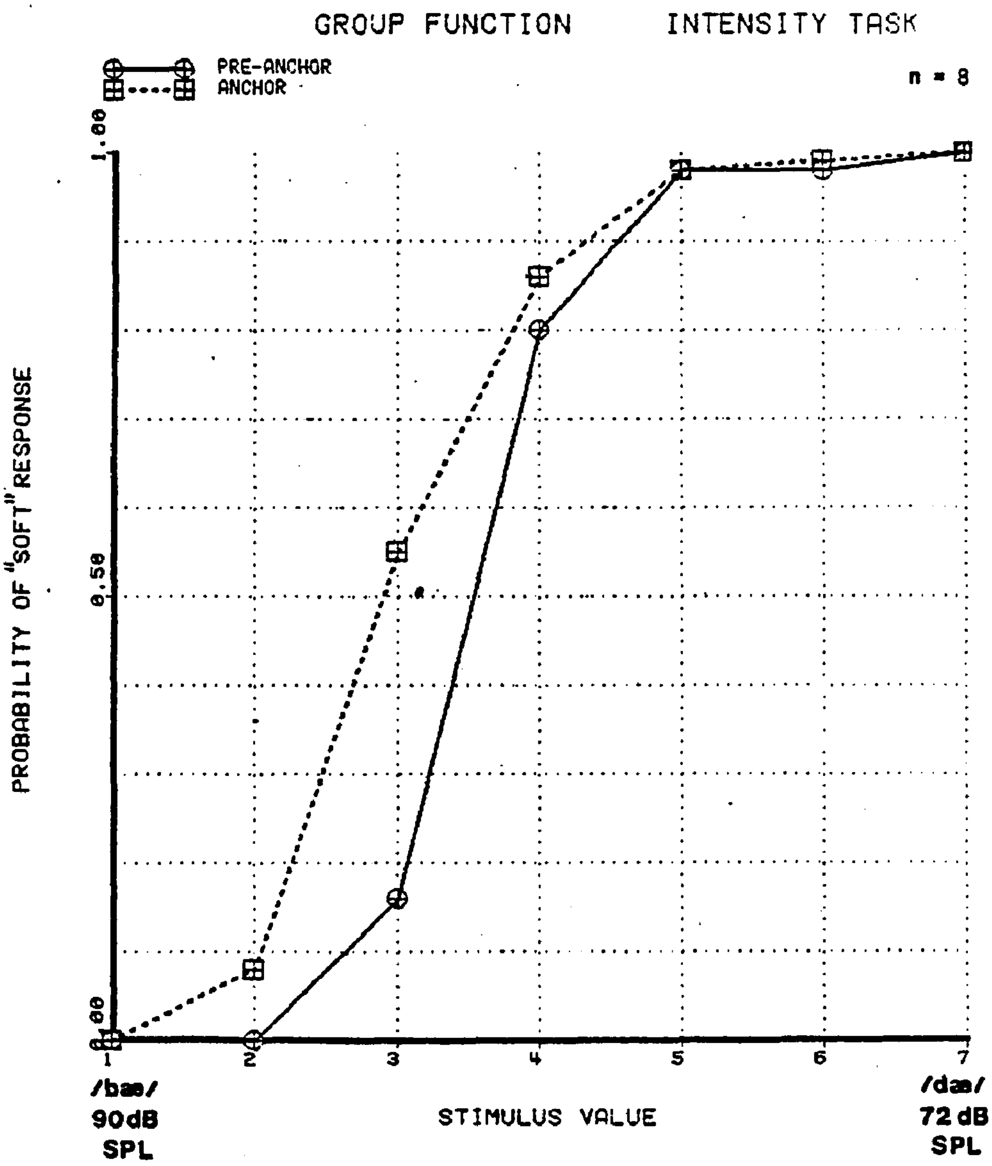

GROUP FUNCTION INTENSITY TASK
PRE-ANCHOR
ANCHOR
n = 8
PROBABILITY OF "SOFT" RESPONSE
1.00
0.50
0.00
1
2
3
4
5
6
7
/bæ/
90dB
SPL
STIMULUS VALUE
/dæ/
72 dB
SPL

Figure 5. Average identification function for phoneme task.

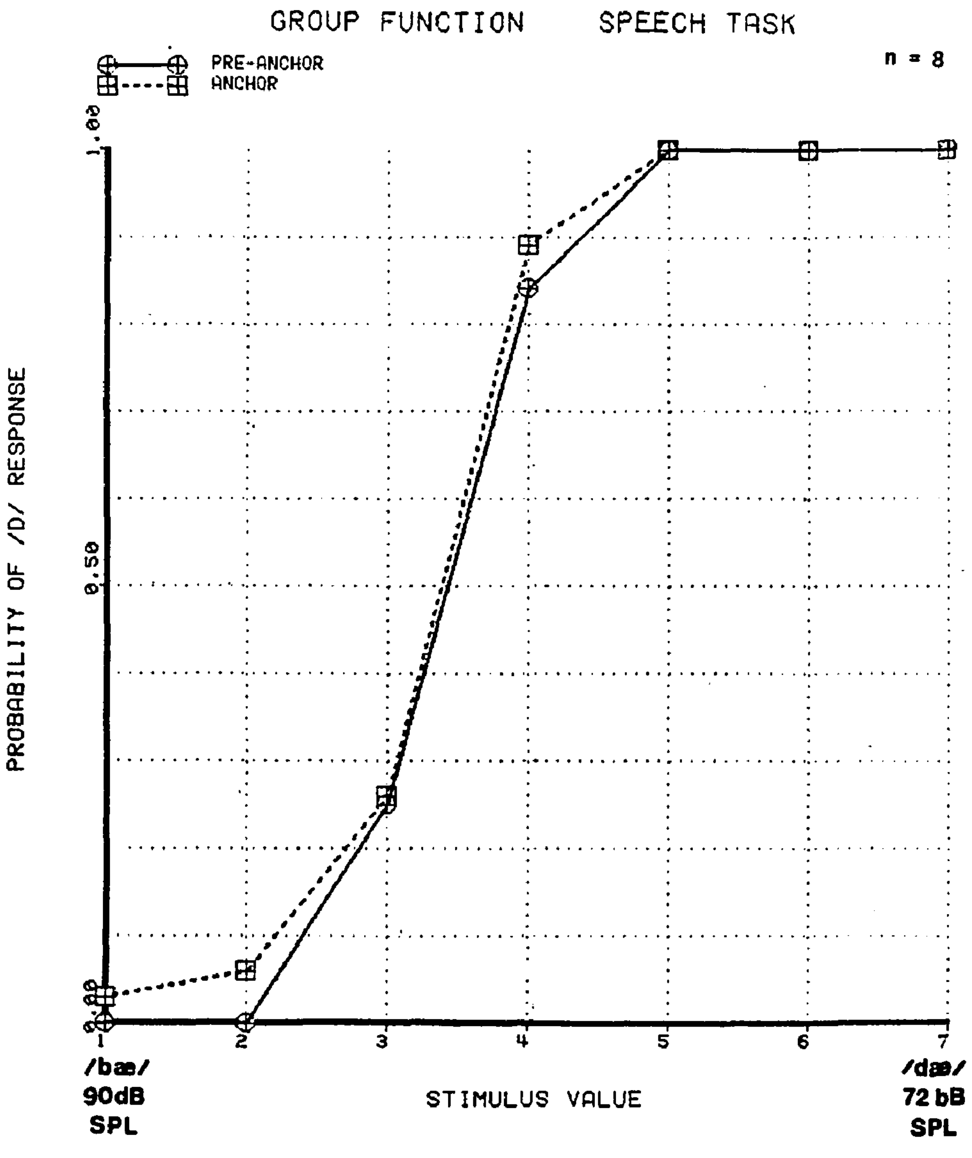

GROUP FUNCTION SPEECH TASK

Table 9. Individual and mean number of "soft" and /d/ responses and shifts for each test condition.

Individual and mean number of "soft" and /d/ responses and shifts for each test condition.

	INTENSITY			SPEECH		
Subject	Control	Anchor	Control-Anchor	Control	Anchor	Control-Anchor
NK	43	50	- 7	40	41	- 1
DB	41	48	- 7	36	37	- 1
RG	39	51	-12	47	52	- 5
MS	39	42	- 3	37	43	- 6
DH	43	47	- 4	40	42	- 2
CW	28	32	- 4	41	34	7
VH	41	44	- 3	46	45	1
AR	39	42	- 3	40	43	- 3
MEAN	39.1	44.5	- 5.4	40.9	42.1	- 1.3

were more "soft" responses in the anchor condition than in the pre-anchor condition as expected. In the phoneme task, there was also a shift in the predicted direction. However, the mean shift for the intensity condition was 4.78 total "soft" responses while for the phoneme condition it was only 2.88.

Group probit analyses were done on these identification functions. As in Experiment I, the slopes of the functions were steeper for the consonant judgment than for the non-speech judgment.

An analysis of variance, Subjects by Ratio by Continua, was performed with the total number of "soft" and total number of /d/ responses in both the pre-anchoring and anchoring conditions. Table 10 shows that the main effect of ratio was significant ($F_{1,7} = 10.60$; p 0.05) indicating that the boundary shifted the anchoring condition.

As in Experiment I, the main effect of continuum was not significant. However, the ratio by continuum interaction was significant ($F_{1,7} = 7.065$; $p < 0.05$), again presumably due to the greater anchor effect on the intensity task than on the phoneme task. Inspection of the individual data of Table 9 shows that all 8 subjects give a shift in the expected direction in the intensity task; in the phoneme task, six subjects shift in the expected direction, two in the opposite direction, away from the anchor. Also, all subjects give a greater shift on the intensity task than on the phoneme task.

A Wilcoxon Signed Ranks Test was used to test the significance of these differences. For the intensity task, a significant shift was found between the pre-anchor and anchor conditions while no significant boundary movement was found in the phoneme condition ($T = 0$; $p < 0.05$ for the intensity task; $T = 9.33$; $p > 0.05$ for the phoneme task).

Table 10. Summary table of analysis of variance for the total number of /d/ and "soft" responses for each ratio.

Summary table of analysis of variance for the total number of /d/ and "soft" responses for each ratio. Ratio (R): 1:1, 4:1; Continua (C): Intensity, Phoneme; Subjects (S).

Source	df	MS	F	*p*
Subjects (S)	8			
R	1	87.78	10.60	< 0.05*
RxS	7	8.281		
C	1	0.781	0.027	NS
CxS	7	28.85		
RxC	1	34.03	7.065	< 0.05*
RxCxS	7	4.82		
TOTAL	24			

On the basis of these results, it appears that there is an effect of anchoring on the intensity continuum, but no effect on the phoneme continuum. These results parallel the findings of Experiment I and the results of Sawusch, Pisoni, and Cutting (1974) where significant shifts were found for the non-speech tasks but not for the phoneme task. However, in this experiment there was a slight, non-significant shift in the consonant task.

CHAPTER IV: EXPERIMENT III

ANCHORING OF CORRELATED VOWEL AND STOP CONSONANT CONTINUUM

The first two experiments in this series suggested that psychophysical anchoring effects on phonetic and non-phonetic dimensions of a simple stimulus continuum can be dissociated. It might be expected from these results that other continua housing both linguistic and non-linguistic dimensions could be manipulated by this paradigm. Since it has been shown extensively in the literature that, in certain environments, vowels and consonants yield basic perceptual differences, analogous in some respects to differences between speech and non-speech stimuli, the next experiment extends the paradigm of the previous two experiments to compare the effects of anchoring on correlated consonant and vowel continua.

Method

Description of 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 the first- and second-formant transition onset and steady-state formant frequencies. These values are shown in Table 11. The third-formant resonator on the parallel resonant synthesizer was turned off during synthesis because of the difficulty encoun-

Table 11. Starting frequencies of the second-formant transitions of stop consonants and the steady-state frequencies of vowels.

Starting frequencies of the second-formant transitions of stop consonants and the steady-state frequencies of vowels.

	Transition Duration	Transition	Steady-State	
	(msec)	Frequency in Hz		
Stimulus Number	F_2	F_2	F_1	F_2
1	40	1385	718	1541
2	40	1468	692	1620
3	35	1541	666	1695
4	30	1620	640	1772
5	45	1695	614	1845
6	25	1772	588	1920
7	20	1845	562	1996

tered in generating a perceptually acceptable three-formant linear vowel and consonant continuum.

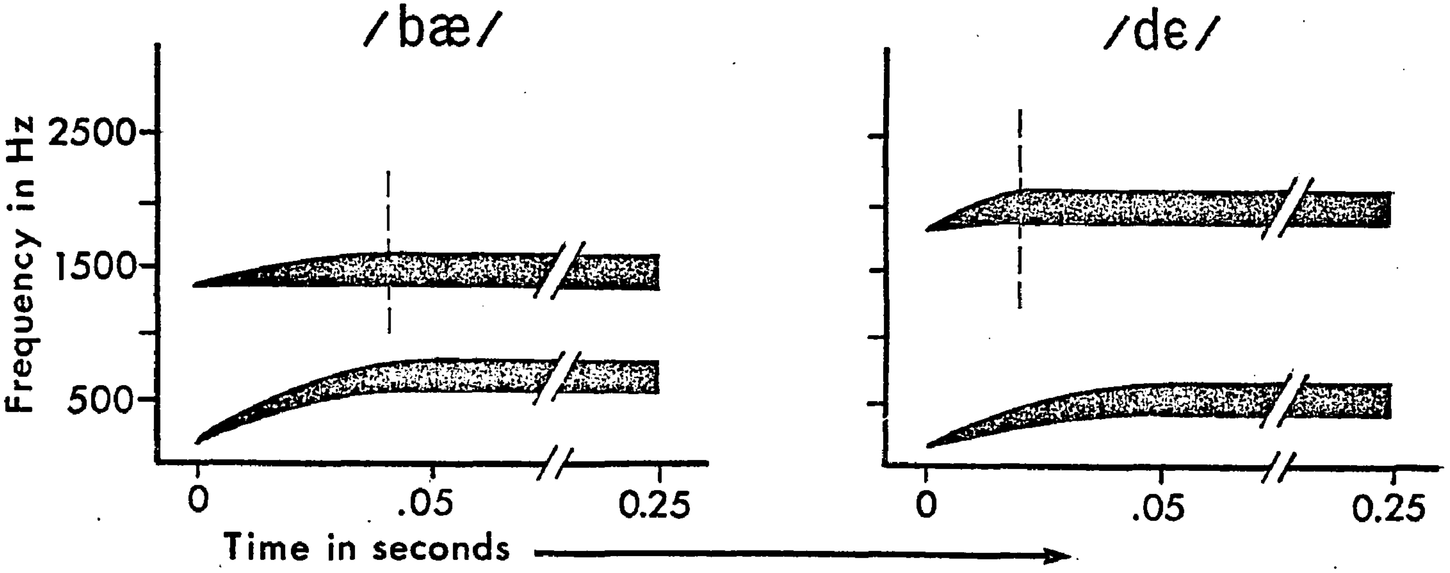
Stimulus 1 had first-and second-formant starting frequencies of 234 Hz and 1385 Hz respectively with steady-state formant frequencies of 718 Hz and 1541 Hz. The steady-state value of F_1 for successive stimuli in the series decreased in systematically equal steps while the steady-state value of F_2 increased to generate an acceptable American English /ɛ/. F_2 transitions increased from stimulus 1 to stimulus 7 while F_1 transitions for each stimulus began at 234 Hz and rose linearly.

The seven stimuli were 250 msec in length with a 50 msec transition in F_1 and transitions ranging from 20 to 40 msec in F_2. The variable transition durations were due to the restricted range of frequencies that could be used to generate an acceptable continuum while maintaining linearity of frequency steps. A pilot experiment established the acceptability of the stimuli and the anticipated phoneme boundary. Fundamental frequency (114 Hz) and overall amplitude were consistent across all stimuli. Schematized spectrographs of the two end-point stimuli, /bæ/ and /dɛ/, are illustrated in Figure 6. The anchor in this experiment was again the end-point /bæ/.

Experimental Tapes

All experimental tapes were produced on the Haskins PCM system as previously described. As before, two test tapes were prepared for each condition: a 70 item identification or pre-anchor tape with each stimulus presented equally often, and a 100 item anchor tape with each

Figure 6. Schematized spectrographic patterns appropriate for the syllables /bæ/ and /dɛ/.



STIMULUS 1

STIMULUS 7

end-point /bæ/ presented 4 times to each presentation of stimuli 2-7. Each tape was recorded twice from the PCM system: once for the consonant task and once for the vowel task.

Procedure

The experiment was conducted in a sound isolated room at Haskins Laboratories. The experimental tapes were reproduced from the output of an Ampex AG 500 tape recorder and presented binaurally through Telephonics (TDH-39) matched and calibrated headphones. The voltage across the headphones as measured with a Hewlett-Packard voltmeter was 80 dB SPL re 0.0002 dyne/cm^2.

As in the previous two experiments, the study required identification of the stimulus series under two conditions: a pre-anchor or control condition and an anchor condition for each of the two tasks. In the first instance, the task was to categorize the vowel as an /æ/ or /ɛ/. Since there was a tendency for some subjects in the pilot study to perceive /ɪ/, any /ɪ/-like vowel was to be categorized into the /ɛ/ category and any /æ/-like vowel into the /æ/ category. Since it was evident that a correlation existed between the vowel and consonant, it was decided to enhance this correlation by informing the subjects about it. However, they were asked to pay no attention to the correlation and to listen for the vowel or consonant at the appropriate task time.

The test order remained consistent across all subjects: vowel pre-anchor, vowel anchor, stop-consonant pre-anchor and stop-consonant anchor. Again, the subjects were instructed that the pre-anchoring and anchoring test tapes were identical save for length. The same pre-anchor and anchor tapes were heard in both the vowel and consonant task conditions. Only the response directions varied as the task changed. Experimental tapes were preceded by listening and practice tapes. Subjects were tested either singly or in pairs and all testing was completed in one session.

As in the previous experiment, there were nine paid listeners.

Results

Figure 7 shows the group probability of /ɛ/ vowel judgments for the seven test stimuli in both the pre-anchor and anchor conditions. This figure shows that subjects made more /ɛ/ and less /æ/ responses in the anchor condition than in the pre-anchor condition as indicated by a shift in the phoneme boundary (50% crossover point) from stimulus 3.8 to stimulus 3.1. Approximately a one-half step shift in the group phoneme boundary function for the stop-consonant task from stimulus 3.5 to stimulus 3.0, towards the anchor stimulus, is shown in Figure 8.

One subject, CA, was eliminated upon inspection of the consonant identification function, since she could not consistently categorize the members of a particular category. Therefore, a total of 8 subjects was included in the data analysis.

The contextual effects evident in the fundamental frequency judgment task are not apparent in the vowel or consonant task. However, inspection of the group identification functions shows, that although subjects partitioned the stimulus continua into two relatively distinct

Figure 7. Average identification function for vowel task.

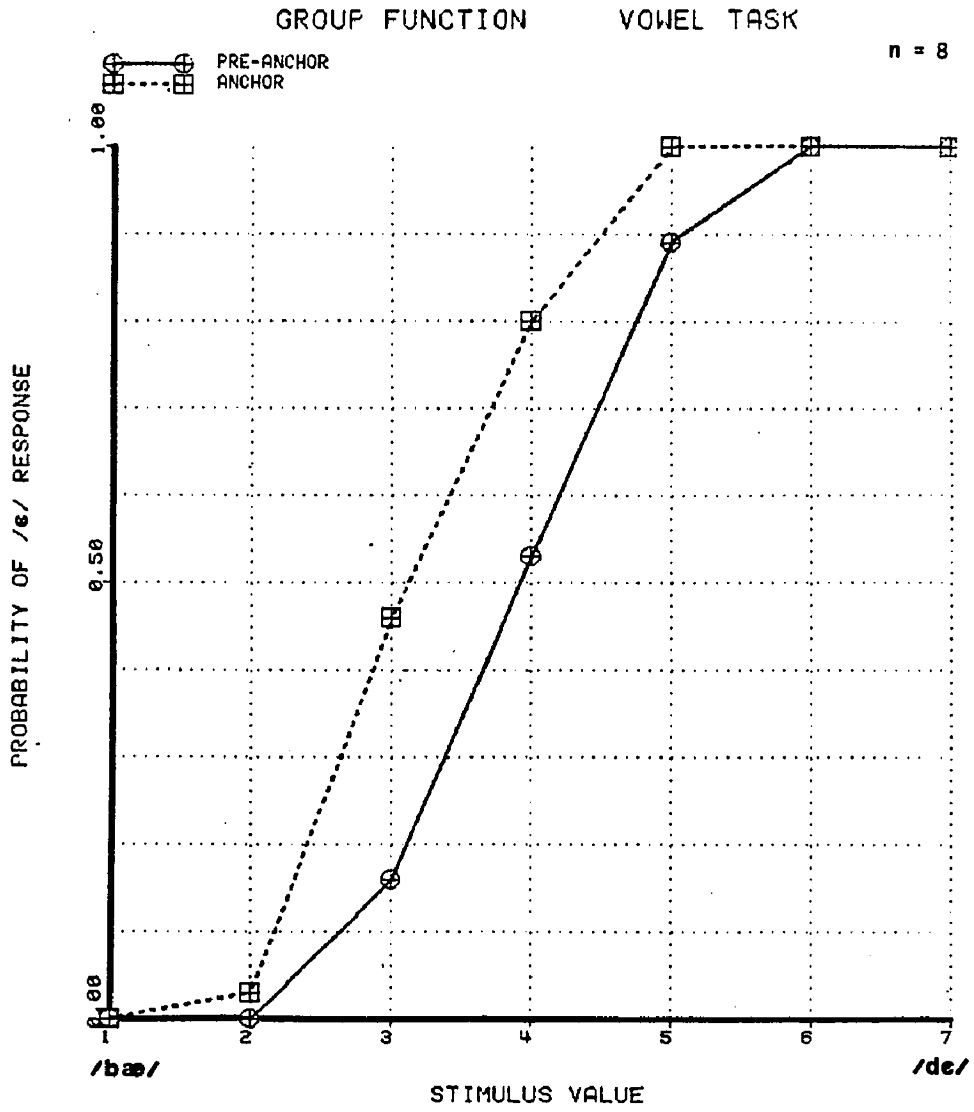
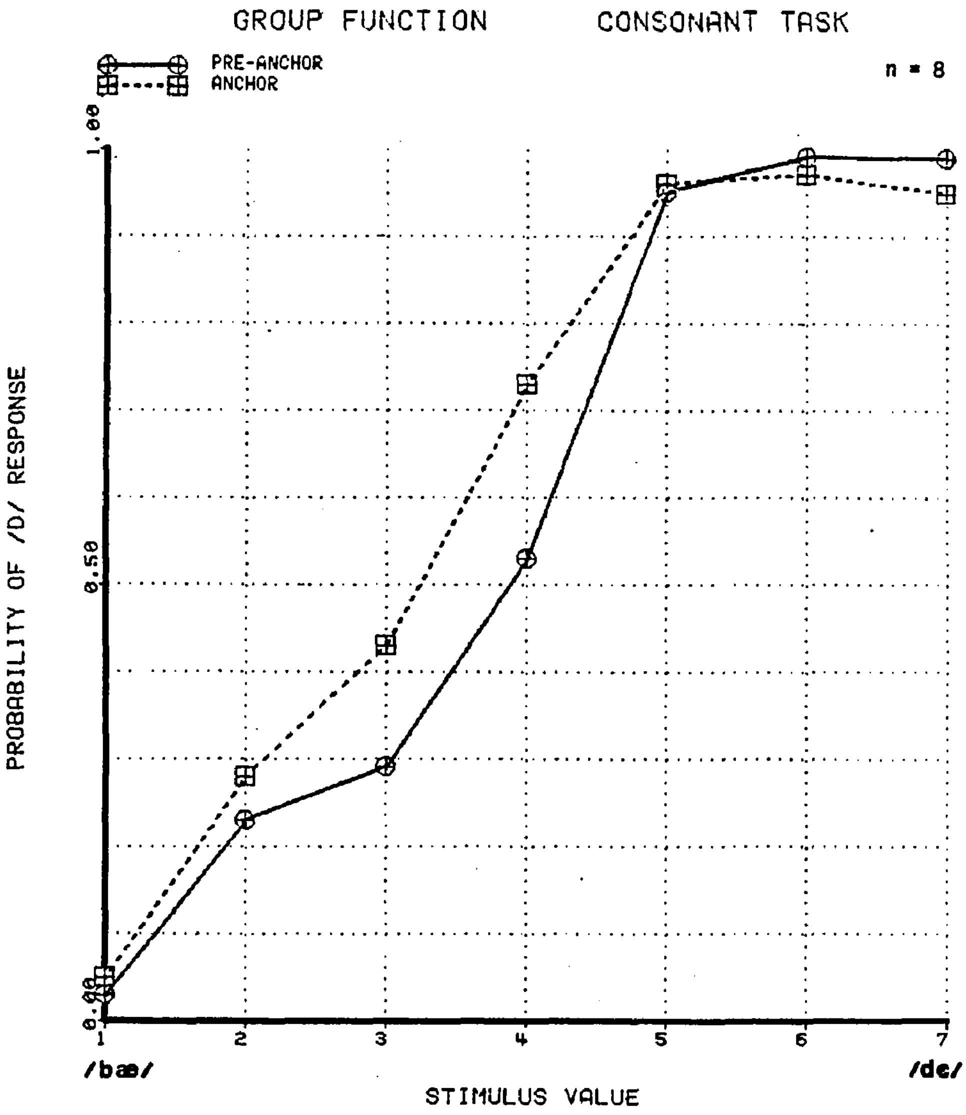

GROUP FUNCTION VOWEL TASK
n = 8
PRE-ANCHOR
ANCHOR
PROBABILITY OF /ɛ/ RESPONSE
1.00
0.50
0.00
1
2
3
4
5
6
7
/bæ/
/dɛ/
STIMULUS VALUE

Figure 8. Average identification function for stop consonant task.


GROUP FUNCTION CONSONANT TASK
PRE-ANCHOR
ANCHOR
n = 8
PROBABILITY OF /D/ RESPONSE
1.00
0.50
0.00
1
2
3
4
5
6
7
/bæ/
/de/
STIMULUS VALUE

categories, the consonant function is less clear-cut than the vowel function. The consonant function is also less sharp than those in Experiment I and II, therefore making more room for contextual effects. Individual identification functions are also generally sharper for vowels than for consonants, especially in 5 of the 8 subjects, showing weak /b/ categories. These findings contrast with previous experiments on the identification of vowels and consonants (Stevens, Liberman, Studdert-Kennedy, and Öhman, 1969; Pisoni, 1971). However, both these studies used steady-state vowel continua rather than CV syllables.

As in Experiments I and II, group probit analyses were obtained on the identification functions. However, in this experiment, the slopes of the functions for the consonant were less steep than those of the vowel. Table 12, which shows the individual and mean number of /ɛ/ and /d/ responses and shifts for each condition, also indicates a difference between the vowel and consonant task. However, in the vowel task, there was a greater boundary shift (-6.63 responses) from the pre-anchor condition to the anchor condition than in the consonant task (-3.50 responses). A negative number indicates a shift towards the anchor stimulus with a concomitant increase in the number of /ɛ/ or /d/ responses.

An analysis of variance, Subjects (8) by Ratio (1:1 and 4:1) by Continua (Vowel and Consonant) was performed with the total number of /d/ and the total number of /ɛ/ responses in both the pre-anchor and anchor conditions. Table 13 shows that the main effect of ratio was significant ($F_{1,7} = 20.012$; $p < 0.01$) indicating that the boundary shifted as a result of the anchoring condition.

Table 12. Individual and mean number of /d/ and /ε/ responses and shifts for each test condition.

Individual and mean number of /d/ and /ɛ/ responses and shifts for each test condition.

	VOWEL			CONSONANT		
Subject	Control	Anchor	Control-Anchor	Control	Anchor	Control-Anchor
DH	34	35	- 1	36	35	1
NK	39	46	- 7	33	36	- 3
NV	42	44	- 2	33	35	- 2
HK	32	44	-12	52	57	- 5
CW	33	41	- 8	30	36	- 6
RG	41	45	- 4	31	42	-11
JK	36	43	- 7	54	52	2
VH	33	45	-12	48	52	- 4
MEAN	36.3	42.9	-6.6	39.6	43.1	-3.5

Table 13. Summary table of analysis of variance for the total number of /d/ and /ɛ/ responses for each ratio.

Summary table of analysis of variance for the total number of /d/ and /ɛ/ responsees for each ratio. Ratio (R): 1:1, 4:1; Continua (C): Vowel Consonant; Subjects (S).

Source	df	MS	f	p
Subjects (S)	7			
R	1	205.031	20.012	< 0.01*
RxS	7	10.246		
C	1	26.281	0.244	NS
CxS	7	107.638		
RxC	1	19.531	2.895	NS
RxCxS	7	6.746		
TOTAL	24			

A Wilcoxon Matched-Pairs Signed Ranks Test was used to test the significance of the difference between the pre-anchor and anchor conditions. A significant shift was found in both the vowel task ($T = 0$; $p < 0.05$) and the consonant task ($T = 3.5$; $p < 0.05$). The difference between the two conditions (Vowel versus Consonant) was not significant ($T = 12$; $p > 0.05$).

Summary and Discussion

The shift in the consonant as well as the vowel phonetic boundary in this experiment contrasts with the results of Experiments I and II where no shift in the consonant boundary from the pre-anchor to the anchor position was shown. Two possible explanations for the boundary shifts in the present experiment suggest themselves. First, subjects were specifically advised that there was a correlation between the consonant and the vowel: /b/ was always paired with /æ/ and /d/ with /ɛ/. Subjects may have made use of this redundant vowel information to facilitate judgment of the consonant, since the vowel condition always preceded the consonant task. A second, not necessarily incompatible explanation, is that subjects were unable to attend selectively to the consonant as the target: the consonant and the vowel could not be separated in the syllable.

Theories of speech perception have claimed that there are inherent differences in the perceptual processing of speech sounds that differentiate it from the processing of other sounds. This difference leads to the distinction between phonetic and auditory phenomena (Liberman, 1970) and to the notion that special perceptual mechanisms may exist for the processing of speech sounds (Studdert-Kennedy and Shankweiler, 1970). Various experimental paradigms have explored these differences.

In terms of investigating the relation between auditory and phonetic processing, these results may be compared to a series of experiments using a two-choice speeded classification procedure in selective attention tasks with two dimensional stimuli (Day and Wood, 1972; Wood and Day, 1974, 1975). When the two dimensions being tested were both linguistic (consonant or vowel identification in synthetic CV syllables), mean reaction time was faster if all test items were identical on the non-target dimension (same vowel on a consonant task or same consonant on a vowel task) than if both the target and nontarget dimensions varied independently. This symmetrical interaction suggests that information about an initial stop consonant and the following vowel is processed as an integral unit since subjects were unable to attend selectively to the target dimension and ignore the irrelevant dimension in the latter condition (Wood and Day, 1975). On the other hand, when only one of the independently varying dimensions was linguistic (place of articulation of the stop consonant versus fundamental frequency), unidirectional or asymmetrical interaction occurred: irrelevant variation in pitch produced interference with the processing of place (Day and Wood, 1972; Wood, 1974, 1975).

The failure of selective attention in the processing of two-dimensional linguistic (Wood and Day, 1975) and nonlinguistic stimuli (Wood, 1975) has also been obtained for visual stimuli (Garner and Felfoldy, 1970; Garner, 1974). Garner and Felfoldy (1970) distinguished between stimulus dimensions that were "separable" (dimensions that can be pulled apart, analyzed, and perceived as unrelated) and those that were "integral," (dimensions that are perceived as a single dimension and cannot be analyzed). For example, pitch and loudness have been found to be integral since discrimination times for both dimensions showed a substantial increase when the second dimension varied as an irrelevant dimension (Wood, 1973), while size (of a square) and lightness (achromatic Munsell lightness varying from white to black) are separable dimensions (Garner, 1974). In selective attention procedures, separable dimensions produce no redundancy gain when correlated and no interference when orthogonal, in contrast to the redundancy gain (as measured by decreased reaction time) and interferences in the correlated condition. Redundancy gain refers to improved performance in constrained-classification tasks as the result of two or more dimensions varying in correlated steps. Where there is such a gain, the dimensions are said to be integral and where there is no gain, the dimensions are said to be separable. Wood and Day (1975) suggest that the inability to attend selectively to the two linguistic dimensions (consonant/vowel identity) in the CV syllables is evidence for their functioning as integral dimensions and lends support to the concept of dimensional integrity in information processing. However, they also conclude that the pattern of asymmetrical interference observed in the linguistic-nonlinguistic conditions must temper a strict dichotomy between integral and separable dimensions.

In the present series of experiments, the difference in the results between the experiments using both linguistic and non-linguistic dimensions (Experiments I and II) and the experiment using only linguistic dimensions (Experiment III) may be interpreted as differences in the processing of integral versus separable dimensions. It also may be that the subjects used the vowel information in Experiment III to aid in judging the consonant, since there was a correlation between the vowel and consonant and the vowel task was always presented first.

Since the vowel was always the same throughout the continuum in Experiments I and II, the subjects did not have to pay attention to it. In fact, due to the parallel processing that characterizes the speech code (Liberman et al., 1967; Liberman, 1970) and the fact that one sound segment carries information about several adjacent segments (Fant, 1962), the first 30-40 msec of the rapid formant transitions that distinguish the place of articulation are also carrying information about the following vowel. Therefore, the subjects are able to make a decision about the consonant alone during the first 30-40 msec when the following vowel is always the same. However, when the following vowel is not always the same, the subjects must make both a judgment about the vowel they are perceiving as well as the consonant. These two judgments may not be separable. Therefore, the perception of the vowel, in this correlated continuum, may help make the decision regarding the consonant. The shifts in the consonant boundary may just be a mirror of the vowel shift.

The shift in the category boundary for the non-speech task with no shift in the boundary for the speech task in Experiments I and II suggests that subjects selectively attend to one dimension rather than the other. Taken a step further, if the CV syllable functions as an integral unit, the consonant effect in the present experiment perhaps indicates that consonants can be driven into showing psychophysical anchoring effects by their integration with vowels that are more liable to anchoring.

Also along these lines, the differences in the decay properties of the two stimuli might be explored. Studies in selective adaptation have suggested adaptation of both acoustic and phonetic properties. Although the stimuli used in the anchoring experiments are the same, regardless of the task (fundamenal frequency or stop consonant, for example), there are acoustic differences between the two dimensions of the stimuli that the subject must pay attention to. In the fundamental frequency task, the information to be monitored exists for the full 250 msec of the syllable while the information in the consonant task is less persistent and only exists for the first 30-40 msec of the rapid formant transitions that distinguish place of articulation. It is possible, therefore, that there is a decay in the anchoring effect of the consonant while there is no decay in the anchoring effect of the fundamental frequency. This would hold true for the intensity continua, and is somewhat analogous to the auditory memory explanations of categorical perception. However, for the vowel and consonant continuum in Experiment III, decay of the consonant may have been prevented due to the parallel processing and the inability to separate the consonant from the vowel.

If anchoring with the vowel and the consonant in this experiment shifts the two boundaries differentially, there is evidence of definite anchoring of the consonant independent of the vowel. If, however, the boundaries for the consonant and the vowel overlap, the outcome is ambiguous. It may mean, as stated above, that the consonant and the vowel could not be separated in the syllable and that the consonant effect was based on its integration with the vowel. But, it may also mean that the two are separable and their boundaries coincide due to averaging of the data. In fact, this is what appears to have occurred. Table 14 shows the category boundaries for both the vowel and consonant responses in both the pre-anchor and anchor conditions. The boundaries for the vowel and the consonant judgments in the group functions are approximately .20 steps away from each other. Although significant boundary shifts were found in both the vowel and the consonant tasks on a Wilcoxon Matched-Pairs Signed Ranks Test ($T=0$; $p < 0.05$ for the vowel; $T=3$; $p < 0.05$ for the consonant), there was no significant difference between the vowel and consonant shifts ($T=12$; $p > 0.05$).

However, the individual boundaries show greater variability as seen in Table 14 and therefore the average or group function may be misleading. The range of differences between the vowel and the consonant anchor boundaries for the individual subjects is between -6 and .7 steps.

A Spearman Rank-Difference Correlation Coefficient between the vowel and the consonant boundary shift was not significant (rho = .25) showing that there was no relationship between the two tasks and that the group data are misleading. This lack of a correlation between the vowel and the consonant boundaries taken with the large differences in the individual boundaries shows that there are no grounds for assuming that the amount of the shift in both boundaries is the same, and, by

Table 14. Category boundaries in stimulus numbers along correlated /æ/ - /ɛ/ and /b/ - /d/ continua.

Category boundaries in stimulus numbers

along correlated /æ/ - /ɛ/ and /b/ - /d/ continua

Subject	/æ/ - /ɛ/ Anchoring Ratio 1:1	4:1	Boundary Shift	/b/ - /d/ Anchoring Ratio 1:1	4:1	Boundary Shift
DH	3.8	4.2	-0.4	3.6	3.8	-0.2
NK	3.7	2.7	1.0	4.2	3.8	0.4
NV	3.2	2.6	0.6	4.4	3.7	0.7
HK	4.3	3.2	1.1	2.2	1.4	0.8
CW	4.4	3.4	1.0	4.5	3.8	0.7
RG	3.3	3.0	0.3	4.4	3.5	0.9
JK	3.8	3.2	0.6	1.5	1.6	-0.1
VH	4.0	2.8	1.2	3.0	2.2	0.8
MEAN	3.8	3.1	0.7	3.5	3.0	0.5

Note: Stimulus 1 was the anchor stimulus so that a reduction in the category boundary value (i.e., a positive boundary shift as tabulated above) reflects a shift toward the anchor stimulus.

implication that the subjects did not make use of the redundant vowel information to process the consonant. Therefore, subjects were evidently judging the two continua independently. The consonant and vowel were dissociated as are the pitch/intensity and consonants in Experiments I and II. This would effectively rule out the consonant shift as due to differences in the processing of integral versus separable dimensions.

Another tack to be explored to explain the shift of the consonant boundaries in this experiment are the stimuli used and the synthesis of these stimuli. If the consonant stimuli are poor exemplars of the two stop consonants to be judged, it is possible that the shifts in the boundaries are due to poor synthesis. As previously discussed, the major cues to place of articulation in initial stop consonants are the direction and extent of the second- and third-formant transitions (Delattre, et al., 1955; Liberman, et al., 1956; Hoffman, 1958) and the presence and frequency locus of an initial noise burst (Cooper, et al., 1952; Hoffman, 1958). In order to effect an acceptable continuum while maintaining linearity of frequency steps, the transition durations were made asymmetrical and the third-formant generator on the parallel resonance synthesizer was turned off during the synthesis. The burst cue was not used in the synthesis of these stimuli and the steady-state frequencies of the boundary stimuli is an ambiguous cue. According to Liberman, et al., (1959), the steady-state frequency appropriate for both the /d/ and /b/ stop-consonants in front of the vowel /ɛ/ is 1800 Hz. The steady-state values of the boundary stimuli (stimuli numbers 3 and 4) in this experiment were 1695 and 1772 Hz respectively. The fact that all the formant transitions in the seven stimuli were rising would predict more /b/ responses with anchoring on the basis of an acoustic account. Although the frequency spectrum at syllable onset is appropriate for alveolar stop consonants (Stevens and Blumstein, 1976), the other cues available to differentiate place of articulation are either absent or would cue the /b/ consonant. However, in spite of these acoustic cues to the /b/ consonant, there was a weak categorization of the /b/ and a shift in the category boundary towards the /bæ/ adaptor (more /d/ responses).

The weak /b/ categorization might be explained by the difference in the slope and duration of the second-formant transitions in this experiment as compared to those in Experiment I, for example. The duration of the formant transitions in this experiment varied from 20 msec to 40 msec while in Experiment I the durations were consistently 40 msec. Since there are variations in the extent of transitions dependent upon the place of articulation and the following vowel, the variation in the duration of the formant transitions in this experiment may be an ambiguous cue. For example, transitions tend to be longer for bilabials and shorter for apicals before front unrounded vowels than before back, rounded vowels. (Dorman, et al, In press.)

The difference between the onset frequency of the second formant transition and the steady-state value may also be a factor in the weak categorization of the /b/ stimulus. This difference for stimulus number 1 was 234 Hz in Experiment I and only 155 Hz in Experiment III. For stimulus number 2 in both experiments, the difference was 155 Hz and 152 Hz, while stimulus number 3 in Experiment I shows only a 79 Hz difference and in Experiment III continues to be similar to the other stimuli in the experiment, as previously mentioned, 154 Hz.

Whatever the correct interpretation of these results, these experiments are among the first to show a purely psychophysical effect on the position of the consonant boundary. This contextual effect, a well-known phenomenon in psychophysical scaling as previously discussed, has been shown to occur in vowel continua (Fry, et al., 1962; Stevens, et al., 1969), but not in judgments of consonant continua (Liberman, et al., 1961).

Since Experiment III showed that anchoring effects can occur for consonants, it becomes especially important to discover the conditions of the anchoring effect and to determine how it is related to the selective adaptation effect in speech perception. The following experiment, the first in a series of parametric studies, investigates the effect of the number of adaptor repetitions on the magnitude of the category boundary shifts in a stop-consonant continuum.

CHAPTER V: EXPERIMENT IV

ADAPTATION AND THE NUMBER OF ADAPTOR REPETITIONS

Introduction

The present experiment considers the effect of the number of adaptor repetitions presented in the adaptation sequence preceding each test sequence, with repetition rate held constant. Experiment V holds the number of adaptor repetitions constant, while varying their presentation rate. The two experiments were designed to examine the effects of varying the timing of the presentation of the adapting syllable.

Bailey (1975) manipulated the number of adaptor repetitions in an adapting sequence in a search for a more efficient adaptation procedure, and Hillenbrand (1975) did the same in order to estimate the range of the obtainable adaptation effects.

Bailey (1975) found no systematic difference in the size of the adaptation effect produced by 32 repetitions versus 8 repetitions, and therefore established justification for using 8 or less repetitions on each trial. The Hillenbrand (1975) study, on the other hand, found systematically larger shifts in the phonetic boundary with greater intensity and with more adaptor repetitions. The discrepancy in the results of these two experiments may be difficult to interpret, as pointed out by Hillenbrand, due to methodological differences between the two experiments.

The determination of the time course of the development of the effect was precluded, according to Bailey (1975) due to the apparent lack of decay of adaptation in his study. The absence of decay was due to three factors: during a single session only one adapting syllable category was heard, blocks of trials were preceded by 60 repetitions of the adaptor, and the interval between both the subject's response on a trial and the start of the repeating sequence for the next trial was less than one second. It was assumed, therefore, that the level of adaptation at the time each test syllable was identified in adaptation was due to the undecayed adaptation from previous trials as well as to the number of adaptor repetitions on that particular trial. Further, the decay of adaptation would have to be exceptionally fast for the influence of previous trials not to occur (Bailey, 1975).

In auditory adaptation and auditory fatigue, as previously discussed (Chapter I), loudness decrement and threshold increase, respectively, are a function of the intensity, duration, and frequency of the stimulating tone (or noise-band), its rate of presentation, and the time interval (cf. Small, 1963; Ward, 1963; Elliot and Fraser, 1970) between stimulus exposure and post-exposure threshold measurement. In general, increases in the intensity, duration, and presentation rate increase the threshold shift or loudness decrement. Cumulative effects may be produced when interruption rate (presentation rate increased) is increased as the time for recovery is shortened (Carterette, 1955). When the interruption rate is increased, the adaptation period is also reduced. However, Elliot and Fraser (1970) state that the adapting processes suffer less from the reduction in time than do the recovery processes. These results suggest that the total amount of stimulus energy is important in terms of receptor (cochlear and neural) functioning.

The Bailey paradigm, typical of the selective adaptation paradigm originated by Eimas and Corbit (1973), consisted of a pre-adaptation identification, run twice monaurally, followed by two blocks of 50 adaptation trials with 4 minutes rest between them; before each block 60 repetitions of the adaptor were presented, followed by a 400 msec silence. An adaptation trial consisted of either 8, 16, 24, or 32 repetitions (depending on the experimental condition) of a 300 msec syllable with an inter-adaptor-interval (IAI) of 200 msec, followed by an 800 msec silence and one test syllable. All testing was done with the adaptor in one ear and the test syllable in the other. Three adapting syllables, /ba/, /da/, and /ga/ were used in separate testing sessions on different days. However, the four conditions of repetition were given in a single listening session.

The Hillenbrand paradigm consisted of a pre-adaptation condition of twenty binaural presentations of the seven test stimuli, /bæ-dæ/, followed by a single block of 28 adaptation trials. An adaptation trial consisted of either 10, 25, or 100 repetitions (depending on the experimental condition) of a 300 msec syllable with an IAI of 225 msec followed, after 2 seconds of silence, by seven test syllables presented at 3 second intervals. Each repetition condition (10, 25, or 100 adaptors) was presented binaurally at three intensity levels (65, 80, and 95 dB SPL). Different subject groups (15 subjects each) listened to each of the nine experimental conditions binaurally.

The two experimental paradigms described above and the paradigm of the present experiment, are illustrated in Table 15. The differences in the two former paradigms, namely, the absolute number of adaptor repetitions, adaptation sequences per block, number of test syllables subsequent to each adaptation sequence, ear of presentation, and presentation schedule, were noted by Hillenbrand and considered as factors in the discrepant results. An additional difference is the pre-adaptation sequence of repetitions (60 repetitions) present in the Bailey paradigm, absent in the other two. The estimated time course for each paradigm, also displayed in Table 15, will be discussed below.

Method

Description of Stimuli

The stimuli used in this experiment were the same series of seven three-formant CV syllables previously identified as /bæ/-/dæ/ but with all the syllables at a fundamental frequency of 114 Hz and of 250 msec duration. The starting frequencies of the second- and third-formant frequencies are presented in Table 1. The single adaptor /bæ/ was taken from the low end-point of the test series (Stimulus number 1). As in the previous experiments, since a differential effect of the two end points was not of significant interest, only one end-point was used as the adaptor.

Again, the stimuli were recorded and tapes were prepared as in the previous experiments.

Table 15. Time course of the number of adaptor repetitions

	Ss	Adaptation Stimuli	Pre-Adaptation No. of Adaptors	No. of Adaptor Repetitions per trial	Syllable Duration (msec)	IAI (msec)	Average Repetition Rate (per sec)	Duration of Adaptation Sequence Per Trial	Duty Cycle
Bailey	6	/ba/	60	8				4.0 sec	
(1975)		/da/		16	300	200	2.0	8.0 "	
		/ga/		24 or				12.0 "	
		/ba/-/da/		32				16.0 "	
		/da/-/ga/							
		continuum							
Hillenbrand	45*	/bæ/	None	10	300	225	1.95	5.25 "	
(1975)		/bæ -dæ/		25				13.12 "	
		continuum		100				52.50 "	
Simon	9	/bæ/	None	8	250	250	2.0	4.0 "	50%
Experiment IV		/bæ-dæ/		16				8.0 "	50%
		continuum		32				16.0 "	50%
Simon	9	/bæ/	None	16	250	250	2.0	8.0 "	50%
Experiment V		/bæ-dæ/		16	250	750	1.0	16.0 "	25%
		continuum		16	250	1750	.5	32.0 "	12.5%

*15 different Ss were run at each of the three intensity levels for each adaptor condition = 45 Ss.

Experimental Tapes

One pre-adaptation and three adaptation experimental sequences were prepared from these stimuli and recorded on magnetic tape at 3 3/4 inches per second. The stimuli were recorded on a Crown 800 tape deck connected directly to the output of the PCM system. A 1000 Hz calibration tone generated on a Hewlett-Packard oscillator was placed at the beginning of each tape.

The pre-adaptation tape consisted of the seven stimuli recorded 25 times each and randomized to give a total of 175 stimuli. There was a 2 sec interval between stimulus items and a 10 sec interval after every 10 stimuli.

Three adaptation test orders consisting of 25 adaptation trials each were prepared. An adaptation trial consisted of a specific number of adaptors (8, 16, or 32), depending on the experimental condition, followed by the 7 test stimuli. The 250 msec adaptor stimuli were repeated with a 250 msec IAI, that is to say, at a rate of two adaptors per second. Following each series of adaptors was a 750 msec pause, a tone pip to alert the subjects to respond, a one-second pause, and the 7 test syllables. The test syllables were separated by 1750 msec.

Subjects

Nine normal hearing (screened at 20 dB HTL re: 1969 ANSI standard) listeners between the ages of 18 and 30 served as subjects for this experiment. All subjects were paid at a rate of 2.00 per hour.

Procedure

All experimental tapes were reproduced as in the previous experiments. A determination of the pre-adaptation phoneme boundary was made before each adaptation test.

Three adaptation sessions were run, one for each adaptor condition, separated by a minimum of 24 hours. Subjects were informed of the number of adaptor repetitions they would hear on a given day. The order of presentation of the adaptation conditions was counterbalanced over the first six subjects; the last 3 subjects received the same order as the first three. Each testing session was preceded by practice and listening tapes for the identification sequence.

Results

Figures 9, 10, and 11 plot the group phoneme boundary functions for subjects in the pre-adaptation and adaptation conditions. Inspection of these figures reveals shifts of the phoneme boundary in the predicted direction, towards the /bæ/ adaptor for all three adaptor repetition conditions. The data indicate that subjects as a group generally identified stimuli 1, 2, 3 as /b/ and stimuli 4, 5, 6, 7, as /d/ in the pre-adaptation condition and shifted approximately 1/2 to 1 step in the adaptation conditions, depending on the number of adaptor repetitions involved. Relatively sharp identification functions typical of stop consonant synthetic continua are evident.

Tabulation of the individual and mean number of /d/ responses for stimuli 2-7 (Table 16) shows that the magnitude of the boundary shift

Figure 9. Average identification function for stop consonant at an 8:1 ratio.

GROUP FUNCTION 8:1 RATIO

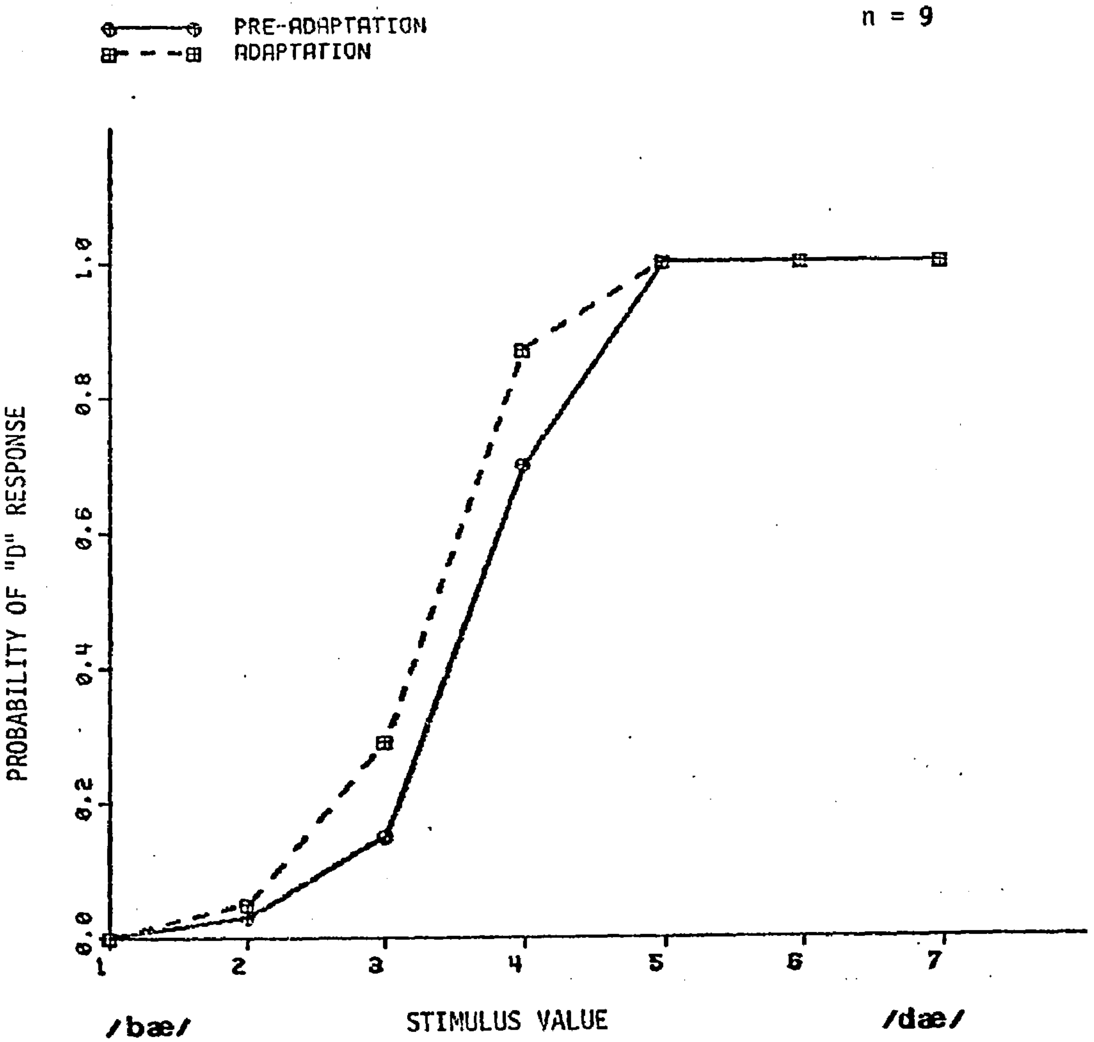


Figure 10. Average identification function for stop consonant at a 16:1 ratio.

GROUP FUNCTION 16:1 RATIO

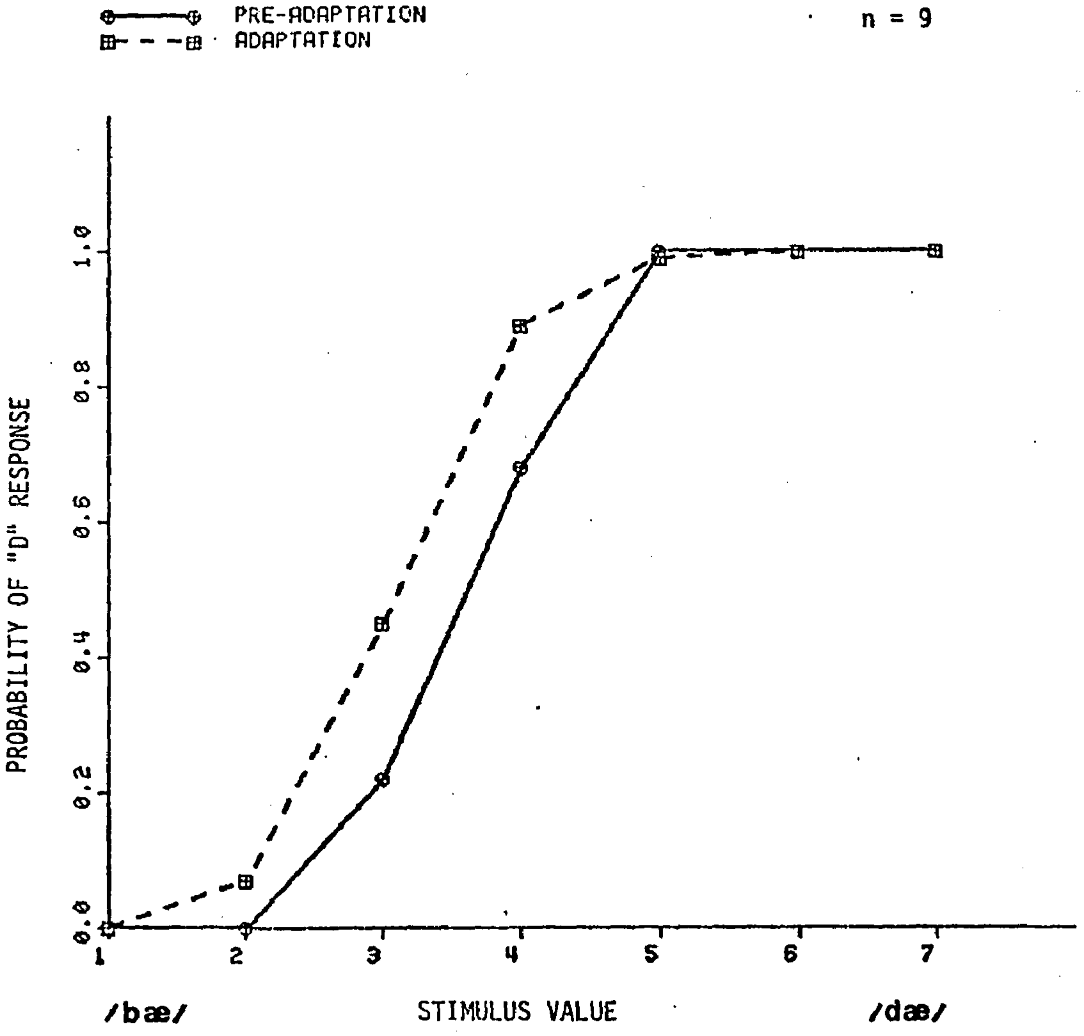


Figure 11. Average identification function for stop consonant at a 32:1 ratio.

GROUP FUNCTION 32:1 RATIO

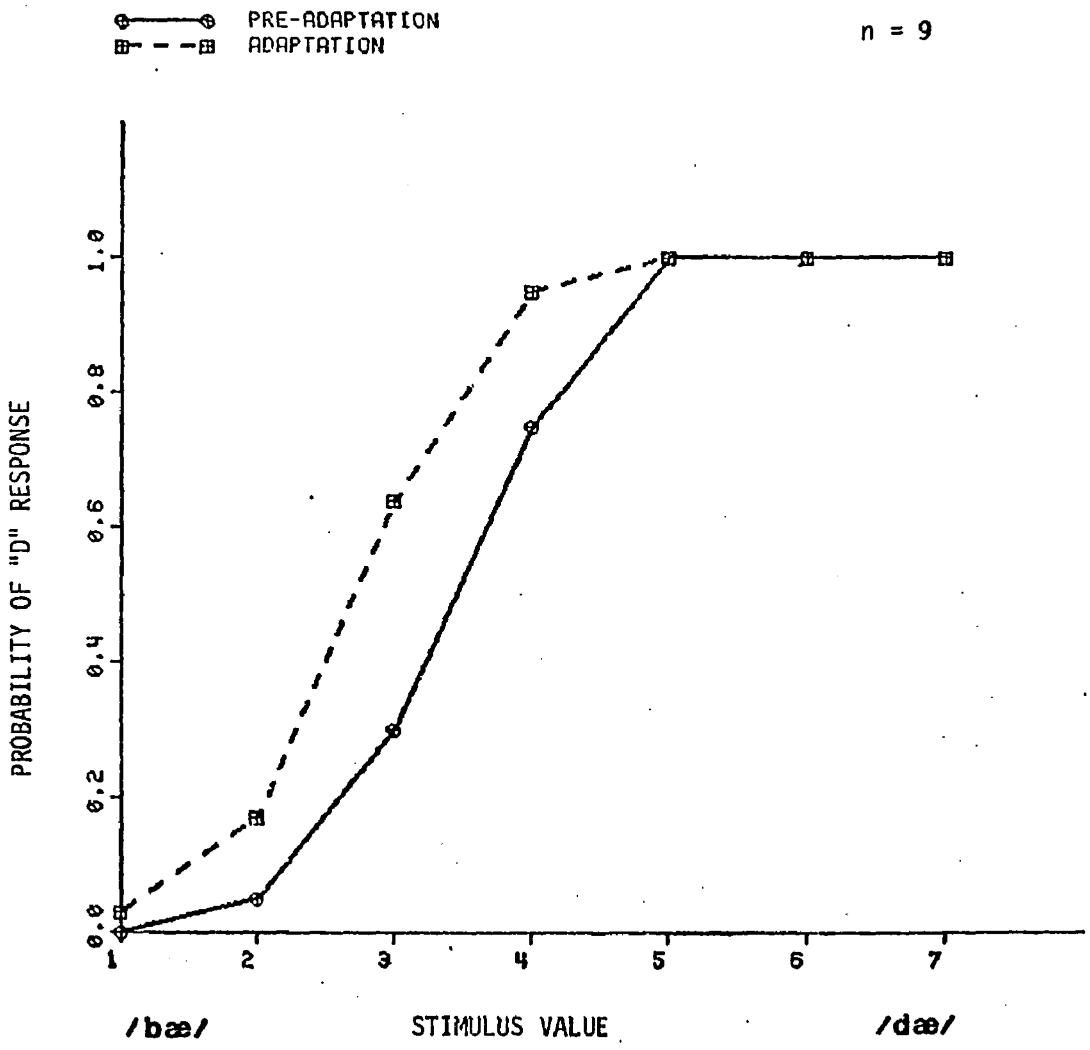


Table 16. Number of /d/ responses before and after adaptation for the three ratios. (Stimuli 2-7 only).

	8:1			16:1			32:1		
Subjects	Pre-Adaptation	Adaptation	Difference	Pre-Adaptation	Adaptation	Difference	Pre-Adaptation	Adaptation	Difference
QS	78	82	- 4	84	95	- 11	80	97	- 17
MS	78	91	- 13	82	90	- 8	82	94	- 12
FB	97	106	- 9	99	115	- 16	103	121	- 18
RG	100	107	- 7	94	110	- 16	90	115	- 25
DH	96	107	- 11	95	120	- 25	97	126	- 26
GK	101	101	0	115	124	- 9	130	130	0
PB	131	133	- 2	124	129	- 5	125	139	- 14
AL	100	113	- 13	97	100	- 3	113	127	- 14
HJ	94	108	- 14	86	108	- 22	102	121	- 19
MEAN	97.2	105.3	- 8.1	97.3	110.1	-12.8	102.4	118.9	-16.4

changes as a function of the number of adaptor repetitions: the mean increase in number of /d/ responses after adaptation increases systematically from 8.1 to 12.8 to 16.4 as the number of adaptors increase from 8 to 16 to 32. Figure 12 shows that the increase over this range is a linear function of the logarithm of the number of adaptors. Notice, however, that the effect is not monotonic for all subjects: two subjects (MS and AL) display reversals between 8:1 and 16:1 ratios; two others (GK and HJ) display reversals between 16:1 and 32:1.

Analysis of variance, Subjects (9) by Ratio of Adaptors to test syllables (8:1, 16:1, 32:1), was performed on the increase in the total number of /d/ responses (maximum:150) for each ratio over stimuli 2-7. Responses to stimulus 1 (i.e., the adaptor stimulus) were not included in the data analysis in order to make the analysis consistent with that to be used in subsequent anchoring studies. The summary table for the analysis of variance is shown in Table 17. A significant main effect of the number of adaptor repetitions ($F_{2,24} = 18.361$; $\underline{p} < 0.01$) was found, indicating that the degree of adaptation increased with the increasing number of adaptor repetitions.

Figure 13 displays the mean number of /d/ (unadapted class) responses per subject for the three adaptor repetition conditions as a function of blocks of 25 responses. Each block represents five responses per test syllable per subject for stimuli 2-7 in order of presentation. Block 1, therefore, contains the mean of the first 5 responses, block 2 the next five, and so on. All adaptor repetition conditions show slight but statistically nonsignificant increases in the mean number of /d/ responses from block 1 to block 5.

Table 17. Summary table of analysis of variance for the total number of /d/ responses for each of the three ratios.

Summary table of analysis of variance for the total number of /d/ responses for each of the three ratios.

Ratio (R): 8:1, 16:1, 32:1; Subjects (S).

Source	df	MS	F	_p_
Subjects (S)	8	715.069		
R	3	625.879	18.361	< 0.01*
RxS	24	34.087		
TOTAL	35			

Figure 12. Mean increase in number of responses in unadapted class as a function of ratio of adapting to nonadapting stimuli.

MEAN INCREASE IN NUMBER OF RESPONSES IN UNADAPTED CLASS AS A FUNCTION OF RATIO OF ADAPTING TO NON-ADAPTING STIMULI

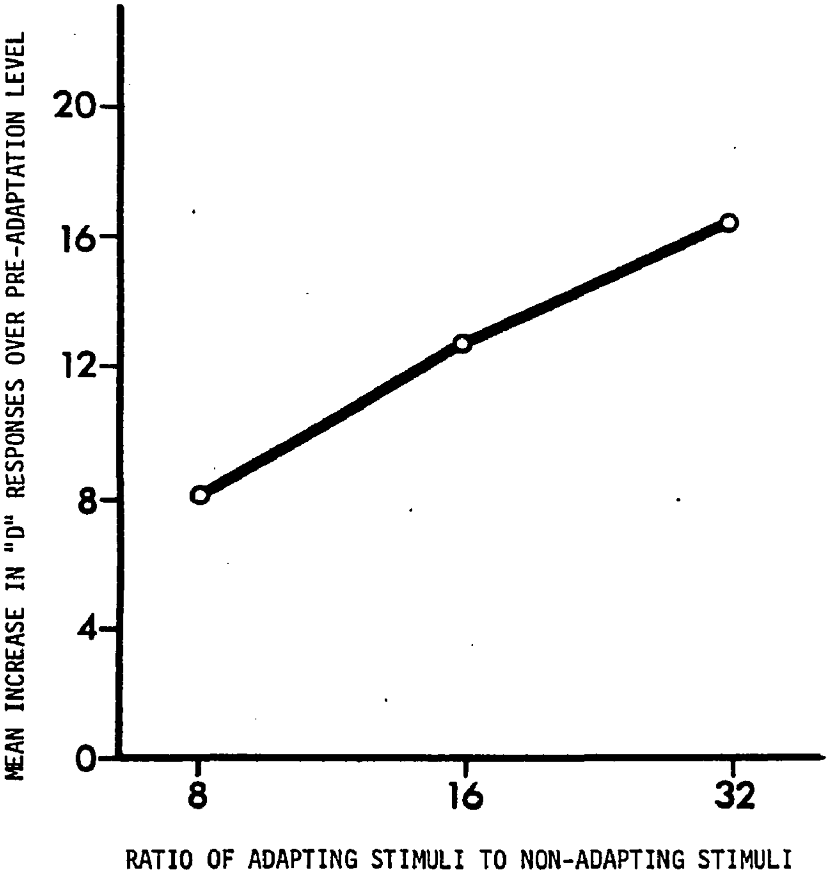


Figure 13. Degree of shift in the mean number of responses in the unadapted class per block of 25 responses.

DEGREE OF SHIFT IN MEAN NUMBER OF RESPONSES IN UNADAPTED CLASS PER BLOCK OF 25 RESPONSES

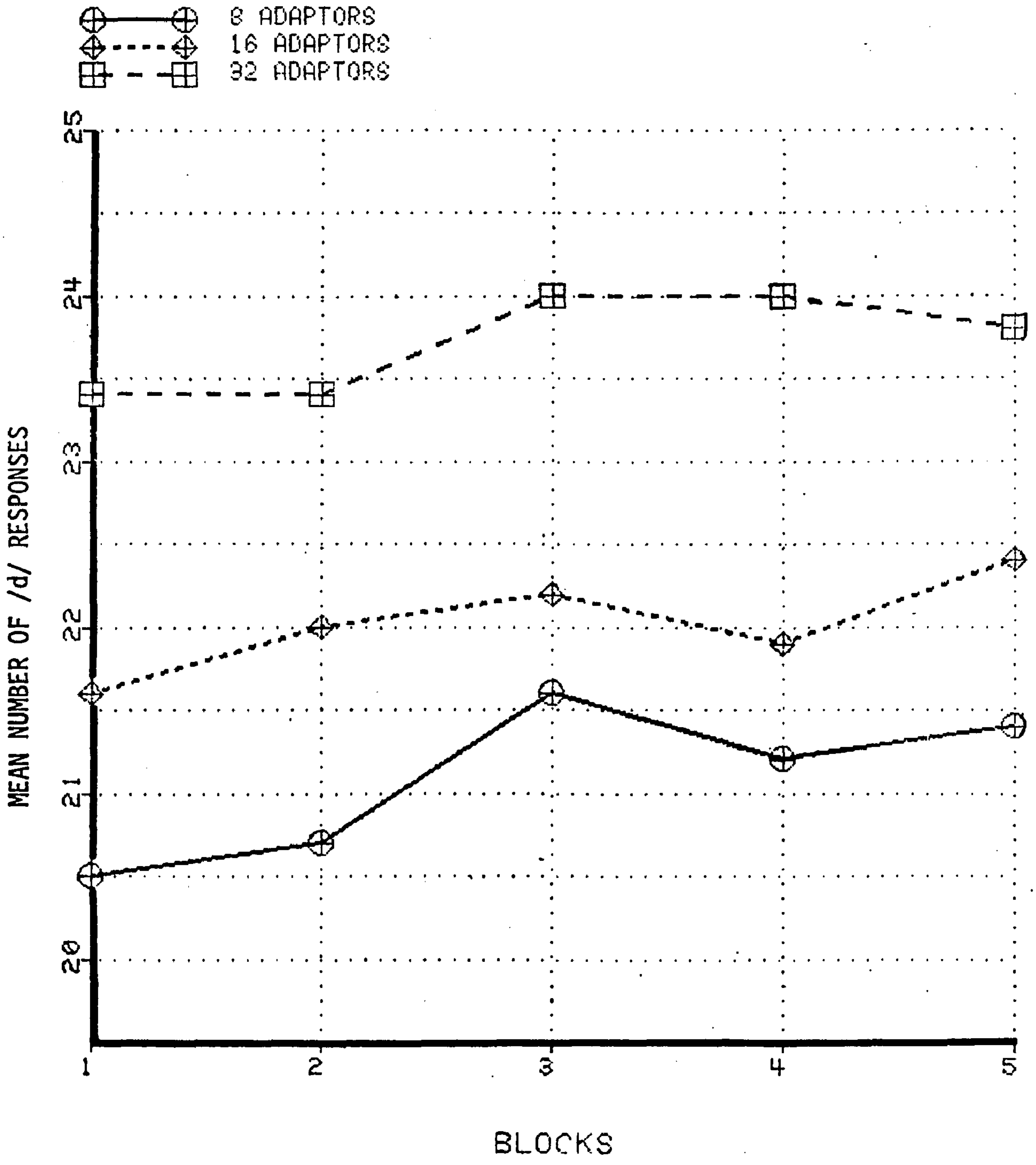


Figure 14 is a graphic representations of the percentage of /d/ responses as a function of the cumulative number of preceding adaptation repetitions ("hits") and of the cumulative total of adaptation duration in seconds under the three experimental conditions. It is evident from this graph that there is little change in the percentage of /d/ responses from 80 to 160 preceding adaptors for either 8 or 16 repetitions, and no change from 160 to 320 preceding "hits" for either 16 or 32 repetitions. In other words, the effect is evidently not cumulative over the test: the maximum effect for each condition is already displayed within the first 40 seconds of adaptation.

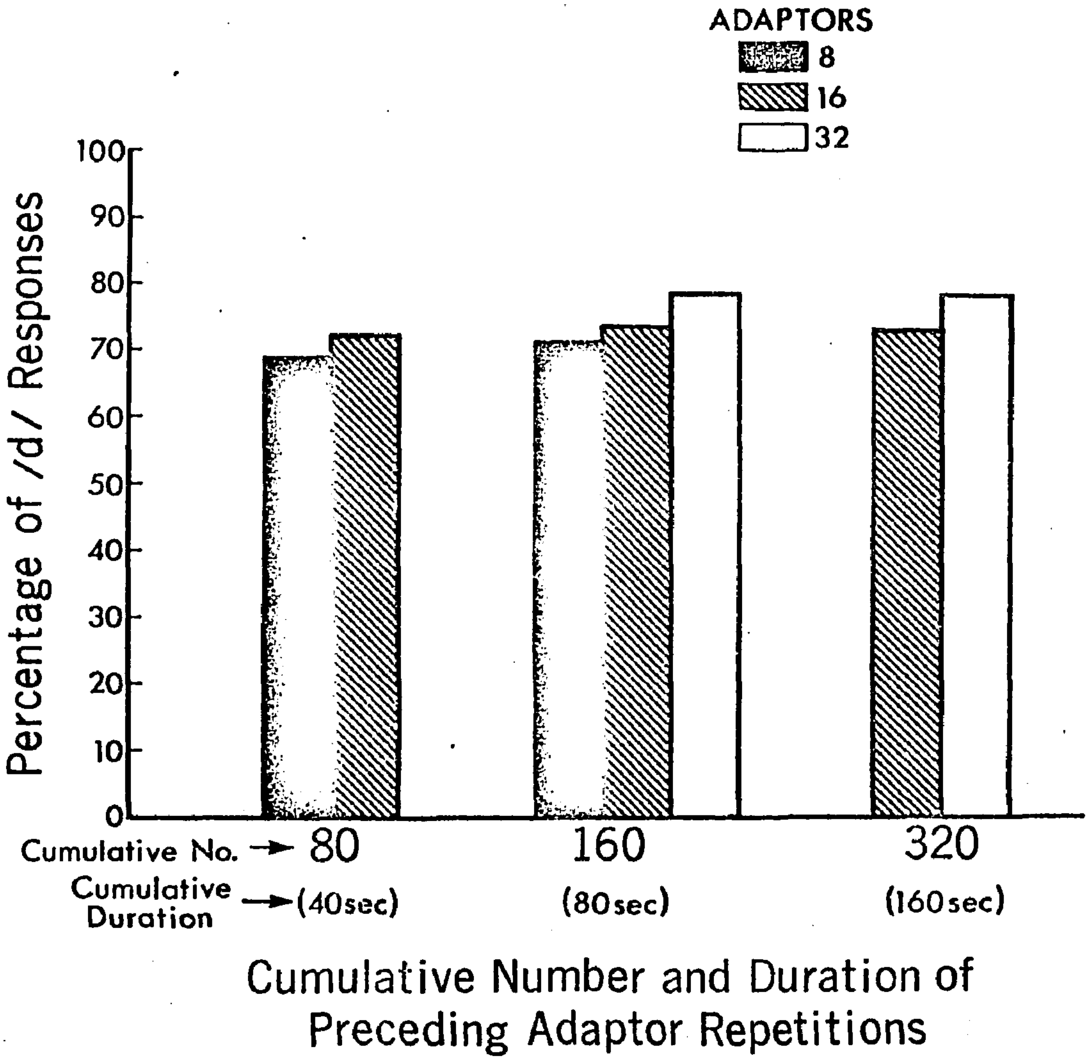
Summary and Discussion

The following general observations concerning the effects of varying the number of adaptor repetitions on the category boundaries of a stop consonant judgment can be made.

1. Over the range studied, adaptation is a linear function of the logarithm of the number of adaptor repetitions per trial. The increase in the number of adaptor repetitions resulted in a systematic increase in the magnitude of the phoneme boundary shift.

These findings contrast with those of Bailey (1975) who found no systematic difference in the size of the adaptation effect produced by 32 repetitions versus 8 repetitions. However, they are in agreement with those of Hillenbrand (1975) who found systematically larger shifts in the phoneme boundary with greater intensity level and more adaptor repetitions. As previously discussed, the methodological differences

Figure 14. Cumulative number and duation of preceding adaptor repetitions.


ADAPTORS
▨ 8
▧ 16
☐ 32
Percentage of /d/ Responses
100
90
80
70
60
50
40
30
20
10
0
Cumulative No. → 80
160
320
Cumulative Duration → (40sec)
(80sec)
(160sec)
Cumulative Number and Duration of Preceding Adaptor Repetitions

among the three studies makes comparisons difficult.

Although there are two points in time that are common to the present study and Hillenbrand (1975) (repetition rates of 1.95 and 2 per second) as seen in Table 15, the number of adaptor repetitions per trial varies, resulting in differences in the duration of the adaptation sequence per trial. Therefore, the total amount of stimulus energy available cannot be compared.

On the other hand, the Bailey (1975) study and the present study have three points in common: the average repetition rate, the duration of the adaptation sequence per trial, and the number of adaptor repetitions per trial. However, they differ in three aspects: that the Bailey (1975) study used 60 adaptor repetitions preceding each block of trials while the present study used none, that the time between the end of the adaptation sequence and test syllables differed (800 msec in the Bailey study and 2 sec in the present study), and that the number of stimuli presented after each adaptation sequence differed, only one test syllable in the Bailey study as compared with all 7 in the present study.

2. Over the range studied, the effect is not cumulative over trials and is present at full strength within the first five-ten trials. Figures 13 and 14 indicate negligible or no shifts in the mean number of /d/ responses as a function of the cumulative number and duration of the preceding adaptor repetitions. It appears that the maximum effect is already displayed within the first block or 40 seconds of adaptor repetitions (80 "hits") and the effect of the number of adaptor

repetitions is not cumulative over the test. From this it would appear that the time course of both the development and the decay of the effect is extremely rapid. These findings will be discussed further in Chapter VI in light of the findings in Experiment V.

As Bailey suggests (1975), the determination of the precise time course of the development of speech adaptation should be attempted with comparisons of the values obtained from studies in non-speech auditory adaptation. Table 18, modeled after Table 3.13, (page 110) in Bailey (1975) shows the time course of the development of adaptation using various measurement procedures for both the visual and auditory modality. As Bailey (1975) points out, the different techniques used to obtain the measurements make comparisons difficult in the auditory modality. However, in this experiment, maximum adaptation occurred by 40 sec of stimulus energy, if not sooner, regardless of the number of adaptor repetitions before each testing sequence. This would be in agreement with the adaptation time for the verbal transformation effect for CV syllables (Lass and Gasperini, 1973) and isolated vowels (Lass and Golden, 1971).

As previously discussed, the results of studies in auditory adaptation suggest that the total amount of stimulus energy is important in terms of receptor (cochlear and neural) functioning. The findings in this study, that the low-level stimulus energy variable (i.e., the number of adaptor repetitions), had an effect on the magnitude of the phoneme boundary shifts, may be evidence for some level of adaptation operating at an auditory or acoustic stage of processing.

Table 18. Time course for the development of adaptation in the visual and auditory modalities.

Technique	Temporal Measurement	Adaptation Time
Auditory Modality		
Loudness balance (Average of the pre-adaptation intensities in the control ear is compared to the perstimulatory and post-stimulatory values.)	Time to maximum observed difference in the intensity necessary to obtain a balance.	Can occur in less than 10 sec. Greatest rate of adaptation occurs during the first 1-2 minutes of stimulation. Asymptote reached from 3-7 minutes after onset of stimulus (Elliot and Fraser, 1970).
Detection threshold of frequency-modulated tones.	Time to maximum threshold shift.	12 seconds (Kay and Matthews, 1972)
Verbal transformation effect	Time for first transformation dependent upon stimuli.	
	1) CV syllables	25-30 seconds (Lass and Gasperini, 1973)
	2) isolated vowels	40 second (Lass and Golden, 1971)
	3) nonspeech sounds	70 seconds (Lass, West and Taft, 1973)
Visual Modality		
Figural after-effects	Time to maximum effect	60 seconds (Hammer, 1949)
Dark adaptation	Time course of the increase in sensitivity of eye after illumination is terminated	1-10 minutes (Reviewed in Haber and Nershenson, 1973)
Size adaptation: (Contrast threshold for grating patterns)	Time for maximum effect	2-3 minutes (Blakemore and Sutton, 1969; Blakemore and Campbell, 1969)

The next experiment will further attempt to investigate the effects of the total stimulus energy in the selective adaptation paradigm by holding the number of adaptor repetitions constant while varying the presentation rate. The rate will be varied by increasing the interval between the multiply-presented adaptors (the inter-adaptor-interval) from 250 to 750 to 1750 msec. The combined results of these two experiments will permit the effects of number of adaptors and rate of repetition (or density) to be separated.

CHAPTER VI: EXPERIMENT V

ADAPTATION AND INTER-ADAPTOR-INTERVAL

Introduction

The primary purpose of this study was to determine the effect on stop-consonant adaptation of variations in the rate (or density) of adaptor presentation, using the lowest effective number of adaptor repetitions as established in Experiment IV. The rationale for varying the rate of repetition by increasing the IAI was two-fold. First, this was a necessary preliminary to the comparison of adaptation and anchoring planned for Experiment VI. In order to equate the two paradigms, consistent temporal intervals must be maintained between the test syllables and anchors and adaptors. This time interval needed to be at least 1750 msec so that subjects might have time to respond to both the anchors and adaptors. (See Chapter VII for a detailed explanation of the paradigms). Therefore, it is important to find boundary shifts for the adaptation paradigm using 16 adaptors (a number sufficient to produce a phonetic boundary shift in Experiment IV) at an IAI of 1750 msec before proceeding to Experiment VI.

Second, Experiment IV shows a clear effect of the number of adaptor repetitions with rate held constant (see Table 16). If the magnitude of the effect depends solely on the number of adaptor repetitions per trial, there should be no difference in the effect of using 16 adaptors, whatever the IAI, at least up to 1750 msec for there must be an outer limit of rate where the system can no longer be adapted. However, if the effect does change with manipulation of the IAI, we may conclude that the variations of Experiment IV were due to rate (density). It is further possible that a trade-off relationship exists between the number of adaptor repetitions and IAI. For example, if the larger IAI results in a decreased phonetic boundary shift, due to decreased repetition rate, an increased number of adaptor repetitions may be necessary to compensate for the reduction. In other words, there may be a time interval of silence between the adaptors where there is no longer an adequate summation of energy either due to the recovery of the system between adaptors or to a reduction in the total stimulus energy available.

Method

Description of Stimuli and Experimental Tapes

The seven stimuli used in Experiment IV also served as stimuli in this experiment. One pre-adaptation and three adaptation experimental tapes were prepared from these stimuli and recorded at 3 3/4 inches per second. The pre-adaptation tape was identical to that used in Experiment IV, using a total of 175 stimuli.

The three adaptation tapes were similar to those in Experiment IV and to each other in all ways save the IAI. The adaptor again was Stimulus 1. The number of adaptor repetitions per trial was held constant (16) while the IAI varied from 250 msec to 750 to 1750 msec. These intervals were not arbitrary but were chosen because, when added to the 250 msec adaptor, a doubling of the onset-to-onset interval was effected. Thus, with the 250 msec adaptor there were 2 adaptor repetitions per second (1000 msec), with the 750 msec adaptor, there was one repetition per second, and with the 1750 msec interval, there was one adaptor every two seconds (see Table 15). There was a total of 175 adaptation trials per IAI condition, yielding 25 judgments of each test stimulus at each IAI.

Procedures

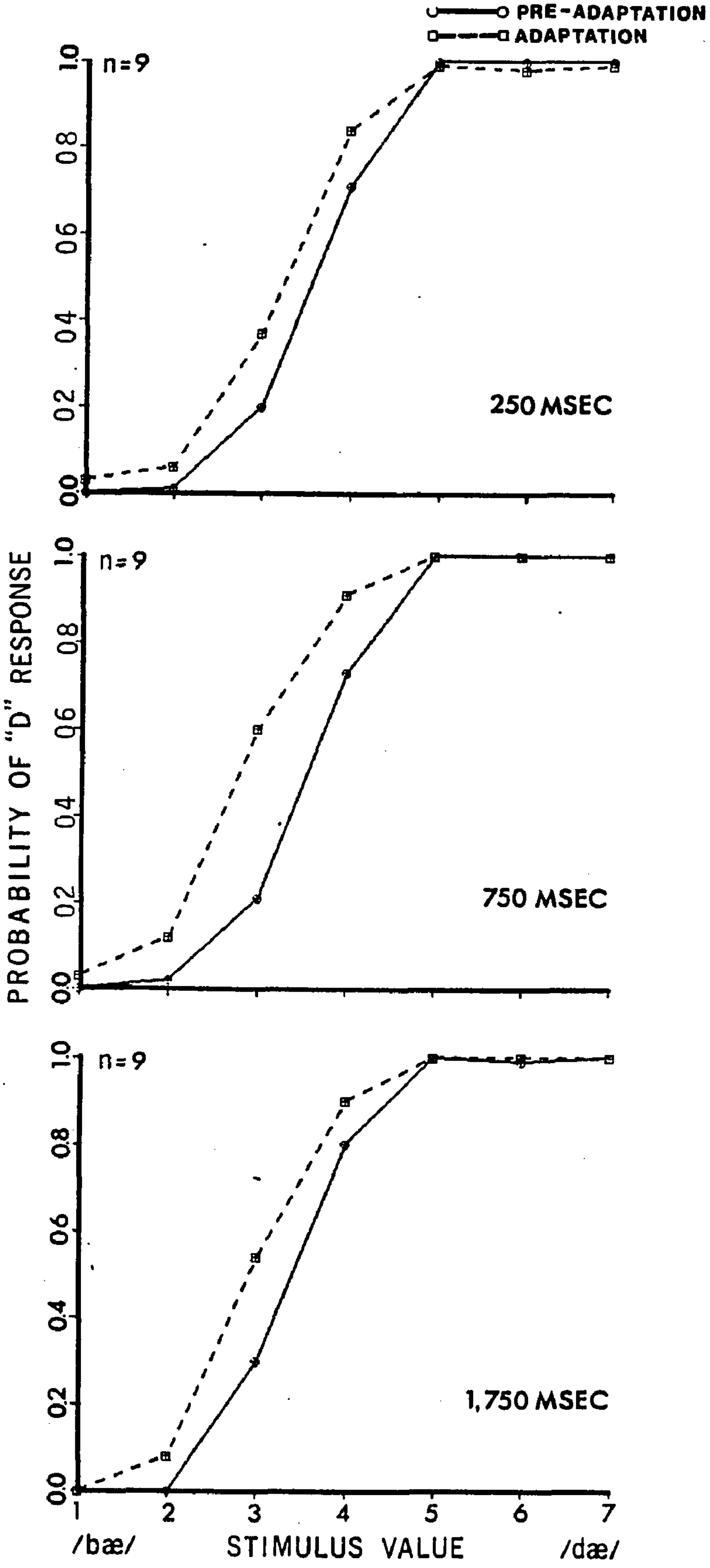
As in the previous experiments, nine subjects were used, including two from the previous experiment. Tapes containing the appropriate stimuli for the experiment were produced as described previously.

Three adaptation sessions were run, one for each IAI, separated by a minimum of 24 hours. Subjects heard a pre-adaptor tape before each experimental adaptor tape. They were informed that the experimental variable would be pauses of differing lengths between the syllables preceding the test stimuli. The order of presentation of the three IAI conditions was counter balanced for the first six subjects with the order of subjects 1-3 repeated for subjects 7-9 . Each testing session was preceded by practice and listening tapes in the pre-adaptation condition.

Results

Group phoneme boundary and functions for subjects in the pre-adaptation and adaptation conditions for all three IAI values are shown in Figure 15 . Inspection of the figures shows a shift in the phoneme

Figure 15. Average identification functions for stop consonants for each of the three inter-adaptor-interval conditions.

PRE-ADAPTATION
ADAPTATION
n=9
250 MSEC
750 MSEC
1,750 MSEC
PROBABILITY OF "D" RESPONSE
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
/bæ/
STIMULUS VALUE
/dæ/

boundary of the continuum towards the adaptor stimulus for all three IAI conditions. For all three pre-adaptation conditions the boundary lies between stimuus 3 and 4, but the magnitude of the phoneme boundary shift evident for all three IAI values changes as a function of the interval between the adaptors. Figures 16 and 17 are clearer indications of the magnitude of the shifts.

Table 19 shows the individual and total number of /d/ responses in the pre-adaptation and adaptation conditions for each IAI value for stimuli 2-7. There is a mean increase of 8.0 /d/ responses between the pre-adaptation and adaptation conditions for 250 msec, 16.2 for 750 msec, and 9.4 for 1750 msec. (Upon inspection of the pre-adaptation totals, it can be seen that there is a difference of 40 /d/ responses between 250 msec and 1750 msec. Since the identical tape was used in each of the three pre-adaptation conditions, it is difficult to explain this discrepancy.) This nonmonotonic relation between the increase in the duration of the IAI and the phonetic boundary shift can be seen in Figure 16, where the degree of shift in the total number of /d/ responses (unadapted class) for the three interval conditions is displayed.

Analysis of variance, Subjects (9) by IAI (250 msec, 750 msec, 1750 msec) by Ratio (1:1, 16:1) was performed with the total number of /d/ responses for each IAI for stimuli 2-7. The summary table for the analysis of variance is shown in Table 20. The main effect of ratio or the change from the pre-adaptation to adaptation condition indicates that there was a significant boundary shift with the adaptation

Figure 16. Degree of shift in the number of responses in the unadapted class per inter-adaptor-interval condition.

MEAN INCREASE IN THE NUMBER OF RESPONSES IN THE UNADAPTED CLASS PER I-A-I CONDITION

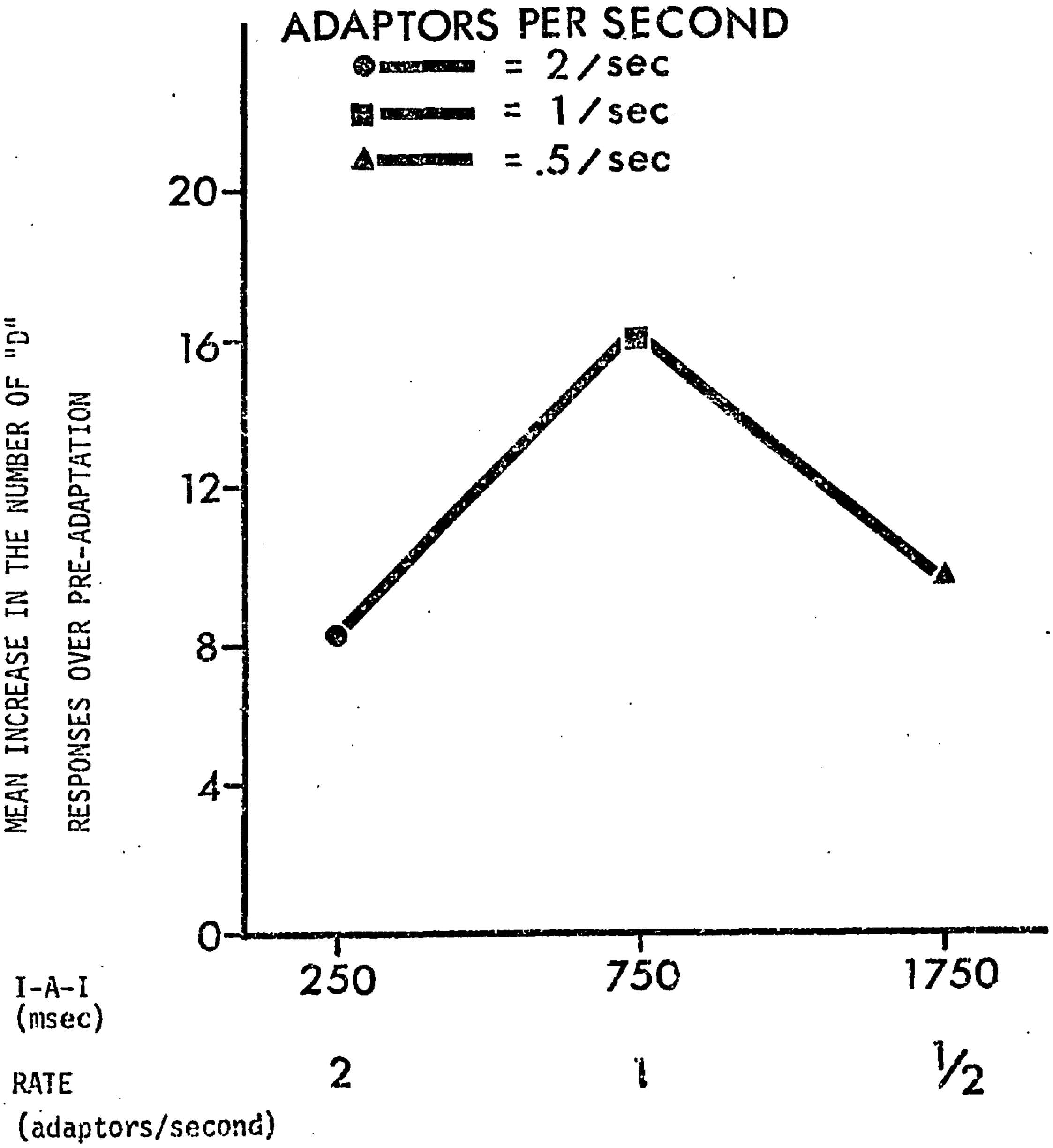
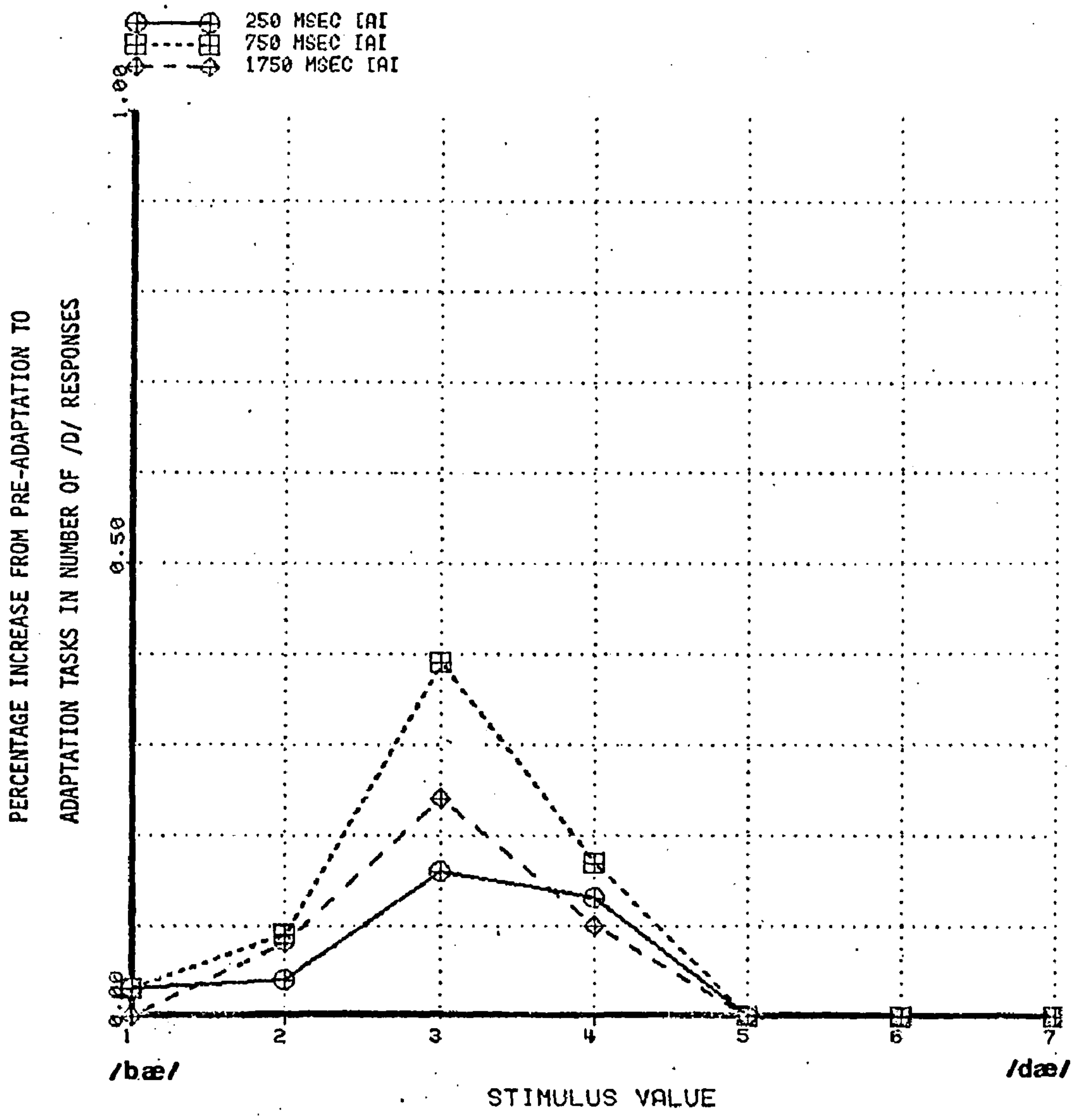


Figure 17. Reinterpretation of the data shown in Figure 16.



REINTERPRETATION OF DATA SHOWN IN FIGURE 16

Table 19. Number of /d/ responses before and after adaptation for the three inter-adaptor-intervals.

Number of /d/ responses before and after adaptation for the three Inter- Adaptor- Intervals. (Stimuli 2-7 only).

	250 msec			750 msec			1750 msec		
Subject	Pre-Adaptation	Adaptation	Difference	Pre-Adaptation	Adaptation	Difference	Pre-Adaptation	Adaptation	Difference
JV	109	106	3	114	118	- 4	114	125	-11
HJ	91	108	-17	94	109	-15	93	109	-16
DB	99	107	- 8	104	120	-16	114	132	-18
DC	81	85	- 4	75	91	-16	87	89	- 2
MR	85	86	- 1	89	100	-11	84	93	- 9
DH	98	121	-23	88	126	-38	99	123	-24
CW	97	105	- 8	98	122	-24	107	109	- 2
VH	100	118	-18	106	122	-16	109	119	-10
JK	123	119	4	126	132	- 6	116	109	7
MEAN	98.1	106.1	-8.00	99.3	115.5	-16.2	102.5	112.1	-9.4

Table 20. Summary table of analysis of variance for the total number of /d/ responses for each of three inter-adaptor-intervals.

Summary table of analysis of variance for the total number of /d/ responses for each of three inter-adaptor-intervals. Ratio (R): 1:1, 16:1; Msec (M): 250, 750, 1750; Subjects (S).

Source	df	MS	F	p
Subjects (S)	8			
R	1	1700.125	15.460	< 0.01*
RxS	8	109.969		
M	2	165.500	3.504	NS
MxS	16	47.230		
MxR	2	86.719	5.48	NS
MxRxS	16	15.824		
TOTAL	45			

condition in all the IAI conditions ($F_{1,8} = 15.460$; $p < 0.01$). The main effect of msec fell short of significance ($F_{2,16} = 3.504$; $p > 0.05$). This lack of a significant difference among the means of these three treatments suggests that the change in the IAI may have no effect in determining the magnitude of the phoneme boundary shift over the range of intervals tested. However, the ratio x msec interaction was significant ($F_{2,16} = 5.48$; $p < 0.05$), presumably due to the greater boundary shift with the 750 msec interval. In fact, a test of the significance between the boundary shifts of the three IAI conditions by means of a Wilcoxon Matched-Pairs Signed Ranks Test revealed a significant difference between the 250 and 750 sec intervals ($T = 3.0$; $p < 0.05$), although not between the 250 and 1750 msec interval nor between the 750 and 1750 msec interval.

A Wilcoxon Matched-Pairs Signed Ranks Test was also used to test the significance of the pre-adaptation boundary and the adaptation boundary in each of the IAI conditions. A significant shift was found in all three intervals (250 msec: $T=5$; $p < 0.05$; 750 msec: $T=0$; $p < 0.01$; 1750 msec: $T=3$; $p < 0.05$). Therefore, there was a significant change in the phoneme boundary even with as much as a 1750 msec interval between the 16 repeated adaptors.

Summary and Discussion

1. Over the range studied, there was a significant effect of adaptation as observed by the significant boundary shifts in the change from the pre-adaptation (1:1) to the adaptation (16:1) condition in each

of the Inter-Adaptor-Intervals.

2. There was a non-monotonic relation between the increase in the duration of the Inter-Adaptor-Interval (IAI) and the phoneme boundary shifts. A greater boundary shift was found for the 750 msec interval than for either the 250 msec interval or the 1750 msec interval. While these differences fall short of significance in the overall analysis of variance, there was a significant difference between the 250 msec and 750 msec intervals on the Wilcoxon Matched-Pairs Signed Ranks Test, but not between 250 msec and 1750 msec nor between 750 msec and 1750 msec. In fact, examination of the Figures 15, 16, and 17 shows that at every stimulus, the 750 msec pre-adaptation and adaptation conditions result in a higher percentage of /d/ responses.

It is obvious that time must have an effect in adaptation. Since the effect would presumably disappear if IAI were increased to infinity, it seems likely that a significant difference between 750 msec and the larger value would have been shown had IAI been extended further than 1750 msec. In other words, the maximum effect for stimuli of this duration lies closer to an interval of 750 msec than to 250 msec or 1750 msec or repetition rates of 1 "hit" and 1/2 "hit" per second. Stimuli presented with an IAI of 250 msec and 1750 msec are not as effective as those with an IAI of 750 msec.

There is certainly good reason to expect a peak in the function due to the time needed for the development and the decay of the effect, so that , although the evidence for a peak at 750 msec is not solid, some speculative interpretations can be made in light of the present data in terms of an optimum range or rate of adaptation.

First, it may be that the time peak in the function does not lie at 750 msec. The peak may lie closer to 1 second, for it is obvious that stimuli presented at 1750 msec (as well as 250 msec) are not as effective as adaptors. Sampling at more frequent intervals is needed to isolate the true peak.

Second, this non-monotonic relationship may be explained by the fact that the system recovered with the 1750 msec IAI. In other words, the 750 msec interval appears to be near optimum for adaptation with the 16 adaptors and it is assumed that there is sufficient energy to fatigue the receptor without sufficient recovery time. At the 1750 msec interval there is also sufficient energy to fatigue the receptor, however, time is allowed between the adaptors for some recovery of the system. This would be in agreement with studies in auditory adaptation, where, according to Elliot and Fraser (1970), the adaptation processes suffer less from a reduction in time due to an increase in interruption rate than do the recovery processes.

However, the smaller effect with the 250 msec IAI is more difficult to rationalize. It is possible that there is a minimum time interval between adaptors needed for an effect to accumulate. If the adaptor repetitions are too close in time, there may be weak adaptation because the system did not have a chance to summate properly. However, the minimum values for summation of energy in psychoacoustics are much smaller than 250 msec (Munson, 1947) and in auditory adaptation, the degree of adaptation increases with an increased interruption rate (Carterette, 1955). In fact, Carterette (1955) suggested that cumulative effects may be produced by the increased interruption rate due to the shortened recovery time, although the adaptation period is also reduced. However, as previously discussed, the adapting processes are assumed to suffer less from the interrupted signal than do the recovery processes. However, in the auditory adaptation studies, the duty cycle was maintained at 50%, therefore keeping total energy equal in all conditions. This concept of duty cycle in selective adaptation will be further discussed in the next section.

Experiments IV and V: Summary and Discussion

The results of Experiment IV showed that an increase in the number of adaptor repetitions from 8 to 16 to 32 with a constant IAI of 250 msec systematically increased the magnitude of the phoneme boundary shift. In fact, over the range studied, adaptation is a logarithmic function of the number of adaptor repetitions per trial. Increasing the number of adaptor repetitions increased the total amount of stimulus energy reaching the receptor. The greatest amount of stimulus energy reaching the receptor in Experiment IV was using 32 adaptors at a 250 msec IAI, and, as expected, the largest boundary shift occurred. The results of Experiment IV also showed that over the range studied, the effect is not cumulative over trials and is present at full strength within the first 40 seconds of adaptor repetitions.

The results of Experiment V showed that, while there was no significant overall effect of increasing the IAI from 250 to 750 to 1750 msec, there were significant boundary shifts in all IAI conditions for the adaptation condition over the pre-adaptation condition. Also, there was a significant difference between the phoneme boundary shifts in the 250 msec condition versus the 750 msec condition.

In fact, the results of Experiment V may partially explain the lack of a phoneme boundary effect in Experiments I and II and the lack of a cumulative effect in Experiment IV. The outcome of Experiment V suggests that the effect of a single "hit" is already decaying within 2 seconds. The anchoring experiments (Experiments I and II) used a maximum of 4 adaptor repetitions with IAI's of 2 seconds which may not have been sufficient to obtain a phoneme boundary shift.

In Experiment IV, the presence of the test syllables with a 1750 msec interval between the syllables following the adaptation sequence did not allow further adaptation to take place during the presentation of the test syllables, and, in fact, probably caused a decay in the adaptation effect. Therefore, following each series of test syllables, the system was either fully or partially recovered and would again be adapted by the next adaptation sequence. Cumulative effects, therefore, would not be produced due to the time allowed for recovery. The fact that there was not a cumulative effect of the experiment may be reflecting relatively weak adaptation as well as a rapid development and decay of the effect.

It is also suggested that since increasing the IAI above 750 msec (and decreasing the repetition rate) resulted in a tendency to decrease the number of /d/ responses, this interval may give sufficient time between the adaptor repetitions to prevent a summation of stimulus energy to adequately fatigue the receptor. In other words, if there is a threshold of adaptation, a certain level of temporal and/or stimulus energy may be necessary to reach this threshold. Lengthening the IAI while maintaining a set number of adaptor repetitions also temporally increased the adaptation sequence, but did not increase the total amount of stimulus energy involved. It is possible that this threshold level may function in a trade-off relationship in that larger IAI values (above 750 msec) may need to be accompanied by a greater number of adaptor repetitions and/or a greater syllable duration to achieve the same magnitude of effect.

In auditory adaptation as well as auditory fatigue, the duration of the fatiguing stimulus has been found to be linearly related to the amount of adaptation or fatigue. It has been found for TTS, that growth is linearly proportional to the log D (i.e. it is negatively accelerated) except for frequencies below 2000Hz. This is particularly true when the fatiguing stimulus is noise or a rapidly interrupted tone (Ward, 1963). Using interrupted pure tones, TTS has also been found to increase with increased durations ranging from 30 seconds to 15 minutes (Elliot and Fraser, 1970). The cumulative aspects of fatigue have also been tested and although the intensity-duration relation is not a simple multiplicative one, greater increases in exposure intensity are needed to produce a given TTS at shorter durations than at longer durations (Elliot and Fraser, 1970).

Inspection of Table 15 illustrates two ways to approach the analysis of the temporal factors in adaptation. For Experiment IV, the table displays the time of each adaptation trial, the average repetition rate (per second), and the "duty cycle" with the number of adaptor repetitions as the variable and a constant IAI. The same time intervals but with varying IAI's are illustrated under Experiment V. In the first example, the number of adaptor repetitions is varied from 8 to 32, the average repetition rate remains constant at 2 per second, while the total time per trial varies. As the IAI is varied, the duration of the adaptation per trial doubles while the average repetition rate per second decreases by half with each increase in IAI.

The second comparison establishes the identical duration of adaptation per trial although both the number of repetitions and the IAI differ in the experiments (32 versus 16, 250 versus 750). Therefore, in the case of the 32 adaptors there are more adaptors with a shorter IAI.

Another way of looking at the data is in terms of the "duty cycle" or the ratio of the time of the event (duration of the syllable) to the total time of the period (total time of one adaptation sequence). Table 15 illustrates that the "duty cycle" remains constant at 50% for each of the adaptor repetitions while it halves at each increase in IAI time. Therefore, in the case of the 32 adaptors, there is a greater time of each adaptation trial than with the other adaptor repetitions, but the duty cycle remains the same as for the other adaptor repetitions. Although the 16 adaptors at 750 msec manifest the identical time values, the duty cycle is only one-half that of the 32/250 msec condition. If one assumes that there is no recovery of the mechanism in the 250 msec allowed between the adaptors, the constant high ratio of "hits" should fatigue the receptor and therefore effect a higher number of /d/ responses than in the 16/750 msec case. On the other hand, the 16/750 msec condition would have the same effect on the receptor if adaptation is dependent upon the adaptation time per trial but would have a lesser effect if the duty cycle hypothesis is correct.

The total amount of /d/ responses in the 32/250 condition is 1070 versus 1040 in the 16/750 condition. The difference between these two conditions is not significant ($T_8=.46$; $p > 0.05$), therefore, suggesting the density or duty cycle effects alone for a duty cycle up to 25% may not be totally responsible for the effect. However, a lack of a significant difference between the 250 msec and 1750 msec conditions on the Wilcoxon also shows that there was no difference in the results with the 12.5% duty cycle.

However, it is still possible that the 25% duty cycle as interpreted in terms of summation of stimulus energy is minimum to achieve a strong adaptation effect. If stimuli of more than 250 msec duration were used to maintain a duty cycle of 25%, a greater shift in the phoneme boundary with the larger IAI's may be anticipated. Therefore, it would be interesting to lengthen the duration of the syllable, as well as the number of adaptor repetitions above 750 msec to see if a trade-off relationship exists or if there is a compensatory effect of the increased duration of the syllable and the number of adaptor repetitions.

Therefore, further experimentation is needed to see if a trade-off relationship exists between the variables discussed: the number, rate (density) and the duty cycle of the stimuli involved. These results suggest that the total amount of stimulus energy is important in terms of receptor (cochlear and neural) functioning. The findings of these studies, that the low-level stimulus energy variable, the number of adaptor repetitions and the rate of repetition, had an effect on the magnitude of the phoneme boundary shifts, may be evidence of adaptation operating at an auditory or acoustic stage of processing.

The results of these two experiments might also be considered in terms of response bias or response organization effects. As previously discussed, the Range-Frequency Theory (Parducci, 1974) states that category judgment is a compromise between the range principle (the division of the stimulus range) and the frequency principle (the frequency of use of the different categories). According to the range principle, the subject divides the range into subranges corresponding to each category. If, for example, there are two categories available for judgment, as in these experiments (/b/ and /d/), according to Parducci (1965, 1974), each category would correspond to half the range. A change in the frequency of the presented stimuli would not affect judgment unless the psychological range were also changed. Parducci (1965) states that the psychological range, however, cannot be directly inferred from the physical range but must be inferred from the judgments of the subject. "Anchoring" has most often been applied to shifts in the scale that are dependent upon extensions of the range.

The frequency principle states that the subject uses each category for a fixed proportion of judgments, usually in equal proportions (Parducci, 1965). The frequency proportions are independent of stimulus conditions, as are the relative sizes of the subranges, but are affected by the stimulus frequencies. Therefore, the two principles conflict when there is an unequal presentation of stimulus frequencies from each of the subranges, and judgments represent a compromise between the range of stimuli and their relative frequencies. The equal-frequency tendency occurs in comparative judgments where the PSE (point of subjective equality) shifts towards the more frequently presented stimulus, but there is no shift in discrimination scaling with changes in frequency (Parducci, 1974). This model also shows a trade-off relationship between the effects of relative spacing of the stimuli (in value, not in inter-stimulus-interval) and the frequency of the stimuli. The shape of the judgment function is steeper when the stimuli are more closely spaced or presented with greater frequency.

According to a response bias interpretation, therefore, equal partitioning of the categories based on the inputs might be expected. With a change in the frequency of the presentations, a shift in the range might be expected with added anchors (or adaptors acting as anchors) in addition to a shift in the scale due to the change in the frequency of the presented stimuli. Therefore, the subject might be expected to respond with a greater number of /d/ judgments when there are a systematicaly greater number of preceding repetitions of the /b/ stimuli based on the range principle and the equal-frequency principle. This is precisely what occurs in Experiment IV with the change from the pre-adaptation level to the various adaptor levels. It might be stated, therefore, that some aspect of the response is interwoven with the adaptation ratio.

However, the temporal effects of Experiment V (with a peak about 1 to 1/2 "hit" per second) are not compatible with a response bias account. A response bias or response orgnaization interpretation would anticipate a shift in the function from the pre-adaptation to the adaptation condition, but would not predict any difference in the three IAI conditions since the frequency of the presented adaptors remained constant. Therefore, the significant difference in the number of /d/ responses from the 250 msec condition to the 750 msec condition is evidence against a response bias interpretation of selective adaptation.

CHAPTER VII: EXPERIMENT VI

COMPARISON OF ADAPTATION AND ANCHORING

Introduction

The purpose of this experiment was to investigate the effects of the selective adaptation and anchoring paradigms as originally conceived by Eimas and Corbit (1973) and Sawusch and Pisoni (1974), respectively, upon a fundamental frequency and stop-consonant continuum. While Experiment IV in this series, an adaptation experiment, demonstrated that manipulation of the experimental variables of the number of adaptor repetitions had a notable effect upon the magnitude of the phonetic boundary shift, Experiments I and II, anchoring experiments, demonstrated that no boundary shift occurred in the linguistic judgment, although a shift did occur in the non-linguistic judgment.

Two questions, previously posed in Experiment I, are immediately apparent. First, why does the boundary shift occur in the non-linguistic task and not in the linguistic task? Second, why does a shift in the phoneme boundary occur in the adaptation procedure and not in the anchoring procedure? The answers to these questions are far from obvious and comparisons of data cannot be made due to inherent differences in the paradigms and to differences between the stimuli and stimulus energy (the amount of signal reaching the receptor) involved.

The basic paradigm differences are easier to quantify than the stimulus energy differences. The most obvious difference between the paradigms is in stimulus density. Stimulus density may be considered in relation to the number of adaptor repetitions, the placement of the adaptor repetitions and the inter-adaptor-interval (IAI), all of which have an effect of reducing or increasing the time of an adaptation sequence and/or causing a reduction or enhancement of the stimulus energy impinging upon the receptor.

The first basic difference between the paradigms is the placement of the anchor or adaptor stimuli that relates to the density of the stimulus energy. Density refers to the amount of stimulus energy per time period. In these experiments, density can be manipulated by varying either the number of adaptor repetitions or the IAI. For example, 8 adaptors at a 250 msec IAI would have a greater density than the same number of adaptors at 750 msec. On the other hand, 8 anchors would produce less density at any IAI since they would be interwoven with the test syllables. The adaptation paradigm offers numerous rapidly repeating presentations of the adaptor immediately preceding the test syllables to be identified, while the anchoring paradigm presents the multiple adaptors interwoven with the test syllables. It is obvious that if the multiple adaptors are placed together preceding the test syllables, a summation of energy may be present that is not present when the same number of adaptors or anchors are randomized with the test syllables. Therefore, the anchoring paradigm did not effectively shift the phonetic boundary may simply be that the density of the anchors was not sufficient to cause an effect. The density effect may be related to the number of anchor repetitions presented, the IAI, as well as the placement of the anchors. If one assumes auditory and phonetic detectors, it is possible that the threshold of adaptation for the detectors governing adaptation at a phonetic level was not reached in Experiments I and II, although the detectors governing the frequency and intensity judgments were stimulated adequately. This density effect may also have been affected by the relatively long interval (two seconds) between the test syllables and the anchor syllables, as well as by the small number of adaptor repetitions. (Experiments I-III used only a 4:1 anchor to test syllable ratio, while Experiments IV and V used from 8 to 32 adaptors, a minimum of 8:1 ratio.) If, therefore, there is a threshold of adaptation that is reached by a certain stimulus energy reaching the receptor, it is possible that the former anchoring paradigm did not contain sufficient energy to reach threshold. Raising the number of anchor repetitions in this paradigm may overcome the longer IAI and low energy density if such a trade-off relationship exists.

The other differences between the two paradigms are displayed in Table 15. It can be seen from this table and other tables in Experiments IV and V that the number and rate of adaptor/anchor repetitions differ widely depending upon the experimental paradigm utilized. There were temporal differences not only between the test syllables and adaptors, but also in the different rest periods, the silences before and after a warning tone, and the use of the 250 msec warning tone itself. Since the amount of adaptation time per sequence may be involved in obtaining a boundary shift, it is important that the temporal variable be controlled.

It is also apparent that the total number of adaptor stimuli far exceeds the number of anchor stimuli. In fact in the pre-adaptation condition (1:1) where stimulus number 1 is interwoven with the other six syllables (making this condition more like the anchor paradigm), there are more stimuli than even in the 4:1 anchor condition. The reason for this is that the test sequence was repeated 25 times in adaptation and only 10 times in anchoring. Also the test syllables in the adaptation sequence were taken from stimuli 1-7, thus repeating stimulus number 1 an additional 25 times. Therefore, there were really 9:1, 17:1, and 33:1 ratios in the adaptation paradigm, if one adds the additional test syllables, stimuli 1, as adaptors. In the anchoring paradigm, only stimuli 2-7 were used as test syllables, truly creating a 4:1 ratio. Other paradigm differences are related to subjects' response (to all presentations in anchoring and only test syllables in adaptation) and the presence or absence of a warning tone.

In summary, these differences between anchoring and adaptation make comparisons of the results of the previous studies difficult. The following experiment again uses a correlated fundamental frequency and stop-consonant continuum, but makes the paradigm (anchor vs. adaptation) a main variable.

It is hoped that using a correlated fundamental frequency and stop consonant continuum that differ only in the placement of the anchor and adaptor stimuli will answer some of the previously posed questions regarding the lack of phonetic boundary shifts in the anchoring paradigm. If the revised paradigms both achieve a phonetic as well as a fundamental frequency boundary shift, it is possible that anchoring and selective adaptation should not be regarded as two distinctly different paradigms. Furthermore, if the two paradigms only differ in the degree of their effectiveness, is anchoring to be explained as adaptation or adaptation as anchoring? If anchoring is to be explained as adaptation, we can readily explain the shifts in the consonant boundary due to sensory fatigue of a detector mechanism. However, it is doubtful that feature detectors for fundamental frequency exist. On the other hand, if adaptation is to be explained by anchoring, the fundamental frequency boundary shifts can easily be explained on the basis of response organization or response bias. However, this would again leave feature detecting mechanisms in question for stop consonants. Therefore, if it can be shown that similar effects can be achieved on continuously perceived variables and on categorically perceived variables, the claim that adaptation effect reflect specialized detectors functioning to extract the relevant information from the acoustic signal in speech perception may be untenable.

Method

Description of Stimuli

The stimuli used in these experiments were the same series of seven three-formant CV syllables previously identified as /bæ-dæ/ in Experiment I but with all the syllables of 250 msec duration. As in Experiment I, the stimuli differed from one another only in the starting frequencies and direction of the second- and third-formant frequency transitions and in the fundamental frequency. The correlated fundamental frequency stop-consonant continuum is described in Table 1. The anchor and adaptation stimuli were the same in both Experiments VI and VII: Stimulus 1 at 114 Hz F_0.

Adaptation Experimental Tapes

Pre-adaptation and adaptation tapes were prepared for this experiment. The pre-adaptation tape consisted of the seven stimuli recorded 10 times each and randomized (a total of 70 test syllables). There was a 1750 msec interval between stimuli and a 5 second pause before every 21st stimulus.

The adaptation tape consisted of the specific number of adaptors (4 and 16) followed by the 6 test syllables. There was a consistent 1750 msec interval between both the adaptors and the test syllables with a 5 second pause after every 20th stimulus. Each adaptation sequence (4 or 16 adaptor repetitions) was recorded 10 times (a total of 100 stimuli in the 4:1 condition and 220 stimuli in the 16:1 condition). Two copies of these adaptation and pre-adaptation tapes were made directly from the PCM system and recorded on magnetic tape at 3 3/4 inches per second on a Crown 800 tape recorder.

Anchoring Experimental Tapes

Pre-anchor and anchor tapes were prepared from the same set of stimuli as in the adaptation experiment. The pre-anchor tape consisted of the seven stimuli recorded 10 times each in a randomized test order (a total of 70 test stimuli). As in the pre-adaptation tape, there was a 1750 msec interval between stimuli and a 5 second pause after every

21st stimulus.

In the anchor tape, the anchor stimulus was recorded 40 (4:1) or 160 (16:1) times and randomized with the other 6 stimuli that occurred 10 times each. A total of 100 stimuli were produced in the 4:1 condition and 220 stimuli in the 16:1 condition. The stimuli were recorded singly with a 1750 msec interval between each stimulus and a 5 second pause after every 20th stimulus. Two copies of each tape were made as previously described.

Procedure

The procedure followed was identical in both the adaptation and anchor experiments. All experimental tapes were reproduced binaurally from the output of an Ampex AG 500 tape recorder over calibrated Telephonic (TDH-39) matched headphones with a circumaural seal and presented to subjects at 75 dB SPL re 0.0002 dyne/cm^2.

Subjects were required to identify all the stimuli, both adaptors (anchors) and test syllables, under two conditions (pre-adaptation/pre-anchor and adaptation/anchor) for each of the two task conditions (identification of the fundamental frequency and identification of the phoneme) at each ratio (4:1 and 16:1). Each subject heard the same pre-adaptation/pre-anchor tape and the same adaptation /anchor tapes for each of the two task conditions. Only the directions to the subjects varied as the task changed. As in Experiments I through III, an identical test order was maintained: fundamental frequency pre-adaptation/pre-anchor, fundamental frequency adaptation/anchor, phoneme pre-adaptation/pre-anchor and phoneme adaptation/anchor presented at each of the two ratios. The pitch task was presented first, as in Experiment I, so that if there was an absence of an anchoring effect in the phoneme task, it would have occurred in spite of the expected anchoring effect in the pitch task and in spite of the correlation between fundamental frequency and phoneme category.

The standard set of instructions, read aloud, alerted the subjects to respond to each of the test stimuli presented with "high" or "low" for the pitch condition and "b" or "d" for the phoneme condition. Each experimental tape was preceded by listening and practice tapes in the pre-adaptation/pre anchor condition. Subjects were tested singly or in pairs, and a minimum of 24 hours and a maximum of a week separated experimental conditions. All subjects received the adaptation sequence before the anchor sequence, while the order of presentation of ratio was varied systematically within the paradigms.

Subjects

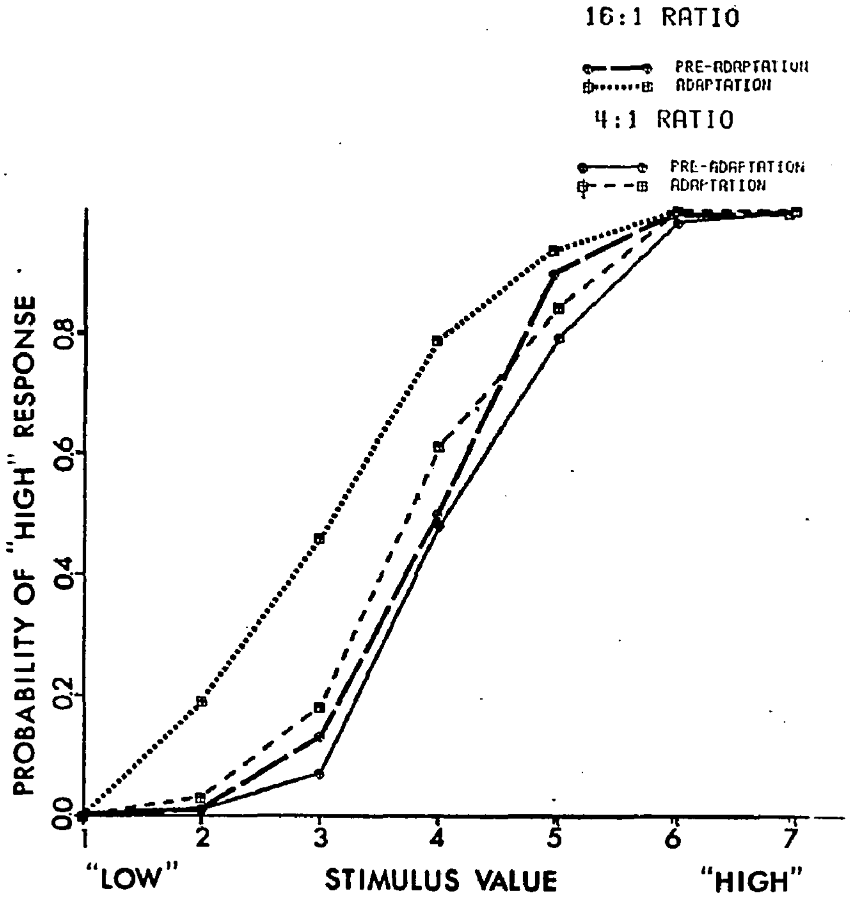
Nine listeners, between 18 and 28 years of age and meeting the previously stated criteria for subjects, participated in the experiments.

Results

The results of the experiment, averaged over all subjects, are shown in Figure 18 for the adaptation paradigm and Figure 19 for the anchor paradigm. The abscissa refers to the stimulus number and the ordinate to the probability of a /d/ or "high" pitch response.

Figure 18. Average identification functions for stop consonants and fundamental frequency for the 4:1 and 16:1 adaptation conditions.

FUNDAMENTAL FREQUENCY ADAPTATION



PHONEME ADAPTATION

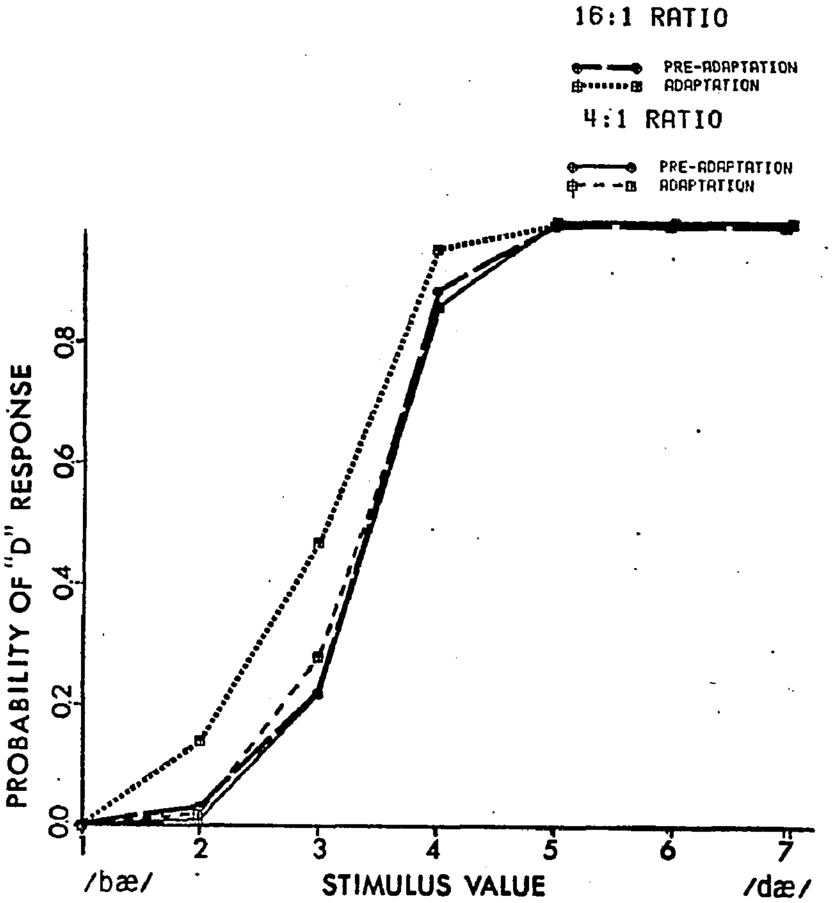
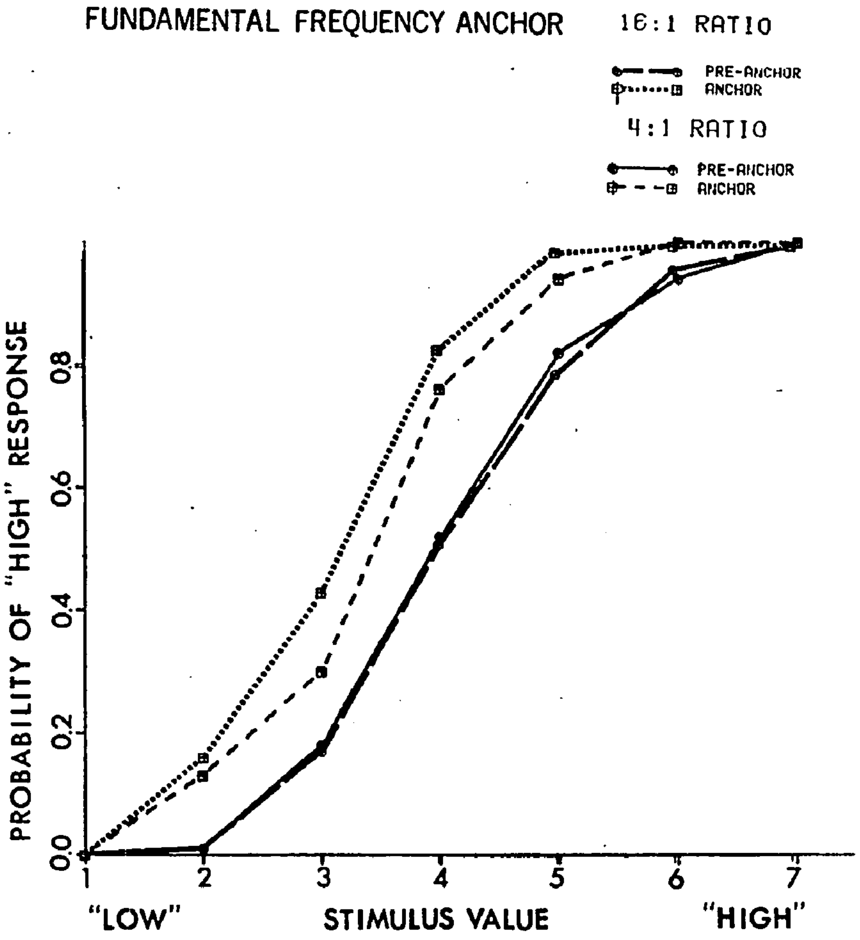
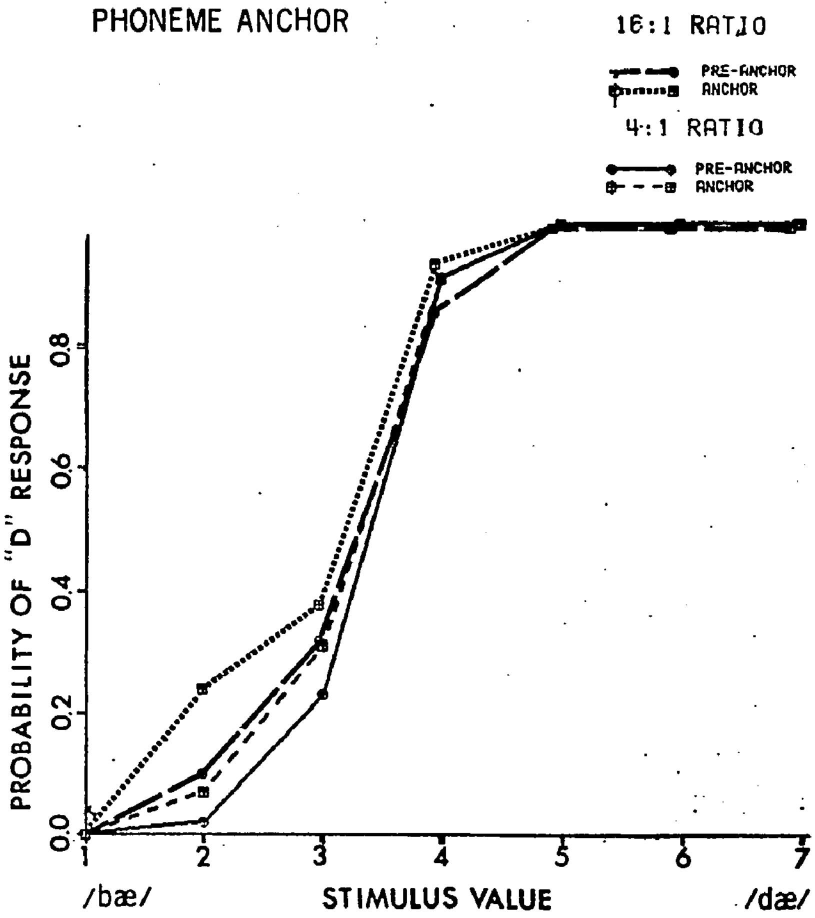


Figure 19. Average identification functions for stop consonants and fundamental frequency for the 4:1 and 16:1 anchor conditions.


FUNDAMENTAL FREQUENCY ANCHOR
16:1 RATIO
PRE-ANCHOR
ANCHOR
4:1 RATIO
PRE-ANCHOR
ANCHOR
PROBABILITY OF "HIGH" RESPONSE
0.0
0.2
0.4
0.6
0.8
1
2
3
4
5
6
7
"LOW"
STIMULUS VALUE
"HIGH"


PHONEME ANCHOR
16:1 RATIO
PRE-ANCHOR
ANCHOR
4:1 RATIO
PRE-ANCHOR
ANCHOR
PROBABILITY OF "D" RESPONSE
0.0
0.2
0.4
0.6
0.8
1
2
3
4
5
6
7
/bæ/
STIMULUS VALUE
/dæ/

Inspection of both the fundamental frequency and phoneme functions in both paradigms at each ratio indicates that the subjects partitioned the stimulus continuum into relatively distinct categories with slightly sharper functions for the phoneme judgment than for the pitch. Contextual effects were not apparent in either function. The fundamental frequency identification functions for both paradigms at both ratios show a consistent shift towards the anchor (Stimulus number 1) relative to the pre-adaptation or pre-anchor condition. The phoneme identification functions, however, show almost no shift of the anchor and adaptation task judgments relative to the pre-anchor and pre-adaptation judgment in the 4:1 condition. Slight shifts of the phoneme boundary for the anchor and adaptation task judgments do occur in the 16:1 condition. Tables 21-22 show the total number of /d/ and "high" responses for each of the conditions for each subject for stimuli 2-7 before and after adaptation.

Group probit analyses were done on the identification functions in this experiment. The slopes of these functions were steeper for the consonant than for the fundamental frequency in both anchoring and adaptation. However, in the adaptation and anchor conditions, there was a reduction in the slope for the consonant in comparison to the pre-anchor and pre-adaptation conditions, while in the fundamental frequency condition the reduction in slope only occurred for the adaptation condition.

A four-way analysis of variance (Table 23) was carried out: ratio (1:1, 4:1, 16:1), by paradigm (adaptation vs. anchoring) by continuum (fundamental frequency vs. phoneme) by subjects. Since there were two

Table 21. Number of "high" and /d/ responses before and after adaptation for the 4:1 and 16:1 ratios.

Number of "high" and /d/ responses before
and after adaptation for the 4:1 and 16:1 ratios.

			4:1 Ratio			
	F_0			SPEECH		
Subject	Pre-Adaptation	Adaptation	Difference	Pre-Adaptation	Adaptation	Difference
NK	35	33	2	48	49	- 1
GG	34	34	0	39	37	2
KH	37	41	- 4	39	40	- 1
KS	32	40	- 8	41	42	- 1
MR	36	40	- 4	37	37	0
DHN	37	38	- 1	46	47	- 1
BT	28	31	- 3	38	41	- 3
VH	26	31	- 5	44	46	- 2
DHS	34	42	- 8	36	35	1
MEAN	33.2	36.6	-3.4	40.9	41.6	- .67
			16:1 Ratio			
NK	35	43	- 8	46	51	- 5
GG	37	55	-18	40	40	0
KH	35	43	- 8	39	43	- 4
KS	41	55	-14	43	46	- 3
MR	40	44	- 4	36	36	0
DHN	35	33	2	40	49	- 9
BT	33	47	-14	40	49	- 9
VH	25	30	- 5	51	52	- 1
DHS	38	44	- 6	36	45	- 9
MEAN	35.4	43.8	-8.4	41.2	45.7	-4.5

Table 22. Number of "high" and /d/ responses before and after anchoring for the 4:1 and 16:1 ratios.

Number of "high" and /d/ responses before and after anchoring for the 4:1 and 16:1 ratios.

	F_0			SPEECH		
Subject	Pre-Anchor	Anchor	Difference	Pre-Anchor	Anchor	Difference
			4:1 Ratio			
NK	38	43	- 5	47	52	- 5
GG	44	49	- 5	39	40	- 1
KH	37	41	- 4	40	40	0
KS	43	48	- 5	43	46	- 3
MR	35	41	- 6	35	34	1
DHN	34	39	- 5	41	40	1
BT	28	42	-14	40	45	- 5
VH	22	28	- 6	51	49	2
DHS	32	41	- 9	39	39	0
MEAN	34.8	41.3	-6.5	41.7	43	-1.3
			16:1 Ratio			
NK	37	51	-14	50	57	- 7
GG	41	51	-10	42	42	0
KH	37	50	-13	39	44	- 5
KS	33	44	-11	41	40	1
MR	39	44	- 5	33	38	- 5
DHN	31	38	- 7	46	40	6
BT	33	46	-13	42	54	-12
VH	24	33	- 9	56	58	- 2
DHS	34	40	- 6	36	38	- 2
MEAN	34.3	44.1	-9.8	42.8	45.7	-2.9

Table 23. Summary table of analysis of variance for the total number of /d/ and "high" responses.

Summary table of analysis of variance for the total number of /d/ and "high" responses. Ratio (R): 1:1, 4:1, 16:1; Paradigm (P): Anchor, Adaptation; Continua (C): F_0, Phoneme; Subjects (S).

Source	df	MS	F	p
Subject	8	56.29		
R	2	395.17	24.78	< 0.01*
RxS	16	15.94		
P	1	41.56	2.09	NS
PxS	8	19.89		
C	1	452.23	2.05	NS
CxS	8	220.04		
RxP	2	16.78	1.69	NS
RxPxS	16	9.93		
RxC	2	71.95	7.32	< 0.01*
RxCxS	16	9.82		
PxC	1	3.34	0.42	NS
PxCxS	8	7.86		
RxPxC	2	9.23	1.24	NS
RxPxCxS	16	7.43		
TOTAL	107			

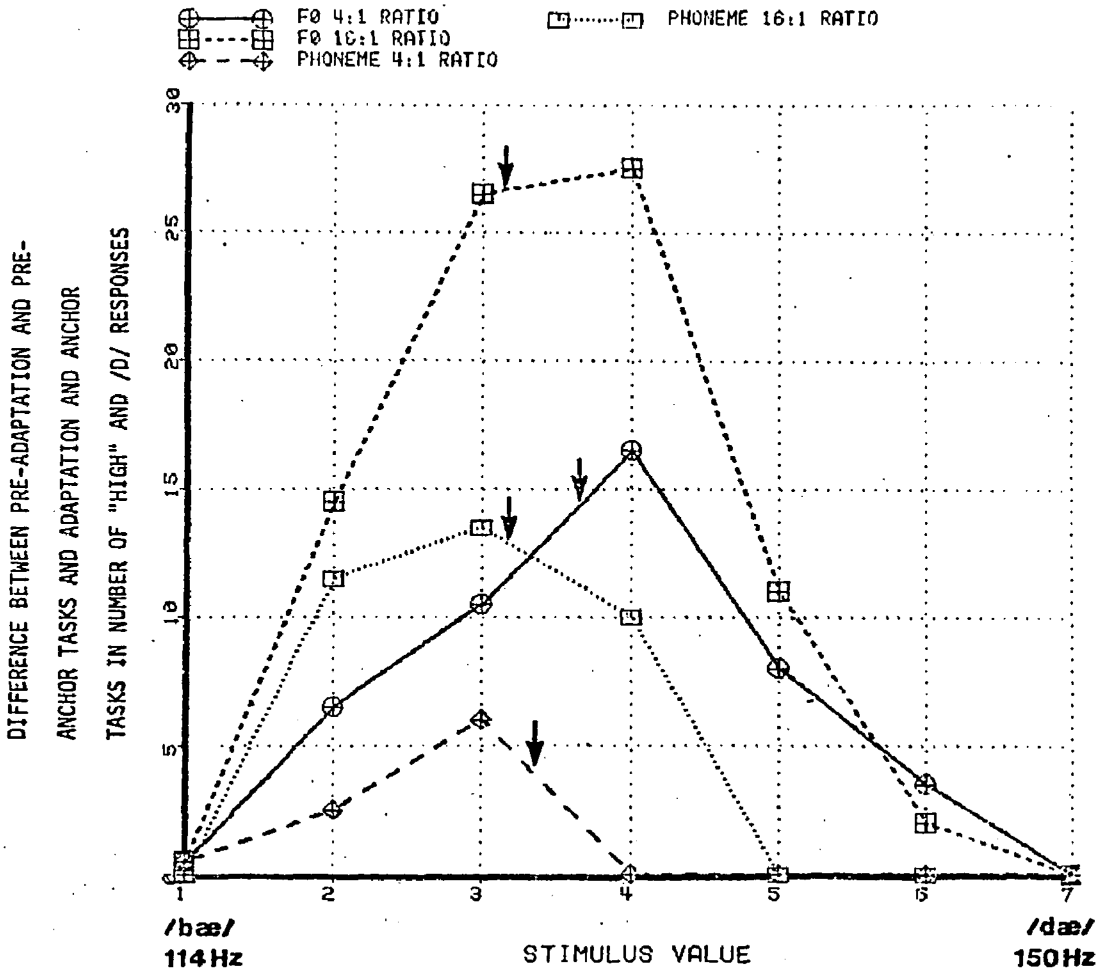
pre-adaptation tests for each paradigm (one before each ratio condition) only the mean of these tests was used. The ratio factor (R) was highly significant, as in the other experiments ($F_{2,16} = 24.78$; $p < 0.01$). This factor assesses the statistical significance of the overall adaptation or anchoring effects. The continuum (C) factor was not significant: there was no overall difference in the effects on the fundamental frequency and the phoneme identification tasks. However, the ratio x continuum interaction (R x C) was significant ($F_{2,16} = 7.32$; $p < 0.01$), evidently because the fundamental frequency task produced the larger shift for each ratio on both paradigms.

The main effect of Paradigm (P) was not significant: Table 23 reveals a similar pattern of ratio and continuum effects for both paradigms. This most interesting lack of a significant difference between these two treatments suggests that the two paradigms are essentially equivalent when all variables, except for the placement of the anchors and adaptors, are matched. In other words, the two designs may engage similar underlying processes when presented at a large enough adaptor or anchor ratio.

Figure 20 shows a reinterpretation of the graphed data. Here the pre-anchor and pre-adaptation conditions are represented as 0 and the anchored and adapted functions are averaged and plotted as the difference between the pre-adaptation and pre-anchor responses and the adaptation and anchor responses. In this figure, the greatest distance from the abscissa (0 point) is for the largest ratio (16:1) in the fundamental frequency condition. The two phoneme conditions show the least difference. In the fundamental frequency conditions, the data peak

Figure 20. Reinterpretation of the data shown in Figures 18 and 19.

REINTERPRETATION OF DATA
SHOWN IN FIGURES 18 AND 19



between stimulus numbers 3 and 4 and at stimulus number 3 for the phoneme, indicating that the greatest shift in the anchor and adaptation conditions occurs there.

The greatest shift is closer to the anchor or adaptor stimulus for the phoneme condition at both ratios than for the fundamental frequency conditions. In other words, the greatest boundary shift for the fundamental frequency occurs at the mid-point of the continuum (between stimulus 3 and 4) while for the phoneme the peak is to the left of the mid-point. There is a symmetry in the fundamental frequency functions not seen in the consonant functions.

The phoneme boundaries are shown by the arrows superimposed on Figure 20. These arrows show that while the boundaries in the fundamental frequency conditions are superimposed on the peak or to the left of the peak (towards the anchor or adaptor), the boundaries for the phoneme conditions are on the other side of the peak towards the non-adapted end of the continuum.

Figures 21 and 22 illustrate the category boundaries as a function of each continuum. The mean of the four pre-adaptation/anchor tests for each continuum was used as the control against which to judge the boundary shift for the four adaptation/anchor tests. A comparison of these figures shows that the pre-adaptation/pre-anchor boundary in the phoneme condition is almost one step closer to the anchor/adaptation stimulus (stimulus number one) than is the fundamental frequency boundary. Also, the adaptation/anchor boundaries are much closer together and show less of a shift in the phoneme condition than in the

Figure 21. Average identification functions for fundamental frequency in all the ratio/ paradigm conditions.

AVERAGE IDENTIFICATION FUNCTIONS FOR FUNDAMENTAL FREQUENCY IN ALL THE RATIO/PARADIGM CONDITIONS

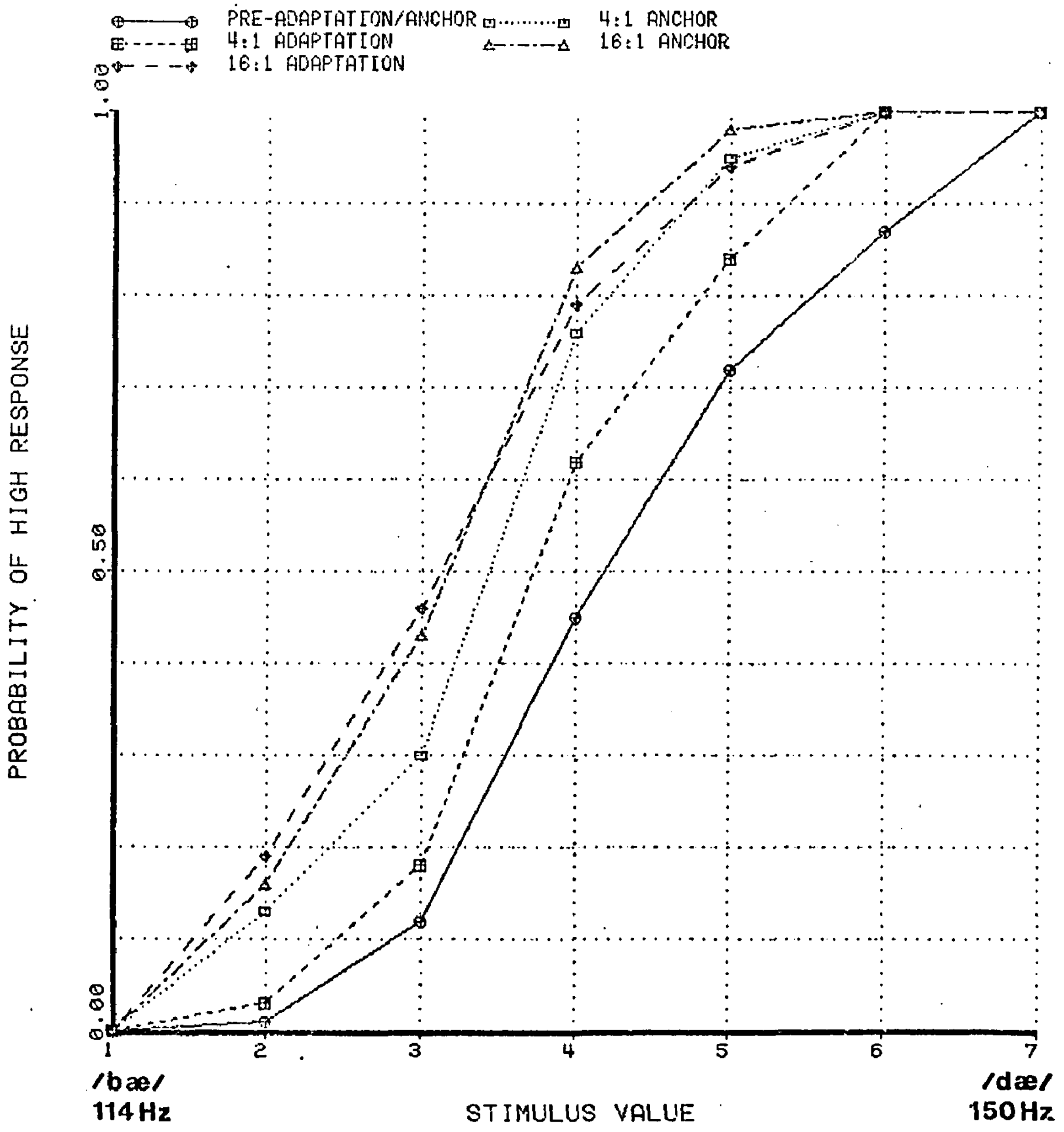
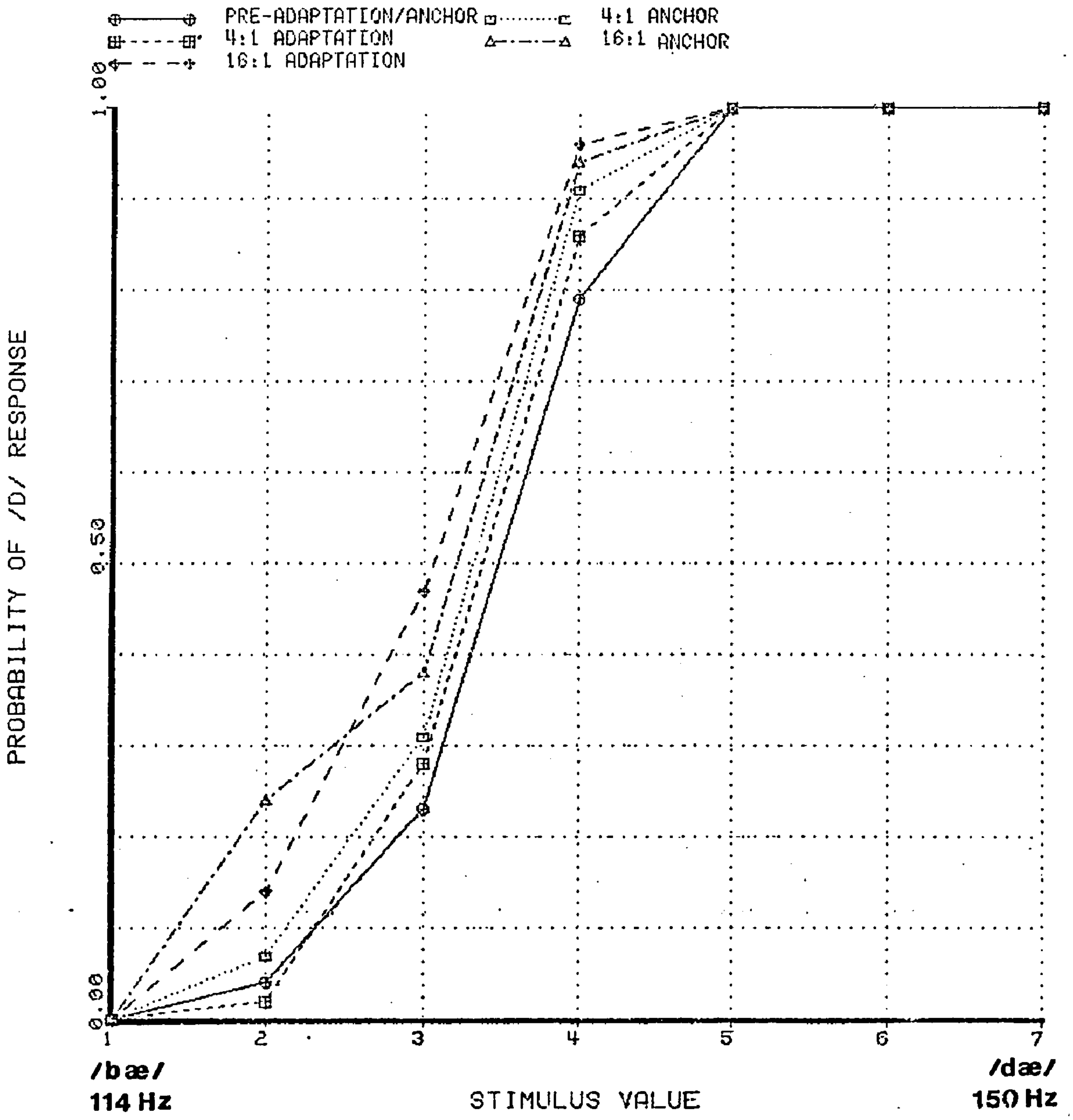


Figure 22. Average identification functions for stop consonants in all the ratio/paradigm conditions.

AVERAGE IDENTIFICATION FUNCTIONS FOR STOP
CONSONANTS IN ALL THE RATIO/PARADIGM CONDITIONS



fundamental frequency condition. For both paradigms and continua, the 16:1 ratios show the greatest shift (approximately one-step for the fundamental frequency and one-half step for the phoneme).

Summary and Conclusions

The following general observations concerning the anchoring and selective adaptation of fundamental frequency and stop consonants can be made.

1. The main effect of ratio was significant, as in the other experiments, showing an overall effect of both adaptation and anchoring. Since there was no significant paradigm effect in this experiment, we may conclude that both paradigms gave rise to the same pattern of effects. Experiments III and VI, therefore, demonstrate that a consonant continuum is subject to manipulation by ratio variation in anchoring, while Experiments IV, V, and VI demonstrate the effects of ratio variation in selective adaptation.

2. There was no overall continuum effect in this experiment. There was, however, a significant ratio x continuum interaction, and Figures 18-22 show that there is, in fact, a consistently greater ratio effect for fundamental frequency in both paradigms. Therefore, it may be concluded that pitch identification and consonant identification do not differ in kind, as Sawusch, Pisoni, and Cutting (1974) supposed, but in degree.

There may be two interrelated explanations for this difference in degree. First, it may be that the anchoring paradigm as presently used, invokes a response bias or response organization effect, similar to that found in certain psychophysical judgments (Helson, 1964; Parducci, 1974). Second, perceptual categories in speech are not arbitrary: they are constrained, within limits, by the physical structure of the signal and so cannot be readily shifted by external conditions.

The rigidity and non-arbitrary nature of the categorically perceived variable relative to the continuously perceived variable can be shown in this experiment, not only by the greater boundary shift for the fundamental frequency continua, but also by the distinctive patterns in the identification functions (Figures 18 and 19) and in the reinterpretation of the functions (Figure 20). All of the figures show that the boundaries of the two continua differ. The boundaries of the fundamental frequency functions are approximately at the midpoint of the continua (between stimulus 3 and 4), while the boundaries for the phoneme functions are to the left of the midpoint (stimulus 3). The boundaries for the fundamental frequency continua are also to the left of the peak (the point of the greatest shift from the pre-anchor/adaptation condition to the anchor/adaptation condition) as seen in Figure 20.

In other words, the fundamental frequency functions show a symmetrical distribution around the middle of the range, while the consonant functions show a slightly asymmetrical distribution. Therefore, the categories for the consonants are determinedby the individual stimulus properties, while the categories for the fundamental frequencies are determined by the stimulus range.

3. The difference between anchoring and adaptation paradigms is primarily, if not entirely, based on stimulus density, that is, on the distribution of "hit" energy. Hit energy refers to the energy of either an adaptor or an anchor. Density is based in part on the placement, as well as upon the ratio of the adaptors and anchors relative to the test syllables. As previously explained, the only difference between the paradigms as they were used in this experiment, was in the placement of the anchors and adaptors. In the adaptation paradigm, the adaptors were placed together before the test stimuli to be judged, while the anchors were randomized within the test syllable sequence. There were 1750 msec between test items, between adaptors, and between anchors. Subjects were instructed to respond to each anchor or adaptor as well as to each test item.

It is possible that anchors interwoven within the test series are not as effective in adapting the sensory system as a comparable number of adaptors placed together at higher repetition rates. It seems probable that by increasing the number of anchors within the test series, the two paradigms would be equalized.

When anchor density is increased in these experiments, a greater anchoring effect is achieved, illustrating the equalization of the paradigms. For example, Figures 20 -22 show that there is a greater shift in the consonant boundaries for anchors and adaptors at the 16:1 ratio than at the 4:1 ratio, in this experiment. Therefore, the density of the adaptor and anchor stimuli is shown as mediating the differences between this experiment and Experiments I and II. In those experiments, only 4 anchor stimuli were interwoven with the test syllables. The categoricity of speech stimuli may have made them immune to the effects of anchoring at this low ratio as it did at the 4:1 ratio in this experiment. However, when the anchors are increased to 16, as many as 10 anchors could be consecutively repeated. Although this did not happen in the randomization, it was shown in Experiment IV, that as few as 8 adaptors would produce a shift in the category boundary. Therefore, it may not be that in the strict sense, anchoring and adaptation are two different paradigms. However, as the number of anchors is increased, and total "hit" energy increases, the anchoring paradigm becomes one of selective adaptation. Therefore, the two paradigms may differ in degree in terms of available energy, rather than in kind.

As seen in Experiment V, the 1750 msec IAI did not produce the greatest boundary shift for the 16 adaptors due either to decay of adaptation during the presentation of the adaptors or to weak adaptation because of an inadequate summation of energy. Therefore, another reason that the anchoring paradigm in Experiments I, II, and VI did not produce as great an effect is due to decay of "adaptation" between presentations that were 1750 msec apart. It is possible, as seen in Experiment V, that the same 16 anchors used in this experiment would be more effective with a 750 msec IAI or that using a greater number of anchors at the 1750 msec IAI would have produced a greater effect. Therefore, in anchoring with 16 anchors, not enough stimulus energy is available to fatigue the detector.

On the other hand, it may be argued that anchoring and adaptation do not engage the same underlying mechanism. A comparison of Table 21 in this experiment and Table 16 in Experiment IV shows that the mean difference between the pre-adaptation and adaptation conditions in the two experiments were different at the 16:1 ratio. In Experiment IV, the mean difference was 12.78 while in the present experiment it was 4.5. This difference between the two experiments could be explained by the fact that the IAI in the present experiment was 1750 msec while in Experiment IV the IAI was 250 msec. However, in Experiment V, the mean differences for the 250 msec and 1750 msec conditions were 8.0 and 9.4 respectively. This difference was not significant, so one would not expect as great a difference between Experiment IV and Experiment VI based on the IAI. Therefore, while the longer IAI in Experiment VI caused some decay of adaptation, the comparison with the results at the same IAI in Experiment V is puzzling. For some reason there was weak adaptation at the 16:1 ratio in this experiment. The fact that the consonant did not give a strong effect in either adaptation or anchoring in this experiment explains the lack of a paradigm x continuum interaction.

Therefore, in spite of the fact that there is no overall paradigm effect, it cannot be argued on the basis of these results that there is no difference in the paradigms. The difference between the paradigms may be concealed by the failure to get a normal strong adaptation effect for the consonant. It cannot be argued that consonants do not give strong adaptation effects, for the literature cited in Chapter I gives many examples of strong consonant effects in selective adaptation. The weak consonant adaptation effect in this experiment may have masked the paradigm effect. Therefore, it cannot be concluded that anchoring and selective adaptation are the same processes.

The conclusion to be drawn from Experiment VI is then, that anchoring and adaptation do not have the same effect at low adaptation (or anchoring) to test item ratios. However, if the "hit" energy in anchoring is increased, especially for the consonant task, the same underlying processes may be engaged. Nonetheless, the consonants are less susceptible to the effects of anchoring at the 16:1 ratio than is fundamental frequency, and are immune to the effects of anchoring at a 4:1 ratio; they are also less susceptible than fundamental frequency in adaptation.

CHAPTER VIII: FINAL DISCUSSION

Research with stop consonants has shown that they are perceived categorically (Liberman, et al., 1967). Furthermore, the phonetic categories involved are natural rather than arbitrary, since they reflect a "range of patterns that the articulatory apparatus can produce and the auditory system can analyze" (Studdert-Kennedy, 1976, p. 17). Although there is now ample evidence for the categorical perception of certain non-speech stimuli (Locke and Kellar, 1973; Cutting and Rosner, 1974; Miller, Pastore, Wier, Kelly, and Dooling, 1974; Cutting and Rosner, 1976; Cutting, Rosner, and Foard, 1976; and Miller, Wier, Pastore, Kelly and Dooling, 1976), the fact that stop consonants are perceived categorically is still of importance, since rapid and unequivocal classification of speech sounds is probably a condition of efficient perception.

If, therefore, consonants are not arbitrarily named, and if they are perceived absolutely rather than relatively, they should also be immune to context effects, whether these effects are response biases or perceptual biases. It was in this belief that Eimas and Corbit (1973) and Eimas, Corbit, and Cooper (1973) pursued a physiological rather than a psychological account of adaptation, and that Sawusch and Pisoni (1973) and Sawusch, Pisoni, and Cutting (1974) argued that consonants were immune to anchoring.

However, the results of two experiments in this series, Experiments III and VI, have shown that consonant continua are not immune to anchoring. Although the consonant boundary showed no anchoring effects at a 4:1 ratio in Experiments I, II, and VI, it did shift in Experiment III at a 4:1 ratio and in Experiment VI at a 16:1 ratio. Therefore, although a consonant is less susceptible to anchoring than, for example, a pure tone, it is evidently not immune, if certain experimental conditions are met.

These conditions may require that the consonant be part of a CV syllable where both phonemes must be processed for a judgment to be made about each of them and/or that the total anchor energy in the anchoring paradigm be above some threshold sufficient to cause a shift in the boundary. That threshold may be higher for consonants than for non-speech stimuli.

The Effect of Parallel Processing on Anchoring

Experiment III was the only one in which both the consonant and the vowel were varied. Although there was no correlation between the consonant and vowel boundary shifts of individual subjects, this does not mean that acoustic information in the following vowel was not used to process the consonant information nor vice versa. In fact, many studies attest to the parallel processing of all portions of a syllable. (See, for example, Harris, 1958; Liberman, et al., 1967; Lindblom and Studdert-Kennedy, 1967; Strange, Verbrugge and Shankweiler, 1976; Dorman, Studdert-Kennedy, and Raphael, In press).

The contingent adaptation effects, discussed in Chapter I, may also reflect the processing of whole syllables and have, in fact, been attributed by some authors to vowel-dependent consonant feature channels (Cooper, 1974c; Miller and Eimas, 1976).

Since we know that synthetic vowel series are more subject to context effects than are consonantal series (Fry, et. al., 1962), and since these effects can only be evidenced when the test contains more than one vowel, it is not unreasonable to suppose that the relative increase in the consonant's sensitivity to anchoring in Experiment III (4:1 ratio) as against Experiment VI (where a 16:1 ratio was needed for the effect) is contingent on a shift in the vowel boundary.

The Effect of Total Anchor Energy on Anchoring

Despite the parallel processing of consonant and vowel, there still may be some effect of signal duration in anchoring and adaptation: anchoring effects may be directly proportional to the energy in the anchoring signal. Since the syllable onset typically has less energy than the syllable nucleus, energy accumulation should be slower for the onset than for the nucleus, and the consonant should be more resistant to boundary shifts than the vowel.

Categorical and Continuously Perceived Variables

Not only do anchoring and adaptation yield the same effects, but they yield them for consonants, vowels, fundamental frequency, and intensity, only to different degrees. Consonants have been shown to be less susceptible than fundamental frequency in both anchoring and adaptation.

Evidence for the relative stability of consonant boundaries can be seen in this series of experiments. In addition to the greater boundary shift for the fundamental frequency in Experiment VI, for example, there is a difference in the placement of the boundary and the point of the peak effect. The boundaries for the fundamental frequency functions are approximately at the midpoint of the range of stimuli while the boundaries for the phoneme functions are to the left of the mid-point. In other words, the fundamental frequency functions shows a symmetrical distribution around the middle of the range, consistent with a standard psychophysical distribution, while the phoneme shows an asymmetrical distribution. Therefore, while there is some shift of the consonant boundary with adaptation and anchoring, this shift is limited by the non-arbitrary nature of the stimulus. However, while there are limits on the boundary movement of the consonant, there is no reason to suppose that the failure to get anchoring effects in Experiments I and II is totally due to the rigidity of the speech categories. It is evidently also due to the low level of stimulus energy with only 4 anchors.

The question arises, therefore, whether the boundary shifts for fundamental frequency can be accounted for by the fatiguing of detectors. It seems unlikely that any process as specialized as "feature analyzing" devices or "feature detectors" would be reflected if adaptation and anchoring engage the same underlying processes, and, if these are engaged only somewhat more readily for pitch or intensity, than for consonant identification.

However, selective adaptation in vision offers numerous examples of adaptation of non-linguistic features (see Chapter 1 and Eimas and

Miller, In Press, for a review) and recent studies show adaptation for non-linguistic sounds (Kay and Matthews, 1972; Cutting, Rosner, and Foard, 1976).

Kay and Matthews (1972) have suggested that adaptation of FM tones is evidence for auditory pathway "channels" that are predominantly central in location. However, they found that these frequency-modulated channels are independent of the modulation frequency, the signal spectrum, and the mean "carrier" frequency. Also, the finding that the carrier frequency is relatively unimportant to the effect has reference to the independence of the important aspects of the speech signal (e.g. the frequency modulation of the vocal tract) from the unimportant aspects of the signal (e.g. the fundamental frequency of the vocal fold vibration) (Kay and Matthews, 1972).

The adapted stimuli in the Kay and Matthews (1972) experiment, therefore, were frequency-modulated tones concerned with instantaneous frequency changes, but independent of periodicity, either in the modulation frequency or in the carrier frequency. Cutting, Rosner, and Foard (1976) on the other hand, found adaptation for the steady-state aspects of sawtooth and sinusoidal waveforms. Information about the rise-time did not influence adaptation as much as did the number of parameters shared between the adaptors and test stimuli. Cutting, Rosner, and Foard (1976) suggest that "feature detecting systems" exist in humans for non-linguistic sounds and that prior information concerning the waveform and the frequency of the stimuli is necessary in order for the feature detecting mechanism to make a decision about the categorical nature of their "pluck and bow" stimuli. Therefore, the greatest adaptation effects in Experiment VI, in this series, and in the Cutting, Rosner, and Foard (1976) experiment were for steady-state information, whether aperiodic or periodic. Information predominantly of short duration, whether syllable onset information or rise-time information, showed less adaptation than did relatively steady-state information such as fundamental frequency. The effect of signal time in terms of total stimulus energy may also be relevant in the non-speech adaptation of rise-time. However, Sawusch (1976) found that the 45 msec non-speech chirp adaptors produced almost as large an adaptation effect as the full syllable.

It should not be assumed, however, that all the non-speech sounds above are processed the same way nor that a positive adaptation effect is direct evidence for feature-detecting mechanisms for all or any non-speech sounds. There is evidence from studies in auditory physiology that sounds that involve temporal change such as the frequency-modulated tones, are processed differently from steady-state sounds such as fundamental frequency. Sinusoidal information is analyzed in the cochlear along the basilar membrane according to frequency (von Bekesy, 1960). This spatial organization of frequency is preserved at least to the level of the inferior colliculus (Whitfield, 1967). Beyond this level, the tonotopic arrangement does not seem to hold, and frequency discrimination is apparently not a function of the auditory cortex (Whitfield and Evans, 1965). However, discrimination of frequency change with time needs at least part of the auditory cortex (Evans, Ross, and Whitfield, 1965; Whitfield and Evans, 1965).

While there is evidence that there is adaptation of all the non-speech stimuli discussed, there is no way of establishing that feature detectors are the mechanisms being adapted. Kay and Matthews (1972) caution against a complete comparison in modulated stimuli between the visual and auditory modalities due to the separable dimensions of temporal and spatial aspects of the stimuli in vision. In audition, these dimensions may not be separable.

Bilger (1973) also cautions against making analogies from vision to audition, since inhibition is a necessary correlate of a feature detecting system. There is no clear evidence of inhibition in the auditory system. Bilger (1973) further cautions against analogies from one species to another based on the variations known to exist in the structure of the auditory system. To an even greater degree than the anatomical variations, are those expected in the location and nature of the specific detectors (Michael, 1969). Bilger (1973) finally suggests that the appropriate experimental technique to isolate feature detectors in the human subject might be with the hard-of-hearing who manifest discrimination losses for speech. He states that a subject with normal perceptual and non-perceptual auditory functioning might not provide as much relevant information about the underlying structures as a subject with impaired processing.

If the results of this series of experiments are reliable indicators of the susceptibility of the phoneme boundary to contextual effects (vowel dependency and anchoring), the notion that adaptation or anchoring effects reflect the operation of fixed feature detectors in speech perception may be untenable. For we must either assume an infinite number of context-dependent detectors or some other generalized analyzer. The original formulation of a theory of feature detectors assumed that a particular detector would be tuned to a particular feature of the complex signal, regardless of the contextual information in which that feature was embedded. That theory, therefore, suggests a finite number of features for which a corresponding number of detectors should exist. However, if it is assumed that there are contextually dependent detectors, the number of features to which these detectors must respond is vastly enlarged and consequently calls for a much larger repertoire of detectors. Eimas and Miller (In press) have proposed this kind of detector, which they label an " object detector." This detector, according to Eimas and Miller (In press) would respond only to the appropriate value of all the dimensions of the objects. They admit that even if there is some limit placed on the number of contextual dimensions to which a specific detector is tuned, the number of detectors needed would still be large.

The Levels of Processing in Selective Adaptation

However, whatever the correct interpretation of these non-speech adaptation studies in terms of feature detectors, it is possible that adaptation is occuring at low auditory levels, responsive to low-level acoustic information such as transition information, steady-state information, frequency-specific information or simply generalized spectral energy information. Evidence for adaptation at the acoustic feature level has already been cited in Chapter I. However, these assumptions rely on one level of adaptation. It is possible that adaptation is operating at more than one level of perceptual analysis, as has been offered by Ades (1974b), Cooper (1974), Tartter and Eimas (1975) and Sawusch (1976).

The findings of this study (Experiments IV and V), that the low level stimulus energy available, the number of adaptor repetitions, had an effect on the magnitude of the phonetic boudary shifts, may be evidence for some level of selective adaptation operating at an auditory or acoustic stage of processing. Another study in selective adaptation by Hillenbrand (1975) also found systematically larger shifts in the phonetic boundary with greater intensity and with a larger number of adaptor repetitions.

However, as Hillenbrand (1975) points out, the effect may be ambiguous since increasing the number of adaptor repetitions, while increasing the amount of stimulus energy at an acoustic level, may also increase adaptation at a phonetic level of processing. It is possible that the repetition effect could be the result of the cumulative effect of the number of times a phonetic feature was presented.

The Importance of the Paradigm Variables in Selective Adaptation

The fact that the magnitude of the boundary shifts can be manipulated by changes in the low-level stimulus energy variables, the number of adaptor repetitions and the inter-adaptor-interval indicates that experiments in selective adaptation should also address themselves to these paradigm effects. It is possible that adapting at high intensities or with a high number of adaptor repetitions with a large IAI, without recognizing the limits of the system, may induce either subsequent fatigue or recovery of the mechanism, that may obliterate the effects of adaptation. Furthermore, if there is adaptation at a low auditory level, the paradigm itself becomes increasingly important in order to investigate these low level effects.

It is well known in the literature that auditory adaptation can be obtained at low intensity levels (close to threshold) while auditory fatigue is demonstrated at high intensity levels (Elliot and Fraser, 1972). Although Hillenbrand (1975) used low intensity levels (65dB SPL) to assess the strength of the adaptation effect, only Kay and Matthews (1972) consistently used comfortable listening levels (40dB above threshold) to assure that their stimuli were not physiologically damaging nor produced "after masking of intensity acuity" (Kay and Matthews, p. 674).

The lack of a cumulative effect in Experiment IV also has implication for adaptation testing. It was found that the maximum adaptation effect was obtained with only 40 sec of adaptor repetitions. Adaptation testing, therefore, can be minimized to avoid subject fatigue without sacrificing an adaptation effect.

Summary of Results

The overall conclusions to be drawn from this series of experiments are as follows:

1. The anchoring and adaptation paradigms, as used in this

experiment, do not differ in kind, but in degree in terms of available energy. As the number of "anchors" is increased, and the "hit" energy increases, the two paradigms may invoke the same underlying processes.

2. The effects of adaptation and anchoring on the categorically perceived variables differ in degree from those on the continuously perceived variables. Anchoring effects have been found with consonants when they are part of a CV syllable where both phonemes must be processed for judgment and when the total anchor energy is above some threshold sufficient to cause a shift in the phoneme boundary.

3. The findings that adaptation and anchoring have effects that are similar for consonants, vowels, fundamental frequency, and intensity call into question the notion that these paradigms reflect the operation of feature detecting devices in speech processing. Although a major purpose of adaptation studies has been to formulate an account of speech processing in terms of feature detecting mechanisms, the generality of adaptation effects must raise doubts. There seems to be no good reason to posit feature detectors for consonant adaptation and "something else" for fundamental frequency adaptation.

4. The low level stimulus energy variable in adaptation, Inter-Adaptor-Interval and the ratio of adaptor repetitions affect the magnitude of the phonetic boundary shifts in what may be a trading relationship.

BIBLIOGRAPHY

Abbs, J. H. and H. M. Sussman. (1971) Neurophysiological feature detectors and speech perception: A discussion of theoretical implications. J. Speech Hearing Res. *14*, 23-36.

Abramson, A. S. and L. Lisker. (1970) Discriminability along the voicing continuum: Cross-language tests. In *Proceedings of the VI International Congress of Phonetic Sciences*, Prague, 1967. (Prague: Academia) 569-573.

Ades, A. E. (1973) Some effects of adaptation of speech perception. Quarterly Progress Report (Research Laboratory of Electronics, MIT) *111*, 121-129.

Ades, A. E. (1974a) Bilateral component in speech perception? J. Acoust. Soc. Amer. *56*, 610-616.

Ades, A. E. (1974b) How phonetic is selective adaptation? Experiments on syllable position and vowel environment. Percept. Psychophys. *16*, 61-67.

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

Bailey, P. J. (1973) Perceptual adaptation for acoustical features in speech. *Speech Perception*, Progress Report. (Department of Psychology, The Queen's University of Belfast), *2.2*, 29-34.

Bailey, P. J. (1975) Perceptual adaptation in speech: Some properties of detectors for acoustical cues to phonetic distinctions. Unpublished Doctoral Dissertation, University of Cambridge.

Bieri, J., A. L. Atkins, S. Briar, R. L. Leaman, H. Miller, and T. Tripodi. (1966) *Clinical and Social Judgment: The Discrimination of Behavioral Information*. (New York: John Wiley and Sons, Inc.).

Bilger, R. C. (1973) Research Frontiers in Audiology. 2nd ed. In *Modern Developments in Audiology*, ed. by J. Jerger. (New York: Academic Press).

Blakemore, C. and F. W. Campbell. (1969) On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. J. Physiol. *203*, 237-260.

Blakemore, C. and P. Sutton. (1969) Size adaptation: A new aftereffect. Science, *166*, 245-247.

Blumstein, S. E. and K. N. Stevens. (1975) Property detectors for bursts and transitions in speech perception. Paper presented at the 89th meeting of the Acoustical Society of America, April 8-11, Austin, Texas.

Capranica, R. R. (1965) *The Evoked Response of the Bullfrog*. (Cambridge, Mass.: MIT Press).

Carterette, E. C. (1955) Perstimulatory auditory fatigue for continuous and interrupted noise. J. Acoust. Soc. Amer. *27*, 103-111.

Chomsky, N. and M. Halle. (1968) *The Sound Pattern of English*. (New York: Harper & Row).

Cole, R. A. and W. E. Cooper. (1975) Perception of voicing in English affricates and fricatives. J. Acoust. Soc. Amer. *58*, 1280-1287.

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

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

Cooper, F. S. and I. G. Mattingly. (1969) Computer-controlled PCM system for investigation of dichotic speech perception. Haskins Laboratories Status Report on Speech Research SR-17/18, 17-21.

Cooper, W. E. (1974a) Adaptation of phonetic feature analyzers for place of articulation. J. Acoust. Soc. Amer. 56, 617-627.

Cooper, W. E. (1974b) Selective adaptation for acoustic cues of voicing in initial stops. J. Phonetics. 2, 303-313.

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

Cooper, W. E. (1975) Selective adaptation to speech. In Cognitive Theory, vol. 1, ed. by F. Restle, R. M. Shiffrin, J. N. Castellan, H. Lindman, and D. B. Pisoni. (Hillsdale, N.J.: Erlbaum Associates), 23-54.

Cooper, W. E. and S. E. Blumstein. (1974) A "labial" feature analyzer in speech perception. Percept. Psychophys. 15, 591-600.

Cooper, W. E., R. R. Ebert, and R. A. Cole. (1976) Perceptual analysis of stop consonants and glides. J. Exp. Psychol.: Human Percept. Perform. 2, 92-104.

Cuddy, L. L., J. Pinn, and E. Simons. (1973) Anchor effects with biased probability of occurrence in absolute judgment of pitch. J. Exp. Psychol. 100, 218-220.

Cutting, J. E. (1976) Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening. Psychol. Rev. 83, 114-140.

Cutting, J. E., and R. S. Day. (1975) The perception of stop-liquid clusters in phonological fusion. J. Phonetics, 3, 99-113.

Cutting, J. E., B. S. Rosner, and C. F. Foard. (1976) Perceptual categories for musiclike sounds: Implications for theories of speech perception. Quart. J. Exp. Psychol. 28, 361-378.

Darwin, C. J. (1976) The Perception of Speech. In Handbook of Perception, vol. 7, ed. by E. C. Carterette and M. P. Friedman. (New York: Academic Press), pp. 175-226.

Day, R. S. and C. C. Wood. (1972) Interactions between linguistic and nonlinguistic processing. J. Acoust. Soc. Amer. 51, 79 (A).

Delattre, P. C., A. M. Liberman, and F. S. Cooper. (1955) Acoustic loci and transitional cues for consonants. J. Acoust. Soc. Amer. 27, 769-773.

Diehl, R. L. (1975) The effect of selective adaptation on the identification of speech sounds. Percept. Psychophys. 17, 48-52.

Diehl, R. L. (1976) Feature analyzers for the phonetic dimension stop vs. continuant. Percept. Psychophys. 19, 267-272.

Dorman, M. R., M. Studdert-Kennedy, and L. J. Raphael. (In press) Stop consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. Percept. Psychophys.

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

Eimas, P. D. (1974) Speech perception in early infancy. In Infant Perception, ed. by L. B. Cohen and P. Salapatek. (New York: Academic Press).

Eimas, P. D., E. R. Siqueland, P. Jusczyk, and J. M. Vigorito. (1971) Speech perception in infants. Science 171, 303-306.

Eimas, P. D. and J. D. Corbit. (1973) Selective adaptation of linguistic feature detectors. Cog. Psychol. 13, 247-252.

Eimas, P. D., W. E. Cooper, and J. D. Corbit. (1973) Some properties of linguistic feature detectors. Percept. Psychophys. 13, 247-252.

Eimas, P. D. and J. L. Miller. (In press) Effects of selective adaptation on the perception of speech and visual patterns: Evidence for feature detectors. In Perception and Experience, ed. by H. L. Pick and R. I. Walk (New York: Plenum Press).

Elliott, D. N. and W. R. Fraser. (1970) Fatigue and Adaptation. In Foundations of Modern Auditory Theory, ed. by J. Tobias. (New York: Academic Press).

Engen, T. (1971) Psychophysics II. Scaling Methods. In Experimental Psychology, ed. by J. W. Kling and L. A. Riggs. (New York: Holt, Rinehart, and Winston, Inc.).

Evans, E. F. (1974) Neural Processes for the Detection of Acoustic Patterns and for Sound Localization. In The Neurosciences: Third Study Program, ed. by F. O. Schmitt and F. G. Worden. (Cambridge, Mass: MIT Press).

Evans, E. F. and I. C. Whitfield. (1964) Classification of unit responses in the auditory cortex of the unanaesthetized and unrestrained cat. J. Physiol. 17, 476-493.

Evans, E. F., H. F. Ross and I. C. Whitfield. (1965) The spatial distributation of unit characteristic frequency in the primary auditory cortex of the cat. J. Physiol. 179, 238-247.

Fant, C. G. M. (1962) Descriptive analysis of the acoustic aspects of speech. Logos 5, 3-17.

Fant, C. G. M. (1968) Analysis and synthesis of speech processes. In Manual of Phonetics, ed. by B. Malmberg. (Amsterdam: North-Holland), pp. 173-277.

Frishkopf, L. and M. Goldstein. (1963) Responses to acoustic stimuli in the eighth nerve of the bullfrog. J. Acoust. Soc. Amer. 35, 1219-1228.

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

Funkenstein, H. H., P. G. Nelson, P. Winter, Z. Wollberg, and J. D. Newman. (1971) In Physiology of the Auditory System: A Workshop, ed. by M. Sachs. (Baltimore, Md.: National Educational Consultants, Inc.).

Ganong, W. F. (1975) An experiment on "phonetic adaptation." Progress Report (Research Laboratory of Electronics, MIT) 116, 206-210.

Garner, W. R. and G. L. Felfoldy. (1970) Integrality of stimulus dimensions in various types of information processing. Cog. Psychol. 1, 225-241.

Garner, W. R. (1974) The Processing of Information and Structure. (Potomac, Md.: Lawrence Erlbaum Assoc.).

Guilford, J. P. (1954) Psychometric Methods. (New York: McGraw-Hill Book Company, Inc.).

Haber, R. N. and M. Hershenson. (1973) The Psychology of Visual Perception. (New York: Holt, Rinehart, and Winston, Inc.).

Hammer, E. R. (1949) Temporal factors in figural after-effects. Amer. J. Psychol. 62, 337-354.

Harris, K. S. (1958) Cues for the discrimination of American English fricatives in spoken syllables. Lang. Speech 1, 1-7.

Harris, K. S., H. S. Hoffman, A. M. Liberman, P. C. Delattre, and F. S. Cooper. (1958) Effect of third-formant transitions on the perception of the voiced stop consonants. J. Acoust. Soc. Amer. 30, 122-126.

Helson, H. (1938) Fundamental problems in color vision. The principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination. J. Exp. Psychol. 23, 439-476.

Helson, H. (1947) Adaptation-level as a frame of reference for prediction of psychophysical data. Amer. J. Psychol. 60, 1-29.

Helson, H. (1964) Adaptation-level theory: An experimental and systematic approach to behavior. (New York: Harper and Row).

Helson, H. (1971) Adaptation level theory: 1970 and after. In Adaptation-Level Theory, ed. by M. H. Appley. (New York: Academic Press).

Hillenbrand, J. M. (1975) Intensity and repetition effects on selective adaptation to speech. Research on Speech Perception Progress Report. (Department of Psychology, Indiana University) 2, 56-137.

Hochberg, J. (1971) Perception I. Color and shape. In Experimental Psychology, ed. by J. W. Kling and L. A. Riggs. (New York: Holt, Rinehart, and Winston, Inc.).

Hoffman, H. S. (1958) Study of some cues in the perception of the voiced stop consonants. J. Acoust. Soc. Amer. 30, 1035-1041.

Hood, J. D. (1950) Studies in auditory fatigue and adaptation. Acta Oto-Lar., Suppl. 92.

Hubel, D. H. and T. N. Wiesel. (1962) Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. J. Physiol. 60, 106-154.

Hubel, D. H. and T. N. Wiesel. (1965) Receptive fields and functional architecture in two nonstriate visual areas (18 & 19) of the cat. J. Neurophysiol. 28, 229-289.

Jerger, J. F. (1957) Auditory adaptation. J. Acoust. Soc. Amer. 29, 357-363.

Johnson, D. M. (1972) Systematic Introduction to the Psychology of Thinking. (New York: Harper and Row).

Johnson, D. M. (1955) The Psychology of Thought and Judgment. (New York: Harper and Row).

Kay, R. H. and D. R. Mathews. (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.

Kohler, W. and H. Wallach. (1944) Figural after-effects: An investigation of visual processes. Proc. Amer. Phil. Soc. 88, 269-357.

Koenig, W., H. K. Dunn, and L. Y. Lacey. (1946) The sound spectrograph. J. Acoust. Soc. Amer. 17, 19-49.

Kozhevnikov, V. A. and L. A. Chistovich. (1965) Rech' Artikuliatsia i vospriiatie. (Moscow-Leningrad). Trans. as Speech: Articulation and Perception. (Washington, D.C.: Clearinghouse for Federal Scientific and Technical Information), JPRS, 30, 543.

Kuhl, P. K. and J. D. Miller. (1975) Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. Science, 190, 69-72.

Kuhn, G. M. (1975) On the front cavity resonance and its possible role in speech perception. J. Acoust. Soc. Amer. 58, 428-433.

Ladefoged, P. (1967) Three Areas of Experimental Phonetics. (New York: Oxford University Press).

Lass, N. J. and S. S. Golden. (1971) The use of isolated vowels as auditory stimuli in eliciting the verbal transformation effect. Canad. J. Psychol. 25, 349-359.

Lass, N. J. and R. M. Gasperini. (1973) The verbal transformation effect: a comparative study of the verbal transformations of phonetically trained and non-trained subjects. Brit. J. Psychol. 64, 183-192.

Lass, N. J., L. K. West, and D. D. Taft. (1973) A non-verbal analogue to the verbal transformation effect. Canad. J. Psychol. 27, 272-279.

Lettvin, J. Y., H. R. Maturana, W. S. McCulloch, and W. H. Pitts. (1959) What the frog's eye tells the frog's brain. Proc. Inst. Rad. Engr. 47, 1940-1951.

Liberman, A. M. (1970) The grammars of speech and language. Cog. Psychol. 1, 301-323.

Liberman, A. M., F. S. Cooper, K. S. Harris, and P. F. MacNeilage. (1963) A motor theory of speech perceeption. Proceedings of the Speech Communication Seminar. (Stockholm: Royal Institute of Technology.) D3

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

Liberman, A. M., P. C. Delattre, and F. S. Cooper. (1952) The role of selected stimulus variables in the perception of the unvoiced-stop consonants. Amer. J. Psychol. 65, 497-516.

Liberman, A. M., P. C. Delattre, and F. S. Cooper. (1958) Some cues for the distinction between voiced and voiceless stops. Lang. Speech 1, 153-167.

Liberman, A. M., P. C. Delattre, F. S. Cooper, and L. J. Gerstman. (1954) The role of consonant-vowel transitions in the perception of the stop and nasal consonants. Psychol. Monogr. 68, 1-13.

Liberman, A. M., P. C. Delattre, J. J. Gerstman, and F. S. Cooper. (1956) Tempo of frequency change as a cue for distinguishing classes of speech sounds. J. Exp. Psychol. 52, 127-137.

Liberman, A. M., K. S. Harris, H. S. Hoffman, and B. C. Griffith. (1957) The discrimination of speech sounds within and across phoneme boundaries. J. Exp. Psychol. 54, 358-368.

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

Liberman, A. M., F. Ingemann, L. Lisker, P. C. Delattre, and F. S. Cooper. (1959) Minimal rules for synthesizing speech. J. Acoust. Soc. Amer. 31, 1490-1499.

Lindblom, B. E. F. (1963) Spectrographic study of vowel reduction. J. Acoust. Soc. Amer. 35, 1773-1781.

Lindblom, B. E. F. and M. Studdert-Kennedy. (1967) On the role of formant transitions in vowel recognition. J. Acoust. Soc. Amer. 42, 830-843.

Lisker, L. (1975) Is it VOT or a first-formant transition detector? J. Acoust. Soc. Amer. 57, 1547-1551.

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

Lisker, L. and A. S. Abramson. (1967) Some effects of context on voice onset time in English stops. Lang. Speeech 10, 1-28.

Lisker, L. and A. S. Abramson. (1970) The voicing dimension: Some experiments in comparative phonetics. In Proceedings of the 6th International Congress of Phonetic Sciences. (Prague: Academia), 563-567.

Locke, S. and L. Kellar. (1973) Categorical perception in a nonlinguistic mode. Cortex 9, 355-369.

Maturana, H. R., J. Y. Lettvin, W. S. McCullouch, and W. B. Pitts. (1969) Anatomy and physiology of vision in the frog Rana pipiens. J. Gen. Physiol. 43, 129-175.

McCollough, C. (1965) Color adaptation of edge-detectors in the human visual system. Science 149, 1115-1116.

Michael, C. R. (1969) Retinal processing of visual images. Sci. Amer. 220, 104-114.

Miller, J. D., R. E. Pastore, C. C. Wier, W. J. Kelly, and R. J. Dooling. (1974) Discrimination and labeling of noise-buzz sequences with varying noise-lead times. J. Acoust. Soc. Amer. 55, 390(A).

Miller, J. D., C. C. Wier, R. E. Pastore, W. J. Kelly, and R. J. Dooling. (1976) Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. J. Acoust. Soc. Amer. 60, 410-417.

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

Moffitt, A. R. (1971) Consonant cue perception by twenty- to twenty-four-week-old infants. Child Develop. 42, 717-731.

Møller, A. R. (1969) Unit responses in the cochlear nucleus of the rat to sweep tones. Acta Physiol. Scand. 76, 503-512.

Morse, P. A. (1972) The discrimination of speech and nonspeech stimuli in early infancy. J. Exp. Child Psychol. 14, 477-492.

Munson, W. A. (1947) The growth of auditory sensation. J. Acoust. Soc. Amer. 19, 584-591.

Nelson, P. G., S. D. Erulkar, and S. S. Bryan. (1966) Response units of the inferior colliculus to time-varying acoustic stimuli. J. Neurophysiol. 29, 834-860.

Öhman, S. E. G. (1966) Coarticulation in VCV utterances: Spectrographic measurements. J. Acoust. Soc. Amer. 39, 151-168.

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

Parducci, A. (1974) Contextual effects: A range-frequency analysis. In Handbook of Perception, Vol. II, ed. by E. C. Carterette and M. P. Friedman. (New York: Academic Press).

Pisoni, D. B. (1971) On the nature of categorical perception of speech sounds. Supplement to: Haskins Laboratories, Status Report on Speech Research. November.

Pisoni, D. B. (1975a) Dichotic listening and processing phonetic features. In Cognitive Theory: Volume I, ed. by F. Restle, R. M. Shiffrin, N. J. Castellan, H. Lindman, and D. B. Pisoni. (Hillsdale, N. J.: Erlbaum Associates).

Pisoni, D. B. (1975b) Stages of processing in speech perception: Feature analysis. Paper presented at the Eighth International Congress of Phonetic Sciences, August, Leeds, England.

Pisoni, D. B. (In press) Identification and discrimination of the relative onset of two component tones: Implications for the voicing perception in stop consonants. J. Acoust. Soc. Amer.

Pisoni, D. B., J. R. Sawusch, and F. T. Adams. (1975) Simple and contingent adaptation effects in speech perception. Paper presented at the 90th meeting of the Acoustical Society of America, November, San Francisco.

Pisoni, D. B. and J. B. Tash. (1975) Auditory property detectors and processing place features in stop consonants. Percept. Psychophys. 18, 401-408.

Pollack, I. and J. Boynton. (1963) Identification of elementary auditory displays: Effects of unbalanced probabilities of occurrence. J. Acoust. Soc. Amer. 35, 1831-1832.

Rogers, S. (1941) The anchoring of absolute judgments. Arch. Psychol. 261, 5-56.

Sawusch, J. R. (1976) The structure and flow of information in speech perception: Evidence from selective adaptation of stop consonants. Technical Report (Department of Psychology, Indiana University), 2. University.

Sawusch, J. R. and D. B. Pisoni. (1973) Category boundaries for speech and non-speech sounds. Paper presented at the 86th meeting of the Acoustical Society of America, November 1973, Los Angeles.

Sawusch, J. R., D. B. Pisoni, and J. E. Cutting. (1974) Category boundaries for linguistic and nonlinguistic dimensions of the same stimuli. Paper presented at the 87th meeting of the Acoustical Society of America, April 1974, New York City.

Sawusch, J. R. and D. B. Pisoni. (1976) Response organizations and selective adaptation to speech sounds. Percept. Psychophys. 20, 413-418.

Small, A. M. (1963) Auditory adaptation. In Modern Developments in Audiology, ed. by J. Jerger. (New York: Academic Press).

Small, A. M. (1973) Acoustics. In Normal Aspects of Speech and Language, ed. by F. D. Minifie, T. J. Hixon, and F. Williams. (Englewood Cliffs, New Jersey: Prentice-Hall, Inc.)

Stevens, K. N. (1973) Potential role of property detectors in the perception of consonants. Quarterly Progress Report (Research Laboratory of Electronics, MIT), 110, 155-168.

Stevens, K. N. (1975) The potential role of property detectors in the perception of consonants. In Auditory Analysis and Perception of Speech, ed. by C. G. M. Fant and M. A. A. Tatham. (New York: Academic Press), 283-330.

Stevens, K. N. and A. S. House. (1972) The perception of speech. In Foundations of Modern Auditory Theory, ed. by J. Tobias. (New York: Academic Press), pp. 3-62.

Stevens, K. N. and D. H. Klatt. (1974) Role of formant transitions in the voiced-voiceless distinction for stops. J. Acoust. Soc. Amer. 55, 653-659.

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

Stevens, K. N. and S. E. Blumstein. (1976) Context-independent cues for place of articulation in stop consonants. Paper presented at the 91st meeting of the Acoustical Society of America, April 5-9, Washington, D.C.

Strange, W., R. R. Verbrugge, D. P. Shankweiler, and T. R. Edman. (1976) Consonant environment specifies vowel identity. *J. Acoust. Soc. Am. 60*, 198-212.

Studdert-Kennedy, M. (1974) The perception of speech. In *Current Trends in Linguistics*, ed. by T. A. Sebeok, vol. 12 (4). (The Hague: Mouton), pp. 2349-2385.

Studdert-Kennedy, M. (1976) Speech perception. In *Contemporary Issues in Experimental Phonetics*, ed. by N. J. Lass. (New York: Academic Press), pp. 243-293.

Studdert-Kennedy, A. M. Liberman, K. S. Harris, and F. S. Cooper. (1970) Motor theory of speech perception: A reply to Lane's critical review. Psychol. Rev. *77*, 234-249.

Studdert-Kennedy, M. and D. P. Shankweiler. (1970) Hemispheric specialization for speech perception. *J. Acoust. Soc. Am. 48*, 579-594.

Suga, N. (1965) Analysis of frequency-modulated sounds by auditory neurons of echo-locating bats. J. Physiol. (Lond.) *179*, 26-53.

Summerfield, Q. (1974) Towards a detailed model for the perception of voicing contrasts. Speech Perception (Department of Psychology, The Queen's University of Belfast) *3*, 1-26.

Summerfield, Q. (1975) Cues, contexts and complications in the perception of voicing contrasts. Speech Perception. (Department of Psychology, The Queen's University of Belfast) *4*, 99-129.

Summerfield, Q. and M. P. Haggard. (1974) Perceptual processing of multiple cues and contexts: Effects of following vowel upon stop consonant voicing. J. Phonetics *2*, 279-295.

Summerfield, A. Q. and M. P. Haggard. (1977) On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Haskins Laboratories Status Report on Speech Research SR-49*.

Tartter, V. C. and P. D. Eimas. (1975) The role of auditory feature detectors in the perception of speech. Percept. Psychophys. *18*, 293-298.

Volkman, J. (1951) Scales of judgment and their implications for social psychology. In *Social Psychology at the Crossroads*, ed. by J. H. Rohrer and M. Sherif. (New York: Harper and Row).

von Bekesy, G. (1960) *Experiments in Hearing*. (New York: McGraw-Hill).

Ward, W. D. (1963) Auditory fatigue and masking. In *Modern Developments in Audiology*, ed. by J. Jerger. (New York: Academic Press).

Waters, R. S. and W. A. Wilson, Jr. (1976) Speech perception by rhesus monkeys: The voicing distinction in synthesized labial and velar stop consonants. Percept. Psychophys. *19*, 285-289.

Wever, E. G. and K. E. Zener. (1928) The method of absolute judgment in psychophysics. Psychological Review *35*, 466-493. (Cited in Johnson, 1972).

Whitfield, I. C. (1965) 'Edges' in auditory information processing. In *Proceedings of XXIIIrd International Congress of Physiological Sciences*, (Tokyo), September, pp. 245-247.

Whitfield, I. C. (1967) *The Auditory Pathway*. (London: Edward Arnold, Ltd.).

Whitfield, I. C. and E. F. Evans. (1965) Responses of auditory cortical neurons to stimuli of changing frequency. *J. Neurophysiol*. *28*, 655-672.

Wood, C. C. (1974) Parallel processing of auditory and phonetic information in speech perception. Percept. Psychophys. 15, 501-508.

Wood, C. C. (1975) Auditory and phonetic levels of processing in speech perception: Neurophysiological and information-processing analysis. J. Exp. Psychol.: Human Percept. Perform. 104, 1-33.

Wood, C. C. and R. S. Day. (1975) Failure of selective attention to phonetic segments in consonant-vowel syllables. Percept. Psychophys. 17, 346-350.

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

Woodworth, R. S. and H. Schlosberg. (1954) Experimental Psychology. (New York: Holt, Rinehart, and Winston).

Wright, H. N. (1960) Measurement of perstimulatory auditory adaptation. J. Acoust. Soc. Amer. 32, 1558-1567.