Adaptation of the Category Boundary Between Speech and Non-speech: A Case Against Feature Detectors

Robert E. Remez†

ABSTRACT

Two experiments were performed employing acoustic continua that change from speech to nonspeech. The members of one continuum, synthesized on the Pattern Playback, varied in equal steps of change in the bandwidths of the first three formants, from the vowel /a/ to a nonspeech buzz. The other continuum, achieved through digital synthesis, varied in the bandwidths of the first five formants, from /æ/ to buzz. The categorical perception of each continuum was established by standard procedures. Perceptual adaptation on these continua then revealed effects on the category boundaries comparable to those reported for speech sounds. The results are interpreted as suggesting that neither phonetic nor auditory feature detectors are responsible for perceptual adaptation of speech sounds.

INTRODUCTION

In their account of the perceptual process underlying phonetic identification, Eimas and Corbit (1973) combined phonetic feature analysis and hypercomplex cells (Hubel and Wiesel, 1965) to yield phonetic feature detectors, understood as special cortical devices tuned to "listen" to the acoustic stream and extract the phonetic building blocks. This claim seemed reasonable on several counts. First, the elusiveness of the acoustic-phonetic correspondence argued that a higher order evaluation would be needed to accomplish the extraction of the meaning from the acoustic stream (Halle and Stevens, 1962; Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967). Second, a century of efforts directed toward delineating the cortical loci of mental faculties established that speech and language are mediated in a restricted cortical area (Penfield and Roberts, 1959; Geschwind and Levitsky, 1968; Witelson and

†Also University of Connecticut, Storrs.

Acknowledgment: During the course of this research the author was supported by an NICHD Pre-doctoral Training Grant in Language and Psychology awarded to the University of Connecticut. It is a pleasure to thank Tom Casola, Harriet Greisser, Alvin Liberman, Ignatius Mattingly, Philip Rubin, Michael Studdert-Kennedy, Quentin Summerfield, Michael Turvey, Hal Tzeutschler, and Robert Verbrugge for their aid, comment, and criticism during the development of this study and report. Also, a special thank you goes to Agnes McKeon, who taught me how to paint Playback-style.

Pallie, 1973). Third, the nativist position of the generative transformational enterprise argued that rather detailed linguistic knowledge might be previred into every human infant (Chomsky, 1965; Chomsky, 1968). Fourth, perceptual experiments on neonates seemed to show that infants were sensitive to some phonetic distinctions before any relevant experience (Eimas, Siqueland, Jusczyk, and Vigorito, 1971). And fifth, developments in electrophysiology suggested that the sensitivities of single cortical cells are both elaborate and correlated with the ecology of the animals studied (see Kuffler, 1973, for a review). The notion of genetically pretuned hypercomplex cells that translate auditory events into phonetic descriptions neatly addressed these issues.

Perceptual adaptation of phonetic category boundaries has been the technique by which the properties of the hypothesized detectors have been examined\(^1\), but this body of research only marginally confirms the original detector description. Although certain adaptation effects have required the explanation to include a phonetic level of analysis at which particular acoustic-auditory values are less perceptually significant (Ades, 1974; Diehl, 1975; Miller, 1975; Remez, Cutting, and Studdert-Kennedy\(^2\), other research has produced evidence of non-phonetic adaptation that is fully compatible with any of the previously obtained phonetic effects (Ades, 1974; Pisoni and Tash, 1975; Bailey, 1975; Diehl, 1976). In short, the present situation is paradoxical. While the underlying detectors are phonetic by the original intention as well as by occasional necessity, some of them may suffer acoustic fatigue, and all suffer from inexplicit specification of their tuning curves. [In tonotopically organized cells, the dimension for measuring sensitivity is frequency (Woolsey and Walzl, 1942); in phonetically organized detectors, the dimensions of analysis must correspond to those of vocal production, and many of these have yet to be defined.] Additionally, and perhaps fatally, the passive filtration method of phonetic feature extraction in speech assumes a simple relation between the acoustic pattern, the phonetic segments, and the ordinal correspondence between them. But it is the fundamental point of many perceptual studies that segmental identity typically is carried by the sound pattern distributed across the entire syllable (for example, Cooper, Delattre, Liberman, Borst, and Gerstman, 1952). This requires, in essence, that each feature detector be a little homunculus, omniscient on the nature of the context conditioned variation of its favorite feature of speech. For these reasons, either the task requires vast multiplexing, well beyond the sensible (Halwes and Jenkins, 1971), or the loss of the appealing simplicity of the device in order to deal with the complexity of the message structure of the speech chain.

---

\(^1\)Haggard (1967) first described the paradigm of after-effect research employing speech sounds, but his rationale was completely independent of neurophysiological claims.

A test of the hypothesis that adaptation effects reveal detectors tuned to linguistic features might attempt to produce adaptation outside the set of effects predictable from the feature inventory. The experiments reported here, which use acoustic continua from speech to nonspeech sounds, satisfy this condition. Further, if acoustic-auditory explanations can be ruled out, there would then be reason to suggest an active, cognitive basis for this effect, one which, if applied to speech, would be compatible with a phonetic level of analysis, but incompatible with phonetic feature detectors.

EXPERIMENT Ia

Methods

Subjects. Sixteen University of Connecticut undergraduates, whose participation fulfilled the introductory psychology course requirement, served as listeners in this part of the study. All were native English speakers with no known speech or hearing disorder or psychopathology. None had any experience with synthetic speech sounds before the listening session.

Stimuli. The Haskins Laboratories Pattern Playback (Cooper, Liberman, and Borst, 1951) was used to synthesize the basic materials. This device uses a tone wheel to generate the harmonics of 120Hz in light intensities arrayed in a frequency scale. A graphic pattern selectively reflects portions of this scale, and this reflection, through capture by a photocell, is transduced to a frequency by amplitude by time acoustic signal. Figure 1 displays the pattern painted (Liquitex Acrylic Titanium White Grumbacher #4) on the acetate belt (Eastman Kodak) and the frequency values of the transduced signal. The pattern changes from a vowel /a/, with formant frequency values of 600 Hz, 1200 Hz, and 2400 Hz, to a nonspeech buzz, by modifying the bandwidths of the formants; initially, the bandwidths are 100 Hz, and they increase to effectively infinite width at the end of the pattern. Figure 2a presents one spectral section through each of the endpoints.

This 1-second sound was transferred to audiotape and then digitized by the Haskins Pulse-Code Modulator (PCM) (Cooper and Mattingly, 1969), sampling at 10 KHz with low-pass filtering at 5 KHz. A ten-step continuum was then made by editing the digitized waveform. Nine cuts were made, one every 100 milliseconds; the oscillographic patterns were equated for amplitude, producing 10 tokens, each of 12 pitch periods which vary in formant bandwidth as does the overall pattern.

Two test tapes were created using the PCM system. An identification sequence contained 10 occurrences of each of the 10 continuum items, for a test of 100 trials, with 5 seconds between trials, and 9 seconds following each decade. A discrimination sequence consisted of ABX triads with 1 second

3Neither the Haskins Parallel Resonance Synthesizer nor the Ove III were suitable for this study because of hardware-imposed limits on formant bandwidth. These devices are devoted speech synthesizers, and this study required a full-frequency synthesizer, that is, one with no such restriction.
Figure 1: The painted pattern which the Playback reproduced.
Figure 2: Spectral sections through the endpoints of the Playback-synthesized continuum (a), and the software synthesized continuum (b).
between items, 5 seconds between trials, and 9 seconds separating the decades. The 4 permutations of each ABX comparison were represented: ABA, ABB, BAB, and BAA. In a one-step discrimination, the comparisons are items 1 and 2, 2 and 3, 3 and 4, and so on; at 4 trials per comparison, and 9 comparisons, there were 36 trials in this test.

Procedure and Apparatus

The sixteen listeners were tested in four groups of four. Sounds were presented binaurally over Grason-Stadler earphones activated by a Crown 820-144 tape recorder through a junction box so that several listeners could listen simultaneously. Each session commenced with a briefing sequence in which the endpoints of the continuum were each repeated ten times, in alternation. At that time, listeners were asked to signify if they had a good idea of the sounds' identity; their instructions were to consider the buzz a machine noise, and the vowel a synthetic speech sound. The identification test was then administered. Identifications were scored on a response sheet as speech or buzz (S or B). After a short intermission, listeners were given sample ABX sequences in which they judged which of the first two sounds was identical to the third; the continuum endpoints were used here to insure clarity of the instructions. The actual test, begun when all agreed that they understood the instructions, consisted of the 36-trial discrimination sequence played twice.

Results and Discussion

Three subjects were dropped because they either failed to follow instructions (two subjects declined to judge difficult items) or responded at chance on the identifications (one subject). Results for identification and discrimination appear in Figure 3a. Each point is the mean of 130 observations in the identification test, and 108 observations in the discrimination test. These functions are reasonably consistent with the criteria for categorical perception (Studdert-Kennedy, Liberman, Harris, and Cooper, 1970), in that a peak in discriminability occurs at the breakpoint between the identification categories. The term "categorical perception" describes a situation in which the judged difference between two entities is contingent on their identities rather than on the physical differences between them.

One anomaly in the discrimination function should be addressed, namely, the troughs of the function. Given two categories, speech and buzz, the peaks should number one, not three. In the function of Figure 3, however, the discrimination peaks between items 1 and 2, and between 9 and 10, are as prominent as the category boundary peak. Examination of the waveforms of these tokens revealed amplitude differences between 2 of the 12 pitch periods in each item of the pairs. No such difference could be discovered between items in the other pairs in the set. One possible account, then, of the spurious peaks is that they result from judgments of amplitude rather than spectrum differences. If this reasoning is correct, the peaks can be discounted in any challenge to categoricity, because the manipulation of interest is spectral, and this particular discrimination is made on a nonspectral basis.
Figure 3: Identification and discrimination plots for Playback (a) and software (b) continua.
EXPERIMENT Ib

Methods

Subjects. Eight University of Connecticut undergraduates were paid to listen in this part of the study. They had all participated in Experiment Ia (five from the original group could not attend the listening sessions for scheduling reasons).

Stimuli. The 10 tokens from Experiment Ia were used. An adaptation sequence consisted of an initial 100 repetitions of the adapting item, one of the continuum endpoints, at 1-second intervals. After a 10-second pause, which cued the listeners that the identification trials were coming up, six items from the continuum were presented for identification, as either speech or buzz (S or B). At the conclusion of the block of six, there was a 10-second pause followed by 50 repetitions of the adapting item, another block of six, and so on for the remainder of the test.

Each of the 10 sounds drawn from the continuum was presented for identification twelve times, with the exception of the four most extreme, the two on each end that were presented six times each. This preserves sensitivity in the midrange of the continuum and shortens the test by two blocks of identifications, to the relief of the listener. With 96 trials (6 twelves and 4 sixes) there were 19 blocks of six trials each. The random order for these items was the same in both the speech and the buzz adaptation sequences.

Procedure. An identification sequence was used to determine a standard identification function in each of the two sessions. This was used for comparison with the adapted identification function. All subjects took part in both conditions; half took the /a/ test first; half took the buzz test first. Several days separated the test sessions. The equipment and test conditions were in all other respects the same as in Experiment Ia.

Results

Each subject contributed two sets of judgments per session, a pretest set and an adaptation set. To each of these a standard ogive was fitted, after Woodworth (1938). Thus, two scores were available for each subject per test, one mean of the fitted ogive, measured in continuum units, for the pretest, and one for the adaptation test.

The curves for the grouped data for each session appear in Figure 4. Each pretest plot represents the means of 80 trials per continuum item; in the adaptation plot, the two extreme points on either end, items 1, 2, 9, and 10, are the means of 48 observations each; the remaining six medial points are the means of 96 judgments each. The change in the ogive mean due to speech adaptation was 1.32 continuum units, and that due to buzz was .638 continuum units in the opposite direction.

A two-factor repeated measures analysis of variance was performed on the ogive means, with two levels of adaptor (Speech or Buzz) and two conditions (Pretest or Adaptation). There were no significant main effects, indicating
Figure 4: Pretest and adaptation test curves for the Playback continuum.
that the means, when collapsed across conditions or across adaptor, did not
depart from the grand mean. However, the interaction of adaptor by condition
was significant \( F(1,7)=12.771, \ p < .01 \), reflecting the different, and
opposite, effect of each adaptor relative to the pretest mean, on the
adaptation test mean.

Discussion

In this experiment the shifts of the category boundary do damage to the
proposal that the units mediating the effect are isomorphic with the primi-
tives of phonetic feature analysis. Miller (1975), for example, has proposed
that feature analyzing devices are arranged so that the fatigue of one leads
to the relatively enhanced strength of its inverse or opponent. However, in
the case of /a/ fatigue, the resulting change requires going outside the
feature set used in phonetic descriptions to capture the distinction between
speech and nonspeech. In other words, a processing device arranged along the
lines of phonetic features could neither predict nor explain this case of
adaptation.

Must we then have recourse to a "lower level" account of the results? Other
authors have found higher order descriptions unwarranted by the adapta-
tion effects (Pisoni and Tash, 1975; Bailey, 1975; Ades\(^5\)), and since higher
order phenomena can be described in lower order physical terms, they have
ascribed their effects to alterations in sensitivity at a lower level within
the auditory system. For example, a receptive unit with a "best" stimulus may
tend to show, under fatigue, a decreased sensitivity to that stimulus; its
decrease leads to the release from inhibition of adjacent, similar receptors,
and consequently, to the relative enhancement of near misses from the best
value. Thus, the auditory system, early on in the course of an analysis, can

\(^4\)Morse, Kass, and Turkienicz (1976) found that an /e/ or /i/ adaptor, but not
/i/, changed both boundaries on an /i/-/ɪ/-/ɛ/ continuum; they concluded that
the binary distinctions "tense" and "high" of Chomsky and Halle (1968), which
would specifically rule out such a finding, had been empirically falsified.
They offered that, because their results had shown continuity rather than
discreteness, the feature system underlying the result was necessarily
continuous, many-valued rather than binary; by extrapolation, so was the
detector. Their approach was like that of Cooper and Blumstein (1974) who
were the first to use adaptation to find perceptual interactions and thereby
to define the phonetic features perceptually, rather than acoustically or
articulatorily. Nevertheless, because no language makes phonetic use of the
distinction [+speech, -speech], it is safe to say that /a/-buzz adaptation
does not require that we posit a new feature; rather, it undermines the
interpretation of speech adaptation effects in terms of phonetic features.

\(^5\)Ades, A. E. Source assignment and feature extraction in speech.
(unpublished manuscript).
yield a mistransformed description which the unsuspecting analyzers in the next step are helpless to reverse. [The actual neurophysiology of this is still open to question. One current topic of investigation is whether the reciprocal inhibition demonstrable at the VIII nerve nucleus arises cochlearly or in the nucleus itself (Mountcastle, 1974).] By this mechanism, then, a unit or units sensitive to a portion of the frequency range, when fatigued, will be less sensitive to the absolute values of stimulation, and will, via disinhibition, effectively amplify departures from the original fatigued values. For example, a receptor that mediates a rising second formant value over the course of 35 msec (which specifies a voiced bilabial in some circumstances) will be less sensitive when fatigued to values that exactly conform to the pattern of fatigue. Disinhibition of neighboring receptors would, in effect, boost receptor responses to second formant transitions that depart by small amounts from the fatigued values. This will be costly in terms of the perceptual outcome only in borderline cases, that is, those in which the fatigue-disinhibition throws the pattern over the line from one category into another. This explanation insists that the auditory transcription which the phonetic system is given to work with has been irretrievably altered.

However, if this reasoning is applied to the adaptation by /a/ and buzz, a curious situation arises. Fatigue caused by buzz should decrease sensitivity throughout the range of frequencies used; no frequency-specific effects should occur. Indeed, the auditory view of adaptation would predict no adaptation at all.

On the other hand, fatigue caused by /a/ should eat holes in the "neural spectrogram" of the buzz, decreasing sensitivity at 600 Hz, 1200 Hz, and 2400 Hz. If a listener were then presented with the buzz, he should hear a pattern the inverse of /a/, with formants at 300 Hz, 900 Hz, and 1800 Hz. In this case, if the listener judges the sound on the basis of the acoustic feature of presence or absence of formant structure, then the auditory point of view predicts that the boundary should move toward the buzz, since the fatigued spectral receptors, even in this extreme case, might be expected to retain a pattern showing acoustic maxima and minima. However, precisely the reverse boundary movement was actually observed.

The present experiment, therefore, produces a situation unique in the adaptation literature. While the conventional approach has been to suspect auditory processes by default whenever a phonetic account fails, this is obviously not possible here. The listener must be judging the sounds on other than a simple acoustic basis. The perception of the novel distinction between /a/ and buzz required here indicates a stable perceptual capacity that can be reconfigured to suit the demands of the particular situation. Neither phonetic feature nor acoustic property detectors can be reconciled with this type of perceptual flexibility.

---

6This pattern, when rendered by the Playback, does not sound like a speech sound.
Based on the foregoing, there are several motives for extending this investigation. First, the anomalous discrimination peaks mildly threaten the claim of categoricity for this perceptual distinction. By implication, the explanation that the adaptation effect is judgmental rather than perceptual, cannot be confidently rejected. Second, the artificiality of the speech synthesized on the Playback may be a factor to consider. Because the Playback, however phonetically identifiable its message may be, has a voice quality unlike that of any person, its use in an experiment of this kind may produce synthesizer artifacts. As a precaution, then, it would be valuable to try this procedure with a more natural sounding synthesizer. Third, the possibility that this effect is restricted to /a/, that /a/ might intrinsically be more nonspeechlike than other vowels, could be assessed by using a different vowel in the same paradigm. On these accounts, Experiment II was performed.

**EXPERIMENT IIa**

**Methods**

**Subjects.** Eight University of Connecticut undergraduates, not those of Experiment I, were paid for their participation. All were naive with respect to synthetic speech.

**Stimuli.** The software synthesizer of Fisher and Engebretson (1975), modified to permit variable parameterization of all five formants, was used to make the acoustic tokens. This program calculates a digital wave from user-determined parameters of source frequency, formant frequency and bandwidth, and overall duration and amplitude. The digital wave is then converted to audio via a PCM system. These programs, implemented by Joe Kupin and Hal Tzetzschler, run on the University of Connecticut Language and Psychology Data General NOVA 2.

A nine-step continuum from /a/ to buzz was made by successive 50 Hz increments in the formant bandwidths, starting from an initial bandwidth of 100 Hz for each formant. Duration was 140 msec; overall amplitude was 45 dB; fundamental frequency was 120 Hz; formant frequencies for the vowel were F1:750, F2:1650, F3:2460, F4:3500, and F5:4500. The audio output was transferred to the Haskins PCM via Ampex tape recording, to permit algorithmic envelope shaping. Each item was sixteen pitch periods (140 msec) long, with ramp on and off of 3 periods (25 msec); overall amplitudes were equated. Spectral sections through the endpoints appear in Figure 2b.

The identification test consisted of ten judgments of each of the nine items in random order. The discrimination test consisted of eight judgments of each of the eight one-step comparisons, in random order.

**Procedure and apparatus.** The outline of Experiment Ia was followed.

**Results and Discussion**

Figure 3b displays the functions for identification and discrimination. The identification plot displays the means of 80 trials per point, the
discrimination plot the means of 64 trials per point. Inspection of the figure will reveal that, relative to Experiment Ia, the peak at the category boundary remains a property of the discrimination function, while the peaks at the extremes of the continuum have disappeared at the buzz end, and all but disappeared at the speech end. It is reasonable to conclude that this more carefully controlled continuum elicited true categorical perception.

EXPERIMENT IIb

Methods

Subjects. The eight listeners from IIa were paid for their participation in this section of the study.

Stimuli. The nine-item continuum from IIa was used to make the adaptation sequences. These tests differed from Ib only in the consequences of using a nine- as opposed to a ten-step continuum. Here, the four most extreme items were presented six times each for identification during adaptation, and the remaining five medial items twelve times each. With 84 trials overall (4 sixes and 5 twelves), there were fourteen blocks of six trials each, which alternated with the repeating adaptation item, either /aw/ or buzz. The random order of identifications during adaptation was the same in the speech and buzz adaptor conditions.

Procedure. As in the previous procedure, each test day began with the identification sequence in order to obtain a standard for comparison with the adapted identification; test days were consecutive. All subjects took part in both adaptor conditions that were counterbalanced for order.

Results

The ogive fitting method was again used on the two tests per day contributed by each subject.

Averaged functions for both adaptation conditions appear in Figure 5. Pretest plots show the means of 80 trials per continuum item; adaptation plots show the means of 48 trials for items 1, 2, 8, and 9 and 96 trials for 3 through 7. The change in the ogive mean due to speech adaptation is .883 continuum units, due to buzz adaptation .692 in the opposite direction.

An analysis of variance was performed on the ogive means, with two levels of each factor, adaptor (Speech/Buzz) and mean (Pre/Post). The interaction is the term of interest here; with F(1,7)=32.842, p < .001. The main effect of adaptor was also significant; with F(1,7)=65.858, p < .001. The statistical significance of the adaptor term was due to the close correspondence of the two pretest means, which, when averaged with the adaptation means, clearly reflect the differential effects of adaptation. (Experiment Ib showed no such significance for this term because the pretest means varied in opposition to the adaptation means, thus cancelling the effect of adaptor upon averaging.)
Figure 5: Pretest and adaptation test curves for the software continuum.
Discussion

This study with software synthesized sounds strengthens the original argument made from Playback data. The results show that it is neither the artificiality of the speech synthesizer nor the particular vowel involved that enables listeners to treat the present acoustic continua as continua of proper speech sounds. It is this comparability which suggests that a speech/nonspeech detector may be responsible here, a detector just like those that putatively underlie phonetic adaptation. However, an auditory feature account is ruled out by the argument presented earlier, leaving either (1) a phonetic-type detector or (2) a detectorless explanation as the readily apparent alternatives to consider.

A phonetic detector explanation here would require the extension of the detector inventory, since a speech/nonspeech feature is not found in linguistic analysis. The distinction, in fact, is not even truly linguistic, in the sense of distinctive feature theory, but it certainly is a feature of human perceptual sensitivity, and on that basis might be seen as potentially detector-mediated. However, the existence of perceptual sensitivity should not be the only criterion for postulating a detector. The very advantage of this style of pattern recognition is that it makes infinite use of finite means; if a new detector is to be added to the set at every new discovery, then the contradiction of an indefinitely expandable finite means reduces the attractiveness of the model. The device required by these data can preserve its economy only if it has a small group of detectors, tuned to speech, set in opposition to a small group tuned to nonspeech. On this account, the search for independent confirmation of this organizational plan does not find the neurophysiology encouraging. Although there have been discussions of single-cell mediation of all perception, along the lines of innate taxa (Stent, 1975), as well as descriptions of arrays of phonetic single units (Miller, 1975), no proposal has yet been made to oppose speech neurons and nonspeech neurons. In fact, some claims for uniqueness of the speech neurology imply that the speech processor, whatever it may be, is separate from the nonspeech processor (Milner, 1962). Speech, in this view, is a mode, like vision or audition, and, by analogy, interacts with other modes but is independent of them. In short, a vast opponent process system for speech/nonspeech is not to be endorsed on the basis of any current view, and it may be presumed, in addition, that such a system is unlikely to exist given what is already known.

Finally, the only direct evidence for feature detectors in speech, as opposed to the invitation to such a conceptualization offered by neurophysiological metaphor, is the selective adaptation work. Boundary shifts occasioned by adaptation are precisely the effects that would permit the perceptual correlates of phonetic feature manipulations to be recast as the products of hypothetical detectors. However, though the hypothesis is reasonable when the endpoints differ by a single feature, it is difficult to imagine that a vowel and a buzz are also distinguished by but a single feature, speech/non-speech. The adaptation technique, the only test for feature detectors, is, ironically, not a demonstration of feature detectors at all. It simply reveals that certain perceptual contrasts, in this particular case of higher order properties, undergo selective alteration following saturation.
This study of vowel-buzz adaptation suggests that because the hypothe-
critical detectors are incapable of handling the result, and because the detectors
required to handle it are implausible, selective adaptation does not depend on
the existence of feature detectors. If the basis for adaptation, and perhaps
speech perception as well, can be understood as sensitivity to the higher-
order values inherent in acoustic pressure fluctuations, without decomposition
into features, then the description of such a process, not mere verification
of analytic features, is the goal toward which further research might well
proceed.

REFERENCES

syllable position and vowel environment. Percept. Psychophys. 16, 61-
66.

Bailey, P. J. (1975) Perceptual adaptation in speech. Unpublished disserta-
tion, Cambridge University.


Chomsky, N. (1968) Language and Mind. (New York: Harcourt, Brace and
World).

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

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., and Gerstman,
L. J. (1952) Some experiments on the perception of synthetic speech

of audible and visible patterns as a basis for research in the perception

investigation of dichotic speech perception. J. Acoust. Soc. Am. 46,
115(A).


Diehl, R. (1975) The effect of selective adaptation on the identification of

Diehl, R. (1976) Feature analyzers for the phonetic dimension stop

Eimas, P. D. and Corbit, J. D. (1973) Selective adaptation of linguistic


Haggard, M. (1967) Figural after-effect phenomena for speech and non-speech
stimuli. Haskins Laboratories Status Report on Speech Research SR-9, 61-
62.


Halwes, T. and Jenkins, J. J. (1971) Problem of serial order in behavior is
not resolved by context sensitive associative memory models.

166


