449

The role of release bursts in the perception of [s]-stop clusters

Bruno H. Repp
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511-6695

(Received 29 July 1983; accepted for publication 13 December 1983)

The role of the release burst as a cue to the perception of stop consonants following [s] was investigated in a series of studies. Experiment 1 demonstrated that silent closure duration and burst duration can be traded as cues for the "say"–"stay" distinction. Experiment 2 revealed a similar trading relation between closure duration and burst amplitude. Experiments 3 and 4 suggested, perhaps surprisingly, that absolute, not relative, burst amplitude is important. Experiment 5 demonstrated that listeners' sensitivity to bursts in a labeling task is at least equal to their sensitivity in a burst detection task. Experiments 6 and 7 replicated the trading relation between closure duration and burst amplitude for labial stops in the "slit"–"split" and "slash"–"splash" distinctions, although burst amplification, in contrast to attenuation, had no effect. All experiments revealed that listeners are remarkably sensitive to the presence of even very weak release bursts.

PACS numbers: 43.70.Dn, 43.70.Ve

INTRODUCTION

A large proportion of speech perception research has been concerned with stop consonants. Nevertheless, there are still gaps in our knowledge of the relevant acoustic cues and their perceptual importance. While much attention has been lavished on the perception of stop consonant voicing and place of articulation, the more basic question of whether or not a stop consonant is perceived at all has been addressed in only a handful of studies. Moreover, nearly all of these studies have used synthetic speech stimuli in which at least one important cue was commonly absent: the release burst that terminates the stop closure. The present series of studies explores the role of this cue in the perception of stop consonants after [s].

A good deal is known about some other cues to stop manner perception, at least in the context of preceding [s] and following vowel or [1]. One very important cue is an interval of silence corresponding to the period of oral closure that characterizes stop consonant articulation. Early research at Haskins Laboratories by Bastian (1959, 1962) as well as the recent thorough investigations of Bailey and Summerfield (1980) have shown that an interval of silence between an [s]-noise and a steady-state synthetic vowel is generally sufficient to elicit a stop consonant percept, given that the silence is longer than about 20 ms (but not excessively long), and that the vowel is not too open. Silence frequently is not only sufficient but also necessary for the perception of a stop, for even when other stop manner cues are present in the signal (neglecting release bursts for the moment), stops are rarely perceived in the absence of an appropriate closure interval (Bailey and Summerfield, 1980; Best *et al.*, 1981; Dorman *et al.*, 1979; Fitch *et al.*, 1980).

Other relevant cues reside in the signal portions adjacent to the closure interval. Changes in spectrum and/or a rapid amplitude drop in the preceding fricative noise signify the approach of the closure and thereby contribute to stop perception (Repp, unpublished data; Summerfield *et al.*, 1981). Similarly, formant transitions and/or a rapid amplitude rise at the onset of the following vocalic portion—a rising transition of the first formant ($F1$) in particular—signify rapid opening and thereby constitute an important stop manner cue (Bailey and Summerfield, 1980; Best *et al.*, 1981; Fitch *et al.*, 1980). There is also evidence that the durations of the acoustic segments preceding and following the closure can influence stop manner perception (Summerfield *et al.*, 1981; however, see also Marcus, 1978). These additional cues engage in trading relations with the temporal cue of closure duration; that is, the stronger they are, the less closure silence is needed to perceive a stop. (For analogous findings for stops in vowel-[s] context, see Dorman *et al.*, 1980.) In general, however, these studies suggest that a minimal amount of silence (about 20 ms) is needed for a stop to be perceived at all.

Nearly all of the above-mentioned studies used synthetic speech stimuli which did not include any release bursts. One reason for this omission was presumably that good bursts are difficult to synthesize. Although most researchers are probably aware of the relevance of release bursts to the perception of stop manner, the importance of this cue has not been sufficiently acknowledged in the literature, which has emphasized the role of the closure duration cue. In an unpublished study, Repp and Mann (1980) took three tokens each of [sta], [ska], [ʃta], and [ʃka], produced by a male speaker, excised the closure period, and replaced the natural fricative noises with synthetic ones of comparable amplitude. In one condition, the stimuli retained the natural release bursts, and the subjects continued to report stop consonants on 100% of the trials, with very few place-of-articulation errors. In another condition, the release bursts were excised, and stop responses fell to 3% (except for two subjects who continued to report stops, but with poor accuracy for place of articulation). These data clearly illustrate the salience of the release burst as a manner cue for alveolar and velar stops following fricatives.[1] Labial stops, on the other hand, are associated with weaker release bursts (see Zue, 1976) that may not be sufficient to cue a stop percept in the absence of an appropriate closure interval.

The present series of studies attempts to answer several questions about the role of release bursts in stop manner perception: (1) Given that an interval of silence is needed to hear an alveolar stop when there is no release burst but not when there is one, how much can the burst cue be weakened before any silence is needed, and will further weakening of the burst result in increasing amounts of silence required? In other words, how sensitive are listeners to burst cues, and is there a regular trading relation between the burst and silence cues? These questions are explored in experiments 1 and 2 by manipulating alveolar burst duration and amplitude. (2) Given an effect of burst amplitude that can be traded against silence duration, experiments 3 and 4 investigate whether it is absolute or relative burst amplitude that matters. (3) Experiment 5 addresses the question of whether the point at which an attenuated release burst ceases to trade with silence coincides with the auditory detection threshold for the burst. (4) The role of burst amplitude is further investigated in experiments 6 and 7 with labial stops, with special attention to the question of whether amplification of a weak labial burst can make it a more powerful manner cue.

I. EXPERIMENT 1

The purpose of experiment 1 was to demonstrate the relative importance of an alveolar release burst as a stop manner cue, and to create a trading relation between burst and silence cues by varying the durations of both in natural-speech stimuli.

A. Method

1. Stimuli

Good tokens of "say" and "stay" were selected from recordings of several repetitions produced by a female speaker in a sound-insulated booth. These two utterances were low-pass filtered (− 3 dB at 9.6 kHz, − 55 dB at 10 kHz), digitized at 20 kHz, and modified by waveform editing. To reduce stop manner cues in the fricative noise, which were not of particular interest in the present study, the [s]-noise of "say," 176 ms in duration, was used in all experimental stimuli. This noise was followed by a variable interval of silence and by one of seven different, "day"-like portions, roughly 550 ms in duration. Six of these were derived from the token of "stay" while the seventh represented the vocalic portion of the "say" token.

Figure 1 shows the waveforms of the onsets of these stimulus portions. On top is the original post-closure portion of "stay," which began with a rather powerful (but, for that speaker, not atypical) release burst of somewhat less than 20 ms in duration. The rms amplitude of the total burst was determined to be 4.6 dB below the vowel onset and 6.8 dB below the vowel peak (135 ms later), with an amplitude decrease of about 10 dB from the initial to the final quartile of the burst.[2] The release burst was cut back in five steps, as indicated in the figure. Successive cuts (versions 2–6) were made at 6.1, 10.6, 13.4, 15.2, and 19.6 ms from the onset. These cutpoints were selected visually on the basis of local dips in the waveform. In each case, the cut was made at the nearest zero crossing. The stimulus portion derived from "say" is shown at the bottom of Fig. 1, aligned so as to show its similarity with version 5 of the "day" portion on top. Despite this similarity of waveforms, however, there were presumably some spectral differences between these two portions, due to the different contexts in which they had been articulated.

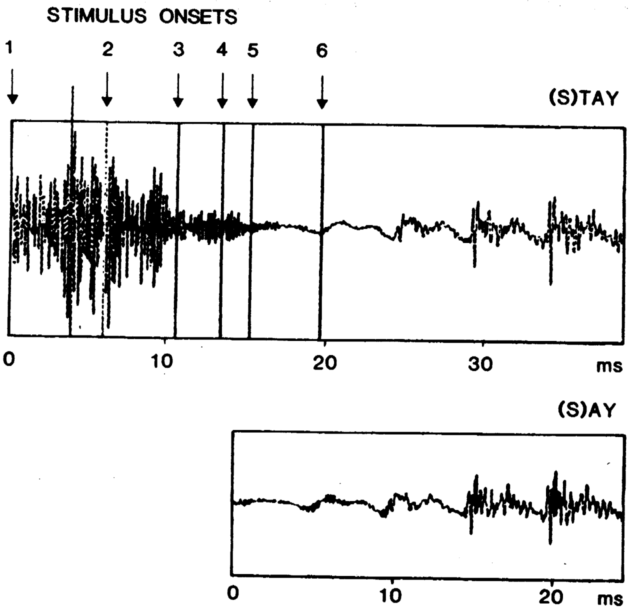

FIG. 1. Onset waveforms of stimuli used in experiment 1. The top panel shows the first 40 ms following the closure in "stay"; the bottom panel shows the first 24 ms following the fricative noise in "say." Arrows in the top panel indicate cut points for release burst truncation.

The silent interval separating the initial fricative noise from the "day" portions was varied from 0 to 60 ms in 10-ms steps. Because tokens with large bursts were expected to be perceived as "stay" even without any silence, a semi-orthogonal design was employed that assigned an increasingly wider range of silence durations to tokens with increasingly shorter bursts. Thus the stimuli with the most powerful burst occurred only with the 0-ms silence, while the stimulus derived from "say" occurred with all seven silent intervals. This led to a total of 28 different stimuli which were recorded on audio tape in ten different randomizations, with interstimulus intervals of 2 s.

2. Subjects and procedure

Ten subjects participated, including nine paid volunteers and one member of the laboratory staff (not a speech researcher). None of the subjects reported any hearing problems, and all had only very limited experience in speech perception experiments. The stimuli were presented binaurally over calibrated TDH-39 earphones in a quiet room.[3] The subjects identified in writing each stimulus as either "say" or "stay."

B. Results and discussion

Average percentages of "stay" responses are shown as a function of silent closure duration in Fig. 2, separately for each of the seven stimulus patterns. It is evident that versions

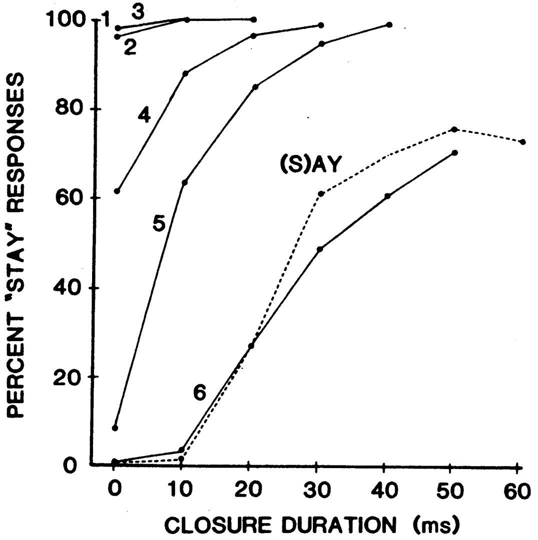

FIG. 2. Trading relation between alveolar release burst duration and closure duration (experiment 1). Numbers refer to cut points illustrated in Fig. 1. The dashed line represents the token derived from "say." Closure duration (abscissa) refers to the *actual* silence in the stimuli.

1, 2, and 3 were invariably identified as "stay," even in the absence of silence. Thus even the remainder of the burst following the initial high-amplitude portion (version 3, see Fig. 1) was a sufficient cue for stop manner. As the burst was cut back further, increasing amounts of silence were necessary to achieve a percept of "stay." The stimulus with the "say"-derived portion yielded results similar to those for version 6, and it appears that neither provided sufficient cues for unambiguous "stay" percepts, even at the longest silences used here.

What is most striking about these results is the large perceptual effect that a small burst cutback had on perception. The change from version 4 to version 5 consisted of the elimination of only 1.8 ms of relatively low-amplitude noise at onset (see Fig. 1); however, listeners needed approximately 10 ms more silence to compensate for this loss and achieve the same average rate of "stay" responses. Similarly, the change from version 5 to version 6 consisted of the elimination of the last 4.4 ms of burst residue. The perceptual effects were dramatic: At least 20 ms of additional silence were needed to compensate for the loss, and several listeners were not able to compensate for it at all, reporting only "say" for version 6. Even those few subjects who did reach a 100% "stay" asymptote for version 6 and had very steep labeling functions showed large effects of the stimulus manipulations.

Thus this study not only demonstrates a perceptual trading relation between burst duration and silence duration but also that listeners are remarkably sensitive to what seem to be rather minute changes in the onset characteristics of the stimulus portion following the silent closure interval. Of course, the truncation of the release burst introduced not only variations in burst duration but also changes in overall burst amplitude, in its onset amplitude characteristics, and perhaps correlated spectral changes. Any of these may have been responsible for the effects observed, but it is still true that relatively small physical changes had relatively large perceptual consequences.

II. EXPERIMENT 2

Experiment 2 examined one parameter that may have played a role in experiment 1—the overall burst amplitude. The purpose of the study was to demonstrate a trading relation between release burst amplitude and closure duration as joint cues to stop manner perception.

A. Method

1. Stimuli

In experiment 1, stimulus version 3 was just on the verge of requiring some silence in addition to the truncated burst, in order for a stop to be perceived on all trials (see Fig. 2). This stimulus was chosen as the starting point. Its residual burst was 9 ms in duration (see Fig. 1), with a total rms amplitude 10.8 dB below the vowel onset and 15.1 dB below the vowel peak. Five additional versions were created by digitally attenuating the burst by up to 30 dB in 6-dB steps. In a seventh version the burst was infinitely attenuated (i.e., it was replaced with 9 ms of silence); thus this stimulus was equivalent to stimulus version 6 in experiment 1.

Silent intervals ranging from 0 to 60 ms in 10-ms steps were assigned to the stimuli using the same design as in experiment 1. Thus, version 1 occurred only with the 0-ms interval while version 7 occurred with the full range of closure durations. The resulting 28 stimuli were recorded in ten different randomizations.

2. Subjects and procedure

Twelve new subjects participated in this study. The data of one had to be discarded because he could not reliably distinguish among the stimuli. The remaining 11 subjects included eight staff members of Haskins Laboratories (including the author) with varying amounts of experience in speech perception tasks, and three paid student volunteers. The procedure was the same as in experiment 1.

B. Results and discussion

Average percentages of "stay" responses are shown as a function of silent closure duration in Fig. 3, separately for each of the seven attenuation conditions. It is evident that there is an orderly progression of labeling functions: As the burst got weaker, more silence was needed to perceive a stop consonant.

The figure suggests that a burst attenuated by as much as 30 dB still led to more stop responses than a stimulus without any burst. This was confirmed in a one-way analysis of variance on the stop responses to these two types of stimuli, summed over closure durations of up to 40 ms, $F(1,10) = 9.8, p < 0.02$. Since, in the 30-dB attenuation condition, the amplitude of the 9-ms residual burst was about 45 dB below the vowel peak amplitude (or at about 38 dB SPL versus about 83 dB SPL for the vowel at the subjects' earphones), this finding again reveals that listeners are remarkably sensitive to burst cues.

Two additional comments are in order concerning Fig. 3. First, it should be noted that, in the infinite attenuation condition, the *nominal* closure ended at the beginning of the nonexistent burst. Therefore, the *actual* duration of the silence in these stimuli was 9 ms longer, as indicated by the arrows in the figure, which makes the results more nearly comparable to those for the same stimulus (version 6) in experiment 1 (see Fig. 1). It would not have been appropriate to plot these data in terms of actual silence duration because the effective silence durations resulting from various degrees of burst attenuation are not known. Note, however, that such a plot would tend to space the functions in Fig. 3 farther apart and thus increase the observed effects. This distinction between nominal and actual closure duration will recur in later experiments.

Second, it will be noted that, contrary to expectations based on experiment 1, the unattenuated stimulus did not receive 100% "stay" responses, while the burstless stimulus did reach this asymptote at the longer silences. Although there was considerable variability among individual subjects with regard to how the unattenuated stimulus was perceived, the pattern of the data suggests that the subjects gave somewhat more weight to closure duration and less weight to the burst in experiment 2 than in experiment 1. The reason for this is not known.

In summary, the present study demonstrated the expected trading relation between burst amplitude and closure duration, and it showed that severely attenuated (and truncated) bursts still can have a perceptual effect.

III. EXPERIMENT 3

Given the finding of the preceding study that burst amplitude is an important parameter, experiment 3 addressed the question of whether the perceptually relevant aspect of burst amplitude is its absolute magnitude or its magnitude relative to the surrounding signal portions.

A. Method

1. *Stimuli*

Taking the data of experiment 2 as a guideline, the stimulus with the 12-dB attenuation of the 9-ms residual burst was selected as the starting point for the present study. Four other stimuli were created by selectively attenuating portions of this original stimulus, as illustrated schematically in the upper-right-hand corner of Fig. 4. In addition to (a) the original stimulus, there were stimuli with attenuation of (b) only the burst, (c) both the burst and the following vocalic portion, (d) the burst and the preceding fricative noise, and (e) the whole stimulus. Attenuation was by 12 dB in all cases.

All stimuli occurred with all closure durations, which varied from 0 to 40 ms in 10-ms steps. The resulting 25 stimuli were recorded in ten different randomizations.

2. *Subjects and procedure*

Ten subjects participated, including six new paid volunteers and four staff members of Haskins Laboratories (including the author). Results were similar for the two groups of subjects and were combined. One subject reported only "say" during the first half of the test, so only the data from the second half were included. The procedure was the same as in experiments 1 and 2.

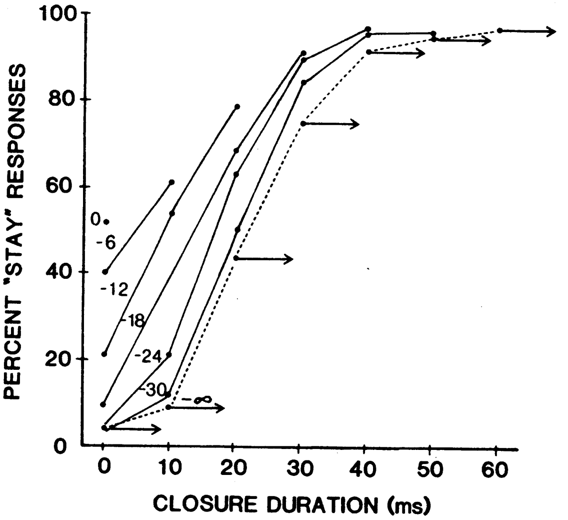

FIG. 3. Trading relation between alveolar release burst amplitude and closure duration (experiment 2). Negative numbers refer to amplitude decrement (in dB). Closure duration (abscissa) is *nominal*; the actual silence durations in the infinite-attenuation condition were 9 ms longer due to the silenced burst, as indicated by the arrows.

B. Results and discussion

The labeling functions for the five conditions are drawn in the top panel of Fig. 4. Clearly, the stimulus manipulations made a difference. This was confirmed by a one-way analysis of variance on the percentages of "stay" responses

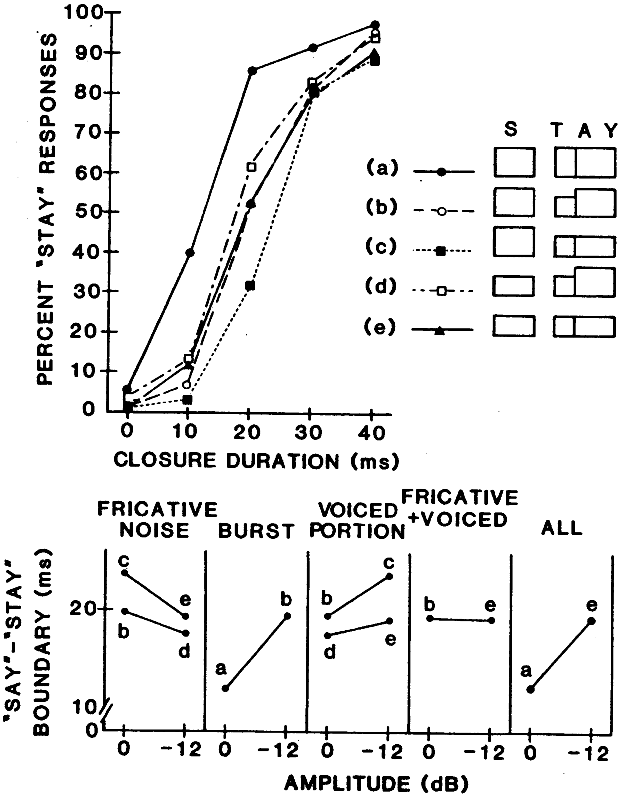

FIG. 4. Design and results of experiment 3. Labeling functions for the five conditions are provided on the upper left, with the key on the upper right. Rectangles in the key represent schematically the "s" fricative noise, the "t" burst, and the "day" voiced portion. The height of the rectangles represents amplitude relative to the base stimulus [condition (a)]. At the bottom, comparisons among the various conditions are presented in terms of average category boundary values (in ms of closure silence). Lower-case letters refer to the key on top.

summed over all closure durations, $F(4,36) = 12.6$, $p < 0.001$. Statistical comparisons among individual conditions were done by post-hoc Newman-Keuls tests. According to these tests, condition (a) differed significantly ($p < 0.01$) from all other conditions, and condition (c) differed ($p < 0.05$) from condition (d).

A graphic comparison among conditions is provided in the bottom part of Fig. 4 in terms of the location of the average "say"–"stay" boundary (obtained by linear interpolation between the data points straddling the boundary) on the closure duration dimension. Proceeding from left to right through the five panels, we see the following: (1) Attenuation of the fricative noise, holding the other stimulus components constant, increased the number of stop responses slightly (i.e., the boundary shifted to a shorter silence duration). (2) Attenuation of the burst decreased stop responses substantially, which replicates experiment 2. (3) Attenuation of the voiced portion resulted in a slight decrease in stop responses. (4) Attenuating both the fricative noise and the voiced portion together had absolutely no effect. (5) Attenuation of the whole stimulus caused a substantial decrease in stop responses equivalent to that resulting from attenuation of the burst alone.

These results point toward absolute burst amplitude as the relevant factor. Clearly, attenuating the burst's environment did not have the same effect as amplifying (more precisely, restoring) the burst by the same amount (see Fig. 3). Contrary to expectations, attenuation of the vocalic portion did not increase stop responses. Perhaps, additional stop manner cues contained in that portion (initial formant transitions and amplitude envelope) were weakened by the attenuation, thus counteracting the gain in burst salience relative to its environment. If so, however, we are forced to conclude that the absolute amplitude of *those* cues matters, which is equally interesting.

Another possibility is that the present study suffered from floor effects due to listeners' inability to detect the burst when it was attenuated. This would explain why the largest difference occurred between condition (a) and all others. Note that conditions (a) and (b) were equivalent to the 12-dB and 24-dB attenuation conditions in experiment 2. The average category boundaries for these conditions were at 9 and 17 ms, respectively, in experiment 2, and at 10 and 20 ms in experiment 3—a rather close agreement. Note also that, in experiment 2, a burst attenuated by 24 dB still had a significant perceptual effect. The agreement between experiments 2 and 3 suggests that the absolute stimulus amplitudes were similar, and that no floor effect occurred. Nevertheless, it seemed advisable to replicate the present results with the burst amplitude set at somewhat higher absolute levels, and with inclusion of a no-burst baseline condition.

IV. EXPERIMENT 4

This replication of experiment 3 used new stimuli in a complete $2 \times 3 \times 2$ orthogonal design. By including burstless stimuli in the design, it was possible to examine the effects of fricative noise and vowel attenuation separate from their effects on the relative salience of the burst—an important control condition.

A. Method

1. Stimuli

Good tokens of "say" and "stay" were selected from among several repetitions recorded by a new female speaker. Both utterances were digitized at 20 kHz. As in experiments 1–3, the fricative noise (170 ms long) was taken from "say." The "day" portion of "stay" was about 450 ms in duration and began with a release burst 13.35 ms long. The overall rms amplitude of this burst was determined to be 5.5 dB below the vowel onset, 11.0 dB below the vowel peak (only 20 ms later), and 4.1 dB above the fricative noise maximum. Informal listening confirmed that this burst, as usual, was sufficient for "stay" to be perceived without any closure silence (see also experiment 5). To be able to trade burst amplitude against silence, the most intense burst used was 15 dB below the original. A total of 12 stimulus versions were created by orthogonally combining three factors: fricative noise attenuation (0 or 10 dB), burst attenuation (15 or 25 dB, or no burst at all),^{4} and "vowel" attenuation (0 or 10 dB). Each of these 12 versions occurred with five closure durations ranging from 0 to 40 ms in 10-ms steps. The resulting 60 stimuli were recorded in five different randomizations.

2. Subjects and procedure

Ten new paid volunteers identified the stimuli as "say," "stay," "spay," or "svay." The last two response alternatives were included because the author, as a pilot subject, had noticed a tendency to hear these additional categories. The tape was repeated once, so that each subject gave ten responses to each stimulus.

B. Results and discussion

Of the ten subjects, three gave only "say" and "stay" responses, while the other seven used one or both of the additional response categories as well. In the initial analysis, all consonant cluster responses were pooled.

Since the burstless stimuli had been created by omitting the burst rather than by infinitely attenuating it, 13.35 ms (the duration of the burst) must be subtracted from their actual closure durations to directly compare results for stimuli with and without bursts. This has been done graphically in Fig. 5, where the arrows point toward the actual closure durations. The figure shows average labeling functions for the three burst conditions, averaged over fricative and vowel attenuation conditions. Clearly, the subjects gave many more cluster responses to the stimuli with bursts than to those without. Elimination of the burst resulted in a flattening of the labeling function; 40 ms of silence was not enough to make a burstless stimulus sound like an unambiguous "stay." The figure also shows the expected effect of the 10-dB burst attenuation. It is clear that this experiment avoided the danger of floor effects; if anything, the burst amplitudes were somewhat too high.

The effects of variations in fricative noise and vowel amplitude, which were smaller than the effects of burst amplitude, are summarized in Table I in the form of response percentages averaged over all closure durations. A three-

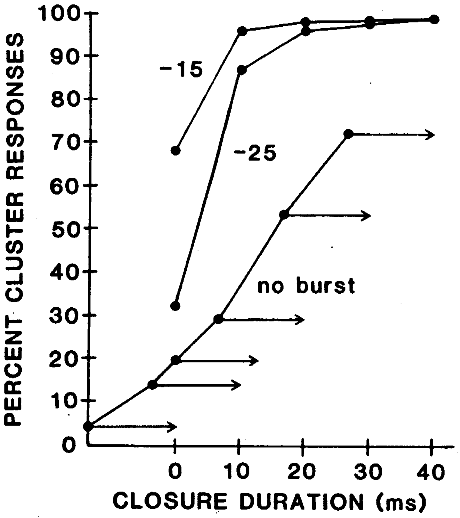

FIG. 5. Effect of burst amplitude in experiment 4, averaged over other amplitude conditions. Numbers refer to burst amplitude (in dB) relative to original burst. Closure durations are nominal; actual silence durations in the no-burst condition are indicated by arrows.

way repeated-measures analysis of variance (with the factors Burst, Fricative, and Vowel) was first conducted on the total cluster responses, ignoring the incommensurability of actual closure durations for stimuli with and without bursts. This analysis revealed, besides the expected Burst effect, $F(2,18) = 64.8$, $p < 0.001$, significant main effects of both Fricative, $F(1,9) = 9.6$, $p < 0.02$, and Vowel, $F(1,9) = 12.4$, $p < 0.01$, as well as significant interactions between Burst and Fricative, $F(2,18) = 10.7$, $p < 0.001$, and between all three factors, $F(2,18) = 12.1$, $p < 0.001$. To clarify the triple interaction, separate analyses of variance were conducted on stimuli with and without bursts. Stimuli with bursts exhibited significant main effects of Burst, $F(1,9) = 32.7$, $p < 0.001$, Fricative, $F(1,9) = 45.2$, $p < 0.001$, and Vowel, $F(1,9) = 10.4$, $p = 0.01$, as well as a marginal Burst by Fricative interaction, $F(1,9) = 5.4$, $p < 0.05$, and a strong triple interaction, $F(1,9) = 30.6$, $p < 0.001$. Thus the triple interaction was not due to different patterns of results for stimuli with and without bursts. The separate analysis of burstless stimuli revealed only a significant effect of Vowel, $F(1,9) = 8.5$, $p < 0.02$, not of Fricative.

Consider now the directions of these effects. The Burst effect, of course, was due to a decrease of cluster responses as the burst was attenuated or eliminated altogether (Fig. 5). The Fricative effect, too, was in the expected direction: Attenuation of the fricative noise *increased* the number of cluster responses. (A similar but nonsignificant trend was observed in experiment 3.) This is the kind of effect that might be expected if the fricative noise reduced the salience of the burst through some form of auditory forward masking (see Delgutte, 1980). This interpretation is supported by the finding that the Fricative effect was absent in burstless stimuli, where there was no burst to be masked (see Table I).

Turning now to the Vowel effect, it can be seen in Table I that attenuation of the vocalic portion, like attenuation of the fricative noise, resulted in an *increase* of cluster responses, contrary to a nonsignificant opposite trend observed in experiment 3. Since this was true regardless of whether a burst was present or absent, the effect was apparently not due to release from a backward masking effect of the vowel on the release burst, or simply to an increase in the salience of the burst relative to the vowel.[5]

The results are complicated by the triple interaction, which was due to the fact that, with the higher burst amplitude, fricative and vowel attenuation seemed to have independent effects whereas, with the lower burst amplitude, only simultaneous attenuation of both produced an effect. An explanation of this complex pattern is beyond reach at the moment.[6]

In summary, this experiment, in conjunction with experiment 3, provides little support for a role of relative burst amplitude in stop manner perception. While the preceding fricative noise may exert a slight masking effect on the burst, the amplitude of the following vocalic portion seems to have its perceptual effects primarily by changing the relative salience of cues contained in that portion itself. While the present data cannot be considered the last word on the issue, the possibility of a fixed perceptual criterion in the amplitude domain deserves further attention, both with regard to the perception of stop manner and to place-of-articulation distinctions in stops (see Ohde and Stevens, 1983) and fricatives (Gurlekian, 1981).

TABLE I. Response pattern in experiment 4, averaged over closure durations.

Stimulus amplitude (dB)			Response (percent)				
Burst	Fricative	Vowel	"say"	"stay"	"svay"	"spay"	Total cluster
−15	0	0	16.6	79.8	3.0	0.6	83.4
	−10	0	6.8	87.4	3.4	2.4	93.2
	0	−10	8.2	90.8	1.0	0.0	91.8
	−10	−10	1.2	97.4	0.8	0.6	98.8
−25	0	0	20.2	57.4	16.8	5.6	79.8
	−10	0	19.6	57.2	17.8	5.4	80.4
	0	−10	20.2	74.8	3.6	1.4	79.8
	−10	−10	10.4	86.0	2.2	1.4	89.6
no burst	0	0	67.6	12.4	10.0	10.0	32.4
	−10	0	70.8	7.0	11.4	10.8	29.2
	0	−10	60.8	25.4	5.2	8.6	39.2
	−10	−10	61.8	20.8	8.8	8.6	38.2

V. EXPERIMENT 5

The preceding experiments, experiment 2 in particular, demonstrate a remarkable sensitivity of listeners to the presence of even very weak release bursts. This suggests the hypothesis that the point at which a burst becomes ineffective and ceases to trade with closure silence actually coincides with the auditory detection threshold for the burst. This hypothesis was tested in the present experiment.[7] In addition, the study examined whether the detectability of the burst is increased when the preceding fricative noise is removed.

A. Method

1. Stimuli

The stimuli were derived from the utterances that also provided the basis for the stimuli of experiment 4. In addition to the original stimulus (full burst amplitude), six levels of burst attenuation were employed: 10, 20, 25, 30, 35, and ∞ dB. In the *identification test*, these seven stimuli occurred with nominal closure durations of 0, 10, 20, and 30 ms. Ten different randomizations of the 28 stimuli were recorded.

In addition, two *discrimination tests* were assembled, which required subjects to detect the presence of a burst. The two tests were identical except that in one the initial fricative noise was omitted from all stimuli while, in the other, the fricative noise was followed by a fixed 10-ms closure interval. A fixed-standard same–different paradigm was employed. The fixed standard was the burstless stimulus; it occurred first in every stimulus pair. After a fixed interval of 500 ms, the comparison stimulus occurred; it either did or did not contain a release burst. Over six successive test blocks, the burst in the comparison stimulus was attenuated by 0, 10, 20, 25, 30, and 35 dB. Each test block consisted of 50 trials, the first 10 of which were practice, with the responses alternating between "same" and "different" and known in advance. Half of the remaining 40 trials were "same" and half were "different," in random order. The intertrial intervals was 2 s.

2. Subjects and procedure

Ten paid volunteers participated in the experiment, six of whom had also been subjects in experiment 4. In the identification test, which was always presented first, they responded "say" or "stay," with "svay" and "spay" as additional options. In the discrimination tests, the responses were "s" ("same") and "d" ("different"). The order of the two discrimination tests was counterbalanced across subjects. Playback amplitude was controlled by adjusting the level so as to achieve a constant maximum deflection on a vacuum tube voltmeter, and by keeping it at that level throughout the experiment. All tapes had been recorded at the same level. The peak amplitude of the vowel (and, hence, of the unattenuated burst as well—see experiment 4) at the subjects' earphones was estimated to be approximately 83 dB SPL.

B. Results and discussion

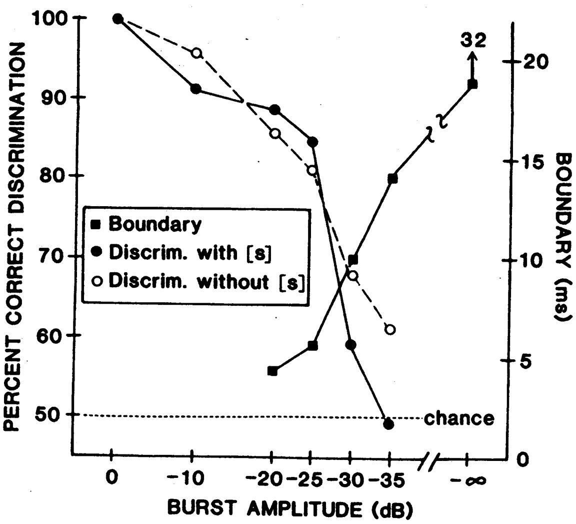

FIG. 6. Identification and burst detection results from experiment 5. Filled squares plot the category boundary in ms of silence (right ordinate) as a function of burst amplitude. Silence duration for the no-burst condition is nominal; actual silence at the boundary was 32 ms, as indicated. Circles show burst detection scores as percent correct (left ordinate) for stimuli with and without initial [s]-noise.

The average data are presented in Fig. 6. The labeling boundary for stimuli with bursts attenuated by 20 dB or more is represented by the filled squares. Stimuli with an unattenuated burst were uniformly identified as "stay," and those with a − 10 dB burst received only 16% "say" responses when no silence was present, so no boundaries could be determined for these stimuli. As expected, the boundary shifted toward increasingly longer values of silence as the burst was attenuated.[8] Note that the boundary seemed to increase beyond the 35-dB burst attenuation, although the difference between this condition and the burstless condition fell short of significance in a t test.

The discrimination (i.e., burst detection) results for the same stimuli are plotted in terms of percent correct as the filled circles in Fig. 6. Performance was perfect for the original burst and declined with increasing burst attenuation, first slowly, and then more rapidly beyond 25 dB. For stimuli with the initial [s]-noise, performance reached chance at the 35-dB attenuation. Note that the category boundary in the identification task continued to shift beyond that point for at least some listeners, suggesting that subjects' sensitivity to the burst was at least as great in phonetic labeling than in auditory discrimination. This result provides strong evidence of the sensitivity of phonetic categorization processes to very subtle changes in acoustic information.

Figure 6 also shows that burst detection was somewhat improved when the initial [s]-noise was removed, but only at the two weakest burst intensities (not a significant difference). Thus there may have been a slight auditory masking effect of the fricative noise on the burst, in agreement with experiments 3 and 4.

VI. EXPERIMENT 6

The purpose of experiment 6 was to demonstrate a trading relation between burst amplitude and closure duration for the perception of a labial stop consonant. Labial bursts are weaker than alveolar and velar bursts (Zue, 1976), and informal observations have suggested that they are generally insufficient cues for stop manner. In other words, some closure silence is usually needed to perceive "sp," even with the original burst in place. This raises the question of whether labial bursts function as manner cues at all; perhaps, they merely add to the effective closure silence. Moreover, labial bursts offer the opportunity of observing not only effects of attenuation but also of amplification. Would an appropriately amplified labial burst become a sufficient stop manner cue?

The "slit"–"split" contrast was selected for the present study for several reasons. First, it has been used extensively in earlier studies (Bastian *et al.*, 1962; Dorman *et al.*, 1979; Fitch *et al.*, 1980; Marcus, 1978; Summerfield *et al.*, 1981). Second, a "p" tends to be heard in this context as long as there are no strong cues to a nonlabial place of articulation in the signal portions surrounding the silent closure interval. That is, listeners report "split" when separately produced "s" and "lit" utterances are joined together with a sufficient interval of silence in between (Dorman *et al.*, 1979). According to limited informal observations, the [l] resonances following a stop closure, unlike those of a full vowel, do not seem to harbor any significant formant transition cues to stop manner and place of articulation, which makes the "slit"–"split" contrast different from the "say"–"stay" contrast employed in experiments 1–5. This fact may be partially responsible for the finding (cf. Fitch *et al.*, 1980; Best *et al.*, 1981) that, in burstless stimuli, the typical "slit"–"split" boundary is located at much longer silent closure intervals (50–80 ms; for an exception, see Marcus, 1978) than the "say"–"stay" boundary (10–30 ms). Differences in place of stop articulation and in phonetic environment may also contribute to this boundary difference, however. One reason for conducting experiment 6 (as well as experiment 7) was to see whether the presence of a labial release burst, amplified to equal the power of an alveolar burst, might shift the "slit"–"split" boundary to the short silences characteristic of the "say"–"stay" boundary.

A. Method

1. Stimuli

A good token of "split" was selected from several utterances produced by a female speaker and was digitized at 20 kHz. In the original utterance, the initial [s]-noise (105 ms) was followed by a silent closure interval (about 150 ms) and a "blit" portion consisting of an initial release burst (16 ms), a voiced portion (about 230 ms), a silent [t]-closure, and a final [t]-release burst. The major energy of the labial release burst was concentrated in the first 4 ms. The rms amplitude of these first 4 ms was determined to be about 14 dB below the [l] maximum, and 20 dB below the [ɪ] vowel maximum. The final 12 ms of the burst were about 13 dB below its initial 4 ms.[9]

Three additional stimulus versions were created by either amplifying or attenuating the 16-ms burst by 12 dB, or by eliminating it altogether. The (actual) silent closure duration in each of the four versions was varied from 40 to 100 ms in 10-ms steps. The resulting 28 stimuli were recorded in ten different randomizations.

2. Subjects and procedure

The same ten subjects as in experiment 3 identified the stimuli as "slit" or "split." Because the author noted that some of the stimuli sounded like "stlit" to him, this response alternative was provided as well. Stimuli without any clear consonant between the "s" and the "l" were to be considered instances of "slit."

B. Results and discussion

Since "stlit" responses were rather infrequent, Fig. 7 shows the combined percentage of "split" and "stlit" responses as a function of closure duration and of burst conditions. Three results are evident. First, attenuation of the burst by 12 dB had a clear effect, especially at the longer silences. Apparently, burst attenuation resulted not so much in a boundary shift as in a flattening of the labeling function. Second, the condition in which there was no burst at all gave results very similar to the attenuated-burst condition, provided the no-burst function is shifted to make the nominal closure durations comparable across the two conditions. (The actual closure durations were 16 ms longer, as indicated by the arrows in Fig. 7.) This result is not surprising, given the initial low amplitude of the labial burst. Third, amplification of the burst by 12 dB had, surprisingly, no effect at all. One side effect of the amplification seemed to be a tendency to hear "stlit" rather than "split," in accord with recent data by Ohde and Stevens (1983) showing that burst amplitude is a cue to the labial-alveolar distinction. However, the present tendency was exhibited only by three of the ten subjects. A bias against the unfamiliar "stl" cluster may have played a role.

The effect of burst attenuation or elimination demonstrates that labial bursts, too, have a function as stop manner cues. The absence of any effect of burst amplification, however, suggests that the "slit"–"split" boundary cannot be

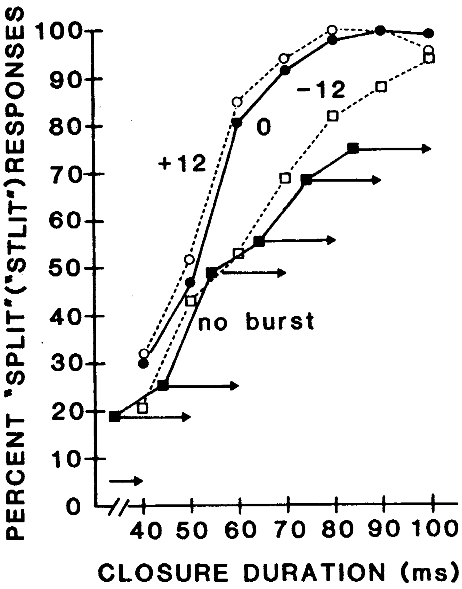

FIG. 7. Effects of labial release burst amplitude in experiment 6. Numbers refer to amplitude in dB relative to the original burst. Closure durations in the no-burst condition are nominal; actual durations are indicated by arrows.

easily pushed toward shorter values of silence. Although one might have expected burst amplification to shift the boundary on purely psychoacoustic grounds, it seems that the amplitude increment was either ignored by listeners or channeled into decisions about stop place of articulation rather than stop manner. This curious and potentially important finding called for a replication experiment.

VII. EXPERIMENT 7

This study was similar to experiment 6, except for differences in stimuli and the ranges of closure durations and burst amplitude values.

A. Method

1. Stimuli

Good tokens of the utterances "slash" and "splash" were recorded by a different female speaker and digitized at 20 kHz. The fricative noise of "slash" (142 ms) was used in all stimuli. The remainder (about 590 ms) was taken from "splash." This portion included an initial 10-ms release burst. (The original closure duration was 66 ms.) The amplitude of the burst was determined to be 7.4 dB below the [1] onset, 11.9 dB below the vowel maximum (75 ms later), and 2.9 dB above the fricative noise maximum (120 ms after noise onset). Six stimulus versions were created by leaving the burst unchanged, amplifying or attenuating it by 10 or 20 dB, or omitting it altogether. Each version occurred with (actual) closure durations ranging from 20 to 60 ms in 10-ms steps. The resulting 30 stimuli were recorded in ten different randomizations.

2. Subjects and procedure

The same ten subjects as in experiment 4 participated. They identified the stimuli as "slash" or "splash," with "stlash" as an additional option. To prepare the subjects for the amplified bursts, the instructions mentioned that some of the stimuli might have "pops" in them, which were to be ignored. The data of one subject had to be discarded because of numerous response omissions.

B. Results and discussion

The results are shown in Fig. 8. The left panel displays the labeling functions for the different burst amplitude conditions. In two respects, the findings replicate the principal results of experiment 6: Attenuation of the burst necessitated a longer interval of silence, whereas burst amplification did *not* have the opposite effect; rather, amplified bursts seemed to function like slightly attenuated ones. In two other respects, the results are different from those of experiment 6: The boundaries were considerably shorter here, and even the 20-dB burst attenuation condition still produced substantially more stop percepts than the burstless condition. These differences may indicate that the present release burst was a more powerful manner cue than that in the previous experiment. In addition, the different range of closure durations, as well as other stimulus characteristics, may have contributed to the boundary difference.

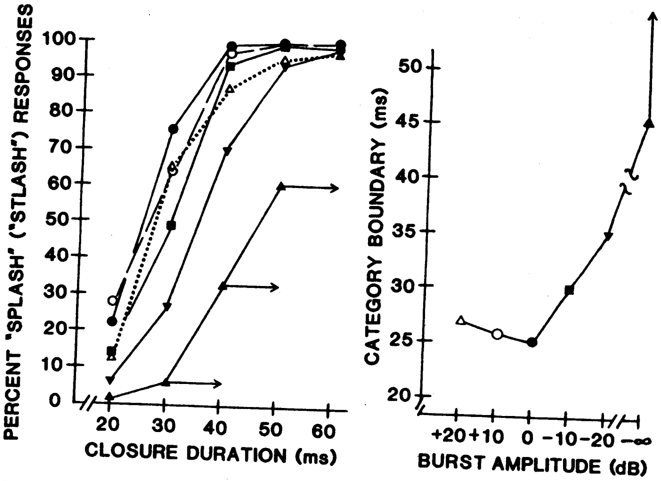

FIG. 8. Effects of labial release burst amplitude in experiment 7. Left panel is analogous to Fig. 7, with legend provided by right panel. Right panel shows category boundary as a function of burst amplitude. Actual silence durations are indicated by arrows.

The right-hand panel in Fig. 8 summarizes the data by plotting the boundary location as a function of burst amplitude. It is plain that burst amplification did not continue the trend established by the burst attenuation results: As soon as the amplitude exceeded that of the original burst, its trading relation with silence duration came to an abrupt end.[10] How is this finding to be explained?

Only four of the nine subjects gave any "stlat" responses. These responses were fairly broadly distributed but tended to occur with the higher burst amplitudes and at short closure durations. However, these weak trends observed in a few subjects are not nearly sufficient to explain the sudden end of the trading relation between burst amplitude and silence duration.

A more relevant observation is that, to the author (and presumably to the subjects as well), the amplified bursts sounded like extraneous pops superimposed on the stimuli. This subjective impression suggests that amplification of the burst destroyed its auditory coherence with the other signal portions and caused it to "stream off." If so, it is particularly interesting that subjects perceived these stimuli *not* as if they had no bursts at all, but rather as if they had a burst of "normal" amplitude (see Fig. 8). This finding thus seems related to two other intriguing phenomena described in the literature: duplex perception (e.g., Liberman *et al.*, 1981) and phoneme restoration (e.g., Samuel, 1981).

In duplex perception, a component of a speech stimulus is heard as a separate nonspeech event while, at the same time, it contributes to phonetic perception. Although the auditory segregation of the component is commonly achieved by dichotic channel separation, monaural duplex perception may occur when an acoustic cue, because of certain extreme properties, loses its coherence with the rest of the stimulus (see also Miller *et al.*, 1983). The present experiment seems to provide such an instance. Its results are also related to phoneme restoration, which is said to occur when a portion of a speech signal is replaced with an extraneous sound without affecting phonetic perception. Samuel (1981) has shown that, for restoration to occur, the extraneous sound must be a potential masker of the replaced portion.

Thus, the so-called phoneme restoration effect may, at least in part, be a "cue restoration effect"; that is, listeners fill in missing *acoustic* information. A particularly relevant study was conducted by Pastore *et al.* (1982): A syllable-initial [p] in one ear was perceived as "t" when a noise burst occurred in the other ear, but only when the noise included the frequencies typical of [t] release bursts. These findings combine aspects of duplex perception and cue restoration, as indeed do the present results. The amplified bursts were, of course, the best possible maskers of spectrally identical "normal" bursts, and because they segregated as "pops" from the rest of the signal, listeners were led to perceptually restore the original burst. If this interpretation is correct, then the data provide a particularly interesting demonstration of the detailed tacit knowledge of acoustic (or, perhaps, articulatory) properties of speech that listeners possess and apply in the course of phonetic perception.

VIII. GENERAL DISCUSSION

The present series of studies fills some gaps in our knowledge of the acoustic cues for stop manner perception. They uniformly show that the release burst is a highly important cue for the perception of stops after [s].

One result that emerges from the experiments is that a natural alveolar release burst is usually sufficient to cue perception of a stop in the absence of closure silence (experiments 1 and 5), whereas a natural labial release burst is usually not sufficient by itself (experiments 6 and 7). Although, in the present studies, alveolar release bursts were followed by pronounced vocalic formant transitions while labial bursts were not, preliminary observations indicate that the generalization holds regardless of following context, and that velar release bursts are similar in salience to alveolar ones. The greater power of alveolar and velar bursts is, in large part, due to their greater amplitude and longer duration, although spectral composition and/or different perceptual criteria for stops at different places of articulation may also play a role.

A second result of the present research is that listeners are extremely sensitive to the presence of even very brief or severely attenuated release bursts (experiments 1, 2, and 5). Experiment 5 showed that, when labeling stimuli phonetically, listeners are at least as sensitive to the presence of such minimal bursts as they are in a low-uncertainty burst detection task. As Nooteboom (1981) has pointed out, "phoneme identification seems to be an excellent way of measuring just noticeable differences" (p. 149). This is not a trivial result, for it suggests that the perceptual criteria employed in phonetic identification are extremely stable and finely tuned, despite the high stimulus uncertainty prevailing in a randomized identification test. Indeed, preliminary data suggest that this stability and sensitivity is maintained even in listening to fluent speech. The operation of stable criteria, internal to the listener and presumably shaped by language experience, is a hallmark of phonetic perception. Nevertheless, these criteria must also be flexible to accommodate natural variability in speech, such as might be due to changes in articulatory rate. In other words, the criteria are stable but not fixed; they are stable in the sense that their variability is not random but controlled by relevant factors.

A third finding is that release bursts, when shortened or attenuated in various degrees, engage in a regular trading relation with closure duration, a second important cue for stop manner: The weaker the burst, the more silence is needed to perceive a stop. There are two contrasting hypotheses about the origin of such a trading relation: It may either be phonetic or psychoacoustic in origin. According to the phonetic hypothesis (see Repp, 1982), the listeners' internal criteria specify the "prototypical" acoustic properties for the relevant phonetic segments, so that a reduction in one relevant property must be compensated for by an increase in another property to maintain the same response distribution. (A similar prediction could be derived from the information integration model of Oden and Massaro, 1978; see also Massaro and Oden, 1980.) According to the psychoacoustic hypothesis, on the other hand, the principal cue for stop manner resides in the onset characteristics of the signal portion (which includes the burst) following the closure silence, and the role of the silence is to prevent a forward masking effect of the preceding fricative noise on the auditory representation of those characteristics, and/or to enable the listener to attend to the critical onset properties. (This hypothesis is also congenial to the acoustic invariance hypothesis of Stevens and Blumstein, 1978.)

The present results are not wholly incompatible with psychoacoustic explanations. For example, the finding that attenuation of the fricative noise resulted in a reduction of the amount of silence needed for stop perception (experiments 3 and 4), but only when a burst was present (experiment 4), could be attributed to auditory forward masking. Effects of burst amplitude on stop manner perception also lend themselves to a psychoacoustic interpretation in terms of burst detectability. Data from other recent studies, however, argue strongly against a psychoacoustic account at least of the role of silence in stop manner perception. Best *et al.* (1981) found that the trading relation between closure duration and the $F1$ transition for the "say"–"stay" contrast was absent in nonspeech analogs of the stimuli. Repp (1983b) demonstrated that this same trading relation, as well as that between closure duration and burst amplitude in "slit"–"split," was restricted to the phonetic boundary region but absent within phonetic categories. Perhaps the strongest result was recently reported by Pastore *et al.* (1983): When the [s]-noise and the vocalic portion of "slit"–"split" tokens were differentially lateralized, so as to reduce peripheral auditory masking and facilitate selective attention, the amount of closure silence needed to perceive "split" remained the same. These results strongly favor a phonetic account of the integration of acoustic cues in stop manner perception, without ruling out certain psychoacoustic interactions in the peripheral auditory system that may, for example, affect burst detectability.

Two findings were unexpected and should provide a stimulus for further research. One result is that, apparently, burst amplitude has its effect on stop manner perception in absolute terms, not relative to the amplitude of the following signal portion (experiments 3 and 4). The role that potential stop manner cues in this voiced portion may have played needs to be examined in a more controlled fashion. The results may suggest, however, that important stop manner cues reside in the first few milliseconds following the closure—that is, in the absolute magnitude and slope of the sudden energy increment.

A second unexpected finding was the absence of a trading relation between *amplified* labial release bursts and closure duration (experiments 6 and 7). This phenomenon was tentatively interpreted as an instance of "cue restoration": The amplified burst was perceived as an extraneous "pop" and thus, instead of functioning as a cue in the speech signal, assumed the role of a masker for the cue expected by listeners—viz., of the "normal" release burst represented in listeners' detailed tacit knowledge of the normative acoustic properties of speech. A relation may exist between this phenomenon and the demonstration by Pols and Schouten (1978) that burstless initial stop consonants are more accurately perceived when preceded by pink noise (a potential masker of an absent burst) and Samuel's (1981) findings on the role of "bottom-up confirmation" in the phoneme restoration paradigm.

In conclusion, the present experiments have yielded factual information on the perception of a little-investigated cue as well as several intriguing effects that should stimulate further research. The results provide a modest challenge to psychoacoustic theories of speech perception. From a psychoacoustic viewpoint, stop manner perception seems a much simpler problem than, for example, perception of place of articulation: All that may be involved is the detection of some critical amount of energy increment or discontinuity in the signal. The eventual success or failure of psychoacoustic theories will rest, of course, on their ability to explain all kinds of phonetic perception, as well as to predict specific results from a model of auditory speech processing. Interesting work along these lines is now in progress (Delgutte, 1980, 1982; Goldhor, 1983), and the present data, being relatively straightforward, may provide a convenient testing ground for new models of peripheral auditory processing.

ACKNOWLEDGMENTS

This research was supported by NICHD Grant HD01994 and BRS Grant RR05596 to Haskins Laboratories. Results of experiments 1 and 2 were reported at a meeting of the Acoustical Society of America in Orlando, FL, November 1982. I am grateful to Joanne Miller, Ralph Ohde, Sigfrid Soli, and Douglas Whalen for helpful comments on an earlier draft. Special thanks go to Joanne Miller for handling the editorial processing of this manuscript.

[1] Dorman *et al.* (1980) found that the presence of an alveolar release burst was not sufficient for perception of a stop in vowel-fricative context (i.e., of an affricate, as in "ditch") in the absence of closure silence. While it is difficult to generalize from results obtained with single tokens of natural speech, it is possible that release bursts are more effective stop manner cues in fricative-vowel than in vowel-fricative environments.

[2] Amplitude measurements were performed after redigitizing the utterance without preemphasis, using a program of the ILS speech analysis system. The powerful appearance of the burst in Fig. 1 is in part due to high-frequency preemphasis. "Vowel onset" refers here to the 20 ms of waveform immediately following the 20-ms burst. The burst, as defined here, may have included a first, extremely weak glottal pulse (between cutpoints 5 and 6 in Fig. 1). No attempt was made to distinguish between transient, fricative, and aspirative phases of the burst (see Fant, 1973).

[3] Playback amplitude was not precisely calibrated but was held constant within a few dB by maintaining a certain setting of the level control on the tape recorder (Ampex AG500) for all subjects. The peak amplitude of the vowel at the subjects' earphones (approximately 83 dB SPL) was estimated postexperimentally by converting the peak deflection of a vacuum tube voltmeter in response to the test syllables into dB SPL, according to a chart prepared by Haskins Laboratories technicians.

[4] For no particular reason, the burst was excised rather than infinitely attenuated. The latter procedure would have been preferable, but there are no serious consequences for the interpretation of the results. (The same applies to experiments 6 and 7.)

[5] The reason for the different effects of vowel attenuation in experiments 3 and 4 is not clear; they may have been due to different strengths of the stop manner cues in the vocalic portions used. In experiment 3, no tendency to hear consonants other than "t" was noted, and a 40-ms silence always yielded close to 90% "stay" responses. In experiment 4, on the other hand, a significant number of "svay" and "spay" responses occurred, and even when these were pooled with "stay" responses, the total percentage for burstless stimuli with a 40-ms silence was only 72. Therefore, the vocalic portion in experiment 4 seemed to contain weaker stop manner cues than that in experiment 3, and this may explain the different effects of attenuation.

[6] Finally, the pattern of "svay" and "spay" responses may be considered (Table I). Attenuation of the burst increased both types of responses, simultaneously decreasing "stay" responses. Total elimination of the burst increased primarily "spay" responses. There was also a consistent Vowel effect, with both "svay" and "spay" responses being less frequent when the vocalic portion was attenuated. Fricative amplitude, on the other hand, had no effect on these responses at all. Closure duration did play a role (not shown in Table I): "svay" responses decreased as closure duration increased in stimuli with bursts, but increased (high vowel amplitude) or remained constant (low vowel amplitude) in burstless stimuli; "spay" responses showed a strong increase with closure duration, provided they occurred at all (stimuli with low burst and high vowel amplitude, and burstless stimuli). The latter trend is in agreement with earlier observations that long closure durations favor perception of a labial place of articulation (Bailey and Summerfield, 1980). "Svay" percepts, on the other hand, may have resulted from either "misinterpreting" the burst as frication when the closure was short, or—in burstless stimuli—they may have taken the place of a possible "sthay" category, which is difficult to perceive but corresponds to the informal observation that burstless "day" portions often resemble "they." In either case, however, attenuation of the vocalic portion favored "stay" over "svay" and "spay," which indicates a role of the vocalic onset envelope in this distinction.

[7] There is a long-standing controversy, familiar from the literature on categorical perception (see Repp, 1983a, for a review), about whether speech perception experiments should be concerned with what listeners *can* do in an optimal situation or with what they do under normal circumstances. Auditory thresholds are often assessed in highly practiced listeners after many hours of training. No strong claim is being made here that these optimal thresholds coincide with the limit of burst effectiveness in phonetic identification, although they obviously define a lower bound. Rather, the hypothesis tested here concerns the burst detection threshold for unpracticed listeners in a brief discrimination test, on the assumption that this threshold is more likely to match the threshold of burst effectiveness in identification. In any case, the hypothesis is that listeners' sensitivity in phonetic identification is no worse than in overt burst detection; if it is better, we would have evidence that the subconscious processes of phonetic identification maximally exploit auditory sensitivities.

[8] Some subjects gave many "svay" and/or "spay" responses; the former occurred most often at intermediate burst amplitudes and silences, the latter at low burst amplitudes and long silences. For the purpose of group boundary determination, these responses were grouped with "stay" responses.

[9] Inspection of the unpreemphasized waveform suggested that a first, very low-amplitude glottal pulse may have been included in the burst as defined here.

[10] It seems likely that amplification of the burst by just a few dB would still have increased its power as a manner cue. However, the present data suggest that the trading relation with silence duration ends well before a 10-dB gain is reached.

Bailey, P. J., and Summerfield, Q. (1980). "Information in speech: Observations on the perception of [s]-stop clusters," J. Exp. Psychol. (HPP) 6, 536–563.

Bastian, J. (1959). "Silent intervals as closure cues in the perception of stop phonemes," in *Speech Research and Instrumentation* (Q. Prog. Rep. No. 33, Haskins Labs).

Bastian, J. (1962). "Silent intervals as closure cues in the perception of stops," in *Speech Research and Instrumentation* (Final Rep. [No. 9], Haskins Labs).

Bastian, J., Eimas, P. D., and Liberman, A. (1962). "Identification and discrimination of a phonemic contrast induced by intervals of silence," in *Speech Research and Instrumentation* (Final Rep. [No. 9], Haskins Labs).

Best, C. T., Morrongiello, B., and Robson, R. (1981). "Perceptual equivalence of acoustic cues in speech and nonspeech perception," Percept. Psychophys. **29**, 191–211.

Delgutte, B. (1980). "Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers," J. Acoust. Soc. Am. **68**, 843–857.

Delgutte, B. (1982). "Some correlates of phonetic distinctions at the level of the auditory nerve" in *The Representation of Speech in the Peripheral Auditory System*, edited by R. Carlson and B. Granström (Elsevier, Amsterdam), pp. 131–149.

Dorman, M. F., Raphael, L. J., and Isenberg, D. (1980). "Acoustic cues for a fricative–affricate contrast in word-final position," J. Phonet. **8**, 397–405.

Dorman, M. F., Raphael, L. J., and Liberman, A. M. (1979). "Some experiments on the sound of silence in phonetic perception," J. Acoust. Soc. Am. **65**, 1518–1532.

Fant, G. (1973). "Stops in CV-syllables," in *Speech Sounds and Features* (M.I.T., Cambridge, MA), pp. 110–139.

Fitch, H. L., Halwes, T., Erickson, D. M., and Liberman, A. M. (1980). "Perceptual equivalence of two acoustic cues for stop-consonant manner," Percept. Psychophys. **27**, 343–350.

Goldhor, R. (1983). "The representation of speech signals in a model of the peripheral auditory system," J. Acoust. Soc. Am. Suppl. 1 **73**, S4.

Gurlekian, J. A. (1981). "Recognition of the Spanish fricatives /s/ and /f/," J. Acoust. Soc. Am. **70**, 1624–1627.

Liberman, A. M., Isenberg, D., and Rakerd, B. (1981). "Duplex perception of cues for stop consonants: Evidence for a phonetic mode," Percept. Psychophys. **30**, 133–143.

Marcus, S. M. (1978). "Distinguishing 'slit' and 'split'—an invariant timing cue in speech perception," Percept. Psychophys. **23**, 58–60.

Massaro, D. W., and Oden, G. C. (1980). "Evaluation and integration of acoustic features in speech perception," J. Acoust. Soc. Am. **67**, 996–1013.

Miller, J. L., Connine, C. M., Schermer, T. M., and Kluender, K. R. (1983). "A possible auditory basis for internal structure of phonetic categories," J. Acoust. Soc. Am. **73**, 2124–2133.

Nooteboom, S. G. (1981). "Speech rate and segmental perception or the role of words in phoneme identification," in *The Cognitive Representation of Speech*, edited by T. Myers, J. Laver, and J. Anderson (North-Holland, Amsterdam), pp. 143–150.

Oden, G. C., and Massaro, D. W. (1978). "Integration of featural information in speech perception," Psychol. Rev. **85**, 172–191.

Ohde, R. N., and Stevens, K. N. (1983). "Effect of burst amplitude on the perception of stop consonant place of articulation," J. Acoust. Soc. Am. **74**, 706–714.

Pastore, R. E., Szczesiul, R., Rosenblum, L. D., and Schmuckler, M. A. (1982). "When is a [p] a [t], and when is it not," J. Acoust. Soc. Am. Suppl. 1 **72**, S16.

Pastore, R. E., Szczesiul, R., Rosenblum, L. D., and Schmuckler, M. A. (1983). "Does silence play a peripheral role in the perception of stops?" J. Acoust. Soc. Am. Suppl. 1 **73**, S52.

Pols, L. C. W., and Schouten, M. E. H. (1978). "Identification of deleted consonants," J. Acoust. Soc. Am. **64**, 1333–1337.

Repp, B. H. (1982). "Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception," Psychol. Bull. **92**, 81–110.

Repp, B. H. (1983a). "Categorical perception: Issues, methods, findings," in *Speech and Language: Advances in Basic Research and Practice, Vol. 10*, edited by N. J. Lass (Academic, New York).

Repp, B. H. (1983b). "Trading relations among acoustic cues in speech perception are largely a result of phonetic categorization," Speech Commun. **2**, 341–362.

Repp, B. H., and Mann, V. A. (1980). "Perceptual assessment of fricative-stop coarticulation," Haskins Labs. **SR-63/64**, 93–120.

Samuel, A. G. (1981). "The role of bottom-up confirmation in the phonemic restoration illusion," J. Exp. Psychol. (HPP) **7**, 1124–1131.

Stevens, K. N., and Blumstein, S. E. (1978). "Invariant cues for place of articulation in stop consonants," J. Acoust. Soc. Am. **64**, 1358–1368.

Summerfield, Q., Bailey, P. J., Seton, J., and Dorman, M. F. (1981). "Fricative envelope parameters and silent intervals in distinguishing 'slit' and 'split,'" Phonetica **38**, 181–192.

Zue, V. W. (1976). "Acoustic characteristics of stop consonants: A controlled study," Lincoln Lab. Tech. Rep. No. 523 (Lexington, MA).