Effects of transition length on the perception of stop consonants

Patricia Keating and Sheila E. Blumstein

Brown University, Department of Linguistics, Providence, Rhode Island 02912
(Received 22 December 1977; revised 7 January 1978)

This study investigated the effects of lengthened transitions on the perception of stop consonants. In experiment I, three continua representing the phonetic categories [da] and [ga] containing transitions of 45, 95, or 145 ms were presented to 20 subjects for both labeling and discrimination. Results indicated that although there was a significant change in identification performance from 95 to 145 ms, the shape of the functions, and the locus and slope of the phonetic boundary did not significantly vary across transition lengths. In addition, discrimination of within-category stimulus comparisons was significantly better at the 95-ms transition length than at 45 or 145 ms. In experiment II, the availability of acoustic information was investigated further with the adaptation paradigm. Eight subjects labeled the 45-ms series before and after adaptation with 45-, 95-, and 145-ms [da] stimuli. No effect of transition length was found. These results suggest that the slope and duration of formant transitions seem to contribute minimally to the perception of place of articulation in stop consonants.

PACS numbers: 43.70.Dn, 43.70.Ve

INTRODUCTION

In the course of research in speech, various paradigms have demonstrated differences in the perception of consonants and vowels. Among other results is the finding that stop consonants are generally perceived more "categorically" than are vowels (Liberman *et al.*, 1957; Fry *et al.*, 1962; Stevens *et al.*, 1969). In the ideal case of categorical perception, stimuli varying in equal steps along an acoustic continuum are perceived as belonging to discrete categories. Discrimination functions for these stimuli show a peak at the phoneme boundary and troughs (near-chance discrimination) within categories. Discrimination can be predicted from the labeling functions since subjects are unable to discriminate stimuli assigned to a single phonemic category. On the other hand, more "continuous" perception is attributed to vowels. Results of such studies have shown that vowel perception is characterized by a lack of sharp discontinuities in identification, and by discrimination that is equally good for all stimulus comparisons (Studdert-Kennedy *et al.*, 1970).

The hypothesis originally offered to explain this dichotomy in consonant and vowel perception was that there are two separate mechanisms used to process speech, one which uses acoustic attributes of the stimulus at all stages, and one specially suited to converting acoustic information from, e.g., a stop consonant into a phonetic label. The specialized speech processor was thought to literally "strip away" the acoustic information from a stimulus in making the identification, so that such information was no longer available (Liberman, Mattingly, and Turvey, 1972).

An alternative hypothesis, offered by Fujisaki and Kawashima (1969), is that the differences attributed to consonant and vowel perception are due to the intrinsic nature of the stimuli and not to the mechanisms which presumably process these stimuli. Fujisaki and Kawashima hypothesized that auditory short-term memory does not retain transient cues such as consonantal formant transitions, and that therefore when such cues occur in speech stimuli, a phonemic (category) identification is stored in phonetic memory. Because the low-level acoustic information extracted from stop consonants deteriorates rapidly, it is generally not available for discrimination judgments. These judgments must be based on a comparison of the stored phonemic labels. In contrast, since vowel stimuli are not transient, they remain accessible in auditory short-term memory. Discrimination judgments can be based on direct comparisons of the stimuli.

Fujisaki and Kawashima tested this hypothesis with respect to categorical perception of consonants in two ways. First, they showed that a /wa/–/ra/ continuum, with formant transition time constants of 30 ms for $F1$ and 35 ms for $F2$, was discriminated more consistently across stimuli than was a /ba/–/da/ continuum with time constants of 10 and 15 ms, respectively (Fujisaki and Kawashima, 1970). Second, they showed, contrary to their hypothesis, that both a /s/–/š/ and a /sa/–/ša/ continuum, which contained sustained fricative noise cues as well as formant transitions, were perceived more categorically than isolated vowels (Fujisaki and Kawashima, 1969).

With regard to vowel perception, several studies have shown that manipulation of vowel stimuli can result in more categorical perception. Stevens (1968) and Sachs (1969) showed that vowels are perceived more categorically (in the sense that discrimination performance is not consistent across all discrimination pairs), in a CVC context; Sachs proposed that one factor might be their relatively short duration. Fujisaki and Kawashima (1969) reported that synthetic vowel stimuli with durations ranging from a single glottal pulse to 100 ms all were perceived categorically. Pisoni (1971), comparing the responses of a single group of subjects to both 50- and 300-ms vowel continua, found that identification of both continua showed sharp boundaries, but that the long vowels were discriminated better than the short. The difference between obtained and predicted discrimination was also greater for the long vowels.

 0001-4966/78/6401-0057$00.80

TABLE I. Second and third formant onset frequencies for each stimulus of the test continua.

Stimulus No.	1	2	3	4	5	6	7
F_2	1715	1700	1685	1670	1655	1640	1625
F_3	2975	2800	2625	2450	2275	2100	1925

The notion that the nature of the obtained identification and discrimination functions for consonants and vowels is a function of the duration of the critical components of the stimuli would seem to predict that, just as the perception of vowels can be modified by shortening the duration of the stimuli, the perception of consonants would also be affected by lengthening the duration of the formant transitions. That is, with the increased duration of formant transitions, stop consonants may be processed in terms of auditory in addition to phonetic information.[1] The perceptual consequences for identification might be less discrete category judgments. For discrimination, on the other hand, performance on within category discriminations should improve. In fact, differences in the discrimination of lengthened transitions have been demonstrated by Tallal and Piercy (1975) in an investigation of the speech perception of normal and dysphasic children. They presented CV syllables with either 43- or 95-ms transitions and found that the dysphasics discriminated like normals at 95 ms, but did worse than normals at 43 ms. These results suggest that it is the brief duration of the transitions rather than their frequency modulation which is the critical factor in the perception of these consonants for the dysphasic children. In addition, it suggests that changes in the duration of formant transitions in stop consonants may affect the nature of their perception.

It is the object of this experiment to explore the effects of lengthened transitions on the perception of stop consonants. To this end, both labeling and discrimination ability of subjects will be compared across three transition lengths.

I. EXPERIMENT I

A. Method

1. Stimuli

We originally attempted to construct stimuli similar to those described by Tallal and Piercy (1975) in which the transitions for formants one to three were lengthened. Pilot work, however, indicated that we were unable to maintain the stoplike quality of the stimuli with the lengthened $F1$ transitions; as transition length increased beyond 65 ms, the consonants were perceived as glides. Because we were interested in studying the effects of lengthened transitions on the perception of place of articulation in stop consonants, it was imperative that the manner of articulation be preserved across transition lengths. There were only two ways that we found which did in fact maintain the stop-quality of the stimuli. The first was the addition of a burst 5–10 ms preceding the onset of the transitions (see Lisker *et al.*, 1975, for discussion of stimuli of this type). The second was to keep the $F1$ transition short, lasting about 25 ms. Because the presence of the burst provided additional place of articulation information which could be used by the subject in performing the task, we opted in this study to use stimuli with no burst and a shortened $F1$. In this way, any differences obtained in either labeling or discrimination could be attributed only to the effects of lengthened transitions.

Twenty-one stimuli forming three continua from /da/ to /ga/ were made using the M.I.T. terminal-analog speech synthesizer developed by Klatt (1972). All stimuli have their first formant rising piecewise linear over 25 ms from 200 Hz to a steady-state value of 720 Hz. Second and third formants had onset frequencies as shown in Table I; steady-state values were 1240 Hz for $F2$ and 2500 Hz for $F3$. Formants 4 and 5 had steady-state values throughout at 3500 and 4500 Hz, respectively. All stimuli were 260 ms long. The three continua differed only in the lengths of the $F2$ and $F3$ consonantal transitions, which were 45-, 95-, or 145-ms duration.[2] Therefore, the duration of the steady-state portions was 215, 165, or 115 ms, respectively. All stimuli were voiced throughout. The fundamental frequency contour was characterized by a rise from 100 to 125 Hz over the first 25 ms of the stimulus followed by a steady contour for 50 ms and then falling to 100 Hz piecewise linear from 75 to 25 ms before the end of the stimulus and then to 55 Hz for the last 25 ms. Figure 1 shows a schematic of the first stimulus in each continuum. It may be worth mentioning briefly the subjective quality of these stimuli. There is no question that the phonetic categories of the end-point stimuli in all three conditions could be unambiguously and clearly determined. However, the

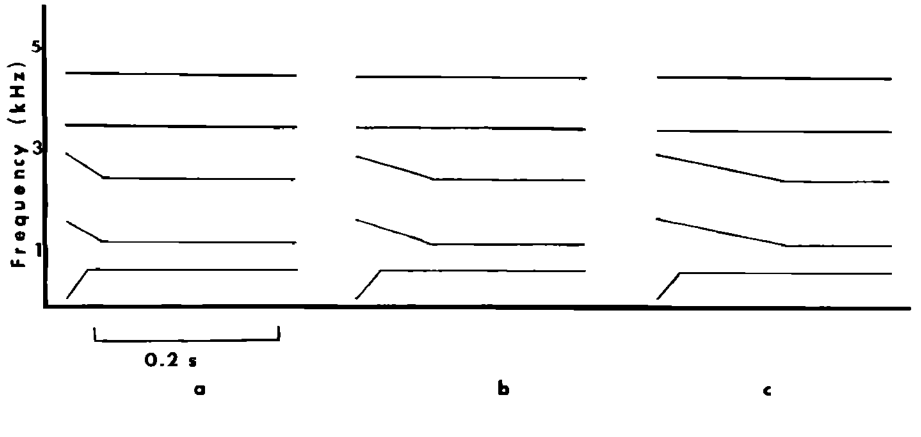

FIG. 1. Schematic of the first stimulus in each of the three continua. The continua varied only in the lengths of the $F2$ and $F3$ consonantal transitions which were (a) 45, (b) 95, or (c) 145-ms duration.

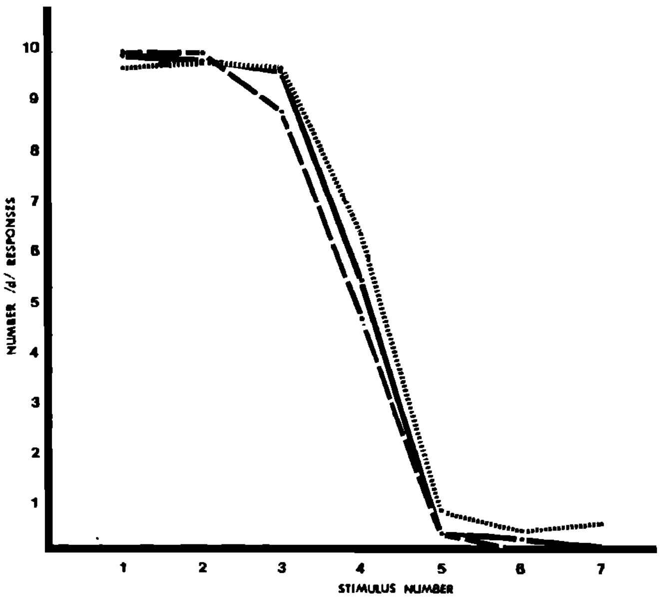

FIG. 2. Mean performance on labeling the test continua across the three transition lengths. The solid lines indicate performance on the 45-ms transition length stimuli, the dotted lines the 95-ms stimuli, and the dashed lines the 145-ms stimuli.

lengthened transition stimuli were perceived as somewhat "strange" in the sense that it sounded like someone was swallowing while the consonant was being released. Nevertheless, the stimuli were perceived as CV syllables.

A labeling and discrimination tape was constructed for each transition length. Each labeling test included the random presentation of ten tokens of each of the seven stimuli, with an ISI of 4 s. The 4IAX procedure was used in the discrimination task. Each trial consisted of a block of 4 stimuli, with 200 ms separating stimuli 1 and 2, 250 ms separating stimuli 2 and 3, and 200 ms separating stimuli 3 and 4. The five 2-step comparisons of the stimulus series, with the additional comparison of stimuli 1 and 6, gave a total of six comparisons.[3] The test tape consisted of eight randomized presentations of each of the six comparisons, with five seconds between each block.

2. *Subjects*

Subjects were native speakers of English associated with Brown University who had no known speech or hearing impairments. Because we were interested in the effects of lengthened transitions on categorical perception, we eliminated all subjects who were unable to label the 45-ms stimuli in terms of two discrete categories. Twenty-five subjects were run; five were eliminated from this group because they did not meet the selection criterion.

3. *Procedure*

Subjects were tested individually or in groups no larger than four in a reasonably quiet room using a Revox A 77 tape recorder and Koss Pro/4*AA* headphones set at a comfortable listening level. Subjects who had never heard synthetic speech before were given a brief listening practice to familiarize them with the stimuli, and all subjects heard the entire continuum at least twice before they took the labeling test for each transition length.

All subjects were presented the 45-ms transition length stimuli first, counterbalanced across subjects for task order (identification and discrimination). The presentation of the remaining tests was counterbalanced for transition length (95 and 145 ms) and task (identification and discrimination) with the limitation that identification and discrimination tasks for a particular transition length were not separated.

For the identification tasks, subjects were instructed to circle the phonetic category heard, i.e., either [d] or [g], on the answer sheet provided. They were told to respond to each item and to guess even if unsure. In the discrimination task, subjects were told that they would hear four stimuli in two pairs, with three of the four stimuli being identical. They were asked to circle on the answer sheet the number of the pair (I or II) containing the two stimuli which were not identical.

4. *Results*

The results of performance on the identification task for the three transition lengths are shown in Fig. 2. A two-way repeated measures analysis of variance for the number of [d] responses at each stimulus number for each transition length showed a significant main effect for transition length ($F_{(2,38)} = 6.92$, $p < 0.01$) stimulus number ($F_{(6,114)} = 326.22$, $p < 0.001$) and a nonsignificant interaction between transition length and stimulus number ($F_{(12,228)} = 1.53$, $p > 0.05$). Thus, although performance differed across transition lengths, the overall shape of the functions remained similar. Subsequent *t* tests based on the analysis of variance error term indicated that the main effect for transition length was due solely to a difference in performance between the 95-ms transition length stimuli and the 145-ms transition length stimuli. This comparison was only marginally significant ($t = 1.938$, 0.10, $p > 0.05$); all other comparisons did not come close to significance. Further analyses revealed that neither the locus of the phonetic boundary (mean boundary values were 3.04 at 45 ms, 3.13 at 95 ms and 2.84 at 145 ms) nor the slopes of the phonetic boundary (− 4.95, − 5.42, and − 4.40 for 45, 95, and 145 ms, respectively) varied across transition lengths.

Results of performance on the discrimination tasks are summarized in Fig. 3. A two-way repeated measures analysis of variance comparing % correct discrimination of each comparison at each transition length for each subject revealed a significant main effect for transition length ($F_{(2,38)} = 6.06$, $p < 0.01$) and discrimination comparison $F_{(4,76)} = 23.39$, $p < 0.01$) and a significant interaction effect ($F_{(8,152)} = 2.00$, $p = 0.05$). *T* tests based on the analysis of variance error term indicated that 95-ms transition length stimuli were better discriminated overall than were the 45 ms transition length stimuli ($t = 2.14$, $p < 0.05$). All other comparisons were nonsignificant. Further analysis revealed that the main ef-

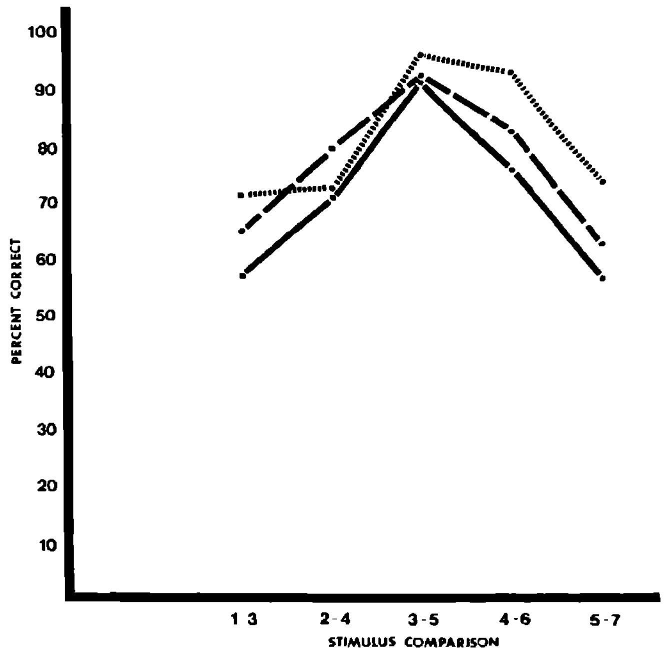

FIG. 3. Mean performance on the discrimination task for the three transition lengths. The solid lines indicate performance on the 45-ms transition length stimuli, the dotted lines the 95-ms stimuli, and the dashed lines the 145-ms stimuli.

fect for discrimination comparison was the result of comparison stimuli 2–4 and 3–5 being significantly more discriminable than the other stimulus comparisons. These results indicate that discrimination comparisons are easier between than within phonetic categories. Post-hoc comparisons revealed that the significant interaction between stimulus comparison and transition length were due to the following comparisons: 95-ms transitions were more discriminable than 45-ms transitions at stimulus comparison 1–3, 4–6, and 5–7 and were more discriminable than 145-ms transitions at stimulus comparison 4–6 and 5–7. Since overall performance in this experiment was extremely good, it is possible that the significant interaction is due to a ceiling effect on discrimination of comparisons 2–4 and 3–5. That is, performance on these comparisons may already be so high at 45 ms that subjects cannot show much improvement at longer transition lengths. All other comparisons were nonsignificant.

In order to determine whether the improved discrimination for 95-ms transition stimuli was due to a learning effect, i.e., all subjects received 95 ms transitions stimuli after at least one session with synthetic speech, we compared the performance of the group that received the stimuli in the order 45–95–145 ms to the group receiving the order 45–145–95 ms. In both cases, 95-ms performance was best, then 145 ms, and worst was 45 ms.

B. Discussion

The results of this study suggest that lengthened transitions minimally affect the perception of stop consonants. In particular, although there was a difference in performance in the labeling task between 95- and 145-ms stimuli, the overall shape of the functions as well as the locus and slope of the phonetic boundaries remained constant across the three transition lengths. Similarly, for discrimination, the differences across transition lengths were due exclusively to the 95-ms transition length stimuli. Thus, they were discriminated overall better than the 45-ms transition length stimuli. Moreover, within category comparisons at both the [d]-end and [g]-end of the continuum were discriminated better for the 95-ms stimuli than were the 45-ms stimuli, and they were better at the [g]-end for the 95-ms stimuli in comparison to the 145-ms stimuli. In spite of these differences, comparison of the performance differences between discrimination and identification across the three transitions reveals no tendency for either identification functions or discriminations to systematically change as transition length increased or for differences in identification or discrimination performance to relate in any systematic way.

The improvement in discrimination performance at 95 ms is consistent with the hypothesis that duration of critical components of the speech signal affects the nature of the perceptual processing of speech stimuli. However, counter to this theory, there was no difference in discrimination between 45- and 145-ms stimuli or between 95- and 145-ms stimuli. This failure to show systematic improvement across the three transition lengths may be due to the "poorer" quality of the stimuli at 145 ms (see p. 3), or it may reflect the limiting conditions for the optimal duration of transition lengths; namely, facilitation of discrimination performance for speechlike stimuli may be a function of certain critical durations.

The results obtained in this study are similar to those found by Mandler and Dechovitz (1977). In their study, they used transition lengths of 30 ms for *F*1 and 30 or 135 ms for *F*2 and *F*3 in two [ba]–[da]–[ga] continua. No significant differences were found for either labeling or discrimination. Like them, we found no differences between the short (45 ms) and longest duration (145-ms) stimuli. All of the differences obtained in the present study are due to the 95-ms duration stimuli.

The finding that discrimination was better at 95 ms than at 45 ms is in agreement with Tallal and Piercy's results with dysphasic children. Thus, despite the fact that the stimuli used in these two studies differed in the duration of the *F*1 transitions, the lengthening of the higher formants seemed to enhance discrimination ability. Nevertheless, as our results showed, although discrimination ability improved between 45 and 95 ms, there was no significant change in identification performance across these two transition lengths. Thus, it is not clear that an improvement in discriminating stimuli with lengthened transitions necessarily implies an ability to label or use them normally. In fact, in a study of the identification and discrimination of a voice-onset time continuum in adult aphasics (Blumstein *et al.*, 1976), a partial dissociation was found between labeling and discrimination ability. In particular, normal labeling performance depended upon an ability to discriminate, whereas discrimination ability did not necessarily entail normal labeling. The results of this study with aphasics suggest that discrimination ability may be a necessary but not a sufficient prerequisite for identification, and may, in addition, set the boundary conditions upon which linguistic categories are ultimately defined.

Finally, the fact that the effects of transition length have minimal effects on the perception of a place of articulation continuum suggests that the slope and duration of the higher formant transitions do not significantly contribute to place of articulation identification. The cues used for such identification may depend upon the locus and direction of the formant transitions (Liberman *et al.*, 1967) or alternatively upon the shape of the acoustic spectrum at stimulus onset (Blumstein and Stevens, 1976; Stevens and Blumstein, 1976). Lengthening of the transitions in this study affected neither the onset frequencies and direction of the formants, nor the shape of the acoustic spectrum at stimulus onset.

II. EXPERIMENT II

The results of the identification and discrimination experiments, which showed significant but not systematic effects of transition length, do not clearly establish the role of transition length in affecting the nature of the perception of a place of articulation continuum. The discrimination task, which by its design is more sensitive to acoustic differences than is the identification task, resulted in better within-category discrimination at 95 ms, indicating that auditory information is more available at longer transition lengths. Therefore we decided to investigate further the effect of transition length by using another paradigm that seems to be sensitive to the acoustic attributes of speech stimuli, i.e., selective adaptation. Selective adaptation involves a comparison of identification performance on a test continuum before and after repeated presentation of a particular adapting stimulus. This comparison usually reveals a shift in the phonetic boundary on the continuum towards the category of the adapting stimulus.

Early adaptation experiments (Eimas and Corbit, 1973; Cooper and Blumstein, 1974) showed that when the adapting stimulus belongs to a different phonetic category from that of the test continuum, the adaptation effect is about 50%–80% of that obtained when the adapting stimulus is a member of the test continuum. In contrast, studies investigating effects of adaptation when the adapting stimulus is a member of the same phonetic category but is not acoustically equivalent to the continuum stimuli have shown varied results (cf. Cooper, 1974; Morse, Kass, and Turdienicz, 1976; Ades, 1974; Pisoni and Tash, 1975; Bailey, 1973; Diehl, 1975; Blumstein, Stevens, and Nigro, 1977). Several of these studies investigated adaptation effects on a place of articulation continuum. Bailey (1973) found that an adapting stimulus in which place was cued by $F2$ with no $F3$ present adapted a test continuum in which $F2$ was fixed and $F3$ varied, but that an adapting stimulus with fixed $F2$ and moving $F3$ would not adapt a test continuum with no $F3$. These results suggest that the test continuum must contain the component that cues the relevant distinction in the adapting stimulus, even if that component is not the cue in the continuum.

On the other hand, Diehl (1975) showed that an adapting stimulus without formant transitions, in which place is cued by burst alone, will adapt a test continuum without bursts, in which place is cued by formant transitions alone. Pisoni and Tash (1975), attempting to reconcile Bailey's and Diehl's results, have suggested that the adapting stimulus and member(s) of the test series sharing the relevant feature must have similar spectra for adaptation to occur. In fact, Blumstein, Stevens, and Nigro (1977) have shown that the presence of full-cue spectral information (i.e., burst and transition) in an adapting stimulus produces a greater adaptation effect than a partial-cue stimulus (transitions only) despite the fact that the test continuum contained only transitions and no burst. These results indicate the importance of spectral similarity between adapting and test stimuli, and perhaps more importantly the dependence of the adaptation effect upon the degree to which the adapting stimulus contains the acoustic properties relevant to a particular phonetic category rather than simply to an isomorphism between adapting and test stimuli.

All of these studies measuring cross-series adaptation effects have used phonetically equivalent stimuli varying in either syllable structure, stimulus structure (i.e., presence versus absence of bursts or transitions) or a particular stimulus attribute (e.g., moving vs fixed transitions). In this study we have maintained all of these characteristics while varying an acoustic component of a cue to place of articulation, i.e., transition length. It might be expected that an adapting stimulus with a different transition length from the test continuum would give less adaptation than a stimulus from the continuum. However, the spectral similarity at stimulus onset for those stimuli varying only in transition length and the similarity between identification functions obtained in the first experiment, might lead us to expect little or no difference in the amount of adaptation obtained across transition length.

A. Method

1. *Stimuli*

The adapting stimuli were the No. 1 [da] stimuli from each of the three continua used in the first experiment, with $F2$ and $F3$ transition lengths of 45, 95, or 145 ms. Presentations of the adapting stimuli were separated by 300 ms of silence. The identification test used in all experiments was the same as that used in the 45-ms transition length task in the first experiment.

2. *Subjects*

Eight subjects who could identify the 45-ms continuum reliably were paid to participate in the experiment; seven of these had participated in the previous experiment.

3. *Procedure*

Subjects were run four at a time in three 1-h sessions. Sessions were separated by at least 24 h. All stimuli

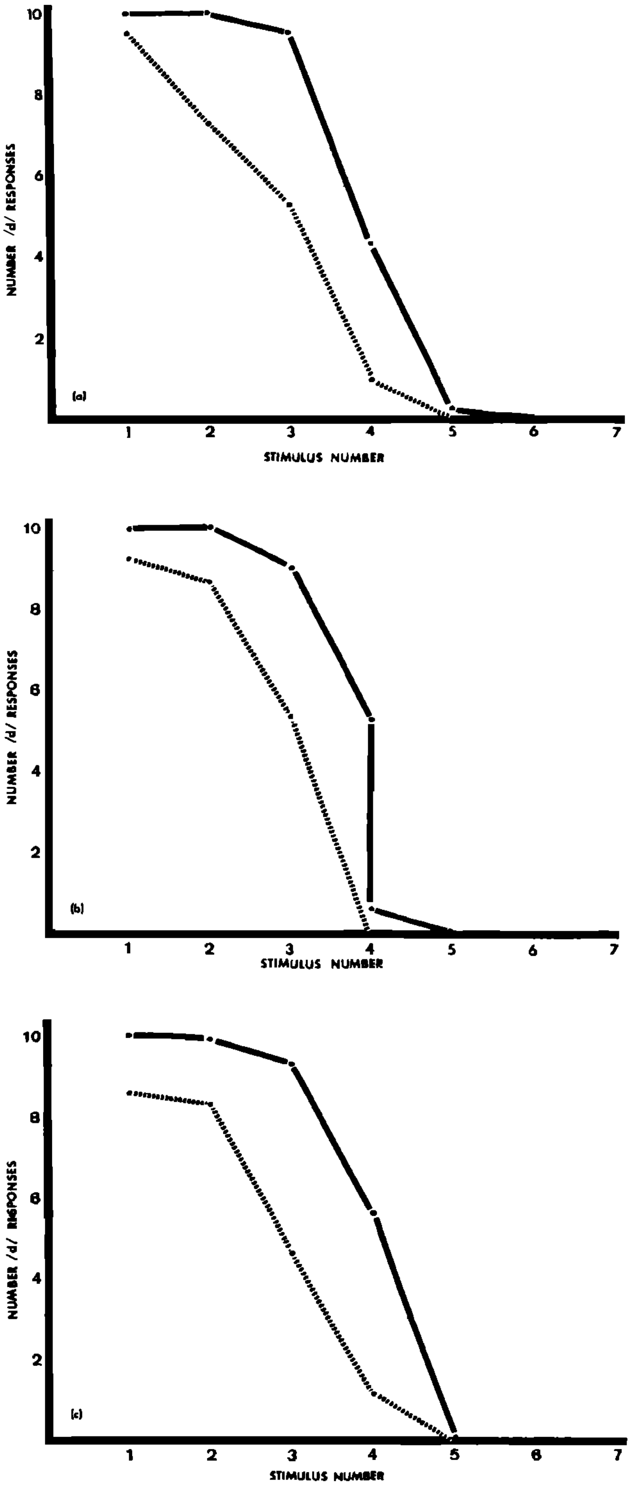

FIG. 4. Mean performance on labeling the 45-ms transition test continuum before and after adaptation with the No. 1 [da] stimuli from each of the continua, with transition lengths of (a) 45, (b) 95, and (c) 145 ms.

were presented on a Revox A77 and a Sony TC-360 tape recorder via four sets of Koss Pro 4-AA headphones. Intensities of adapting and test stimuli were calibrated at 70 dB SPL using a Ballantine voltmeter.

In each session, a baseline function was obtained for each subject, prior to the presentation of the adaptation condition. As in the previous experiment, subjects were required to classify the stimulus as either [d] or [g] and to record their answers by writing the appropriate initial consonant on the answer sheet provided.

Following the baseline series, identification functions were obtained for the same test series after selective adaptation with one of the three adapting stimuli. For half of the subjects, the order of presentation of the adapting stimuli was 45, 95, and 145 ms; for the other half, the order was reversed. The adaptation test began with the repeated presentation of the adapting stimulus for 3 min. Subjects then listened to five practice stimuli from the identification series which they were asked to label. Immediately following this initial period, 14 adaptation trials were administered. On each trial S's listened to the adapting stimulus for 1 min followed by five randomly selected stimuli from the test series.

4. Results

Mean identification functions for all eight subjects in the three adapting and baseline conditions are seen in Fig. 4. Boundary values were computed for each subject using the procedures outlined earlier. A two-way analysis of variance (transition length by adaptation condition) with repeated measures indicated a significant main effect for adaptation condition ($F(1,7)=75.65$, $p < 0.01$), a nonsignificant main effect for transition length ($F<1$), and a nonsignificant interaction of transition length and adaptation condition ($F<1$). Subsequent t tests based on the analysis of variance error term indicated significant adaptation in all adapting conditions. For 45 ms, $t(7)=48.9$, $p<0.01$; for 95 ms, $t(7)=49.24$, $p<0.01$; for 145 ms, $t(7)=60.12$, $p<0.01$.

B. Discussion

The results of this experiment indicate that the duration of the $F2$ and $F3$ transitions does not seem to be a factor in the perception of the place of articulation distinction, at least as measured by the adaptation paradigm. The different transition lengths of the three adapting stimuli used did not affect the amount of adaptation obtained. It is interesting that the adapting stimulus taken from the test continuum gave no more adaptation than the two stimuli which were taken from the other series. In most studies, the cross-series effect has been shown to be quite sensitive to nonphonemic acoustic differences between adapting stimulus and test continuum. However, the results of this study suggest that the size of the adaptation effect does not depend on point-by-point acoustic isomorphism of adapting and test stimuli when nonphonemic variation of transition length is introduced in the adapting stimuli. The results also suggest that, although transitions seem to serve as primary cues to place of articulation, the slope and duration of transitions contribute little information to this phonetic dimension. In fact, recent research has suggested that invariant cues for place of articulation in stop consonants are derivable from the overall shape of the spectrum at stimulus onset (Fant, 1960; Blumstein and Stevens, 1976; Stevens and Blumstein, 1976, 1977). The spectrum shape for [d] is characterized by a diffuse upward sloping spread of energy, in contrast to [g], which has a compact midfrequency spectral peak. In the present study, although the slope and duration of *F*2 and *F*3 of the adapting stimuli change across the three transition lengths, the onset frequencies are identical and the overall shapes of the spectrum at onset are virtually the same. Owing to the near-identity of these spectral cues for the adapting stimuli, one might expect to find no differences in adaptation effects.

III. CONCLUSIONS

The results of the two experiments conducted in this study indicate that *F*2 and *F*3 transition length has little systematic effect on the perception of the place of articulation dimension. In experiment I, performance on the identification task was significantly different only between the 95- and 145-ms transition length series. Experiment I also showed that within-category stimulus pairs with 95-ms transitions are discriminated significantly better than those with 45- and 145-ms transitions.

These results neither support nor refute Fujisaki and Kawashima's cue-duration hypothesis. The hypothesis is partially supported by the significant differences found between 45 and 95 ms especially by the finding that the discrimination improvement was due to performance on within-category comparisons. However, there was no systematic improvement at progressively longer transition lengths. In fact, the significant differences found are all due to the 95-ms transition length stimuli, and performance at 145 ms is not significantly different from that at 45 ms. Yet 145 ms is still much shorter than the duration typically used for vowel stimuli in experiments with synthetic speech—about 250 ms and longer. Thus it seems that the long durations which result in more continuous perception for steady-state vowels do *not* produce a similar effect when used for formant transitions of stop consonants.

Owing to the results obtained in experiment I, where subtle differences were found for discrimination, we attempted to investigate, in experiment II, the effects of lengthened transitions with the adaptation paradigm, a paradigm that seems especially sensitive to acoustic differences among stimuli. Results, however, showed that transition length had little effect on the perception of place of articulation. Thus, the acoustic differences demonstrated perceptually in experiment I were not shown using the adaptation paradigm.

In conclusion, slope and duration of formant transitions seem to contribute minimally to the perception of place of articulation in stop consonants.

[1]It has recently been hypothesized that the differences between consonant and vowel perception are due to range effects intrinsic to the stimulus sets used (Ades, 1977). If this is the case, then perception of lengthened transitions may also be conditioned by range effects. However, as there is no operational definition at this time of what constitutes a so-called range effect for any set of stimuli, we have attempted to choose several stimulus values which have been shown to have perceptual effects in distinguishing classes of speech sounds (cf. Liberman, Delattre, Gerstman, and Cooper, 1956).

[2]The 145-ms transition length was chosen because it was a natural progression from the 45-ms transition length typically used in studies of categorical perception, and the 95-ms transition length used by Tallal and Piercy (1975). It is not clear what the upper limits are for lengthened transitions which would maintain the stop quality of the stimulus.

[3]The additional stimulus comparison was included only because of limitations on the computer program being utilized, and subjects' responses on these discriminations were not scored.

Ades, A. E. (1974). "How phonetic is selective adaptation?", Percept. Psychophys. 16, 61–66.

Ades, A. E. (1977). "Vowels, consonants, speech, and nonspeech," Psychol. Rev. 84, 524–530.

Bailey, P. (1973). "Perceptual adaptation for acoustical features in speech," Ph.D. dissertation (The Queen's University of Belfast) (unpublished).

Blumstein, S. E., Cooper, W., Zurif, E., and Caramazza, A. (1976). "The perception and production of voice-onset time in aphasia," Neuropsychologia 15, 371–383.

Blumstein, S. E., and Stevens, K. N. (1976). "Perceptual invariance and onset spectra for stop consonants in different vowel environments," J. Acoust. Soc. Am. 60, S 90(A).

Blumstein, S. E., Stevens, K. N., and Nigro, G. N. (1977). "Property detectors for bursts and transitions," J. Acoust. Soc. Am. 61, 1301–1313.

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

Cooper, W. E., and Blumstein, S. E. (1974). "A labial feature analyzer in speech perception," Percept. Psychophys. 17, 48–52.

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

Eimas, P. D., and Corbit, J. D. (1973). "Selective adaptation of linguistic feature detectors," Cogn. Psychol. 4, 99–109.

Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, The Hague).

Fry, D. B., Abramson, A. S., Eimas, P. D., and Liberman, A. M. (1962). "The identification and discrimination of synthetic vowels," Lang. Speech 5, 171–188.

Fujisaki, H., and Kawashima, T. (1969). "On the modes and mechanisms of speech perception," Ann. Rep. Eng. Res. Inst. 28, Fac. Eng., Univ. Tokyo, 67–73.

Fujisaki, H., and Kawashima, T. (1970). "Some experiments on speech perception and a model for the perceptual mechanism," Ann. Rep. Eng. Res. Inst. 29, Fac. Eng., Univ. Tokyo, 207–214.

Klatt, D. (1972). "Acoustical theory of terminal analog speech synthesis," *Proc.* 1972 *Int. Conf. Speech Commun. Proc.* (Boston).

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

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

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

Liberman, A. M., Mattingly, I. G., and Turvey, M. T. (1972). "Language codes and memory codes," in *Coding processes in human memory*, edited by A. W. Melton and E. Martin (Winston, Washington, DC), pp. 307–334.

Lisker, L., Liberman, A. M., Erickson, D., and Dechovitz, D. (1975). "On pushing the voice-onset-time (VOT) boundary about," Haskins Labs. Stat. Rep. Speech Res. SR-42/43, 257–264.

Mandler, R., and Dechovitz, D. (1977). "Effects of transition length on identification and discrimination along a place continuum," J. Acoust. Soc. Am. 61, S46 (A).

Morse, P. A., Kass, J. E., and Turdienicz, R. (1976). "Selective adaptation of vowels," Percept. Psychophys. 19, 137–143.

Pisoni, D. B. (1971). "On the nature of the categorical perception of speech sounds," Ph.D. dissertation (Univ. Michigan) (Supplement to Haskins Labs. Stat. Rep. Speech Res.).

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

Sachs, R. M. (1969). "Vowel identification and discrimination in isolation vs word context," Q. Progr. Rep. Res. Lab. Electron., MIT 93, 220–229.

Stevens, K. N. (1968). "On the relations between speech movements and speech perception," Z. Phon., Sprachwiss., U. Komm. Fschg. 21, 102–106.

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

Stevens, K. N., and Blumstein, S. E. (1976). "Context-independent properties for place of articulation in stop consonants," J. Acoust. Soc. Am. 59, S40 (A).

Stevens, K. N., and Blumstein, S. E. (1977). "Onset spectra as cues for consonantal place of articulation," J. Acoust. Soc. Am. 61, S48 (A).

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

Tallal, P. and Piercy, M. (1975). "Developmental aphasia: the perception of brief vowels and extended stop consonants," Neuropsychologia 13, 69–74.