Perception & Psychophysics
1982, 31 (4), 307-314

Phonetic prototypes

ARTHUR G. SAMUEL
Bell Telephone Laboratories, Murray Hill, New Jersey

Speech perception may be viewed as a phonetic categorization task in which the listener assigns incoming sounds to various phonetic categories. The present experiment tests two classes of models of phonetic categorization: (1) models in which the listener has a threshold or boundary between alternative categories vs. (2) models in which the listener compares the input to prototypical representations of the alternative phones. In a pretest, listeners located on a VOT continuum the /ga/ that they thought was prototypical. Selective adaptation was then conducted using both the selected prototype and adaptors nearer and further from the phoneme boundary. The prototype adaptor produced more adaptation than the other adaptors. This result supports a prototype-based representation for phonetic categorization; several process models using such a representation are considered.

Most current theories of speech perception can be characterized as multilevel models in which the acoustic input undergoes several transformations or reencodings. Although there is some dispute over the exact nature of the first transformation, in most models the acoustic information is first encoded into units of approximately phonetic size. The present study is concerned with the psychological representation of units at this level.[1] Following Samuel and Newport (1979), these units will be referred to as *phonetic prototypes.*

This choice of terminology derives from Rosch's (e.g., 1975) work on people's categorization of colors and real-world objects. Previous approaches to the categorization problem focused on the boundary between two categories. Rosch, on the other hand, argued that people tend to categorize on the basis of *foci*; each category (e.g., "bird") is represented by some prototypical member (e.g., "robin") or a collection of prototypical properties (e.g., "wings," "feathers"), and categorization is based upon some distance metric from alternative prototypes. The boundary between categories thus becomes epiphenomenal.

The present study explores the possibility of applying this theoretical framework to the perception of speech; one way to view the speech perception task is as a series of phonetic categorizations. In fact, the identification task which has dominated speech research since its inception is an explicit categorization task. The almost exclusive concern with phoneme boundaries in these studies reflects an implicit acceptance of a categorization process based on boundary criteria. If Rosch's framework is applicable to phonetic categorization, this emphasis must be misplaced, since category boundaries are really epiphenomenal.

Experimental support for prototype-based phonetic processing is presently quite equivocal. Miller (1977) and Repp (1976) provided some support with the dichotic listening paradigm. Both investigators found that CV syllables whose consonant was a "good" (i.e., central) category exemplar competed more effectively than poor exemplars (i.e., boundary stimuli) in the dichotic situation. This result would occur if poor exemplars were engaging prototypes on both sides of the phoneme boundary. However, a boundary-based model can also predict the observed result. If categorization is based on distance from a critical (boundary) value or threshold, boundary exemplars would have less phonetic "strength" than more extreme exemplars. Thus, the dichotic data are consistent with a prototype model, but not decisive.

The data from Samuel's (1977) study of voice onset time (VOT) discrimination provide somewhat stronger support for a prototype model. Samuel trained subjects extensively on the ABX discrimination task. Initially, the usual inverted V pattern of discrimination was observed—discrimination within a phonetic category was near chance, and between category was quite good. With training, the discrimination functions took on a W shape—performance improved substantially at the category ends, but hardly improved at all near the category centers. Samuel concluded that the remaining discriminability dips were due to exemplars near the prototype being assimilated by the prototype; since all three members of the ABX triad would be heard as the prototype in this case, discriminability would be poor. With training, items further from the prototype would be discriminable from the prototype, and thus discriminable from each other.

I would like to thank Donna Kat, Dave Rumelhart, and Mark Liberman for their helpful discussions of the issues considered in this paper, and Osamu Fujimura for his suggestions on how to improve an earlier draft. Requests for reprints should be sent to Arthur G. Samuel, Department of Psychology, Box 11A Yale Station, Yale University, New Haven, Connecticut 06520.

0031-5117/82/040307-08$01.05/0

In a recent study, Oden and Massaro (1978) also argued for prototypical representation. They factorially varied voicing (voice onset time) and place of articulation of synthetic syllables, and showed that in listeners' identification of these stimuli, the two features could be processed independently. From this, they inferred a prototypical representation, with each feature being an independent component of the prototype. However, boundary-based decisions at the feature level could also produce the observed results. As in the dichotic studies, the data are in accord with a prototype-based system, but do not dictate one; many models, including boundary-based ones, can account for the Oden and Massaro results.

We are thus in the position of having a model that is consistent with the known data, but not dictated by them. The present study is intended to provide a more definitive test of the prototype model. The basic idea is that if phonetic prototypes exist, then, with certain processing assumptions, they should behave differently from nonprototypical items in the selective adaptation paradigm. If they do, support is generated for both the prototype model and the processing assumptions.

Several different models of selective adaptation have been suggested. The fatigue model of Eimas and Corbit (1973) is one version of a prototype theory. Eimas and Corbit argued that the voiced and voiceless categories should each be represented by a processing mechanism whose sensitivity varies as a function of VOT. In particular, they suggested that each mechanism's sensitivity was normally distributed, and that their ranges overlapped; the phoneme boundary was taken to be the crossover point. In their model, adaptation reduces the sensitivity of the affected processor across its range, changing the crossover point of the distributions. Since the processor's sensitivity is normally distributed, this model predicts that adaptation with the modal value (the prototype) should be more effective than adaptation with any other token.

Two different adaptation models with similar prototypical representations have been considered by Ainsworth (1977) and Cole and Cooper (1977). In one model, repeated presentation of a syllable "retunes" the location of the appropriate distribution so that it is centered on the adaptor. Ainsworth rejected this model because it incorrectly predicts that adaptors between the phoneme boundary and the prototypes should act like members of the opposing category; furthermore, it predicts no effect of adaptation with the prototype itself. Cole and Cooper tested a retuning model with similar problems. They considered a model in which adaptation produces a narrowing of the distribution of the relevant processor. This model predicts no adaptation effect with the prototype, and appropriate effects for other category members. Cole and Cooper found that adaptation with a combination of different category members was effective, and thus rejected the version of this model class that requires a single repeated stimulus. However, versions without this constraint have not been eliminated empirically.

Helson's (1964) adaptation-level theory is a clear example of a boundary-based rather than prototype-based theory. In this model, subjects establish a boundary between two categories by taking some weighted average of the stimuli that they hear and classify stimuli on the basis of the current boundary. In the adaptation paradigm, many exemplars from one category are presented, shifting the computed boundary. This boundary model predicts that the further the adaptor is from the boundary, the stronger the adaptation should be.

A final class of models attributes adaptation to contrast effects. In Diehl, Elman, and McCusker's (1978) version, marginal stimuli are categorized differently (contrastively) in the context of a clear member of either category. It is not clear from their description whether a prototypical item or one more extreme than it should produce more adaptation; both should be more effective than an adaptor near the boundary.

If one assumes that prototypes provide the best contrast, then the contrast model, along with the fatigue model, predicts more adaptation with the prototype than with less central category members. The retuning models predict *less* adaptation by the prototype; all of the prototype models predict that the results produced by the prototype will be *different* from those produced by nonprototypical items. In contrast, a boundary-based model predicts no special role for a prototypical adaptor. The simplest possible boundary model would predict no difference between any adaptors that are on one side of the boundary. For example, a 40 msec VOT /ga/, a 20 msec VOT /ga/, and a 0 msec VOT /ga/ would be equally good as adaptors if the boundary were 58 msec VOT. The more sophisticated adaptation-level theory predicts that the 0-msec /ga/ should be the best adaptor, followed by the 20-msec and then the 40-msec tokens. Thus, under both process models, the boundary-based system predicts no special role for the central or prototypical adaptor.

Miller's (1977) dichotic study included several adaptation conditions which bear on these predictions. In this study, a "good" voiced consonant (/ba/ or /da/) was dichotically presented with voiceless ones (/pa/ or /ta/) of variable "goodness"; the voiceless stimuli had VOTs between 30 and 60 msec. The finding cited earlier was that voiceless stimuli near the phoneme boundary (the 30-40-msec items) did not compete very well dichotically. In the adaptation conditions, marginal (35-40 msec VOT), medium (45-50), and extreme (55-60) members of the voiceless stimuli served as adaptors. Miller found the degree of adaptation directly followed the VOT of the adaptor—the most extreme adaptor worked best. On the face of it, this might be taken as support for a boundary model. However, the stimuli used do not nearly span the voiceless range; the most extreme members (55-60 msec) are not quite as extreme as the *average* values for voiceless stops (Lisker & Abramson, 1964: /pa/ = 58 msec, /ta/ = 70 msec VOT). Thus, the strong effect of the extreme adaptors was interpreted by Miller as supporting the fatigue of a unit centered in the 60-msec range. She also noted that, for a few subjects, the most extreme voiceless CVs did not compete dichotically quite as well as slightly less extreme items. This could be due to those subjects' having somewhat lower prototypes.

McNabb's (Note 1) study complicates the picture somewhat. In this study, the test continuum varied in place of articulation (/ba/ to /da/) rather than voicing. The three tokens on the /ba/ side of the seven-item continuum were used as adaptors. In addition to the simple identification responses, McNabb collected confidence ratings. She found that although all three /ba/s produced equivalent labeling shifts (measured by identification), they differed on the confidence ratings. The most extreme /ba/ produced the largest rating shift, followed by the one closest to the /ba/-/da/ boundary; the middle /ba/ produced the smallest effect. This pattern is the one predicted by the retuning (sharpening) version of the prototype theory.

The two studies which might resolve the prototype-boundary issue thus yield opposite results. The problem with the studies that makes their interpretation difficult is that they have no external measure of where each subject's prototype should be, or even where on the continuum the "group prototype" lies. In the present study, this problem is overcome by conducting a pretest to measure where each subject's putative prototype is. For each subject, adaptation is then conducted with the prototype, with a token closer to the phoneme boundary and with a token further from the boundary. The goal is to see whether the prototype produces an adaptation effect different from that of its neighbors, and if so, whether the effect is larger or smaller.

METHOD

Subjects

Ten women from the Murray Hill, New Jersey, area served as subjects. All were native English speakers (average age approximately 45 years), and none reported any hearing problems. The subjects were paid for their participation.

Stimuli

A synthetic /ga/-/ka/ continuum was generated on a five-formant cascade/parallel terminal analog synthesizer (Klatt, 1980). Voice onset time was varied by delaying the onset of the first formant and energizing the upper formants with a band-passed noise source rather than a buzz source for the desired VOT. Forty-one different tokens were generated, varying in 3-msec steps from 0 to 120 msec VOT. All stimuli were 300 msec long.

All tokens ended with the vowel /a/ (F1-F5 = 770, 1,230, 2,500, 3,100, and 3,700 Hz). The consonantal transitions were as follows: F1—200-620 in the first 30 msec, 620-770 in the next 30 msec; F2—1,700 for 10 msec, 1,700-1,324 in the next 30 msec, 1,324-1,230 in the next 30 msec; F3—2,100 for 10 msec, 2,100-2,500 in the next 70 msec; F4—3,300 for 10 msec, 3,300-3,100 in the next 60 msec; F5—3,700 throughout.

The fundamental frequency for all stimuli was 125 Hz at voicing onset, and remained so through the first 175 msec. From that point, it fell linearly to 95 Hz at syllable offset. The nominal voicing amplitude was 60 dB from voicing onset through the first 190 msec of the syllable, fell linearly to 56 dB at 270 msec, then fell linearly to 0 at offset.

For the unvoiced portion of each syllable, the bandpassed noise amplitude was 60 dB at onset, and fell linearly to zero at voicing onset. The bandwidth of the first formant was set to 100 Hz during the unvoiced portion, and changed to 50 Hz at voicing onset. The bandwidths of F2 and F3 were fixed at 70 and 110 Hz, respectively.

Apparatus

Stimuli were stored digitally on disk file. They were output through a 12-bit D/A converter (10 kHz), amplified, low-pass-filtered (5 kHz), and played binaurally over stereo headphones. The subjects heard the stimuli in a small computer room and responded by pushing buttons on a keyboard. All experimental events were controlled by an SEL 75 computer.

Procedure

Each subject participated in six sessions of approximately 45 min. The first three sessions constituted a pretest, which was used to estimate each subject's prototypical /ga/. The last three sessions constituted the experiment proper.

The pretest. Since people presumably vary in the exact location along the VOT continuum of their prototypical /ga/, it is necessary to obtain an estimate of each subject's prototype. The first three sessions were used to obtain these estimates. Each of these sessions included two different experimental tasks, the *identification (ID) task* and the *overt prototype task.*

On the identification task, a subset of the /ga/-/ka/ continuum was used; the lowest VOT stimulus was 3 msec, and the highest was 117 msec. Between these endpoints, tokens were in 6-msec steps (3, 9, 15, 21, . . . 111, 117). This arrangement provided a finer grid than is usually used, as well as a larger range than is customary. The test consisted of 10 random permutations of the 20 test syllables.

On each trial, the subjects heard one syllable and pressed one of two keys to indicate whether the syllable sounded more like /ga/ (left key) or /ka/ (right key). They were informed that reaction times would be measured, and told to respond quickly and accurately. After the subject responded, the computer waited 1 sec and then presented the next item. The identification test took approximately 7 min.

On the overt prototype task, subjects were asked to locate the "best" /ga/ several times. To clarify these instructions, the "best" was explained as "which of the syllables should be presented to someone else as an example of /ga/, if only one was to be presented."

Each trial began with the presentation of a randomly chosen starting syllable, selected from the 0-57-msec VOT range. The subject used three buttons to move through the continuum: The left button played a syllable 3 or 6 msec VOT less than the one just heard, the middle one played the syllable just heard, and the right button played a syllable 3 or 6 msec VOT greater than the one just heard. For the right and left buttons, it was randomly determined whether the jump would be 3 or 6 msec VOT; this was to interfere with a simple button-push counting strategy from a recognizable point, such as the phoneme boundary. Using these buttons, the subject could step through the continuum, listening to the syllables until the "best" /ga/ was found. At the continuum endpoints, movement in only one direction was possible: At 0 msec VOT, the middle and left buttons both produced a repetition of the 0 stimulus; at 120, the middle and right buttons were similarly equivalent. When satisfied, the subject ended the trial by pushing a button on the keyboard to indicate that the last syllable heard was the "best." A recorded message told the subject to begin the next trial by pushing the space bar. The overt prototype task consisted of 15 such trials, and took approximately 20 min.

The first session of the pretest consisted of the identification test followed by the overt prototype test. In the second and third sessions, an additional pass through the identification test was included after the overt prototype task. By comparing performance on the two ID tests, we can assess whether potential selective adaptation from the prototype test itself caused a systematic underestimation of subject's prototypes.

The main experiment. The main experiment consisted of three adaptation sessions conducted at weekly intervals. For each subject, one adaptor was chosen to be near the prototype, one was nearer to the /ga/-/ka/ boundary, and the third was further from the boundary than the prototype. On the basis of the pretest, the subjects were divided into two groups, those with "best" /ga/s between 0 and 12 msec VOT and those with prototypes between 15 and 24 msec VOT. For the lower prototype subjects, the three adaptors were at 0, 9, and 18 msec VOT; for the others, the adaptors were at 9, 18, and 27 msec VOT. Within each group, three subjects' sessions were in the order: (1) most extreme adaptor, (2) prototype adaptor, and (3) boundary adaptor; the other two subjects' sessions were run in the reverse order.

Each adaptation session consisted of a *baseline ID* test and an *adaptation* test. The baseline ID was identical to the identification test of the pretest, except that it included 15 passes through the stimuli rather than 10. The adaptation test was identical to the baseline, except that an adapting syllable was interspersed. The adaptor was played 100 times at the beginning of the test and 45 times (at 1.5 repetitions per second) before each block of 10 test syllables. On both the baseline and adaptation tests, the subjects were instructed to label the test syllables quickly and accurately. Each session lasted approximately 45 minutes.

RESULTS

The Pretest

In each of the three sessions of the pretest, the subjects made 15 choices of their "best" /ga/ by converging on it. The medians for each session were computed, and each subject's prototypical /ga/ was operationally defined as the median of these three scores. Using this measure, the estimated prototypes for the 10 subjects fell between 0 and 24 msec VOT. The five subjects with prototypes between 0 and 12 msec VOT were put into the *low-prototype* group; the others constituted the *high-prototype* group. The second column of Table 1 presents the subjects' prototypes.

The third column of Table 1 presents estimates of each subject's /ga/-/ka/ boundary. These values are the (interpolated) 50% /ka/ points from the identification test given at the beginning of Session 3 of the pretest. Two points should be made with respect to these values. First, the subjects' /ga/-/ka/ boundaries are somewhat higher than those reported by Lisker and Abramson (1964), no doubt due to different characteristics of the synthesis. More importantly, there is *no* correlation between the estimated prototypes and boundaries. This is very important, because if there are differences observed between the low and high prototype subjects in the main experiment, they cannot be attributed to confounded boundaries.[2]

Table 1
Prototypes and Phoneme Boundaries, Pretest

Subject	Prototype	Boundary
L1	0	59
L2	3	62
L3	6	73
L4	9	62
L5	12	65
H1	15	63
H2	18	62
H3	18	62
H4	21	64
H5	24	62

Note—The prototype and boundary values are in milliseconds VOT.

The Main Experiment

Table 2 presents the data from the three adaptation sessions for each subject in the main experiment. Each value is the difference between the percentage of stimuli labeled /ka/ on the adaptation and preceding identification tests. The parenthesized values include data from stimuli throughout the test series; the unparenthesized numbers are based on the labeling of stimuli near the phoneme boundary (stimuli with VOTs from 45 to 75 msec). As the table shows, the prototype produced more adaptation than did the adaptors on either side of it. Two-way analyses of variance (subject group × adaptor) for the two mea-

Table 2

Subject	Change in Percent Labeled /ka/		
	Low Adaptor	Prototype	High Adaptor
L1	4(1)	24(8)	9(3)
L2	18(6)	30(10)	26(8)
L3	19(7)	22(7)	20(7)
L4	5(1)	17(5)	11(3)
L5	8(2)	14(4)	7(3)
H1	16(5)	30(9)	26(9)
H2	7(2)	13(4)	6(1)
H3	21(5)	17(5)	23(7)
H4	10(3)	23(6)	29(9)
H5	5(0)	12(3)	23(7)
Mean	11.3(3.2)	20.2(6.1)	18.0(5.7)

Note—The shifts are the differences (in percent labeled /ka/) between the adaptation and baseline tests in each condition. For the low-prototype subjects, the adaptors were 0, 9, and 18 msec VOT. For the high-prototype subjects, the adaptors were 9, 18, and 27 msec VOT. The parenthesized values are based on all of the test items; the unparenthesized numbers only include data from stimuli near the phoneme boundary (45-75 msec VOT).

sures confirmed the main effect of adaptor [overall, $F(2,16) = 9.44$, $p < .005$; boundary, $F(2,16) = 4.60$, $p < .03$]. Neither the main effect of group [both overall and boundary $F(1,8) < 1$] nor the interaction [overall, $F(2,16) = 2.56$, $p > .10$; boundary, $F(2,16) < 1$] was significant. The difference between the effect of the prototype and the "low" adaptor (0 msec VOT for the low group, 9 msec VOT for the high) was highly reliable [overall, $t(9) = 4.41$, $p < .001$; boundary, $t(9) = 4.19$, $p < .005$]. Although it was in the right direction, the difference between the prototype and the high adaptor (18 or 27 msec VOT) did not reach significance [overall, $t(9) = .45$, n.s.; boundary, $t(9) = .90$, n.s.]. This failure to reach significance is not very problematic, however. There is no doubt that if the "high" adaptor were made high enough, its efficacy would drop (as in Miller, 1977); eventually it begins to sound like /ka/. Thus, the critical comparison is the one between the prototype and the "low" adaptor. None of the boundary models outlined in the introduction can predict that an adaptor *further* from the boundary will have a greater effect than one closer, while exactly that result is expected given the fatigue (and possibly the contrast) prototype theories.

The similarity of the results for the boundary and overall measures suggests that the stimuli far from the boundary are contributing little to the adaptation effects. Table 3 presents the adaptation shifts for these stimuli, broken down into stimuli from the /ga/ range (3-33 msec VOT) and the /ka/ range (87-117 msec VOT). As suggested, these shifts are negligible; all of the adaptation effect shows up in the boundary stimuli (Table 2). A three-way analysis of variance (adaptor × continuum range × subject group) confirmed that the shifts differed as a function of range [$F(2,16) = 75.67$, $p < .001$].

It would be tempting to conclude that adaptation operates only upon the boundary stimuli, since all of the labeling shifts are observed at the boundary. In fact, Simon and Studdert-Kennedy (1978) make just such a claim: "Most speech adaptation studies, including the present one, have found that response shifts are confined to a few boundary stimuli toward the adaptor end of the continuum. Certainly none has shown a shift of the entire response distribution" (p. 1352). The problem with this claim is that in most studies, including the present one, identification of items away from the boundary is at ceiling, obscuring any adaptation effects that may occur. The data in Table 4 show quite clearly that such effects do, in fact, occur. The values in the table are the differences in reaction time to label stimuli in the three ranges just discussed; the differences have been normalized such that each subject's times sum to zero.[3] The cen-

Table 3
Adaptation Shifts for the Three Adaptation Conditions, Broken Down by VOT Continuum Range

	VOT Continuum Range					
	Low Adaptor		Prototype		High Adaptor	
Subject	ga	ka	ga	ka	ga	ka
L1	0	0	0	0	0	0
L2	0	3	0	3	2	−1
L3	0	5	0	1	1	2
L4	−1	0	1	0	0	0
L5	0	0	0	0	0	0
H1	0	0	0	0	0	0
H2	0	0	1	0	−1	0
H3	−1	−2	0	−2	0	2
H4	0	0	0	−2	0	0
H5	−2	0	0	0	0	0
Mean	−.4	.6	.2	.0	.2	.3

Note—The shifts are the differences (in percent labeled /ka/) between the adaptation and baseline tests in each condition. For the low-prototype subjects, the adaptors were 0, 9, and 18 msec VOT. For the high-prototype subjects, the adaptors were 9, 18, and 27 msec VOT. The ga range was 3-33 msec VOT, and the ka range was 87-117 msec VOT.

Table 4
Reaction Time Changes as a Function of Adaptation Condition and Continuum Range

	Low Adaptor			Prototype			High Adaptor		
Subject	ga Range	Boundary	ka Range	ga Range	Boundary	ka Range	ga Range	Boundary	ka Range
L1	42	−6	−37	20	33	−53	54	12	−67
L2	39	−36	−3	37	−44	7	30	−34	3
L3	99	7	−107	37	74	−111	−52	125	−74
L4	−42	57	−15	59	−65	6	37	9	−46
L5	−19	8	11	9	−16	8	−10	48	−39
H1	28	5	−32	27	−16	−11	43	5	−47
H2	26	16	−42	44	14	−58	21	33	−55
H3	−29	21	9	15	−12	−4	−25	16	8
H4	19	−11	−7	16	−25	9	41	−33	−9
H5	26	37	−64	18	−27	9	58	−9	−50
Mean	19	10	−29	28	−8	−20	20	17	−38

Note—The values are reaction time differences between the adaptation and baseline tests in each condition, normalized such that the changes sum to zero.

tral result is that after adaptation subjects take about 50 msec longer to label the /ga/ stimuli than they do to label the /ka/ stimuli; boundary stimuli are relatively unchanged. An analysis of variance analogous to the one performed on the identification shifts confirmed the reliability of this range effect [$F(2,16) = 5.90$, $p < .02$]. It is clear that these data rule out any model in which adaptation operates only on boundary stimuli. Taken together, the data in Tables 2, 3, and 4 indicate that adaptation has effects along the entire VOT range (or at least throughout the /ga/ range).

The Pretest Revisited

The overt prototype task used in the pretest has the virtue of having high face validity; it is the most direct method of estimating a subject's prototype. It also proved to be the most reliable of several tasks tested in pilot work. However, it does have a potential problem: Selective adaptation could be occurring, biasing the estimated prototypes. The nature of the adaptation is complex, because the "adaptors" are the syllables that are heard in zeroing in on the prototype. As such, they will almost all be /ga/s, with a frequency distribution centered around the prototype. Since models of adaptation are not well specified at this point, it is difficult to predict the impact of this procedure on the prototype choice.

The identification tests run before and after the prototype task provide a means for testing whether any adaptation occurred as a result of the prototype task. If adaptation did occur, the subjects should have reported fewer /ga/s after doing the prototype task. Table 5 presents the data that bear on this question. The numbers in the table are the differences between the percentage of items labeled /ka/ after the prototype task and before it (the data are taken from the last pretest session). The overall effect (column 2) is small, but fairly consistent (8 of 10 subjects, $p < .06$ by a sign test). A two-factor analysis of variance was conducted on the difference scores in the last three columns (subject group: low vs. high prototype $\times$ stimulus range: /ga/ vs. boundary vs. /ka/). The results of this analysis were marginal: for both the overall change [$F(1,8) = 2.77$] and the effect of stimulus range [$F(2,16) = 2.34$], the p value was .13. Thus, although there was a reduction in /ga/ responses, and the reduction was limited to /ga/s and boundary stimuli, the change was not statistically reliable, suggesting that not much adaptation occurred. If some adaptation did occur during the prototype task, the estimated prototypes may be slightly lower on the VOT continuum than the true values. Such an underestimate would occur if adaptation involved a criterion shift that was applied across the entire VOT range. Note that if the prototypes were, in fact, underestimated slightly, the "high" adaptors would be almost as close to the true values as the "prototype" adaptors. This would account for their producing an effect almost as large as the prototype adaptors.

Table 5
Adaptation Caused by Doing the Overt Prototype Task

	Change (in Percent Labeled /ka/)			
Subject	Overall	ga Range	Boundary	ka Range
L1	−3	0	−10	0
L2	3	2	7	0
L3	3	1	5	4
L4	−2	0	−5	0
L5	10	0	30	0
H1	3	10	5	−12
H2	1	13	0	−10
H3	1	1	1	0
H4	2	0	5	0
H5	4	3	15	−4
Mean	2.2	3.0	5.3	−2.2

Note—The values are differences in the percentage of stimuli labeled "ka" before and after doing the prototype task. The ga range was 3-33 msec VOT, the boundary range was 45-75 msec VOT, and the ka range was 87-117 msec VOT.

DISCUSSION

The central finding of the present study is that adaptation effects are greater with an adaptor chosen to be near a subject's putative prototype than with nonprototypical adaptors. In particular, the prototypical adaptor was more effective than one that was even further from the phoneme boundary. This result provides the strongest evidence currently available for the view that phonetic categories are represented in some sort of prototypical format; in Rosch's (1975) terms, phonetic categories have foci, and phonetic categorization is accomplished by computing a distance function from these foci.

The data are clearly inconsistent with criterion shift models such as adaptation-level theory (Helson, 1964). Similarly, both versions of "retuning" models discussed in the introduction were not supported. Although the data do constrain the nature of phonetic representation, they do not provide strong constraints for models of the adaptation process. For example, both fatigue and criterion shift models can be formulated to operate upon the prototypical representation to produce the observed pattern of results. In a fatigue model, the prototype adaptor engages the phonetic unit most strongly (cf. Eimas & Corbit, 1973), producing maximal adaptation. Similarly, a prototype adaptor is the best exemplar of the phonetic category, and, in one interpretation of Diehl, Elman, and McCusker's (1978) contrast effect model, it would produce the most adaptation. The data from the present study do place two important constraints on adaptation process models. First, as the reaction time data show, adaptation models cannot limit effects to boundary stimuli; identification of quite good category members must also be degraded (or at least delayed). This constraint is inconsistent with simple contrast effect models. Second, adaptation effects may persist for at least a few minutes. This conclusion is based upon the marginal adaptation effect (caused by the prototype task) observed in the pretest. If this persistence constraint is reliable, it rules out simple contrast models in which identification shifts are entirely attributed to a contrast between successive stimuli. A more sophisticated contrast model, perhaps a somewhat modified version of Diehl, Elman, and McCusker's model, could account for the data.

Since the reaction time data rule out models in which adaptation is limited to boundary stimuli, only processes that operate on the whole range of stimuli need be considered. The original fatigue model put forth by Eimas and Corbit (1973) satisfies this constraint—adaptation was attributed to the fatigue of a detector across its entire range. Since Eimas and Corbit conceptualized the response characteristics of the voiced/voiceless detectors as a pair of overlapping normal distributions, fatigue was represented as a reduction of the appropriate distribution. A key feature of this model is that the shape and location of the distribution are not changed, but only its size relative to the opposing distribution.

Sawusch's (1977) model of adaptation includes a peripheral processing level that is susceptible to fatigue and a central level that is sensitive to contrast effects. Simon and Studdert-Kennedy (1978) similarly argue for both processes. Clearly, since both fatigue and contrast-effect versions of prototype theories can be formulated, a model with both can be formulated as well. The results of the present study indicate that whether fatigue or contrast (or both) are invoked, the underlying representations should be prototypical rather than boundary-based.

An interesting question is whether all speakers of a language have essentially the same prototype for a given phone. The results of the pretest suggest that listeners vary slightly in the exact location of their prototypes. These small variations were apparently large enough psychologically to produce quite striking effects on the pattern of selective adaptation. This point is best illustrated by comparing the effect of the 9- and 18-msec VOT adaptors for the two groups of subjects (see Table 2). For all five low-prototype subjects, the (prototypical) 9-msec VOT adaptor was more effective than the 18-msec VOT adaptor. The pattern is almost exactly reversed for the high-prototype subjects. The (prototypical) 18-msec VOT adaptor was more effective for four of the five. An adaptor (9 vs. 18 msec VOT) by subject group (low vs. high prototypes) analysis of variance produced the interaction these data suggest [$F(1,8) = 12.83$, $p < .01$]. Thus, identical physical stimuli produce radically different results; the difference is apparently due to small individual differences in prototype location.

These results should make it clear that some measure of each individual's prototype is critical if data from several subjects are to be combined in experiments of this type. Miller (e.g., Miller & Connine, 1980) is currently conducting an extensive set of adaptation studies which avoid this problem by considering each subject's data individually. Miller reports that, for both a place of articulation continuum (/dae/-/gae/) and a manner continuum (/bae/-/wae/), the efficacy of adaptation first increases as the adaptor moves away from the phoneme boundary, then peaks, and finally decreases. This pattern is completely consistent with the results of the present study. However, the situation is not completely clear-cut. On the voiceless side of a VOT continuum, Miller finds *no* decrease in adaptation efficacy, even when the adaptor is far beyond the subject's prototype. This result is clearly inconsistent with the place and manner data and the results of the present study. The resolution of this difference must involve a complication of the simple prototype model. One possibility is that voicing classifications involve a comparison of the input stimulus to only a voiced prototype; if the input exceeds the prototype's VOT value by some threshold, it is classified as voiceless. This model would produce the observed pattern of adaptation results; it is, however, post hoc, and independent confirmation is clearly needed before its acceptance is warranted.

REFERENCE NOTE

1. McNabb, S. *Must the output of the phonetic detector be binary?* (Research on Speech Perception, Progress Report No. 2). Bloomington: Department of Psychology, Indiana University, 1975.

REFERENCES

Ainsworth, W. A. Mechanisms of selective feature adaptation. *Perception & Psychophysics*, 1977, **21**, 365-370.

Cole, R. A., & Cooper, W. E. Properties of friction analyzers for [j]. *Journal of the Acoustical Society of America*, 1977, **62**, 177-182.

Diehl, R., Elman, J., & McCusker, S. Contrast effects on stop consonant identification. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, **4**, 599-609.

Eimas, P., & Corbit, J. Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, 1973, **4**, 99-109.

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

Klatt, D. H. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 1980, **67**, 971-995.

Lisker, L., & Abramson, A. A cross-language study of voicing in initial stops: Acoustic measurements. *Word*, 1964, **20**, 384-422.

Miller, J. Properties of feature detectors for VOT: The voiceless channel of analysis. *Journal of the Acoustical Society of America*, 1977, **62**, 641-648.

Miller, J. L., & Connine, C. M. Psychophysical tuning curves for phonetically relevant acoustic information. *Journal of the Acoustical Society of America*, 1980, **67**, S52.

Oden, G., & Massaro, D. Integration of featural information in speech perception. *Psychological Review*, 1978, **85**, 172-191.

Repp, B. Identification of dichotic fusions. *Journal of the Acoustical Society of America*, 1976, **60**, 456-469.

Rosch, E. The nature of mental color codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, **1**, 303-322.

Samuel, A. The effect of discrimination training on speech perception: Noncategorical perception. *Perception & Psychophysics*, 1977, **22**, 321-330.

Samuel, A., & Newport, E. Adaptation of speech by nonspeech: Evidence for complex acoustic cue detectors. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, **5**, 563-578.

Sawusch, J. R. Peripheral and central processing in speech perception. *Journal of the Acoustical Society of America*, 1977, **62**, 738-750.

Simon, H. J., & Studdert-Kennedy, M. Selective anchoring and adaptation of phonetic and nonphonetic continua. *Journal of the Acoustical Society of America*, 1978, **64**, 1338-1357.

NOTES

1. The actual size of the unit being investigated could be as small as an acoustic/phonetic feature or as large as a whole syllable, or anything in between. For the present purpose, the exact size does not matter; the logic of the experiment holds at any of the possible levels.

2. Although the lack of correlation is advantageous methodologically, it raises an interesting theoretical question. If the boundary is determined by the "sphere of influence" of the prototype, one might expect a correlation to appear. There are at least two ways to account for the lack of correlation. First, in a prototype model, the boundary is determined by the interface of two prototypes. If the /ka/ prototype does not covary with /ga/, there will be little systematic correlation observed. Second, because the interface determines the boundary, the size of each "sphere of influence" is critical. If, for example, subjects with high /ga/ prototypes tend to have smaller /ga/ categories, no correlation will be found. These possibilities clearly warrant further study.

3. The normalization for each task was done by subtracting each subject's mean reaction time from her /ga/, boundary, and /ka/ reaction times. The normalization is necessary because the conditions of the baseline and adaptation tests are rather different (e.g., the stimuli are presented continuously in one and broken up in the other), making absolute comparisons meaningless. However, the *differences* between parts of the continuum may be meaningfully compared after normalization.

**(Manuscript received July 15, 1981;
revision accepted for publication December 6, 1981.)**