

Auditory and phonetic processes in place perception for stops

JAMES R. SAWUSCH

State University of New York, Buffalo, New York

and

HOWARD C. NUSBAUM

Indiana University, Bloomington, Indiana

Use of the selective adaptation procedure with speech stimuli has led to a number of theoretical positions with regard to the level or levels of processing affected by adaptation. Recent experiments (i.e., Sawusch & Jusczyk, 1981) have, however, yielded strong evidence that only auditory coding processes are affected by selective adaptation. In the present experiment, a test series that varied along the phonetic dimension of place of articulation for stops ([da]-[ga]) was used in conjunction with a [ska] syllable that shared the phonetic value of velar with the [ga] end of the test series but had a spectral structure that closely matched a stimulus from the [da] end of the series. As an adaptor, the [ska] and [da] stimuli produced identical effects, whereas in a paired-comparison procedure, the [ska] produced effects consistent with its phonetic label. These results offer further support for the contention that selective adaptation affects only the auditory coding of speech, whereas the paired-comparison procedure affects only the phonetic coding of speech. On the basis of these results and previous place-adaptation results, a process model of speech perception is described.

A recurring issue in speech perception research is the distinction between auditory and phonetic processes. Which aspects of listening to and recognizing speech are the result of language-specific and speech-specific processing capabilities and which reflect our general auditory processing of sound? One position on this issue is that speech is handled by a specialized speech-specific subsystem (Liberman, 1982; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Repp, 1982). According to this position, the auditory processing of nonspeech events is mediated by mechanisms that are distinctly different from the mechanisms responsible for speech perception. An alternative view is that phonetic categorization reflects the process of labeling the output of an auditory analysis of the stimulus and that this analysis reflects general auditory processing capabilities (Pastore, 1981; Schouten, 1980). Between these two relatively extreme positions are models of speech processing that incorporate both general auditory processes and language-specific phonetic processes (see Cutting & Pisoni, 1978; Pisoni & Sawusch,

1975; Sawusch, 1977a). However, regardless of the model one chooses, it is necessary to develop experimental procedures for exploring the nature of the auditory and phonetic coding of speech.

One such experimental procedure is selective adaptation. When this procedure was first used with speech stimuli (Eimas, Cooper, & Corbit, 1973; Eimas & Corbit, 1973), it was claimed that exclusively phonetic processes were being tapped. The basic results looked something like those shown in Figure 1. This figure shows the effects of adaptation on the identification of a continuum of synthetic speech stimuli. The solid line shows the categorization in the baseline condition in which subjects identified the test stimuli, presented in random order, without adaptation. The first three stimuli are rated as good examples of one category, and the last three are rated as good examples of a different category. Stimulus 4, in the middle, receives a boundary rating. When Stimulus 1, on the left, is presented repeatedly as an adaptor, the identification of the series changes, as shown by the dashed function on the left. Conversely, when Stimulus 7, on the right, is used as an adaptor, we get the dashed identification function on the right. This type of contrast effect has been found repeatedly in selective adaptation studies (see Cooper, 1979, and Eimas & Miller, 1978, for reviews). Since the original Eimas et al. (1973) work, however, a number of studies have cast doubt upon their conclusion that selective adaptation affects a

This work was supported by NIMH Grant R01MH31468 to the State University of New York at Buffalo. The authors would like to thank Thomas H. Nochajski for his assistance in running the subjects. Some of the present data were presented at the 22nd meeting of the Psychonomic Society, November 1981. Reprint requests may be sent to the first author at the Department of Psychology, SUNY/Buffalo, 4230 Ridge Lea Road, Buffalo, New York 14226.

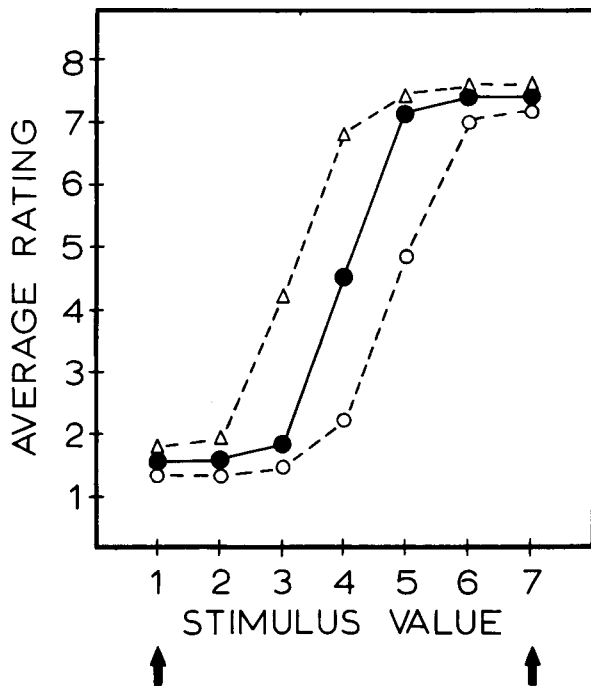


Figure 1. Baseline rating function (solid line) and typical adaptation effects for Stimulus 1 adaptation (open triangles) and Stimulus 7 adaptation (open circles).

phonetic level of processing. Several investigators have reported results that show adaptation varying as a function of the spectral overlap between adapting and test syllables (Bailey, 1975; Sawusch, 1977a). Furthermore, a number of experiments have found no adaptation effects for an adaptor and test series that share a phonetic category but do not have any common spectral characteristics (Ades, 1974; Bailey, 1975; Roberts & Summerfield, 1981; Sawusch & Jusczyk, 1981). These results have led some investigators to propose that adaptation has an entirely auditory locus for its effects (see Ades, 1976; Sawusch, 1977a).

The contrast effects found with selective adaptation can also be produced with a number of other procedures. Diehl and his co-workers (Diehl, Elman, & McCusker, 1978; Diehl, Lang, & Parker, 1980) presented stimuli in pairs to subjects. One of the stimuli in each pair was a good example of a phonetic category and the other was a stimulus near the phonetic category boundary. When subjects categorized both of the stimuli in a pair, the boundary stimulus was often placed in the phonetic category away from or opposite to the category of the exemplary stimulus in the pair. This effect was termed "response contrast" by Diehl et al. (1978), who pointed out the similarity between this result and the results of selective adaptation experiments. In terms of Figure 1, pairing Stimulus 1 with Stimulus 4 would cause the subjects to assign a rating of 6 or 7 to

Stimulus 4, whereas pairing Stimuli 7 and 4 would yield ratings of 1 or 2 to Stimulus 4. Consequently, both Diehl's procedure (which we will term the paired-comparison procedure) and the selective adaptation technique produce analogous results, making interpretation of selective adaptation results problematic.

Recently, the effects from selective adaptation and paired-comparison procedures have been dissociated for stimuli varying along the phonetic dimension of voicing. Sawusch and Jusczyk (1981) generated a syllable with an acoustic (spectral) structure that matched one end of their [ba]-[p^ha] test series and a perceived phonemic identity that matched the other end of their test series. One of the adapting syllables was a /spa/, which was labeled by subjects as containing a /p/.¹ This syllable was formed by placing an [s], followed by 75 msec of silence, in front of a +10-msec VOT [ba]. Consequently, this /spa/ syllable had a spectral structure that was identical to that of a stimulus from the [ba] end of the test series while at the same time it was identified by subjects as perceptually (phonemically) similar to the [p^ha] end of the test series.

The adaptation effects found with this /spa/ adaptor were governed by its spectral overlap with the test series. The /spa/ and the +10-msec VOT [ba] adaptors produced identical shifts in the labeling of the [ba]-[p^ha] test series (toward the [ba] end of the series). By comparison, when the /spa/ syllable was paired with an ambiguous (+30-msec VOT) test item in the paired-comparison procedure used by Diehl et al. (1978, 1980), it produced an effect similar to that of a [p^ha] syllable. Thus, in the paired-comparison procedure of Diehl et al., the identity of the /spa/ as a /p/ governed the direction of the results, whereas in the adaptation procedure, the results were found to depend upon the spectral structure of the /spa/. These results provide evidence for separate auditory and phonetic processes in speech perception. Furthermore, selective adaptation seems to affect the auditory processing of speech, whereas the paired-comparison procedure has its effects on the phonetic coding of speech.

Given the similarity of the [p^ha] and /spa/ results in the paired-comparison procedure, it seems reasonable to conclude that if any phonetic component were present in the adaptation effects of /spa/, then /spa/ should have produced smaller effects than [ba]. In the case of the [ba] adaptor, the phonetic identity and spectral structure would act together, whereas for the /spa/, effects at phonetic and auditory levels of coding would act in opposite directions. This should have reduced the adaptation effects of /spa/, relative to [ba]. The results of Sawusch and Jusczyk were very clear in that no difference between the [ba] and /spa/ adaptors was found. Thus, it would appear that there was no pho-

netic component in selective adaptation to voicing in speech.

Although the effects of selective adaptation and paired-comparison on voicing seem to be clear, their effects on the phonetic dimension of place of articulation in stops are still in need of clarification. Previous experiments on place of articulation using the selective adaptation procedure have provided evidence for the involvement of two levels of processing in selective adaptation (see Sawusch, 1977a). This result has been interpreted as indicating either that there are two distinct auditory levels of processing involved in speech perception (Sawusch, 1977a) or that both auditory and phonetic levels of processing are affected by selective adaptation to speech (Cooper, 1975, 1979). Combining the adaptation results on place of articulation reported by Sawusch (1977a) with the adaptation and paired-comparison results for voicing reporting by Sawusch and Jusczyk (1981), it would appear that the two processing operations for place of articulation that were identified by Sawusch are auditory and that there is no phonetic component to selective adaptation with speech. This conclusion is bolstered by the general failure to find any transfer of adaptation from syllable initial stops (CV) to syllable final stops (VC), reported by Ades (1974) and Sawusch (1977b), even though both CV and VC syllables contained the same stop phonemes. While this line of reasoning suggests that adaptation effects on place of articulation in stops are confined to the auditory processing of speech, it is not conclusive. Cooper (1975, 1979) has suggested that position-sensitive phonetic processes are involved in selective adaptation so that CV to VC (and VC to CV) transfer of adaptation is not expected to occur. The present experiment was designed to test this possibility directly, using the combination of selective adaptation and paired-comparison procedures previously used by Sawusch and Jusczyk (1981).

In order to distinguish between these two views, we need a stimulus with a spectral structure matched to one end of a test series and a phonetic identity that corresponds to the other end of the test series along the dimension of place of articulation in stops. Previous work by Mann and Repp (1981; Repp & Mann, 1981), on phonetic trading relations, provides us with just such a set of stimuli. On the far left and far right of Figure 2 are schematic spectrograms of syllables identified by subjects as [da] and [ga]. The difference between the [da] and [ga] syllables is in the extent of the change in the frequency of the third formant (the third formant transition). In the middle of the figure is a syllable constructed from the [da] on the left. The friction appropriate for an [s], followed by 90 msec of silence, is abutted to the beginning of the [da] syllable. Mann and Repp (1981) found that this syllable, which contains the formant transitions of an alveolar stop (i.e., [d]), was generally identified by subjects as containing the velar stop [k], which shares the velar

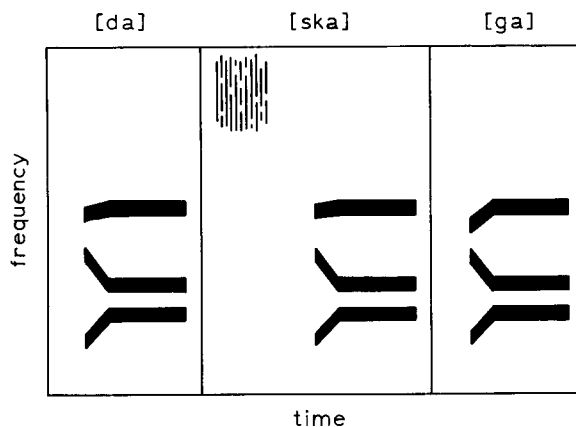


Figure 2. Schematic spectrograms for the Stimulus 3 [da] (left), Stimulus 3 [ska] (center), and Stimulus 7 [ga] (right) used in the present experiment.

place of articulation feature with [g]. In pilot testing, we also found that subjects identified this syllable as [ska] better than 70% of the time in a two-alternative, forced-choice procedure (replicating Mann & Repp, 1981). Consequently, while the spectral structure of the [ska] matches the [da] on the left of Figure 2, its phonetic feature of place of articulation matches the [ga] on the right. This [ska] syllable was used, along with [da] and [ga] syllables, in both adaptation and paired-comparison procedures, to determine the nature of the stages of processing involved in the auditory-to-phonetic coding of speech.

In addition to attempting to distinguish whether the effects of selective adaptation to place of articulation were confined entirely to the auditory processing of speech or represented effects on both auditory and phonetic processes, this experiment had three other goals. One was to attempt to reproduce, on a different phonetic continuum, the pattern of results found by Sawusch and Jusczyk (1981). If this type of pattern of results was found, it would substantially extend the generality of Sawusch and Jusczyk's conclusion that adaptation has its only influence on the auditory coding of speech. The second goal was to determine whether or not the results of the paired-comparison procedure occur at a phonological stage of processing. This question arose because the /spa/ syllable used by Sawusch and Jusczyk can be described as producing a /p/ percept on the basis of a phonological rule of English. This rule states that, in the initial position of an utterance, the stop consonant following a voiceless fricative must be voiceless. There is no phonological rule in English to change the place of articulation of the stimuli used in this study. Mann and Repp (1981) have described this change in the perception of place of articulation as a phonetic trading relation, possibly reflecting knowledge of the co-articulatory influences in speech production.

Finally, the adaptation procedure itself was slightly modified in the present experiment. Sawusch and

Jusczyk used long sequences of the adapting syllable (75 repetitions) with a short interadaptor interval (300 msec). Under these circumstances, it is possible for "streaming" to occur, perceptually dissociating the fricative [s] from the voiced [ba]. This would produce two separate streams of sound (see Bregman, 1981). If this were to happen, it would potentially destroy the perception of /p/ in the adaptor and make the experimental results difficult to interpret. While Sawusch and Jusczyk (1981) did question their subjects after the experiment and found that no subjects reported a streaming effect, it remains a possible explanation of their results. In the present experiment, the number of repetitions of the adapting syllable was reduced (to 30) and the interval between adaptors was increased (to 800 msec, substantially longer than the duration of the adapting syllables) in an effort to eliminate the possibility of the fricative's becoming dissociated from the rest of the syllable and forming a separate stream. If, under these modified conditions, which are not as conducive to auditory stream segregation (see Bregman, 1981), the pattern of results found was the same as that reported by Sawusch and Jusczyk (1981), then it would be unlikely that streaming and its associated breakup of the adaptor were playing any role in these adaptation effects.

METHOD

Subjects

The subjects were 36 undergraduates at the State University of New York at Buffalo, who participated in partial fulfillment of a course requirement. All were native speakers of English with no reported history of either a speech or a hearing disorder.

Stimuli

A nine-stimulus, [da]-[ga], test series was generated using the software cascade/parallel synthesizer described by Klatt (1980) and implemented by Kewley-Port (Note 1). This series was constructed to closely parallel the stimuli used by Mann and Repp (1981). All stimuli were 200 msec in duration and consisted of an initial 50-msec period, during which the formant transitions occurred, followed by a 150-msec steady-state vowel [a]. The vowel had formant center frequencies of 700, 1200, 2500, 3600, and 4200 Hz and bandwidths of 90, 100, 140, 250, and 300 Hz for the first through fifth formants, respectively. For all nine stimuli, the first formant had an onset frequency of 350 Hz followed by a 45-msec linear transition to the vowel target frequency. The second formant had an onset frequency of 1650 Hz followed by a 50-msec linear transition to its vowel target value. The fundamental frequency for all syllables rose from 115 to 126 Hz over the first 50 msec, remained at 126 Hz over the next 50 msec, and then fell to 110 Hz over the last 100 msec of the stimulus. In addition, all nine stimuli were generated with a +15-msec VOT. Over the first 15 msec of each syllable, the first-formant bandwidth was set to 300 Hz, and aspiration (AH) replaced voicing as the excitation source for the formants. Between 15 and 150 msec, the amplitude of voicing (AV) was constant. Over the last 50 msec, the voicing amplitude was linearly ramped off.

The only difference among the nine stimuli was in the initial transition of the third formant. For Stimulus 1, at the [da] end of the series, the third formant had an onset frequency of 2430 Hz, followed by a linear 50-msec transition to the vowel target

(2500 Hz). For each succeeding stimulus in the series, the onset frequency of the third formant was decreased by 50 Hz. Consequently, for Stimulus 9, at the [ga] end of the series, the third formant had an onset value of 2030 Hz, followed by a linear 50-msec transition to the vowel target.

A second series of nine syllables was generated by adding 150 msec of friction, appropriate for [s], followed by 90 msec of silence, to the front of each of the nine [da]-[ga] stimuli. The [s] friction was produced using the parallel branch of the Klatt (1980) synthesizer. The friction source (AF) was ramped on over the first 50 msec, remained at a steady value for 75 msec, and then was ramped off over the last 25 msec of the [s]. The fourth, fifth, and sixth formants, with center frequencies of 3600, 4200, and 4900 Hz and bandwidths of 250, 300, and 1000 Hz, had their amplitudes set to 60, 60, and 50 dB, respectively. These values were taken from spectrograms and LPC analyses of the [s] friction in natural [ska] syllables produced by the two authors.

Procedure

The stimuli were stored on computer disk in digital form and were presented to subjects under the real-time control of a DEC PDP-11/34 computer in the Speech Perception Laboratory at SUNY/Buffalo. The stimuli were converted to analog form at a 10-kHz sampling rate by a 12-bit digital-to-analog converter, then low-pass filtered at 4.8 kHz, amplified and presented to subjects binaurally through TDH-39 matched and calibrated headphones. Subjects responded by pushing the appropriately labeled button on a computer-controlled response box. All responses were recorded by the computer. In all the conditions, sessions were conducted with small groups of two to five subjects.

Pilot data. Ten subjects identified the two series described above. One half of these subjects listened to the [da]-[ga] series first and then the [sta]-[ska] series; the other subjects received the reverse order. For both groups, each series was presented in two blocks of 90 trials. Each block contained 10 randomizations of the nine stimuli in one of the two test series. Subjects rated the stimuli using a 6-point scale that varied from definite [d] or [t] (1), through guessing [d] or [t] (3) or guessing [g] or [k] (4), to definite [g] or [k] (6). Across subjects, Stimuli 1, 2, and 3 were rated as good examples of [d] and Stimuli 6, 7, 8, and 9 as good examples of [g] for the [da]-[ga] series. The category boundary between [d] and [g] fell between Stimuli 4 and 5. For the [sta]-[ska] series, only Stimulus 1 was rated as a good [t] (alveolar, like [d]), whereas Stimuli 4 through 9 were all rated as good examples of [k] (velar, like [g]). This result replicates Mann and Repp (1981). Stimulus 3, which was rated as a good [d] in the [da]-[ga] series, was given a 70% identification of [k] in the [sta]-[ska] series. On the basis of these results, Stimulus 3 from the [sta]-[ska] series was chosen for use as an adapting syllable. We will refer to this as the [ska] adaptor. The corresponding [da]-[ga] stimulus (Stimulus 3), which was rated as a good [da], and Stimulus 7, rated as a good [ga], were also chosen for use as adapting syllables.

Adaptation. Three groups of six subjects each were run in the adaptation conditions. Each group listened to a different adapting syllable from the set {[da], [ga], [ska]}. All of these subjects participated first in a baseline rating condition. In this condition, as in the pilot testing, the subjects listened to two blocks of 90 presentations of the [da]-[ga] stimuli. In addition, one presentation of the [ska] stimulus was included with each of the 10 randomizations of the nine test stimuli. For all stimuli, the subjects were asked to rate the stop in each syllable. Rating scales for both [d]-[g] and [t]-[k] were provided for subjects in the baseline condition. If the trial consisted of a stop-vowel syllable, subjects used the [d]-[g] rating scale. When a fricative-stop-vowel syllable was presented, the subjects were instructed to use the [t]-[k] rating scale. From the baseline condition, we obtained 20 rating responses to each of the nine test syllables and 20 ratings of the [ska] adaptor for each subject.

After the baseline condition, each subject listened to two blocks of 10 adaptation trials each. Each adaptation trial consisted of 30 repetitions of the adaptor ([da], [ga], or [ska]) with 800 msec be-

tween repetitions. This was followed by the nine stimuli of the test series presented in random order. The subjects provided rating responses to each of the test stimuli. Thus, 20 adapted rating responses were also obtained from each subject for each of the nine [da]-[ga] test stimuli.

Paired comparison. In the paired-comparison procedure, the subjects listened to pairs of stimuli, presented sequentially with 500 msec of silence between stimuli in a pair. Eight different pairs were used. One of these pairs consisted of a stimulus from the middle of the series (Stimulus 5, hereafter termed the test stimulus) paired with itself. Stimulus 5 was chosen as the test stimulus for this procedure because it was the closest of the nine [da]-[ga] stimuli to the category boundary in the pilot testing. Six of the pairs consisted of a "good phonetic exemplar" paired with our test item. Two of these were syllables labeled by pilot subjects as good [d]s: Stimuli 1 and 3. Another two were stimuli labeled as good [g]s: Stimuli 7 and 9. The last two were labeled by pilot subjects as [ska]s: Stimuli 3 and 7 from the [sta]-[ska] series. These six stimuli, which we will term the comparison stimuli, included the three stimuli used as adaptors, described above. The eighth pair contained Stimuli 3 and 7, which were included as catch trials to check the subjects' categorization of stimuli within both the [d] and the [g] categories.

Each of the eight subjects who participated in the paired-comparison procedure listened to four blocks of trials. Within a block, each of the eight pairs was presented 10 times—5 times in each order (e.g., comparison-test or test-comparison). The order of pairs within a block of trials was random. After the pair of stimuli was presented, the subjects categorized the first stimulus and then the second stimulus. The subjects recorded their category responses by pushing one of four buttons, labeled D, T, G, and K, for each stimulus.

RESULTS

For each of the adaptation groups, a rating function was determined for each subject from the average rating of each of the nine stimuli in both the baseline and the adapted conditions. The data of one subject from the Stimulus 3 [da] adaptation group were then omitted from further analysis because the subject could not consistently identify the stimuli in the baseline condition. Category boundaries were determined by linear interpolation between the two stimuli on either side of the category boundary for each of the remaining individual rating functions.

The results of the adaptation sessions are shown in Table 1. The [da] and [ga] adaptors produced the expected effect of shifting the [da]-[ga] category boundary toward the end of the series from which the adaptor was taken. The row, labeled, "boundary shift," shows the change in the location of the cate-

gory boundary for each of the adaptation conditions in stimulus units. A positive value indicates a shift toward the [da] end of the series, and a negative value indicates a shift toward the [ga] end of the series. Both the [da] and the [ga] adaptation effects were significant [$t(4) = 2.87, p < .05$; $t(5) = -4.69, p < .01$].² The shift in the category boundary for the [ska] adaptation group was also significant [$t(5) = 2.67, p < .05$]. Furthermore, the effects of the [da] and [ska] adaptors were nearly identical [$t(9) < 1, p > .4$].

The change in the percentage of [da] responses to the test series as a whole is also shown in the second row of Table 1. The [da] adaptor caused fewer of the stimuli to be consistently identified as [da]. The [ga] adaptor had the opposite effect, causing more of the test stimuli to be identified as [da] [$t(4) = 3.04, p < .05$, and $t(5) = -6.55, p < .002$, respectively]. The [ska] adaptor also produced a significant decrease in the percentage of [d] responses to the test series [$t(5) = 4.51, p < .01$]. The effect of the [ska] adaptor was virtually identical to that of the [da] adaptor [$t(9) < 1, p > .4$]. The [da] and the [ska] adaptors produced identical results for both measures of the effects of adaptation (the change in the phonetic category boundary and the change in percentage of [da] responses). The effects of these two adaptors were not statistically different. However, the effects of the [da] and [ska] adaptors were opposite to the effects of the [ga] adaptor. It should be noted that all of the subjects in the [ska] adaptation condition were asked to categorize the [ska] stimulus during baseline testing. One of the subjects labeled this stimulus as containing a [t], while the others all indicated that a [k] was present. All the subjects, however, showed a shift in their category boundary toward the [da] end of the series.³ Thus, for the place of articulation dimension, adaptation seems to follow the spectral overlap between the adaptor and the test series, independent of the phonetic identity of the adaptor.

The results for the paired-comparison procedure are shown in Table 2. These values represent the difference in the percentage of [da] responses to the test item when paired with itself versus when paired with the various exemplar comparison stimuli. Both of the [da] comparison stimuli (Stimuli 1 and 3) caused fewer [da] responses to be given to the test item, whereas both of the [ga] comparison stimuli (Stimuli 7 and 9) caused more [da] responses to be given to the test item. All four of these [da] and [ga] comparison stimuli produced significant contrast effects [$t(7) = 3.96, p < .01$; $t(7) = 2.97, p < .05$; $t(7) = -4.58, p < .01$; and $t(7) = -4.33, p < .01$, for Stimulus 1 [da], Stimulus 3 [da], Stimulus 7 [ga], and Stimulus 9 [ga], respectively]. The effects of the two [ska] comparison syllables (based on Stimuli 3 and 7) are shown in the middle of Table 2. The Stimulus 7 [ska] produced a significant contrast effect [$t(7) = -4.44, p < .01$],

Table 1
Mean Shift in the Category Boundary and the Change in the Percentage of [d] Rating Responses for the Entire Test Series for Each of the Adaptation Groups

	Adaptor		
	Stim 3	Stim 3	Stim 7
	[da]	[ska]	[ga]
Category Boundary Shift	.38	.40	-.87
Change in Percent [d] Response	10.4	8.6	-7.9

Table 2
Mean Change in the Percentage of [d] Category Responses to the Test Item for Each of the Six Comparison Stimuli in the Paired-Comparison Procedure

	Comparison Stimulus					
	Stim 1	Stim 3	Stim 3	Stim 7	Stim 7	Stim 9
	[da]	[da]	[ska]	[ska]	[ga]	[ga]
Change in Percent [d] Response	17.7	13.1	-4.5	-17.3	-15.0	-15.4

which was not significantly different from that of either of the [ga] comparison syllables (both $p > .25$). The Stimulus 3 [ska] produced a very small effect, averaged over subjects, in the direction of a phonetic contrast. The effect of the Stimulus 3 [ska] on the number of [da] responses to the test stimulus was not significantly different from the baseline of the test stimulus paired with itself ($p > .25$). However, from inspection of the subjects' ratings of the Stimulus 3 [ska], it appeared that this [ska] syllable was not always identified as containing a [k]. Rather, four subjects identified this syllable as [ska] on better than 60% of the trials, whereas the other four subjects identified the syllable as [sta] on 60% or more of the trials.

The data for each of the eight subjects for baseline and Stimulus 3 [ska] comparison trials are shown in Table 3. To the left of the vertical dashed line are the percentages of [d] responses to the test item when paired with itself and when paired with the Stimulus 3 [ska], as well as the difference between these two values. To the right of the dashed line are the percentage of [t] responses for each subject to the Stimulus 3 [ska] and a label indicating the dominant category (t or k) for this stimulus. For each of the four subjects who identified the [ska] syllable as containing [k], the ambiguous test item was categorized as more [d]-like (see the bottom half of Table 3). That is, a phonetic contrast effect was found. For each of the subjects who classified the [ska] syllable as containing [t], the ambiguous item was categorized as more

[g]-like. Again, for these four subjects, a phonetic contrast effect was found (see the top half of Table 3). Thus, knowing the phonetic category of the comparison stimulus allows us to predict the effect of the comparison stimulus on the test item. This result stands in sharp contrast to the results of the adaptation experiment, which showed the effects of adaptation to be independent of the phonetic class of the adaptor.

DISCUSSION

On the basis of these results, it seems reasonable to conclude that the effects of selective adaptation on the phonetic dimension of place of articulation arise at an early, auditory level (or levels) of processing. By comparison, the response contrast produced by the paired-comparison procedure seems to reflect a representation of the stimulus that is based on its phonetic category and not its spectral structure. For the adaptation group that listened to the Stimulus 3 [ska], the effects of adaptation were identical to those of the Stimulus 3 [da]-adapted group. As Figure 2 shows, these two stimuli are virtually identical in their spectral structure. The [ska] was formed by adding an [s] plus silence to the front of the [da]. However, when given the alternatives of [t] or [k], five of the six [ska]-adaptation subjects labeled this adaptor as containing a [k]. At the same time, subjects labeled the [ska] counterpart from the test series (Stimulus 3) as containing a [d]. Thus, the identical nature of the [ska] and [da] adaptation effects is strong evidence that the effects of selective adaptation to place of articulation are governed solely by the spectral overlap between the adaptor and the test series. This result, which is identical in its pattern to the results of Sawusch and Jusczyk (1981) for a voicing continuum, implicates auditory level processes as the locus for selective adaptation effects for speech (see also Roberts & Summerfield, 1981).

The results of the paired-comparison procedure were, in many respects, just the opposite of the effects of adaptation. In this case, the direction of the contrast effect produced by the Stimulus 3 based [ska] was determined by the subjects' perception of the [ska] as containing either a [t] or a [k]. For those subjects who labeled the [ska] as containing a velar stop (i.e., [k]), a phonetic contrast effect was found;

Table 3
Percentage [d] Responses to the Test Item Paired With Itself, the Stimulus 3 [ska], and Their Difference (on the Left) for Each Subject. On the Right are the Percentage [t] Responses to the Stimulus 3 [ska] and Its Dominant Response Category

S	Comparison Stimulus				[ska] Category	
	Test Stim	[ska]	Diff		[ska]	Category
1	15	11	4	100	t	
2	60	50	10	60	t	
3	75	56	19	65	t	
4	75	60	15	100	t	
5	50	80	-30	10	k	
6	90	100	-10	30	k	
7	45	70	-25	0	k	
8	40	60	-20	0	k	

the test item was categorized as more alveolar (i.e., [d]) than in the control condition. This effect is opposite to the effects of adaptation where the [ska] was again predominantly labeled as containing a velar [k], but had an alveolar [d]-like adapting effect in yielding more velar [g] responses. For those subjects who labeled the [ska] as containing an alveolar stop (i.e., [t]) in the paired-comparison procedure, phonetic contrast effects were also found. In this case, the test item was categorized as more velar (i.e., [g]) than in the control condition. Consequently, the effects found with the paired-comparison procedure seem to involve a phonetic level of processing. This dissociation between adaptation and paired-comparison results clearly indicates that no phonetic processes are involved in selective adaptation to speech. Even though previous adaptation studies have shown that multiple levels of processing may be affected by adaptation for place of articulation in stops (see Sawusch, 1977a), there does not appear to be any phonetic involvement in selective adaptation.

The present experiment also casts doubt on any alternative explanations of the earlier Sawusch and Jusczyk (1981) study that might invoke a "streaming" argument concerning the presentation of the adaptor. The decreased number of adaptor presentations with a larger interval between adaptors should have inhibited any tendency toward streaming. This conclusion is supported by the verbal report of the subjects during debriefing after the experiment. None of our subjects spontaneously reported the occurrence of streaming and, even when directly questioned about whether the adapting syllables broke up during the repeated presentation, none of our subjects reported that this had occurred. Given the rather compelling subjective impression that is present when streaming does occur (see Bregman, 1981), the explanation of either our data or that of Sawusch and Jusczyk in terms of streaming seems remote at best.

The present data also bear on the question of whether the presence of a phonological rule is necessary to produce a dissociation between the spectral structure and the phonetic identity of a stimulus. While the conversion of the /b/ in /ba/ to a /p/ in /spa/ in the experiment of Sawusch and Jusczyk can be described as the result of a phonological rule of English, no such similar rule exists to describe the change in perceived place of articulation from alveolar [d] to velar [k] in the present stimuli. The one substantial difference between the earlier paired-comparison results of Sawusch and Jusczyk and the present results is that their subjects overwhelmingly identified the stop in /spa/ as a /p/. In the present experiment, one half of the paired-comparison subjects identified the stop in [ska] as [k], while the other half identified this stop as [t]. The presence of a phonological rule in English, which specifies that the

stop in /spa/ is a /p/, may have been responsible for this difference. However, in both our results and those of Sawusch and Jusczyk, the phonetic category used by subjects to label the exemplar (comparison) stimuli still determined the direction of any paired-comparison effects. Thus, a phonological conversion rule is *not required* to dissociate the spectral structure of a stimulus from its phonetic identity.

Based on the present results (and those of Sawusch & Jusczyk, 1981), the claim of Diehl et al. (1978; Diehl et al., 1980) that response-contrast can account for both paired-comparison and adaptation effects can be rejected. While it is true that both of these experimental procedures produce contrast effects in phonetic identification tasks, the present results indicate that these contrast effects occur at distinct processing stages. Contrast effects with stops that are similar to the results produced in the paired-comparison procedure have been reported by Eimas (1963) and Healy and Repp (1982). In all of these experiments, stop consonant-vowel syllables were presented in close temporal proximity to one another. We suggest that in all of these cases the contrast effects are due to the subjects' use of a memory trace of the phonetic quality of the stimuli. Borderline stimuli, whose memory trace indicates only a weak phonetic quality, would be labeled by subjects as belonging to the category opposite to any temporally proximate stimuli that produced a strong phonetic quality. A similar description of this type of contrast effect has been offered by Healy and Repp (1982) in a discussion of categorical perception results.

If we combine the results of the present experiments with previous work on selective adaptation to place of articulation (see Sawusch, 1977a), we are led to the conclusion that three levels of processing are involved in speech perception. Figure 3 shows an outline of one possible organization of these levels of processing. After transduction by the ear and the peripheral auditory system, a stimulus undergoes two levels of auditory coding. One of these is spectrally specific and is labeled "local auditory analysis." The other level of auditory coding is not spectrally specific and is labeled "integrative auditory analysis."

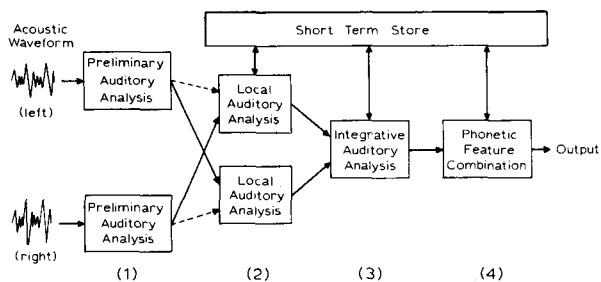


Figure 3. Outline of an information processing model of the earliest stages of speech processing.

This process is responsible for integrating acoustic cues across different frequency loci. Previous adaptation studies indicate that the local processes are primarily monaurally driven, whereas the spectrally integrative process operates binaurally (Sawusch, 1977a). The two auditory processes are shown in a hierarchical (serial) organization only for convenience. We have no evidence for either a serial or a parallel type of organization at this time. Finally, the two auditory levels of processing are followed by a phonetic level of processing which integrates the outputs of the two auditory processes over time to form phonemes (cf. Repp, 1982). Although two auditory levels of processing are influenced by adaptation, the paired-comparison procedure primarily affects the phonetic level (or possibly some later response stage of processing which used the information in phonetic STM).

Although the present experiments were not directly designed to investigate the nature of the mechanisms that mediate the auditory and phonetic processing stages outlined above, they do have implications regarding these mechanisms. With regard to the phonetic coding of speech, it seems highly unlikely that feature detectors are involved in this specialized, language-specific stage of processing. The total lack of a phonetic component in the selective adaptation results suggests that there is no fatigue of the phonetic processing mechanism. With no evidence for fatigue, feature detectors are an unlikely candidate for the processing of phonetic information in speech (see also Remez, 1979). With regard to the auditory coding of speech, selective adaptation effects were found. Thus, these data leave open the possibility that feature detectors may be involved in the auditory processing of speech (see also Eimas & Miller, 1978). A second possibility that has been suggested is that in selective adaptation, the adaptor acts as an anchor, or referent, that leads to a retuning of the auditory processing operations underlying phonetic categories (see Sawusch, 1977a; Simon & Studdert-Kennedy, 1978).

In summary, our results favor neither of the two extreme views of speech perception that were outlined previously. Instead, the phenomenon of phonetic perception seems to result from both multiple auditory levels of processing and uniquely phonetic processes. The phonetic categorization of speech sounds is the result of at least three distinct levels of processing. Of these, the earliest two are auditory in nature and probably represent general auditory processing capabilities. The third process is phonetic and in all likelihood represents a language-specific, highly specialized perceptual subsystem.

REFERENCE NOTE

1. Kewley-Port, D. *KLTEXC: Executive program to implement the KLATT software synthesizer* (Research on Speech Perception,

Progress Report, Vol. 4, pp. 235-246.) Bloomington: Indiana University, 1978.

REFERENCES

- ADES, A. E. How phonetic is selective adaptation? Experiments on syllable position and vowel environment. *Perception & Psychophysics*, 1974, 16, 61-67.
- ADES, A. E. Adapting the property detectors for speech perception. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland, 1976.
- BAILY, P. J. *Perceptual adaptation in speech: Some properties of detectors for acoustical cues to phonetic distinctions*. Unpublished doctoral thesis, University of Cambridge, Cambridge, England, 1975.
- BREGMAN, A. S. Asking the "what for" question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, N.J.: Erlbaum, 1981.
- COOPER, W. E. Selective adaptation to speech. In F. Restle, R. M. Shiffrin, N. J. Castellen, H. Lindman, & D. B. Pisoni (Eds.), *Cognitive theory*. (Vol. 1). Hillsdale, N.J.: Erlbaum, 1975.
- COOPER, W. E. *Speech perception and production: Studies in selective adaptation*. Norwood, N.J.: Ablex, 1979.
- CUTTING, J. E., & PISONI, D. B. An information processing approach to speech perception. In J. F. Kavanagh & W. Strange (Eds.), *Speech and language in the laboratory, school, and clinic*. Cambridge, Mass: M.I.T. Press, 1978.
- DIEHL, R. L., ELMAN, J. L., & MCCUSKER, S. B. Contrast effects in stop consonant identification. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, 4, 599-609.
- DIEHL, R. L., LANG, M., & PARKER, E. M. A further parallel between selective adaptation and response contrast. *Journal of Experimental Psychology: Human Perception and Performance*, 1980, 6, 24-44.
- EIMAS, P. D. The relation between identification and discrimination along speech and non-speech continua. *Speech and Language*, 1963, 6, 206-217.
- EIMAS, P. D., COOPER, W. E., & CORBIT, J. D. Some properties of linguistic feature detectors. *Perception & Psychophysics*, 1973, 13, 247-252.
- EIMAS, P. D., & CORBIT, J. D. Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, 1973, 4, 99-109.
- EIMAS, P. D., & MILLER, J. L. Effects of selective adaptation of speech and visual patterns: Evidence for feature detectors. In H. L. Pick & R. D. Walk (Eds.), *Perception and experience*. New York: Plenum, 1978.
- HEALY, A. F., & REPP, B. H. Context independence and phonetic mediation in categorical perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1982, 8, 68-80.
- KLATT, D. H. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 1980, 67, 971-995.
- LIBERMAN, A. M. On finding that speech is special. *American Psychologist*, 1982, 37, 148-167.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-461.
- MANN, V. A., & REPP, B. H. Influence of preceding fricative on stop consonant perception. *Journal of the Acoustical Society of America*, 1981, 69, 548-558.
- 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.
- PISONI, D. B. Speech perception. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6). Hillsdale, N.J.: Erlbaum, 1978.
- PISONI, D. B., & SAWUSCH, J. R. Some stages of processing in speech perception. In A. Cohen & S. G. Nooteboom (Eds.), *Structure and process in speech perception*. New York: Springer-Verlag, 1975.

- REMEZ, R. E. Adaptation of the category boundary between speech and nonspeech: A case against feature detectors. *Cognitive Psychology*, 1979, 11, 38-57.
- REPP, B. H. Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, 1982, 92, 81-110.
- REPP, B. H., & MANN, V. A. Perceptual assessment of fricative-stop coarticulation. *Journal of the Acoustical Society of America*, 1981, 69, 1154-1163.
- ROBERTS, M., & SUMMERFIELD, Q. Audiovisual presentation demonstrates that selective adaptation in speech is purely auditory. *Perception & Psychophysics*, 1981, 30, 309-314.
- SAWUSCH, J. R. Peripheral and central processes in selective adaptation of place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 1977, 62, 738-750. (a)
- SAWUSCH, J. R. Processing of place information in stop consonants. *Perception & Psychophysics*, 1977, 22, 417-426. (b)
- SAWUSCH, J. R., & JUSCZYK, P. Adaptation and contrast in the perception of voicing. *Journal of Experimental Psychology: Human Perception and Performance*, 1981, 7, 408-421.
- SCHOUTEN, M. E. H. The case against a speech mode of perception. *Acta Psychologica*, 1980, 44, 71-98.
- 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 use of //s denotes systematic phonemes. This includes phonemes whose identity depends upon the phonological rules of a language. Brackets ([s]) are used to indicate a phonetic string (phones). For further information on the difference between phonetic ([s]) and phonemic (/s) representations, see Pisoni (1978).
2. All probability levels were two-tailed.
3. The one subject in the Stimulus 3 [ska] adaptation group who labeled the adaptor as predominantly a [t] showed a shift in his categorization function (and percentage [d] responses to the whole test series) that was neither the largest nor the smallest for the group. Thus, the labeling of the adaptor seems to be unrelated to the adaptation effect it produced.

(Manuscript received February 10, 1983;
revision accepted for publication September 13, 1983.)