Journal of Phonetics (1992) **20**, 143–165

Vowel duration and closure duration in voiced and unvoiced stops: there are no contrast effects here

Carol A. Fowler

Dartmouth College, Hanover, NH 03755, U.S.A. and Haskins Laboratories, New Haven Ct 06511, U.S.A.

Received 24th May 1990, and in revised form 13th November 1990

The experiments test directly a claim in the literature [Kluender, Diehl & Wright, *Journal of Phonetics* (1988), **16**, 153–169] that the generally inverse relationship between durations of vowels and of closures for following voiced and voiceless obstruents occurs because language communities exploit an ostensible property of the auditory system identified as "auditory durational contrast". In this account, the long vowel before voiced, as compared to voiceless, consonants makes their already short closures sound even shorter than they are, thereby enhancing the perceptual distinctiveness of voiced and voiceless obstruents. The first experiment of the present series shows that contrast does not affect judgments either of closure durations or of vowel durations in VCV synthetic-speech disyllables. To the contrary, long vowels are associated with increased judgments that a closure is long, and, compatibly, long closures are associated with increased judgments that a vowel is long. A second experiment fails to replicate findings by Kluender *et al.* showing that classification judgments of non-speech square-wave VCV analogs varying in the duration of the first (vowel-like) tone and in the duration of the (consonant-like) gap between tone pairs exhibit durational contrast. Judgments of these signals, like judgments of VCVs themselves, show assimilation, rather than contrast, effects. Whatever the reason may be for vowel durations and closure durations to vary inversely in voiced and voiceless obstruents, it is not because language communities are exploiting durational contrast in the auditory system. For stimuli in these durational ranges and with these acoustic characteristics, there is no durational contrast effect to exploit. Moreover, if the assimilation effects that durational judgments did exhibit do, in fact, reflect distorting properties of the auditory system for signals such as VCVs, language communities are evidently snubbing those properties entirely in settling on durational properties of vowels before voiced and voiceless obstruents.

1. Introduction

In a great many languages, but perhaps not all (see, e.g., Keating, 1985), and perhaps not in all contexts (see Davis & Summers, 1989), vowels are longer before voiced than voiceless consonants, while consonant closures are shorter for voiced

0095-4470/92/010143 + 23 $03.00/0 © 1992 Academic Press Limited

than for voiceless obstruents (e.g., Luce & Charles-Luce, 1985). In a recent paper, Kluender, Diehl & Wright (1988) review and reject previous accounts of the so-called "vowel-length effect" and propose and test an account of their own.

Kluender *et al.* ascribe not only the vowel duration difference, but also its inverse relation with the following closure duration difference, to an exploitation by languages of a durational contrast effect in the auditory system. As they point out, contrast effects are found to be pervasive in perceptual judgments, so that, for example, a hefted weight is judged heavier if a lighter rather than a heavier weight has just been hefted (Guilford & Park, 1931), a tone is judged higher in pitch if it follows a lower than a higher pitched tone (Cathcart & Dawson, 1928–1929), and so on (see Warren, 1985, for a review). A durational contrast effect would manifest itself for example as a decrease in the judged duration of an interval preceded by a longer than a shorter sound. Kluender *et al.* speculate that, in the context of a preceding long vowel, a consonant closure sounds shorter than in the context of a preceding shorter vowel. Languages may exploit the auditory system's ostensible tendency to exhibit durational contrast in order to enhance the somewhat subtle differences between voiced–voiceless obstruent pairs. The long vowel before a voiced consonant makes its already short closure interval sound even shorter than it is; the short vowel before a voiceless consonant makes its already longer closure interval sound even longer.

Kluender *et al.* offer a test of their account. Two groups of listeners participated in their experiment, one group listening to /aba/-/apa/ disyllables, and the other listening to square-wave analogs of the disyllables. The purpose of the square-wave condition was to obtain an assessment of auditory-system judgments of non-speech intervals. If judgments of speech disyllables and non-speech stimuli pattern similarly, the comparison would lend support to a view that judgments of the speech intervals do not arise in a special-to-speech perceptual system (for research carried out using similar logic see, for example, Diehl & Walsh, 1989; Parker, Diehl & Kluender, 1986; Pisoni, Carrell & Gans, 1983). There were two continua in both stimulus sets that differed in the duration of the first vowel (or first tone in the square-wave set). The short and long vowel (or tone) continua had 155 and 245 ms initial vowels (tones) respectively. Members of each continuum differed in stop-closure (gap) duration, which ranged from 20 ms to 100 ms in 10 steps. Short and long continua were blocked in the experiment and counterbalanced over subjects. In each half session, subjects first heard the endpoints of the appropriate continuum and were trained, using feedback, to hit one response button to one continuum endpoint and a second button to the other endpoint. Following an 80-item identification test of the endpoints, subjects took a 100-item test in which each continuum member occurred 10 times. They were to hit one or the other response button to classify each sound they heard.

Findings were that subjects hit the response button corresponding to the short-gap end of the continuum more frequently in the long-first-vowel (or -tone) condition than in the short-first-vowel (-tone) condition. Kluender *et al.* conclude as follows:

> We suggest that the principle of durational contrast provides a natural explanation both of the speech and non-speech results. Specifically, a long initial segment makes a given medial gap seem shorter by contrast and hence, in the case of speech, more like a voiced segment. A short initial segment, on the other hand, makes a

> medial gap seem longer (i.e. in speech, more voiceless). Thus vowel-length differences are a means of enhancing the perceptual distinctiveness of the closure-duration cue for consonant voicing contrasts (p. 161).

Their conclusions are not yet warranted, however. One reason why the conclusions are not warranted concerns the "principle" of durational contrast. In their references to contrast effects, Kluender *et al.* do not cite a single investigation of auditory durational contrast. In my search of the literature, I uncovered a few investigations that do report contrast (e.g., Behar & Bevan, 1961; Walker, Irion & Gordon, 1981), but these used stimuli longer than the vowels and consonants of spoken languages. Moreover, the one discussion I found of effects of the duration of a filled interval (such as the vowel or tone in the stimuli used by Kluender *et al.*) on a following open interval (such as the closure or gap) reported a positive not a negative relation between filled interval duration and judged open interval duration (Woodrow, 1928). However, this study, like the others I found, used stimuli somewhat longer than the vowels and consonants in the experiments by Kluender *et al.* Finally, there are grounds provided by the literature on contrast effects for guessing that there should be no contrastive effect of a vowel on a following closure duration.

Kluender *et al.* cited Warren's (1985) review and theoretical paper on contrast effects (which Warren calls "criterion shifts") to make the point that such effects are extremely pervasive in perceptual judgments. However, they did not mention that Warren's interpretation of these effects does not predict a criterion shift in closure duration judgments due to variation in vowel duration. Warren proposed that criterion shifts are consequences of the way that perceivers use past experience with relevant objects or events of some type to make judgments about present tokens; that is, a criterion shift is a consequence of the way that perceptual learning works. He suggested that perceivers make perceptual judgments by assessing a characteristic of an object or event (for example, the weight of an object) relative to their past experiences with other relevant objects of the same type, and, rationally, they weight very recent experience disproportionately (see Bjork, 1974, and Freyd & Johnson, 1987, for similar ideas about "memory averaging"). In an experiment in which stimuli vary along some dimension, subjects behave as if they are maintaining a running average of the values along the varying dimension, but recent trials contribute disproportionately to the average. In an experiment on weight judgments, a recent encounter with a very heavy weight will temporarily strongly affect the assessment of the typical weight of objects in the experiment (Guilford & Park, 1931); accordingly a lighter weight encountered next is judged, relatively, very light indeed. Analogously, in judgments of vowel duration, a very long vowel judged recently should strongly affect a judge's running assessment of typical vowel duration, and a following shorter vowel will be judged even shorter than if it had not been preceded by the long vowel. This kind of effect should give rise to trial-to-trial contrast effects in judgments of speech stimuli (as it does, for example, in judgments of vowel color—see, e.g. Repp, Healy & Crowder, 1978). However, it should not give rise to effects of vowel duration on following closure duration. Warren reviews evidence suggesting that criterion shifts are restricted to influences of past experience with objects or events of the same type. This is consistent with the idea that these effects are not blind sensory distortions of perceptual information, but rather reflect adaptive uses of past experience in perceptual judgment. In the example Warren uses, Johnson (1944) found contrast effects of past judged weights on current judgments; however, if, during the experiment, subjects happened to lift an object, such as a book, that was not one of the experimental stimuli, that object had no effect on subsequent weight judgments. Analogously, successive vowel durations may contribute to a running assessment of typical vowel durations in some experiments, and consonant durations may contribute to a running assessment of typical consonant durations, but a long vowel should not change the running assessment of consonant duration, at least not via the criterion shift proposed by Warren (1985).

All of this would be moot, of course, had the experiments of Kluender *et al.* convincingly demonstrated a durational contrast effect for VCVs and squarewave stimuli. However, their test of the effect was indirect in failing to ask subjects to judge closures (gap durations) explicitly. While it is true that, within a stimulus set, closure or gap duration was the only acoustic dimension to undergo variation, that in itself does not ensure that subjects made closure- or gap-duration classifications. Subjects were not told how the stimuli differed in a test; rather, they had to determine that for themselves during the feedback trials with endpoint stimuli. Naive listeners may not even be aware that stops include a silent or near-silent closure interval; if they were unaware of the closures in the speech stimuli, they would have responded on some other basis in the speech condition. Indeed, possibly they made /b/–/p/ classifications, with the short-gap signals sounding like /aba/ and the long-gap signals sounding like /apa/. If so, then the speech condition itself simply replicated previous findings of a vowel-length effect on voicing classifications (e.g., Raphael, 1972).[1]

With respect to the speech/non-speech comparison, Kluender *et al.* argued that they demonstrated, in addition, that the basis for the vowel-length/voicing relationship is to be found in the auditory system of the listener. Elsewhere (Fowler, 1990) I have attempted to challenge the logic by which such inferences are drawn based on speech/non-speech comparisons of the sort made by Kluender *et al.* I will make an analogous attempt here in Section 4. However, the first experiment focuses specifically on the vowel-length/closure-duration relation in speech and asks whether there is a durational contrast effect there.

2. Experiment 1

In this experiment, three groups of listeners made classification judgments of a set of VCV disyllables. Two of the three groups made discrimination judgments as well.

[1] Two reviewers of this paper understood this argument as applying considerably more broadly than I intend it to. It is not meant to advocate collection of explicit judgments of acoustic variables in speech signals by researchers—*au contraire* (see Fowler, 1990). Following is a clarification. Generally, to test the role of some acoustic variable in phonetic perception, it is sufficient and appropriate to cause the variable to vary and to ask subjects to make phonetic classification judgments. Generally, the interpretation of results of experiments like that in relation to the hypothesis under test is clear. That is not the case specifically for the hypothesis proposed by Kluender *et al.* in relation to their test and findings. It was known before their research that long vowels before an obstruent promote judgments that the consonant is voiced. Kluender *et al.* proposed that the effect is mediated by auditory durational contrast. To test this hypothesis, clearly, it is necessary to show that there *is* a contrast effect of vowel duration on judgments of consonant closure. Kluender *et al.* did not tell their subjects how the endpoint stimuli differed and did not tell them the basis on which they were to make their classification responses. Accordingly, if subjects heard one endpoint consonant as /b/ and the other as /p/—a not unlikely possibility—they were free to respond as if the experiment were a /b/–/p/ classification test, leaving the contrast hypothesis untested.

The VCV disyllables differed in closure duration, vowel duration, and post-release aspiration/voicing. One group was instructed to classify the closure durations as long or short (and to discriminate pairs based on closure duration); a second group was instructed to classify vowels as long or short (and to discriminate pairs based on vowel duration); the third group classified the consonants as "b" or "p".

Judgments of closure duration provide a direct test of the view that a durational contrast effect occurs in VCV disyllables. Judgments of vowel duration were included as a matter of interest to determine whether a contrast effect also might work in the opposite direction. Results from the three groups will perhaps permit inferences as to the bases for the button-press classifications made by subjects in the experiment of Kluender *et al.*

2.1. *Methods*

2.1.1. *Subjects*

Subjects were 36 students at Dartmouth College who participated for course credit. All were native speakers of English who reported normal hearing. Twelve students participated in each of the three groups in the experiment.

2.1.2. *Stimulus materials*

The VCVs were not designed originally to address the claims of Kluender *et al.* Accordingly, they were not designed after the stimuli used in that research. They were, however, designed to be consistent with measured vowel and closure durations characteristic of vowel-/b/ or -/p/ sequences reported in the literature (e.g., Luce & Charles-Luce, 1985). Moreover, the closure-duration range falls almost entirely within that tested by Kluender *et al.*, and the vowel duration range falls in between the short and long vowel durations they used.

Eighty 'VCəs were used in the identification tests of Experiment 1. The disyllables reflected the orthogonal combination of four initial-vowel durations, four closure intervals, and five combinations of post-release aspiration/voicing. The 'VCəs were synthesized on the serial resonance speech synthesizer at Haskins Laboratories. Vowel durations ranged from 160 to 190 ms in 10 ms steps; closures ranged from 90 to 120 ms in 10 ms steps. At the /b/ end of the aspiration/voicing continuum, the AH parameter (noise excitation for formants) was set to zero during formant transitions at consonant release, while the AV parameter (glottal pulsing excitation for formants) was set to 20 dB, rising to a value of 30 dB by the end of the transitions. At the opposite end of the continuum, the AH parameter was set to 20 dB, while the AV parameter was set to zero during the release transitions. Three intermediate stimuli were created by varying the AH and AV parameters between their endpoint values in equal steps.

Steady-state formants for the first vowel (bandwidths 50, 100 and 200 Hz) were 730, 1090 and 2440 Hz. Transitions into the closure were 40 ms in duration. F_1 fell to 400 Hz, F_2 to 760 Hz and F_3 to 2110 Hz. Rising transitions out of the closure were also 40 ms in duration; starting frequencies were the ending values of the falling transitions, and steady states for the final schwa vowel were 640, 1190 and 2390 Hz. The second vowel had a duration of 100 ms. Amplitude rose over the first 40 ms of the signal and remained steady until the closure, where it was set to zero; during release, the amplitude rose from 20 to 30 dB; it fell to zero over the final 40 ms of the vowel. Fundamental frequency was steady at 120 Hz during the steady-state of the first vowel; it fell during closing transitions to 104 Hz. During release and much of the steady-state of the second vowel, fundamental frequency was set at 120 Hz; it fell over the last 40 ms of the vowel to 100 Hz.

A single identification test was designed for presentation to all three experimental groups. The test presented each of the 80 disyllables ((four vowel durations) × (four closures) × (five values of the voicing/aspiration variable)) three times each in random order with the constraint that each item occur once in each third of the test. There were 3.5 s between trials and 7 s between blocks of 24 trials.

It was impractical to include all 80 disyllables in the discrimination test. For that test, the aspiration/voicing continuum values were fixed at the /b/ end of the continuum, leaving 16 disyllables in which vowel and closure durations were varied orthogonally. In the discrimination test, all possible pairings of the 16 disyllables were included, yielding a 256-item test. There were 750 ms between disyllables of a pair, 3.5 s between trials and 7 s between blocks of 32 trials.

2.1.3. *Procedure*

Subjects were run in groups of three or fewer in a quiet room; they listened over headphones to stimuli presented at a comfortable listening level. They were assigned to experimental conditions on a rotating basis, depending on order of arrival. All subjects took the identification test first. Subjects in the closure- and vowel-duration groups took a following discrimination test.

Subjects in the closure duration group were shown a pair of spectrographic displays of VCV disyllables differing in closure duration. One disyllable had the largest closure duration and one the smallest of the values used in the experiment. The disyllables had identical values of the vowel duration variable (the second to shortest value) and of the aspiration/voicing variable (the middle value). I explained to subjects the relation between the visual display and the disyllables they would be hearing; in particular, the formant bands were identified with the vowel and the empty gap with the consonant. (I explained that the interval was silent for /b/ and /p/ because the speaker's lips were closed). The reason for showing the displays was to convince subjects that there was a silent gap in the syllables that might be judged.

Listeners were told that they would be classifying the duration of the silent interval between the two syllables of each disyllable as long or short. They listened to a demonstration sequence consisting of the stimuli displayed in the spectrograms played in alternation five times each. Listeners were told that the first disyllable of each pair had the long closure and the second the short closure. They were instructed to listen to the alternating pairs attempting to attend to the relevant difference so that they could classify disyllables by closure duration in the test proper. They were allowed to hear the sequence additional times if they wished, and subjects in this group frequently did request a replaying. In the identification test, listeners were told that they would hear one disyllable on each of 240 trials and they were to classify the closure as long or short by writing L or S on the answer sheet, guessing if necessary.

After the identification test, listeners participated in the 256-item discrimination test. They were instructed to write "1" on their answer sheet if the first disyllable of each pair had the longer closure and "2" if the second disyllable had the longer closure. They were instructed to guess if necessary, not leaving any answers blank.

Subjects in the vowel duration group were shown a pair of spectrographic displays of VCV disyllables differing in vowel duration. One disyllable had the longest vowel duration and one the shortest of the values used in the experiment. The disyllables had identical values of the closure duration variable (the second to shortest value) and of the aspiration/voicing variable (the middle value). The relation between the display and the disyllables they would be hearing was explained to this group as it was explained to the closure duration group. Listeners were told that they would be classifying the initial-vowel duration of the VCV disyllables as long or short. They listened to a demonstration sequence consisting of the stimuli displayed in the spectrograms played in alternation five times each. Listeners were told that the first disyllable of each pair had the long vowel and the second the short vowel. They were instructed to listen to the alternating pairs attempting to attend to the relevant difference so that they could classify disyllables by vowel duration in the test proper. They were allowed to hear the sequence additional times if they wished; however, in this group, such requests were rare. In the identification test, listeners were told that they would hear one disyllable on each of 240 trials and they were to classify the initial vowel as long or short by writing L or S on the answer sheet, guessing if necessary.

Listeners then went on to the discrimination test, writing "1" if the first disyllable of each pair had a longer initial vowel than the second disyllable and "2" if the second disyllable's initial vowel was longer. Again, listeners were instructed to guess if necessary, not leaving any blank responses.

Listeners in the "b"/"p" group participated only in an identification test. Their task was to classify the medial consonant as "b" or "p", guessing if necessary. Subjects in this group first heard alternating b/p disyllables having endpoint values on the voicing/aspiration continuum and identical values (next to shortest) on the other continua. Listeners heard the demonstration pairs five times each before proceeding to the identification test. They were told that the first disyllable of each pair was /abə/ and the second was /apə/. They were to listen to the pairs attempting to learn to label the consonants as "b" and "p", respectively. They were permitted to hear the demonstration sequence again if they wished; however, requests to do so were uncommon. On the identification test, using the same test sequence as the listeners in the other two groups heard, subjects' task was to identify the consonant of each disyllable as "b" or "p", guessing if necessary.

2.2. *Results and discussion*

The aspiration/voicing continuum is not of particular relevance here; accordingly for analyses following the present one, identification data are collapsed over that continuum. However, pre-collapsed data are presented in Fig. 1 from the condition in which subjects identified the consonants as "b" or "p" to illustrate that stimulus variables contributed in the expected way to voicing identification. Figure 1 presents a subset of the data from that condition. In Fig. 1 (top) percent "b" responses are compared across the aspiration continuum for stimuli having the longest of the four vowel durations (averaged over closure duration differences) and for stimuli having the shortest of the four vowel durations. In Fig. 1 (bottom), analogous data compare

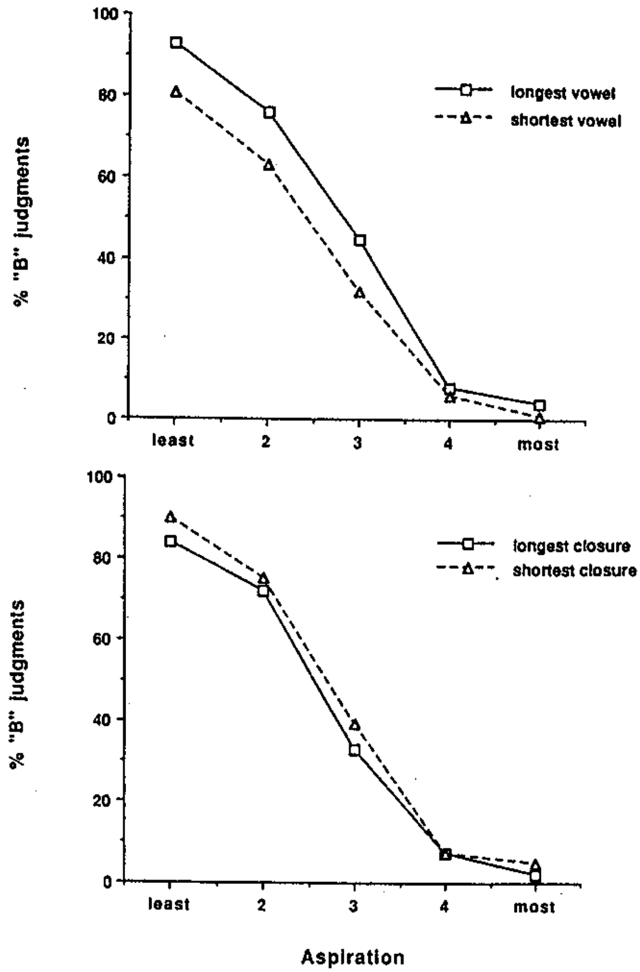

Figure 1. Data from the "b"/"p" identification condition of Experiment 2. (Top) Percent "b" responses across the aspiration continuum for stimuli having the longest and shortest vowel durations (averaged over variation in closure duration). (Bottom) Analogous plot for stimuli having the longest and shortest closure durations.

the effect of closure duration across the aspiration continuum. As expected consonants with little or no aspiration, with long vowels and with short closures are reported as "b"; those with opposite values are reported as "p". In an analysis of variance on all of the data from the "b"–"p" identification condition with factors Aspiration, Vowel duration and Closure duration, all main effects were highly significant (Aspiration: $F(4, 44) = 134.45$, $p < 0.001$; Vowel duration: $F(3, 33) = 17.16$, $p < 0.001$; Closure duration: $F(3, 33) = 6.63$, $p = 0.0013$). "B" judgments decrease significantly with decreases in aspiration, with decreases in vowel duration and with increases in closure duration. Just one interaction reached significance, that between Aspiration and Vowel duration ($F(12, 132) = 4.07$, $p < 0.001$). The interaction was significant because the effect of vowel duration on "b" judgments was diminished for aspiration continuum members 4 and 5 where "b" judgments approached the floor even with long vowels.

2.2.1. *Closure duration judgments*

I next consider the identification judgments of the group of major interest in the experiment—listeners who classified closure intervals as long or short. If there is a durational contrast effect of the sort that Kluender *et al.* propose, then closures should be judged shorter when preceded by longer vowels than by shorter vowels.

Figure 2 shows the two significant effects obtained in an analysis of variance on average "long" classifications of closure durations. The figure plots the percentage of "long" closure judgments for stimuli differing in closure duration, but averaged over vowel duration (dark bars) and for stimuli differing in vowel duration averaged over closure durations (striped bars). The figure shows that "long" closure judgments increased both as closure duration increased *and* as vowel duration increased. In an analysis of variance with repeated measures factors closure duration and vowel duration, both main effects are significant while the interaction does not approach significance (closure: $F(3, 33) = 11.63$, $p < 0.001$; vowel: $F(3, 33) = 36.05$, $p < 0.001$). The vowel factor, in fact, accounts for considerably more of the variance in judged closure durations than does the closure variable itself (31.3% *vs.* 16.6%). Tests for linear decreases in "long" judgments are also highly significant on both variables (closure: $F(1, 33) = 34.39$, $p < 0.001$; vowel: $F(1, 33) = 194.44$, $p < 0.001$). The effect of vowel duration on judged closure duration is the reverse of a durational contrast effect.

The results of the discrimination test reveal a similar effect. Table I presents proportions of trials on which the first presentation of the disyllable was judged to have the longer closure in the different trial types of the experiment. Vowel duration

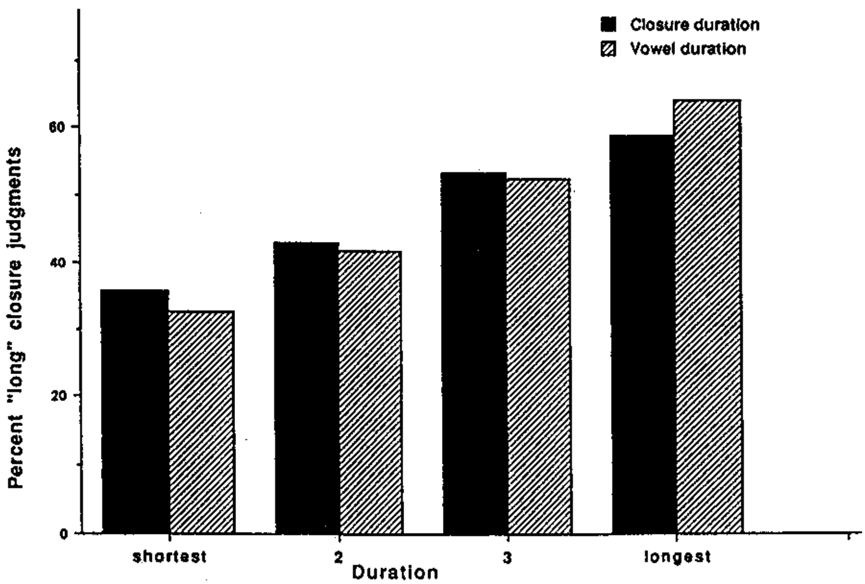

Figure 2. Classifications of closure durations as long or short in Experiment 1 as a function of variations either in closure duration itself or in vowel duration.

TABLE I. Percent judgments that the first disyllable had the longer closure in the discrimination test of Experiment 1. Across the rows is the vowel duration difference (V1 – V2) in ms; down each column is the closure duration difference (Cl1 – Cl2) in ms

	V1 – V2 (ms)						
Cl1 – Cl2 (ms)	–30	–20	–10	0	10	20	30
–30	08	04	28	29	42	54	**58**
–20	08	19	26	43	63	**58**	83
–10	11	17	39	48	**61**	72	69
0	21	33	40	**51**	67	76	83
10	19	35	**51**	60	68	85	69
20	29	**40**	61	65	72	81	83
30	**42**	42	53	65	86	92	100

differences (duration of the first vowel in the disyllable pair [V1] minus duration of the second disyllable's first vowel [V2]) are displayed horizontally in the table while closure duration differences (Cl1–Cl2) are displayed vertically. Generally, percentages of judgments that the first disyllable has the longer closure increase as the closure duration difference increasingly favors the first disyllable (that is, top to bottom in the table). The difference in the percentage of "first longer" judgments between extreme pairs (Cl1–Cl2 = 30 ms *vs.* Cl1–Cl2 = –30 ms) averages 37% overall in the table and 36% on disyllabic pairs having the same vowel duration (the middle column of Table I).

Looking left to right in the table reveals the effect of vowel duration on closure-duration discrimination judgments. There is a large effect, analogous to that in the identification test. If the first disyllable has a longer vowel than the second, subjects are biased toward judgments that the first *closure* is longer than the second. That the effect of the vowel duration difference on closure duration discriminations is, in fact, larger than the effect of the closure duration difference itself can be seen by looking at the diagonal printed in boldface in the table. Along this diagonal, the vowel duration difference and the closure duration difference between the disyllables of a pair cancel out, leaving disyllables in each pair that are identical in duration. If the effects of vowel and closure duration were equal, judgments along this diagonal would all be about 50%. However, there are, in fact, more judgments that the first closure is longer than the second when the closure difference is 30 ms in the wrong direction, but the vowel duration difference offsets it (top right in the table), and fewer such judgments when the closure duration difference is 30 ms in the right direction, again offset by an opposite direction of difference in vowel durations. Moreover, when there is no closure duration difference at all (the middle row of the table), judgments that the first closure is longer than the second increase from 21% to 83% as the vowel duration difference increasingly favors the first vowel. This compares to a smaller shift from 29% to 65% in the middle column of the table where vowel durations are the same in disyllables of a pair.[2] In an analysis

[2] A reviewer of this manuscript raised the concern that subjects could not do the task as instructed and, instead, made their judgments based on overall disyllable duration. However, the findings just described that subjects respond differently to a given duration difference between to-be-discriminated disyllables depending on whether the difference is in the vowel or in the closure disconfirms this idea. (Again, compare the row at 0 ms difference with the column at 0 ms difference.)

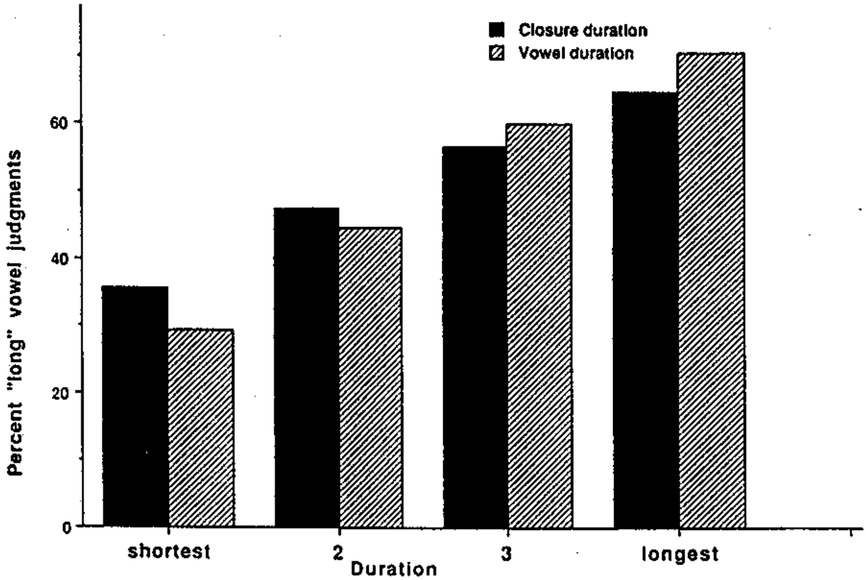

Figure 3. Classifications of vowel durations as long or short in Experiment 1 as a function of variations either in vowel duration itself or in closure duration.

of variance on judgments that the "first disyllable's closure is longer" with factors closure duration difference (Cl1–Cl2) and vowel duration difference, both main effects are significant ($F(6, 66) = 15.87$, $p < 0.001$; $F(6, 66) = 55.79$, $p < 0.001$, respectively) with the vowel duration difference again accounting for more of the variance (37.2%) than the closure difference (9.9%). The interaction was non-significant.

Clearly in these stimuli there is no contrast effect of vowel durations on judgments of following closures. It is unlikely that subjects in the speech condition of the experiment of Kluender *et al.* were judging gap duration.

2.2.2. *Vowel duration judgments*

Comparable analyses were done on the vowel-duration classification and discrimination judgments. Relevant findings are presented in Fig. 3 and Table II.

TABLE II. Percent judgments that the first disyllable had the longer vowel in the discrimination test of Experiment 1. Across the rows is the vowel duration difference (V1 – V2) in ms; down each column is the closure duration difference (Cl1 – Cl2) in ms

	V1 – V2 (ms)						
Cl1 – Cl2 (ms)	–30	–20	–10	0	10	20	30
–30	08	21	33	35	47	63	**83**
–20	21	29	26	42	68	**67**	71
–10	14	29	41	50	**57**	72	81
0	15	35	39	**60**	65	69	92
10	47	29	**46**	56	77	89	81
20	50	**31**	54	67	85	94	88
30	**25**	46	61	65	94	83	92

Figure 3 shows the effects of closure duration and vowel duration on judgments of vowel duration. As in the previous analysis of classification judgments, both main effects are significant, while the interaction is not (closure duration: $F(3, 33) = 24.49$, $p < 0.001$; vowel duration: $F(3, 33) = 47.31$, $p < 0.001$) with the vowel effect stronger (19.6% and 41.3% of the variance accounted for respectively). Tests for linear decreases in judgments that the vowel was long were significant for both the closure-duration variable ($F(1, 33) = 72.5$, $p < 0.001$) and the vowel-duration variable ($F(1, 33) = 141.30$, $p < 0.001$). In short, vowels were more likely to be judged long if they were long or if they were followed by a long closure than if they were short or were followed by a short closure. The positive effect of long closures on judgments of the vowel durations as long was clearly audible in this test.

Once again the discrimination test bears out the findings on the identification test. Both the relative closure durations and the relative vowel durations of each disyllable pair affect the percentage of judgments that the first disyllable of a pair has the longer first vowel. In an analysis the effects of relative closure duration ($F(6, 66) = 11.81$, $p < 0.001$; 6.8% of variance accounted for) and of relative vowel duration ($F(6, 66) = 56.17$, $p < 0.001$; 37.7%) were significant; the interaction was non-significant.

Analogously to the closure duration findings, there is an influence on judged vowel duration of the duration of a following closure, but it is a positive effect, not a contrast effect. Accordingly, it is unlikely that listeners in the speech condition of the experiment of Kluender *et al.* were judging relative vowel durations even though the vowels would be perceptibly different in duration due to differences in duration of the following closures.

2.2.3. "*b*"/"*p*" *judgments*

The purpose of obtaining "b"/"p" judgments was largely to ascertain that, in the disyllables of Experiment 1, long vowels and short closures were associated with increased "b" judgments. As reported above, an analysis of variance with factors aspiration, closure duration, and vowel duration, all main effects were significant, while just one interaction, that between aspiration and vowel duration, was significant. Figure 4 shows the effect of vowel duration and closure duration on "b" judgments. Each variable shows the predicted effect, each with one deviation from monotonicity across variation in duration. Despite these individual departures from monotonicity, tests for linear increases in "b" judgments as closure duration decreased and as vowel increased yielded highly significant results ($F(1, 33) = 10.03$, $p = 0.0034$; $F(1, 33) = 37.44$, $p < 0.001$, respectively).

This, then, is the only test in which the dependent measure (b/p judgments) showed the patterning of the findings of Kluender *et al.* As closure duration increased, "b" responses fell; as vowel duration increased, "b" responses increased (see also Fig. 1). Perhaps, rather than responding to closure duration, the subjects of Kluender *et al.* made b/p classifications in the speech test.

This leads of course to the question of what perceived variable governed button pressing to the non-speech stimuli in the research of Kluender *et al.* where subjects were unlikely to be making b/p classifications, but where their classifications of short- and long-gap stimuli patterned as the speech stimuli did. That issue is addressed in Experiment 2.

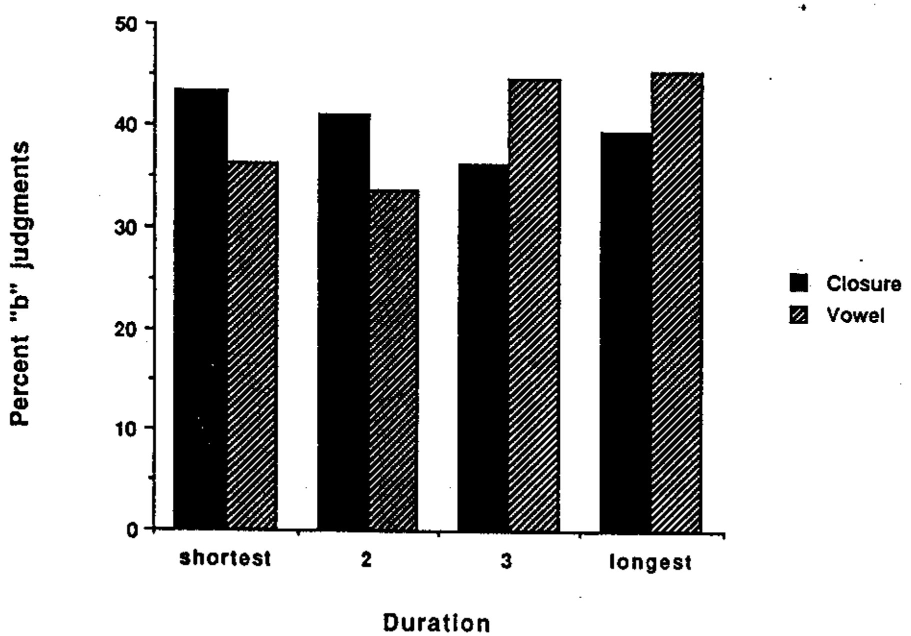

Figure 4. Identifications of consonants as "b" in Experiment 1 as a function of variations either in vowel duration or in closure duration.

3. Experiment 2

The most obvious basis on which subjects in the non-speech condition of Kluender *et al.* might make their button press responses is, of course, the acoustic difference in gap duration. Within a test, that is the only acoustic variable to undergo change. However, if gap duration is the basis for the button press classifications, then listeners apparently do not respond in the same way to variation in the duration of a preceding vowel and tone in their closure- and gap-duration classifications. Whereas Experiment 1 demonstrated a positive relation between vowel duration and judged closure duration, subjects in the non-speech condition of Kluender *et al.* demonstrated a negative relation between initial tone duration and judged gap duration.

In the present experiment, I used square-wave stimuli designed after those of Kluender *et al.* and attempted to determine the conditions under which such stimuli give rise to a durational contrast effect.

3.1. *Methods*

3.1.1. *Subjects*

Subjects were 56 students at Dartmouth College, who received course credits for their participation. They reported normal hearing. Subjects were run until there were 18 students in each condition of the experiment. Data from two subjects were eliminated from one of the experimental conditions for reasons given below.

3.1.2. *Stimulus materials*

Stimuli were designed after the description provided by Kluender *et al.* They were square-wave analogs of the stimuli used by Kluender *et al.* in their speech condition. Accordingly, two sets of 10 signals were created, each signal consisting of two tones separated by a silent gap. The sets were identical except that the first tone in one set was 155 ms long while that in the other was 245 ms. Amplitudes were invariant over the tones except for 10 ms rise and decay intervals. The fundamental frequency was set at 256 Hz except for a fall to 175 Hz during the last 40 ms of the first tone and a symmetrical rise from 175 to 256 Hz over the first 40 ms of the second tone. Within a set of 10 stimuli, the gap duration varied from 20 ms to 110 ms in 10 ms steps. Figure 5 displays a sample nonsense signal. (The figure does not resemble the schematized spectrograms in Figure 1 of Kluender *et al.* in certain respects. However, the spectrograms depicted in Kluender *et al.* also disagree with the description of the signals in the text. Keith Kluender (personal communication, 17 January 1990) has told me that the description in the text is accurate.)

The stimuli were digitized (filtered at 10 kHz and sampled at 20 kHz) and stored in a computer for presentation to listeners. Three test orders were devised for each stimulus set. A demonstration sequence presented the two continuum endpoints in alternation five times. Next, an 80-item test presented each endpoint 40 times in random order. Finally, a 100 item test presented all 10 continuum members ten times each with the constraint that each member occur once in each successive tenth of the test. There were 3.5 s between trials in the 80-item and 100-item tests.

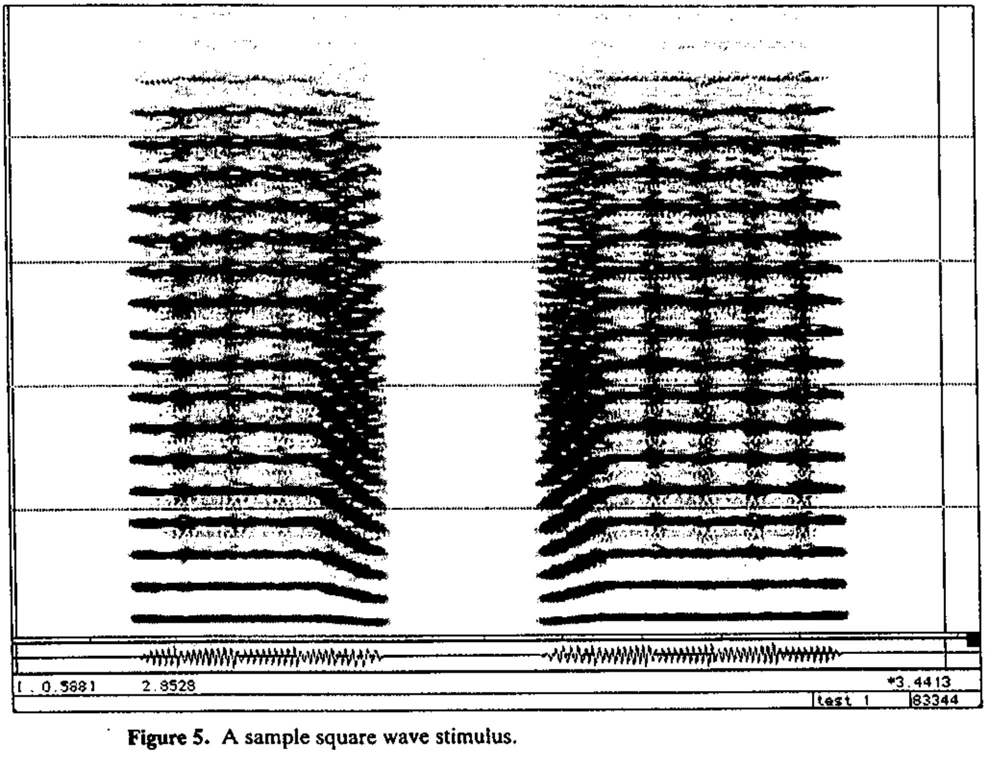

Figure 5. A sample square wave stimulus.

3.1.3. *Procedure*

Subjects listened over headphones in a quiet room. Three groups of subjects were run. Each group heard both sets of test orders. Half of the subjects in each group took the tests involving the short-tone stimuli followed by the long-tone tests; the remaining subjects took the tests in the reverse order. Each test sequence took about 20 min; subjects took a 5 min break between the half-sessions.

The A/B group provided a pencil-and-paper approximation to the button press condition of Kluender *et al.* That is, subjects in this group were not told how the endpoint stimuli differed. They were told that each nonsense sound they would hear consisted of two tones separated by a gap. Initially, they would be hearing just two such sounds that differed audibly in a way that they would have to figure out for themselves. They were to call the first sound of the alternating pair in the demonstration sequence "A" (in fact, the short-gap endpoint) and the second sound "B". Subjects were invited to listen to the demonstration sequence as many times as they wished until they were comfortable that they could label the endpoints accurately. They went on to take the 80-item test on the endpoints. They were told to label each sound either A or B, guessing if necessary. (This procedure contrasts with that of Kluender *et al.* not only in that Kluender *et al.* obtained button-press responses, rather than having subjects write their responses, but, more importantly, they gave their subjects feedback after each trial. Our findings below suggest that neither procedural difference was important.) On the 100-item test, I told them that now they would be hearing not only the sounds they had been labeling but some new sounds as well. Their task was to write A if the sound they heard was more like A than like B and to write B if the sound was more like B; they were instructed to guess if necessary. After a 5 min break, subjects went on to take the same series of tests on the other stimulus set.

The L/S group provided an explicit test of the durational contrast hypothesis. Subjects in this group were told that the nonsense sounds in the demonstration sequence consisted of two tones separated by a silent gap and that the difference between the sounds was in the duration of the gap between the tones of a pair. They were told that the first nonsense sound in the alternating sequence had a considerably shorter gap than the second. Their task was to learn to label the first nonsense sound "S" and the second one "L". Subjects were invited to listen to the demonstration sequence as many times as they wished until they were comfortable that they could label the endpoints accurately. They went on to take 80-item test on the endpoints. They were told to label each sound either S or L, guessing if necessary. On the 100-item test, I told them that now they would be hearing not only the sounds they had been labeling but some new sounds as well. Their task was to write S if the sound they heard was more like the short-gap sound than like the long-gap sound and to write L if the sound was more like the long-gap sound; they were instructed to guess if necessary. After a 5 min break, subjects went on to take the same series of tests on the other stimulus set.

The B/P group provides another kind of control. Kluender *et al.* report that the "square-wave stimuli, which replicate neither the harmonic nor the formant structure of speech, were judged by all three experimenters to bear little perceptual resemblance to the corresponding /aba/–/apa/ stimuli" (p. 158). Although I concur with this assessment, I ran a group anyway in which the task was to label the nonsense sounds as /aba/ (by writing "B") or as /apa/ ("P"). The reason for running such a group was that the corresponding group in Experiment 1 provided the only data conforming in pattern to both the speech and non-speech data of Kluender *et al.* I presumed that the A/B group above would replicate the non-speech findings of Kluender *et al.*, but that the L/S group might not (in view of the finding in Experiment 1 that listeners judging closure duration did not). There would, then, remain the question as to what perceived variable provided the basis for the A/B (and button-press) classifications. One possibility, in view of certain abstract similarities between the acoustic speech and non-speech signals (in tone/vowel-duration variation, in the suggestion that constrictions are achieved, and in gap/closure duration variation) is that the non-speech signals might suggest a sound-producing source similar to the vocal tract in only those abstract ways. That is, the perceived producer of the nonsense signals might have a sound source that produces square-waves rather than glottal-source waves and it might have the capability of producing constrictions as the vocal tract can. If those similarities were striking enough, would the listener behave as if the same negative relation should hold between tone duration and gap duration as holds between vowel duration and closure duration in speech? The only response I could think of in order to address the question was a B/P classification.

Subjects in this group were told that they would initially be listening to some nonsense sounds each consisting of a pair of tones separated by a gap. However, these nonsense sounds had, in some sense, been modeled after the speech disyllables /apa/ and /aba/. Subjects' first task was to listen to the pair of sounds in alternation and to try to decide which sounded more like /aba/ and which more like /apa/. Shortly they would be taking tests in which they would label the /aba/-like sound with a "B" and the /apa/-like sound with a "P". Subjects were invited to listen to the demonstration sequence as many times as they wished until they had decided which pair member they would label "B" and which "P". They went on to take the 80-item test on the endpoints. They were told to label each sound either B or P, guessing if necessary. On the 100-item test, I told them that now they would be hearing not only the sounds they had been labeling but some new sounds as well. Their task was to write B if the sound they heard was more like /aba/ than like /apa/ and to write P if the sound was more like /apa/; they were instructed to guess if necessary. After a 5 min break, subjects went on to take the same series of tests on the other stimulus set.

3.2. *Results*

3.2.1. *A/B and L/S conditions*

On the 80-item tests of the endpoints, data from two subjects were eliminated from the A/B condition, one because the subject had confused the response labels in the first 80-item test and one because the subjects' responses were random. Performance for the remaining 18 subjects averaged 90.5% correct (range 71–100%). In the 100 item test, performance on the endpoints averaged 88%, and eight subjects would not have met the 90% performance level required of their subjects by Kluender *et al.* (I did not apply their criterion, because it caused them to eliminate a high proportion of their subjects, half of their speech subjects and one-third of their non-speech subjects.) However, the responses of these subjects averaged 78% and were all well above chance. In the L/S condition, performance averaged 94.0% across the two 80-item tests (range 80–100%). No subjects' data were eliminated. On the 100 item tests, performance on the endpoints averaged 93.3%, and data from three subjects failed to meet the 90% criterion of Kluender *et al.*[3]

Figure 6 presents the data from the 100-item tests in the A/B condition (top panel) and the L/S condition (bottom panel). Performance shows nearly the same pattern in the two panels. "A" and "short" gap classifications decrease as the gap increases in duration and there is generally a separation between the long- and short-first-tone conditions. The separation is generally opposite in direction to that predicted on the basis of auditory durational contrast, and hence it is also generally opposite to the direction of effect reported by Kluender *et al.* That is, a short first tone is associated with more short-gap classifications than a long first tone. This direction of difference is present numerically for every continuum member in the L/S group and for six of the 10 continuum members in the A/B group. Only three continuum members in the A/B group and none in the L/S group show numerical differences in the direction reported by Kluender *et al.* and predicted from a durational contrast hypothesis.

An analysis of variance was performed with between-subjects factor Group (A/B, L/S) and within-subjects factors Gap duration and First-tone duration. The main effects of Gap duration and First-tone duration were significant ($F(9, 306) = 167.39$, $p < 0.0001$; $F(1, 34) = 7.93$, $p = 0.008$). The effect of Gap duration reflects the decrease in short-gap classifications as gap duration increases; the effect of duration reflects a larger number of short-gap classifications in the short-first-tone condition than in the long-first tone condition. The interaction of Gap duration and First tone duration was also significant ($F(9, 306) = 5.94$, $p < 0.001$). The interaction is significant, because the First-tone main effect holds for eight of the ten gap durations. It is slightly reversed on 90 and 100 ms gap durations, owing to the reversal of the A/B group on those stimuli.

The main effect of Group was not significant, and the only significant interaction in which Group participated was the two-way interaction with Gap duration ($F(9, 306) = 2.68$, $p = 0.0052$). The interaction is significant because the response pattern for the L/S group is somewhat more steeply sloping than that for the A/B group as Fig. 7 shows. Perhaps as a result of being told explicitly what to listen for, subjects made more short-gap classifications than the A/B group on the five shortest gap durations and fewer short-gap classifications than the A/B group on the five longest gap durations.

Finally, to assess whether the inclusions of subjects who did not meet the performance criterion of Kluender *et al.* (90% correct on the training test on the endpoints affected the outcome), I correlated performance on the 80-item practice test with the magnitude of the assimilation effect shown by each subject. Correlations performed separately for subjects in the A/B and L/S groups did not approach significance.

[3] Kluender *et al.* ran 16 subjects in their speech condition and 12 in their non-speech condition. They eliminated half of their speech subjects and one third of their non-speech subjects, because performance on the endpoints on the 100-item training test fell below 90%. Thus, of 28 subjects, 43% were eliminated from the experiment. Even though subjects in the present experiment did not receive feedback during training, just 34% failed to meet the 90% criterion of Kluender *et al.* Accordingly, I conclude that the lack of feedback during practice trials did not noticeably affect subjects' ability to learn the classification.

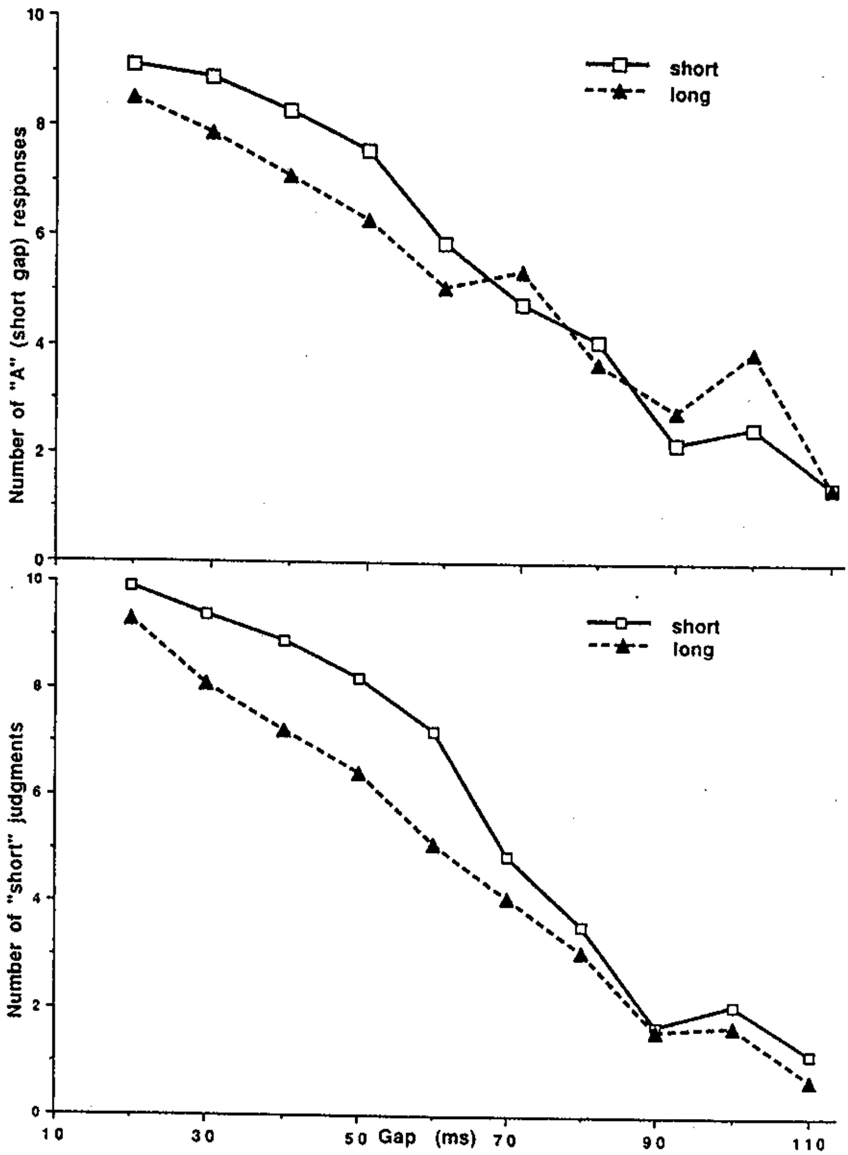

Figure 6. Percentage of short-gap classifications in the A/B and L/S tasks of Experiment 2 as a function of first tone duration.

3.2.2. *B/P Group*

On the 80-item tests, 10 of the 18 subjects consistently identified the short-gap endpoint as "B" and the long-gap endpoint as "P" (87.4% on average). Seven subjects made the opposite assignment consistently (83.7%). The remaining subject made the one classification on one 80-item test and the other on the second test. A t-test testing the proportion of "B" assignments to the short-gap endpoint against 50% (averaging across the two 80-item tests) was not significant ($t(17) = 1.06$, $p = 0.15$).

On the 100-item test, listeners responded true to their endpoint classifications.

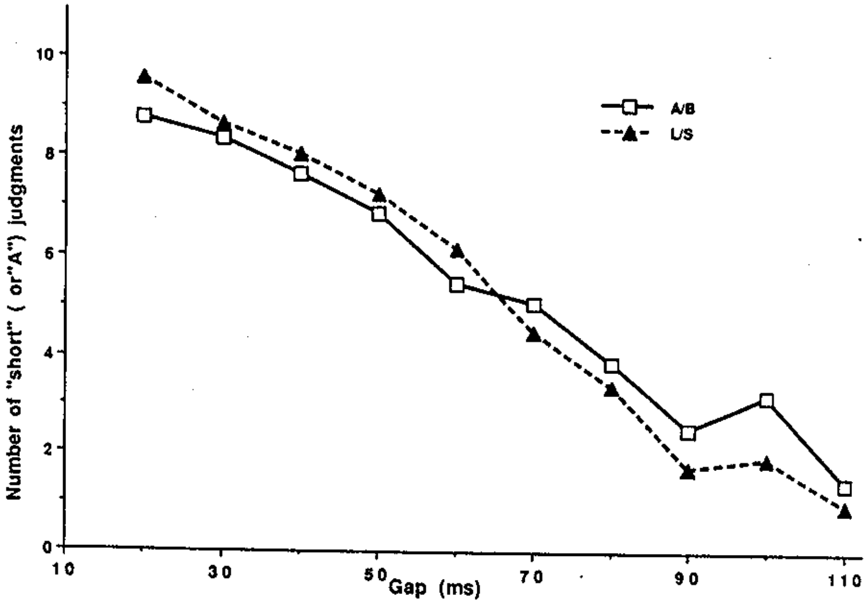

Figure 7. Comparisons of performance levels of A/B and L/S subjects in Experiment 2.

That is, subjects who labeled the short-gap endpoint "B" gave decreasing "B" classifications as gap duration increased; subjects who labeled that endpoint as "P" showed increasing "B" classifications as gap duration increased. Accordingly, the data were analysed separately for the 10 subjects consistently using the one mapping rule and for the seven subjects consistently using the other. Both groups showed a significant effect of Gap duration ($F(9, 81) = 15.95$, $p < 0.0001$; $F(9, 63) = 4.68$, $p = 0.0001$), but no significant effect of first-tone duration and no interaction.

3.3. *Discussion*

Perhaps the most salient aspect of the outcome of Experiment 2 is the failure to replicate the findings of the square wave condition of Kluender *et al.* Those investigators reported a contrastive relation between first tone duration so that gap durations preceded by a long-duration tone were more likely to be classified with the short-gap endpoint than were the same gap durations preceded by a short tone. The findings of Experiment 2 were just the reverse even though the stimuli were modeled directly after those described by Kluender *et al.* For several reasons, I am confident that findings of no contrast are accurate. First, they were obtained under two distinct instructional sets, one inexplicit, following the procedures of Kluender *et al.* and one explicitly instructing subjects to listen for the gap duration. Second, the results replicate findings of Experiment 1 in which listeners made closure-duration (and vowel-duration) judgments of VCV syllables. The VCVs were not speech analogs of the square waves of Experiment 2, and the findings were obtained using different experimental tests than were used to obtain the data in Experiment 2. Accordingly, the finding of assimilative rather than contrastive effects have generality across a variety of stimuli and procedures. Third, the results are compatible with findings reported by Woodrow (1928) on judgments of empty intervals preceded and followed by varying duration filled intervals. Finally, Keith Kluender (personal communications, 15 and 16 May 1990) has reported to me that he has run seven new subjects using his own square-wave stimuli and his procedure and now finds no effect of initial tone duration on gap-duration classifications; three of his subjects show numerical contrast, three show larger numerical assimilation effects, and one shows no difference at all. While seven subjects is a small number, it is just one fewer than the number of subjects whose data were retained in the non-speech condition of Kluender *et al.* These findings are similar to those of the A/B condition of the present Experiment 2. In that condition, 12 of 18 subjects showed numerical effects favoring assimilation, and the assimilation effect for this group alone is not significant if their data are analyzed separately from those of the L/S group. (However, the significant findings for the L/S group, told explicitly what to listen for, suggests that the non-significance for the A/B subjects [and perhaps for Kluender's new subjects as well] has to do with the subjects' uncertainty as to what they should be listening for.) For all of the foregoing reasons, I conclude at least that there is no durational contrast effect of filled on following open intervals in the durational ranges of stimuli in Experiments 1 and 2 and hence in the research of Kluender *et al.* Possibly, there is an assimilation effect.

If findings of assimilation in the present experiments were in fact to reflect auditory-system influences on perception, as Kluender *et al.* suppose such findings to do, the influences they reflect do not explain the inverse relation between vowel duration and consonant closure in voiced and unvoiced stop-consonant contrasts that is so commonplace across languages. To the contrary, they provide an example in which language communities almost universally adopt a durational patterning that, far from exploiting distorting auditory-system influences, instead flagrantly ignores them. Looked at in this way, the findings weaken a notion that languages exploit distorting properties of the auditory system. As much as I am attracted to a conclusion such as this, I will suggest in Section 4 why it is probably unwarranted; however, recognizing why it is involves rejecting the approach that Kluender and his colleagues adopt for examining the role of auditory system processes in speech and non-speech perception.

4. General discussion

It is highly likely that the character of the human auditory system shapes the character of the sound patterns of human languages. Lindblom (1986) has shown convincingly that it does. Language communities develop inventories of sounds that are easy to perceive and to distinguish one from the other. Moreover, analogous constraints quite generally shape the form that human artifacts take—we do not generally construct transparent floors, for example, presumably because they are hard to see. The flaws in the research by Kluender *et al.* do not lie in the general idea that the character of the auditory system imprints itself on the form that language sound systems take, but rather in the logic they use to pursue the idea.

As I have argued elsewhere (Fowler, 1990), procedures like those used in this research and used rather widely in speech research do not necessarily provide interpretable information about any auditory-system processing subserving perception of speech or of speech-like signals. The reason they do not is that they leave out of consideration an important source of patterning in listener's responding.

The logic behind the speech–non-speech comparisons in this and other research, I believe, is as follows. If the responses in a classification task exhibit patternings that are not obviously explainable by structure in the acoustic signal, then they are ascribed to auditory-system processes applied to the signal. This ascription is especially likely if the acoustic signals in question are nonsense signals such as the square waves of the present research. Moreover, if responses to speech stimuli pattern in the same way as responses to somewhat analogous nonsense stimuli, then both response patterns are ascribed to general auditory processes. The responses to speech signals in particular are held not to involve special-to-speech processes that presumably would impose their own distinctive patterns on responding.

In my view, the logic is mistaken. Most importantly, it fails to consider an important source of variation in responding, and that is the character of the perceived sound-producing source of the signal. Our perceptual systems have evolved to put perceivers in contact with the properties of the environment with which they interact, and that is how perceivers strive to use their perceptual systems (cf. Warren, 1981). Visual perceivers see surfaces, objects and events in the world, not the reflected light from these surfaces, objects and events. Of course it is reflected light and not the environment that stimulates the visual system; however, even so, visual systems use reflected light not as a perceptual object, but as information for the sources of its structure, and so viewers see the (visible) world. Haptic perceivers feel surfaces and objects; they use the skin deformations and joint-angle changes that felt objects cause not as perceptual objects, but as information for their causes. Indeed, the properties of objects that we both see and feel are not just compatible; they are the same properties—properties of the objects themselves, not of the perceptual systems that inform about them.

We know that, in at least some respects, the auditory system works in abstractly the same way as the haptic and visual systems in yielding environmental perceptual objects rather than percepts of the acoustic signal itself. We hear the location in space of a sound-producing object; we hear neither the location of the acoustic signals that report about the object, nor the time-of-arrival and intensity differences in the signal themselves.

No one is surprised when, in an experiment on visual perception, subjects' classification responses pattern in a way that is largely explained by patterning in the distal causes of the optical stimulation. We should not be surprised either if there is an analogous source of patterning in responses to acoustic signals. That is, the distal cause of acoustic structure is likely to affect subjects' response patterns, because that will be the listener's perceptual object. It is not necessarily the case, either, that "nonsense" acoustic signals evade that source of influence on response patterns and, therefore, directly reflect the structure in the acoustic signal itself as processed by the auditory system. Listeners certainly strive to pull information from such signals about a source in the world. In particular, they localize the sound in space, and they hear highly impoverished signals as something real (for example, as speech) given only small provocation (e.g., Remez, Rubin, Carrell & Pisoni, 1981).

If listeners' response patterns are determined in part by properties of the perceived cause of the acoustic signal, then, when subjects' responses to speech and non-speech signals pattern similarly, the reason for the similar pattern may as well be similarities in the perceived causes of the signals as similarities in the auditory system processes applied to the signal. I have shown (Fowler, 1990) that response patterns mimicking those of durational contrast (and of the effects of vowel duration on /b/–/w/ classifications as studied by Miller and Liberman (1979) and ascribed to durational contrast by Diehl and Walsh (1989)) can be obtained for non-speech signals, or the reverse pattern can be obtained for similar non-speech signals. Which outcome obtains depends on the character of the physical event that is perceived as giving rise to the signals. Here, response patterns are driven by the character of the distal sound-producing events, not by the auditory system processes applied to the signals. If response patterns to acoustic signals are to be used to draw inferences about auditory system processes, means must be developed to distinguish distal-source-produced response patterns from auditory-system-produced patterns. As far as I can tell, for example, there is no way to tell yet which source underlies the assimilation effects found in the present research.

A final comment on the logic of the research conducted to uncover a role for auditory-system processes in speech perception is that, sometimes, even where inferences to such processes seem warranted, drawing the inference does not, in itself, explain the listeners' behavior patterns. If the character of the auditory system leaves its mark on listeners' response patterns to speech utterances, one possibility is that the mark reflects properties of the auditory system that have evolved to serve a function. If Warren's (1985) account of criterion shifts is apt, it is not enough or even accurate to conclude that we exhibit contrast effects because that is what the auditory system does to stimulus inputs. Rather, criterion shifts reflect a generally adaptive strategy for bringing relevant past experience to bear on assessment of current perceptual experience.

I thank Douglas Whalen for making the square wave stimuli for me at Haskins Laboratories. This research was supported by NICHD Grant HD01994 to Haskins Laboratories.

References

Behar, I. & Bevan, W. (1961) The perceived duration of auditory and visual intervals: Crossmodal comparison and interaction, *American Journal of Psychology*, **74**, 17–26.

Bjork, R. (1974) The updating of human memory. In *The psychology of learning and motivation*, vol. 12 (G. Bower, editor), pp. 235–259. New York: Academic Press.

Cathcart, E. P. & Dawson, S. (1928–1929) Persistence (2), *British Journal of Psychology*, **19**, 343–356.

Davis, S. & Summers, W. V. (1989) Vowel length and closure duration in word-medial VC sequences. *Journal of Phonetics*, **17**, 339–353.

Diehl, R. & Walsh, M. (1989) An auditory basis for the stimulus-length effect in the perception of stops and glides, *Journal of the Acoustical Society of America*, **85**, 2154–2164.

Fowler, C. A. (1990) Sound-producing sources as objects of speech perception: Rate normalization and non-speech perception, *Journal of the Acoustical Society of America*, **88**, 1236–1249.

Freyd, J. & Johnson, J. Q. (1987) Probing the time course of representational momentum, *Journal of Experimental Psychology: Learning Memory and Cognition*, **13**, 259–268.

Guilford, J. P. & Park, D. G. (1931) The effect of interpolated weights upon comparative judgments, *American Journal of Psychology*, **43**, 589–599.

Johnson, D. M. (1944) Generalization of a scale of values by the averaging of practice effects. *Journal of Experimental Psychology*, **34**, 425–436.

Keating, P. (1985) Universal phonetics and the organization of grammars. In *Phonetic linguistics: Essays in honor of Peter Ladefoged*, (V. Fromkin, editor) pp. 115–132. Orlando: Academic Press.

Kluender, K. R., Diehl, R. L. & Wright, B. A. (1988) Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics*, **16**, 153–169.

Lindblom, B. (1986) Phonetic universals in vowel systems. In *Experimental phonology* (J. Ohala & J. Jaeger, editors) pp. 13–44. Orlando: Academic Press.

Luce, P. A. & Charles-Luce, J. (1985) Contextual effects on vowel duration, closure duration and the vowel/consonant ratio in speech production, *Journal of the Acoustical Society of America*, **78**(6), 1949–1957.

Miller, J. & Liberman, A. (1979) Some effects of later-occurring information on the perception of stop consonant and semivowel, *Perception and Psychophysics*, **25**, 457–465.

Parker, E. M., Diehl, R. L. & Kluender, K. R. (1986) Trading relations in speech and non-speech. *Perception and Psychophysics*, **39**, 129–142.

Pisoni, D. B., Carrell, T. D. & Gans, J. S. (1983) Perception of the duration of rapid spectrum changes in speech and non-speech signals, *Perception and Psychophysics*, **34**(4), 314–322.

Raphael, L. (1972) Preceding vowel duration as a cue to the perception of the voicing characteristics of word-final consonants in English, *Journal of the Acoustical Society of America*, **51**, 1296–1303.

Remez, R., Rubin, P., Carrell, T. & Pisoni, D. (1981) Speech perception without traditional speech cues. *Science*, **212**, 947–950.

Repp, B., Healy, A. & Crowder, R. (1979) Categories and context in the perception of isolated steady-state vowels, *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 129–145.

Walker, J., Irion, A. & Gordon, D. (1981) Simple and contingent after effects of perceived duration in vision and audition, *Perception and Psychophysics*, **29**, 475–486.

Warren, R. M. (1981) Measurement of sensory intensity, *The Behavioral and Brain Sciences*, **4**, 175–189.

Warren, R. M. (1985) Criterion shift rule and perceptual homeostasis, *Psychological Review*, **92**, 574–584.

Woodrow, H. (1928) Behavior with respect to short temporal stimulus-forms, *Journal of Experimental Psychology*, **11**(3), 167–193.