

Recognition thresholds for a speech continuum following selective adaptation

LAURA B. HARRIS and RICHARD E. PASTORE
State University of New York at Binghamton, Binghamton, New York

The effect of selective adaptation by end-point stimuli on a synthetic /bi-di/ continuum was investigated in terms of the change in masked recognition thresholds for stimuli within the continuum. Adaptation with either /bi/ or /di/ produced typical shifts in the identification boundary, while adaptation only with /bi/ produced shifts in the ABX discrimination peak. Both adaptors also produced significant decreases in the recognition threshold of all stimuli within the non-adapted category and, for the /di/ adaptor, reduced thresholds also for the adapted category. These results cannot be explained by assuming the action of only a simple, single factor such as fatigue, stimulus contrast, or response contrast.

The research described in this paper was undertaken to help assess the nature of selective adaptation for synthetic speech stimuli. In a selective-adaptation procedure, a single stimulus (adaptor), selected from a continuum of stimuli varying along a specific dimension, is presented repeatedly to an observer just prior to the presentation of continuum members for identification or discrimination. Use of such an adaptor with synthetic speech continua has been found to temporarily, but significantly, alter an observer's perception of stimuli from the continuum (Eimas & Corbit, 1973). Two tasks are typically employed with this paradigm: an identification task to establish the category boundary and a discrimination task to determine the discriminability of pairs of stimuli with constant physical differences drawn from within and across categories. The following study represents a somewhat atypical approach to the investigation of selective adaptation effects with a synthetic speech place continuum; in addition to the typical identification and discrimination tasks, it employs a method-of-adjustment psychophysical task to estimate the recognition threshold for intensity of each stimulus both before and after adaptation.

The method of limits, another psychophysical task, previously has been employed to investigate adaptation in the visual system. Visual-motion after-effects have been conjectured to be the result of adaptation of direction-sensitive cells. Research suggests that the threshold for these cells is elevated as a func-

tion of the degree of adaptation (Sekuler & Ganz, 1963; Weisstein, 1969). An analogous mapping of the effects of standard selective adaptation on the recognition of each stimulus in a speech continuum might provide some critical information concerning the nature of the adaptation phenomenon for speech stimuli. For instance, a systematic increase or decrease in threshold limited roughly to identification categories would imply a differential alteration of separate mechanisms, whereas no effect on threshold would imply far more limited effects of adaptation.

The rationale for measuring absolute thresholds in an adaptation situation is relatively more straightforward than the actual measurement of the thresholds. The measurement of thresholds requires the repeated presentation of a stimulus which, in turn, can lead to additional adaptation. With an approximation to asymptotic adaptation, this additional adaptation due to the measurement procedure probably is most critical primarily when measuring thresholds for stimuli relatively remote from the adaptor. Therefore, in measuring the threshold, we need to employ a psychophysical procedure that will minimize the number of stimulus presentations. With a method of adjustment, we can estimate the threshold while repeating a stimulus no more often than typical of a standard identification or ABX discrimination task; that would not be the case with a method of constant stimuli or with an adaptive psychophysical procedure. However, in order to protect against the errors inherent in this procedure, we must examine both interobserver trends in thresholds and the intra-observer repeated measurements both within and across the ascending and descending trials. Also, we shall variably bias the starting attenuation of each trial. While such threshold data alone may not be strong, direct evidence for or against a given explanation of selective adaptation, the results, when viewed

This research was supported in part by NSF Grant BNS8003704, NINCDS Grant 5 R01 NS10995-06 to the second author, and NIMH Grant MH08570-01 to the first author. Thanks are extended to the staff of Haskins Laboratories for the use of their facility in the preparation of the stimuli employed in this research. The authors' mailing address is: Department of Psychology, State University of New York, Binghamton, New York 13903.

in conjunction with other published reports, should further clarify the nature of selective adaptation.

METHOD

Subjects (Observers)

A within-subjects design was employed. Three females and one male participated in this experiment. All observers were undergraduates at SUNY-Binghamton with no known hearing deficits. All observers received extensive practice with the given stimuli, with the adaptation procedures, and with the method of adjustment prior to the collection of data. All observers received an hourly wage.

Stimuli

The stimuli consisted of synthetic CV syllables produced at Haskins Laboratories with an OVE parallel synthesizer and initially recorded on tape. The stimulus parameters were based upon those used by Miller and Eimas (1976) and Sawusch and Pisoni (1978). An /i/ vowel context was used with a nine-syllable /b/-/d/ continuum. The vowel portion of each syllable was 250 msec in duration preceded by a 50-msec transition. Table 1 provides a description of the various parameters manipulated and the values used for these stimuli. Stimulus value 0 was the most /bi/-like stimulus, while stimulus value 7 was the most /di/-like stimulus.

Procedure

All stimuli were digitized (12 bit x 20 kHz sampling), stored, randomly ordered, resynthesized, and presented to subjects by a North Star Horizon microprocessor. All stimuli were low-pass-filtered at 4 kHz and were presented in a continuous, broadband, 60-dBA white noise. Observers listened to binaurally presented stimuli over TDH-39(300) headphones in a commercial, sound-attenuated booth.

Baseline Conditions

An initial set of standard categorical perception conditions was run to establish that the unadapted observers were perceiving the members of the stimulus continuum in a categorical manner and to provide unadapted identification and discrimination "baselines" for comparison with results obtained following adaptation. There were five separate blocks of identification trials, each consisting of five randomized sequences of continuum members presented with a 2-sec interstimulus interval, thus providing 25 iden-

tification responses per stimulus. Each block of trials was approximately 2 min in duration. The identification task also was conducted at the start of every session as a "warm-up" for the observers and as a check against practice effects; no practice effects were evident during the course of the study.

The second categorical perception task was a random ABX discrimination procedure. Each ABX trial was 4.9 sec in duration: 300-msec stimuli, 1-sec interstimulus intervals, and a 2-sec response interval. Members of each discrimination pair were two step sizes apart. A step size is defined as the distance between two adjacent stimuli along the continuum (see Table 1). The discrimination baseline data for each subject consisted of six blocks of trials of randomized sequences of test triad positions (ABA, ABB, BAB, and BAA) for each discrimination pair for a total of 24 responses per stimulus pair.

Following the above identification and discrimination conditions, unadapted masked recognition thresholds were determined for all members of the stimulus continuum using both an ascending and a descending method of adjustment. With the ascending series, observers were required to increase the intensity of the repeatedly presented stimulus from a just-detectable intensity to an intensity at which the observers could identify the stimulus. In an attempt to guard against errors of anticipation, the observers were instructed to be as accurate as possible and to label the stimulus as soon as they could identify it. With the descending method of adjustment, the observers were asked to decrease the intensity of an easily identifiable stimulus until that stimulus could no longer be heard as an identifiable CV syllable. Each observer generated 10 recognition thresholds for each stimulus for both the ascending and descending method-of-adjustment tasks. These thresholds were averaged to obtain an average unadapted recognition threshold for each stimulus.

Experimental Conditions

Following completion of the above conditions, the standard categorical perception conditions (identification and discrimination tasks as described above) were run following adaptation to determine if the adaptation procedure did affect observers' identification and discrimination functions. The adaptation trial consisted of 184 adaptor repetitions presented at 500-msec intervals, a 2-sec delay, and then a test series. Both the /bi/ and /di/ adaptors were employed, providing a total of 10 identification responses per stimulus and 12 responses per discrimination pair for each adaptor condition.

Following completion of the adapted labeling and discrimination conditions, postadaptation recognition thresholds were measured using the ascending and descending method of adjustment tasks. Each threshold determination was preceded by an adapting sequence of 184 stimuli. The /bi/ adaptation condition was conducted first. Ten ascending recognition thresholds were obtained for each stimulus. Upon completion of the ascending method of adjustment for all stimuli, 10 descending recognition thresholds were obtained for each stimulus. In sessions following completion of the /bi/ condition, the same thresholds were obtained following /di/ adaptation.

RESULTS

The labeling data (percent /bi/ identification) under the no-adaptor, the /bi/-adaptor, and /di/-adaptor conditions for each of the four subjects are shown in Table 2. For all subjects,¹ adaptation by endpoint stimuli produced significant shifts in the general location of the category boundary toward the end of the continuum containing the adapting stimulus. Although not shown, the peaks in discriminability for the unadapted condition corresponded in location to the category boundary for each ob-

Table 1
Frequency Parameters for Stimuli /bi/-/di/

Stimulus Value	F1	F2	F3
Starting Frequencies			
0	200	1467	2181
1	200	1543	2345
2	200	1623	2520
3	200	1695	2690
4	200	1770	2870
5	200	1848	3019
6	200	1916	3199
7	200	2001	3365
8*	200	2075	3540
Ending Frequencies			
	361	2313	3019

Note—Fundamental frequency fell linearly from 154-115 Hz; all stimuli had an initial burst of noise of 24 dB; Stimulus Value 0 was the most /bi/-like stimulus, and Stimulus Value 7 was the most /di/-like stimulus. *Stimulus Value 8 was not used following pilot work that indicated subjects did not perceive this stimulus as /di/.

Table 2
Percent of "bi" Label as a Function of Subject, Stimulus, and Adaptor

Stimulus	Observer 1			Observer 2			Observer 3			Observer 4		
	Adaptor			Adaptor			Adaptor			Adaptor		
	None	bi	di	None	bi	di	None	bi	di	None	bi	di
0	100	100	100	100	100	100	100	100	100	100	100	100
1	100	100	100	100	100	100	100	90	100	93	80	100
2	95	30	100	100	90	90	100	80	100	93	40	100
3	73	40	100	100	40	100	98	80	100	80	50	100
4	48	10	80	88	0	100	53	70	80	40	0	100
5	0	0	10	0	0	60	8	10	30	0	0	30
6	0	0	10	0	0	0	8	0	20	0	0	30
7	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	10	0	0	0	0	0	0	0	0	0

server, with significantly decreased discriminability within categories. The locations of these unadapted identification boundaries and discrimination peaks were similar to those typically claimed to demonstrate categorical perception. Adaptation with the /bi/ stimulus (but not with the /di/ stimulus) produced shifts in the discrimination peak corresponding in location to the shifted category boundary for the same condition.²

The recognition thresholds obtained both prior to and after adaptation with both /bi/ and /di/ adaptors for the observers are shown in Table 3. The recognition thresholds obtained under each condition are the result of averaging the ascending and descending method of limits data for the observer under that given condition. The within-observer differences in mean thresholds across the ascending and descending series for each condition was never greater than 5.9 dB and, for the vast majority of the stimulus conditions, the data were all within ± 1 dB. In addition to the high intrasubject reliability, the trends in the data across subjects were quite similar.

The data in Table 3 indicate that, in the unadapted condition, the recognition thresholds for Observers 1 and 2 appear to rise slightly as one progresses from

stimulus value 0 to stimulus value 7 along the continuum, while the remaining two observers produced essentially flat threshold functions. For the /bi/ adaptor condition, the thresholds for all stimuli labeled "di" [and some stimuli labeled "bi" (Observer 3 only)] following adaptations are lower than the threshold observed in the unadapted condition. For the /di/ adaptor condition (which was run after the /bi/ threshold condition), all postadaptation thresholds for all stimuli are lower than the baseline thresholds. Both of these trends were significant. Separate analyses of variance by observer revealed significant differences among adaptor conditions and, for three of the four observers (Observers 1, 2, and 4), significant differences among stimuli. An analysis of variance of thresholds averaged across subjects indicated significant differences due to adaptor condition [$F(2,72) = 47.44$ ($p < .01$)] and due to stimulus value [$F(7,72) = 4.31$ ($p < .025$)] but no significant interaction.

DISCUSSION

The effects of adaptation on the recognition thresholds of the stimuli employed suggest that the selective

Table 3
Masked Thresholds (in Decibels SPL) Following Adaptation for Each Subject, Stimulus, and Adaptor

Stimulus	Observer 1			Observer 2			Observer 3			Observer 4		
	Adaptor			Adaptor			Adaptor			Adaptor		
	None	bi	di	None	bi	di	None	bi	di	None	bi	di
0	37.6	36.5	34.4	35.5	35.3	33.9	37.7	39.1	34.6	38.8	40.5	37.3
1	38.2	38.4	34.1	36.5	36.9	34.6	39.3	40.9	35.5	41.4	42.0	36.7
2	40.5	37.7	36.6	37.3	37.3	34.1	40.8	41.5	36.3	43.9	40.1	36.3
3	41.1	37.9	36.4	37.1	38.9	34.9	41.2	41.1	35.4	44.5	38.9	37.7
4	42.9	37.7	37.8	40.9	39.0	34.8	43.0	42.5	37.0	45.7	38.9	37.9
5	42.4	37.1	37.6	43.3	40.3	36.5	42.4	41.3	38.6	42.8	36.3	39.4
6	41.4	38.9	38.8	44.2	40.3	37.4	41.7	40.1	38.8	41.7	36.2	39.1
7	43.1	38.4	36.4	46.3	39.3	37.3	39.8	38.5	37.2	40.1	36.2	36.6
8	43.7	38.9	37.8	47.5	37.1	37.5	37.8	37.3	35.0	38.8	35.9	35.3

Note—Data are averaged across 10 ascending and 10 descending estimates per subject and condition.

adaptation procedure functions to lower thresholds for most, if not all, the members of the continuum, indicating that our observers have become more sensitive to the parameters of the stimulus continuum. These findings are not consistent with the predictions made by the major classes of current selective adaptation models. If selective adaptation involved significant fatigue effects, whether in general at a peripheral level (Simon & Studdert-Kennedy, 1978) or specific to a feature detector (Eimas & Corbit, 1973), then we would expect to find a decreased sensitivity (at least for stimuli near the adaptor) and thus an elevation in threshold. Instead, we found either minimal change in threshold near the adaptor and a significantly reduced threshold for other stimuli (following adaptation with /bi/) or overall decreased thresholds (following adaptation with /di/) for the stimuli within the labeling category of the adaptor. Response contrast explanations are based upon changes in the decision criteria and, in signal detection terminology, involve only changes in β (Diehl, Elman, & McCusker, 1978). If selective adaptation effects were due to only a type of response contrast, we would not expect to find any change in sensitivity following adaptation and, clearly, not an overall improvement in sensitivity as exhibited by significantly decreased thresholds. If selective adaptation was due to the establishment of a stimulus reference or anchor at the adaptor (Pastore, 1981), we would expect to find improved sensitivity (and thus a lowering of threshold), primarily for stimuli near the adaptors. Instead, we found significantly lower thresholds for the nonadapted category. Thus, response contrast, fatigue, and even stimulus contrast, acting alone, cannot be easily reconciled with our results.

Although our results would seem to be inconsistent with these current notions concerning the nature of selective adaptation, we can offer several untested hypotheses concerning the relationship between our threshold results and the results for identification and discrimination conditions in our study (and more typical of the selective adaptation literature) which might allow for the validity of those current notions. It may be that the perceptual and decision processes that determine recognition threshold performance are not the same as those which determine identification and discrimination performance. This separate process hypothesis must assume significant orthogonality among these underlying processes to account for the given pattern of obtained results.

A weak version of this hypothesis is based upon the idea that selective adaptation effects are not due to a single process, but rather are the result of several processes; this multiple-cause notion has been proposed by a number of researchers, including Diehl et al. (1978). As noted by Simon and Studdert-Kennedy (1978), the various processes may contribute, to dif-

ferent degrees, to measured performance, depending upon the given experimental paradigm and continuum. Thus, the weighted importance of the various processes may be quite different for labeling and recognition thresholds; this conjecture, however, is at the expense of parsimony.

A specific version of the weak hypothesis is based upon the notion that most processing of natural speech stimuli is in the context of a high degree of variability among the different stimuli which typically all represent a single speech category. Most speech tasks therefore represent high uncertainty situations. With a synthetic CV continuum based upon a single vowel, the variability among stimuli is significantly reduced. The adapting stimulus then provides the subject with information concerning the exact nature of the vowel for all stimuli plus the nature of the specific initial consonant defined by the adaptor. Thus, selective adaptation reduces uncertainty about the parameters of a specific consonant and about the parameters of the vowel defining the continuum. The vowel typically is more intense than the initial consonant; in a masked recognition threshold task, the reduced uncertainty about the parameters of the vowel probably are most important in extracting the CV syllable embedded in the noise. The reduced uncertainty concerning the vowel should apply to the whole stimulus continuum. The labeling and discrimination tasks both involve comparison based primarily upon the consonant portion of the stimuli; the vowel is constant and there is minimal background noise. With uncertainty in the labeling and discrimination tasks limited largely to the consonant, the major effect of the adapting procedure could be to reduce uncertainty concerning the adapting consonant. This would reduce stimulus uncertainty for the portion of the continuum which includes the adaptor, but not for the remainder of the continuum. This stimulus uncertainty model is basically a stimulus contrast explanation which assumes that the recognition threshold may not be tapping exactly the same underlying processes as the more standard labeling and discrimination tasks. The notion of stimulus uncertainty effects is not new (e.g., Cole & Scott, 1974; Pastore & Sorkin, 1971; Tanner, 1958). The specific uncertainty hypothesis might be tested by adapting with C-V1, then measuring recognition thresholds with C-V2; we would expect no adaptation effects on the recognition thresholds. However, we shall leave the evaluation of these and other possible hypotheses to the future.

REFERENCES

- COLE, R. A., & SCOTT, B. Toward a theory of speech perception. *Psychological Review*, 1974, **81**, 348-374.
 DIEHL, R. L., ELMAN, J. L., & MCCUSKER, S. B. Contrast effects

- on stop consonant identification. *Journal of Experimental Psychology: Human Perception & Performance*, 1978, 4, 599-609.
- EIMAS, P. D., & CORBIT, J. D. Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, 1973, 4, 99-109.
- MILLER, J. L., & EIMAS, P. D. Studies on the selective tuning of feature detectors for speech. *Journal of Phonetics*, 1976, 4, 119-127.
- PASTORE, R. E. Possible psychoacoustic factors in speech perception. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech*. Hillsdale, N.J.: Erlbaum, 1981.
- PASTORE, R. E., & SORKIN, R. D. Adaptive auditory signal processing. *Psychonomic Science*, 1971, 23, 259-260.
- SAWUSCH, J. R., & PISONI, D. B. Simple and contingent adaptation effects for place of articulation in stop consonants. *Perception & Psychophysics*, 1978, 23, 125-131.
- SEKULER, R. W., & GANZ, L. Aftereffect of seen motion with a stabilized retinal image. *Science*, 1963, 139, 419-420.
- SIMON, H. J., & STUDDERT-KENNEDY, M. Selective anchoring and adaptation of phonetic and nonphonetic continua. *Journal of the Acoustic Society of America*, 1978, 64, 1338-1357.
- TANNER, W. P. What is masking? *Journal of the Acoustic Society of America*, 1958, 30, 919-921.
- WEISSTEIN, N. What the frog's eye tells the human brain: Single

cell analyzers in the human visual system. *Psychological Bulletin*, 1969, 72, 157-176.

NOTES

1. The labeling function for Observer 3 with the /bi/ adaptor exhibited an overall shift in the labeling function toward the /bi/ adaptor. If the 50% labeling boundary is estimated by interpolating only between two points that bridge 50%, the labeling boundary for the /bi/ adaptor is 4.3, which is shifted toward /di/ relative to the 4.1 unadapted labeling boundary. If the labeling boundary is estimated from a least squares fit of the z-transform of the labeling probabilities between 0.01 and 0.99, the adapted and unadapted boundaries are 3.6 and 4.4, respectively, and thus exhibit the expected adaptation effect. While the former boundary estimation method is typical of the categorical literature, the latter is typical of the psychophysical literature and has greater statistical reliability.

2. These data may be obtained by writing the second author.

(Manuscript received December 13, 1982;
revision accepted for publication April 11, 1983.)