Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production

Paul A. Luce and Jan Charles-Luce[a)]
Speech Research Laboratory, Department of Psychology, Indiana University, Bloomington, Indiana 47405

(Received 19 December 1983; accepted for publication 5 September 1985)

Acoustic measurements were conducted to determine the degree to which vowel duration, closure duration, and their ratio distinguish voicing of word-final stop consonants across variations in sentential and phonetic environments. Subjects read CVC test words containing three different vowels and ending in stops of three different places of articulation. The test words were produced either in nonphrase-final or phrase-final position and in several local phonetic environments within each of these sentence positions. Our measurements revealed that vowel duration most consistently distinguished voicing categories for the test words. Closure duration failed to consistently distinguish voicing categories across the contextual variables manipulated, as did the ratio of closure and vowel duration. Our results suggest that vowel duration is the most reliable correlate of voicing for word-final stops in connected speech.

PACS numbers: 43.70.Fq

INTRODUCTION

The finding that vowels before voiced consonants are longer than before voiceless consonants has been well documented in the literature (Chen, 1970; House, 1961; House and Fairbanks, 1953; Klatt, 1973; Lisker, 1978; Malécot, 1970; Peterson and Lehiste, 1960), and perceptual studies have shown that vowel duration plays a role in signaling voicing of syllable-final consonants (Denes, 1955; Lisker, 1978; Port and Dalby, 1982; Raphael, 1972). Although a number of studies have recently called into question the primacy of vowel duration in the perception of voicing (Barry, 1979; Hogan and Rozsypal, 1980; Lisker, 1981; Raphael, 1981; Wardrip-Fruin, 1982), it is nevertheless clear that the duration of the vowel preceding syllable-final consonants constitutes one of a set of cues to voicing (see Hillenbrand *et al.*, 1984). In addition, recent evidence has suggested that vowel duration may be the primary cue to voicing for infants (Eilers *et al.*, 1984). Thus, despite recent evidence that other cues such as voicing during closure may actually be superior to vowel duration in signaling voicing, this temporal correlate of voicing appears to play some role in adult and infant perception of voicing of syllable-final consonants, especially when other cues to voicing are ambiguous.

Also well documented in the literature are the findings that closure durations for syllable-final stops are longer for voiceless stops than voiced stops (Lisker, 1957; Port, 1978, 1979, 1981a) and that closure duration may cue voicing in perception (Lisker, 1957, 1978; Port, 1979; Port and Dalby, 1982), although, again, only when other cues are ambiguous. Thus both vowel duration and closure duration have been shown to be potential cues to the voicing distinction of word-final stops in the absence of or in concert with other temporal and spectral cues.

Recent production studies examining vowel duration as an acoustic correlate of voicing in connected speech have suggested that the utility of this temporal interval in signaling voicing is mediated by at least two factors. First, both Klatt (1976) and Umeda (1975) have shown that the difference in duration between vowels preceding voiced and voiceless stops is smaller in nonphrase-final and nonprepausal positions than in phrase-final and prepausal positions. Indeed, Klatt (1976) claims that "a large difference in vowel duration is only seen in phrase-final environments, so it is only in these cases that the durational cue has primary importance" (p. 1219).

Another factor that has been shown to affect vowel duration as a correlate of voicing in word-final stops is the identity of the vowel. Crystal and House (1982) have shown that short vowels preceding voiced and voiceless stops in connected speech exhibit almost no differences in duration (on the order of 5 ms), whereas longer vowels exhibit a small, but more marked, effect of voicing of the final stop.

Closure duration has also proven to be a somewhat fickle correlate of voicing in connected speech. Although Umeda (1977) has demonstrated small differences in stop duration between voiced and voiceless word-final stops, Crystal and House (1982) have presented data that show virtually no differences in closure duration as a function of voicing.

Taken together, the results from these studies suggest that vowel and closure duration may have limited applicability as correlates of voicing in connected speech. Thus one purpose of the present study was to explore in some depth the utility of vowel and closure duration as correlates of voicing in connected speech under more systematic conditions than have previously been employed. We reasoned that vowel and/or closure duration may prove to distinguish voicing categories when a number of factors known to influence these temporal intervals are carefully controlled. We hypothesized that larger and more consistent effects of voicing on these temporal intervals may be uncovered once the phonetic and sentential contexts in which the stops are produced are precisely manipulated. Such manipulations may, of

[a)] Also in the Department of Linguistics, Indiana University, Bloomington, IN 47405.

TABLE I. Test words and sentence frames produced by the talkers for experiments 1 and 2.

	Test Words		
Experiments 1 and 2:			
Velars:	/pɪg/–/pɪk/	/pig/–/pik/	/kag/–/kak/
Bilabials:	/dɪb/–/dɪp/	/dib/–/dip/	/kab/–/kap/
Alveolars:	/pɪd/–/pɪt/	/kid/–/kit/	/kad/–/kat/
	Sentence frames		
Experiment 1:			
(1) When Mark read ___, Elaine made a checkmark.			
(2) When Mark read ___ aloud, Elaine made a checkmark.			
(3) If Ted says ___, Tom will leave the room.			
(4) If Ted says ___ today, Tom will leave the room.			
Experiment 2:			
(1) When Mark reads ___, instantly he makes a checkmark.			
(2) When Mark reads ___ inside, instantly he makes a checkmark.			
(3) If Ted says ___, Cindy will leave the room.			
(4) If Ted says ___ sincerely, Cindy will leave the room.			

course, result in speech that is in some sense less "connected" or "fluent" than has previously been examined. However, our approach is advantageous in at least two ways: First, it allows us to compare voice cognates in identical environments and, second, it enables us to apply standard statistical analyses to the data in order to assess the reliability of any differences that we may find.

The present study, then, was conducted to address in a systematic fashion the issue of phonetic and sentential effects on vowel and closure duration as correlates of voicing of word-final stops. We investigated these temporal intervals under a variety of manipulations that were predicted to affect their absolute durations. In particular, we examined the effects of: (1) sentence position (nonphrase-final versus phrase-final); (2) local phonetic environment following the word-final stop; (3) place of articulation of the word-final stop; and (4) inherent duration of the vowel preceding the stop. We were interested in determining under what conditions, if any, vowel and closure durations would distinguish voicing when phonetic and sentential factors are carefully controlled.

In addition to examining in a precise manner the effects of context on vowel and closure durations, we were interested in testing a claim recently made by Port (1981a,b; Port and Dalby, 1982; see also Kohler, 1979) that the consonant–vowel (C/V) ratio (the ratio of closure and vowel duration) is an abstract correlate of voicing that may be independent of certain contextual variables. Port (1981a) has shown that the C/V ratio distinguishes voicing across changes in speaking rate, the number of syllables in the test word, and the tenseness of the vowel preceding the syllable-final stop—all three of which affect the absolute durations of both the vowel and closure intervals. Therefore, although speakers may produce vowel and closure durations within a wide range of durations, the C/V ratio appears to distinguish voicing consistently in speech production as a consequence of temporal covariation of vowel and closure durations. According to Port, because the C/V ratio is a relational cue, and therefore not dependent on the absolute durations of the two relevant segments, it may therefore prove more robust to contextual modifications of the durations of vowels and closures in signaling voicing than either vowel or closure duration considered alone.

Port and Dalby (1982) and Fitch (1981) have provided evidence supporting the perceptual relevance of the C/V ratio. Port and Dalby demonstrated that when other cues to voicing are ambiguous, the C/V ratio is apparently the primary cue for perception of voicing of syllable-final stop consonants. Both they and Fitch (1981) have furthermore shown that the critical C/V ratio remains relatively constant across changes in speaking rate. Thus data from studies of both speech production and perception suggest that the compensatory relationship between the temporal factors of vowel duration and closure duration provides a cue to voicing independent of the effects of speaking rate, whereas either of these factors alone does not. (See, however, Massaro and Cohen, 1983, for an opposing viewpoint.)

Given this evidence, it is possible that the C/V ratio would also distinguish voicing across the various sentential and phonetic contexts examined in the present study. Although existing research suggests that the C/V ratio may undergo substantial modifications across sentence position due to a lack of lengthening of closure duration when vowel duration increases in phrase-final position (Klatt, 1975, 1976; Oller, 1973), we were nonetheless interested in determining whether the C/V ratio better accounts for voicing of word-final stops than the absolute durations of the vowel and/or closure.

I. EXPERIMENT 1

A. Method

1. Subjects

Three male and three female volunteers recruited from the laboratory staff served as unpaid subjects, although the data for one of the subjects were subsequently discarded (see below). All subjects were native English speakers and reported no history of speech or hearing disorders. Subjects were naive to the purpose and design of the study.

2. Materials

Nine minimal pairs of consonant–vowel–consonant test words were used. Words within a minimal pair differed only in the voicing of the final consonant. Three of the pairs ended in a bilabial stop, three ended in an alveolar stop, and three ended in a velar stop. Each word-final consonant was preceded by one of three vowels: /ɪ/, /i/, or /a/. The vowels were chosen to examine the intersection of differences in inherent vowel duration with our dependent temporal measures. On the average, /a/ tends to be longer in duration than /i/, and /i/ longer than /ɪ/ (Crystal and House, 1982; House, 1961; House and Fairbanks, 1953; Peterson and Lehiste, 1960).[1] The test words are shown in Table I.

Previous research by Port (1981a; see also Peterson and Lehiste, 1960) has shown that place of articulation, manner, and voicing of the initial stop consonant in a consonant–vowel–consonant syllable have little or no effect on the durations of the intervals relevant to the present study. The initial stop consonants were thus selected only to facilitate measurement of vowel duration. Furthermore, in order to construct such minimally contrastive pairs, it was necessary to include some nonsense words in the set of test words. All nonsense words, however, were phonologically permissible English sequences.

Each test word was embedded in one of four sentence frames, which are also given in Table I. In sentence frames (1) and (3), the test word occurred phrase finally, whereas in sentence frames (2) and (4) one word occurred between the test word and the phrase boundary. The *sentential contexts* thus made the test words either phrase-final or nonphrase-final. In addition, in sentence frames (1) and (2), the test word was followed by the unstressed, midcentral vowel /ə/. In sentences frames (3) and (4), the test word was followed by the voiceless alveolar stop /t/. The stop /t/ was chosen to contrast with the reduced vowel /ə/ in terms of its possible acoustic effects on the vowel and closure durations preceding it, given the differences in vocal tract constriction required for the two segments.[2] The *local phonetic environments* of the test words were thus either a stop or a vowel. (Local phonetic environment here refers only to the initial phoneme of the word *following* the test word.)

3. Procedure

Two repetitions of each word in each sentence frame were read by the speakers. This resulted in 144 test sentences (18 test words times four sentence frames times two repetitions) for each speaker. Each subject read two blocks of 144 stimuli, only one block of which comprised the materials for the present experiment. The order of presentation of the blocks was balanced across subjects. Each subject received a different randomization of the test sentences.

At the beginning of a session, the subject was instructed to read each sentence in as natural a manner as possible and to avoid placing any undue stress on any of the words. He or she then read a short practice list of sentences to familiarize himself or herself with the materials and to allow the experimenter to adjust the levels on the tape recorder. The utterances were recorded in a sound attenuated booth (IAC Model 401A) using an Electro-Voice D054 microphone and an Ampex AG-500 tape recorder.

4. Measurements

All 144 test sentences for each of the six speakers were low-pass filtered at 4.8 kHz and digitized at a 10-kHz sampling rate via a 12-bit analog-to-digital converter. Measurements were made from a visual waveform display using a digital waveform editor (see Luce and Carrell, 1981). For each test word, vowel duration and closure duration for the final stop consonant were measured. Vowel duration was measured from onset of periodic energy to a marked decrease in amplitude in the periodic energy in the waveform. Closure duration was measured from this decrease in amplitude to the onset of the release burst of the word-final stop. Only words with clearly identifiable release bursts belonging to the final stop of the test word were included in our analysis. Data from one subject, who failed to release over 15% of her final stops, were excluded. Of the remaining five subjects, 2.4% of the test words were excluded due to unreleased final stops.

B. Results and discussion

Vowel durations and postvocalic closure durations were averaged across repetitions of each test word produced by each speaker and entered into two separate five-way (sentence position $\times$ local phonetic environment $\times$ place of articulation $\times$ inherent vowel duration $\times$ voicing) repeated measures analysis of variance. A five-way analysis of variance was also performed on C/V ratios computed from the vowel and closure durations averaged across repetitions. Because the main effect of voicing constituted the variable of primary interest in the present study, only those interactions involving voicing will be reported for the vowel duration, closure duration, and C/V ratio data. Such interactions indicate what effects, if any, the variables manipulated in this study had on vowel duration, closure duration, or the C/V ratio as correlates of word-final voicing. (An extended discussion of those interactions involving sentence position, local phonetic environment, and inherent vowel duration—where not directly relevant to the voicing feature—may be found in Luce and Charles-Luce, 1983.) The results will be discussed separately for the three dependent variables.

1. Vowel durations

Mean vowel durations and between subject standard deviations are shown in Table II for each test word produced in each local phonetic environment and sentence position. Significant main effects were obtained for sentence position $[F(1,4) = 18.61,\ p < 0.02]$, place of articulation $[F(2,8) = 7.28, p < 0.02]$, inherent vowel duration $[F(2,8) = 103.00, p < 0.0001]$, and voicing $[F(1,4) = 171.48, p < 0.0002]$. The effect of local phonetic environment failed to reach significance $[F(1,4) = 7.23, p > 0.05]$.

The significant main effect of sentence position revealed that vowel durations were longer for test words produced in phrase-final than in nonphrase-final position, thus replicating the already well-documented evidence that vowels lengthen phrase finally (Cooper, 1975; Klatt, 1975, 1976; Oller, 1973). On the average, vowel durations were 69 ms longer in phrase-final than in nonphrase-final position (mean duration for phrase-final = 184 ms; mean duration for nonphrase-final = 115 ms). In addition, durations of vowels produced before bilabials (mean duration = 155 ms) were longer than those produced before alveolars (mean duration = 147 ms) and velars (mean duration = 146 ms). Finally, /a/ (mean duration = 179 ms) was significantly longer than /i/ (mean duration = 149 ms), which was significantly longer than /ɪ/ (mean duration = 121 ms).

Of particular interest, however, is the main effect of voicing. On the average, vowel durations were 55 ms longer for test words ending in voiced stops (mean duration = 177 ms) than those ending in voiceless stops (mean duration = 122 ms).

Voicing for the vowel durations entered into two two-way interactions and one three-way interaction. A significant two-way interaction of sentence position and voicing $[F(1,4) = 19.33, p < 0.02]$ revealed that the difference in du-

TABLE II. Vowel durations and (between subject) standard deviations for test words produced in each local phonetic environment and sentential context.

	Vowel environment (/ə/)					
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	106.3	62.2	99.1	84.9	116.5	83.2
	(24.3)	(16.6)	(23.5)	(27.5)	(29.4)	(21.8)
/i/	149.1	83.2	130.7	91.2	166.1	108.6
	(26.6)	(17.4)	(16.6)	(14.3)	(32.0)	(19.8)
/a/	162.6	129.4	165.6	133.3	162.9	128.0
	(22.0)	(22.1)	(27.5)	(16.9)	(30.5)	(18.9)
Phrase-final:						
/ɪ/	187.1	129.6	177.3	131.0	186.9	142.8
	(42.1)	(36.6)	(41.1)	(36.6)	(46.7)	(41.4)
/i/	214.9	138.1	223.3	123.7	226.2	164.7
	(65.3)	(29.8)	(46.2)	(29.0)	(35.9)	(34.2)
/a/	246.9	188.1	249.6	170.6	250.0	173.0
	(57.1)	(42.8)	(55.7)	(40.4)	(50.2)	(58.8)
	Stop environment (/t/)					
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	94.7	52.9	89.6	61.0	91.5	74.3
	(23.8)	(9.5)	(26.0)	(12.3)	(18.7)	(19.1)
/i/	138.2	72.2	130.5	86.1	147.7	99.5
	(26.2)	(14.5)	(22.6)	(17.6)	(19.3)	(23.3)
/a/	164.1	118.1	166.5	120.5	160.0	111.0
	(16.5)	(15.5)	(16.5)	(23.3)	(15.1)	(22.6)
Phrase-final:						
/ɪ/	182.1	113.8	184.1	127.7	187.9	132.2
	(37.3)	(35.7)	(41.2)	(37.3)	(47.8)	(33.1)
/i/	213.1	130.0	220.5	133.5	226.7	155.3
	(50.3)	(36.8)	(55.4)	(24.6)	(47.9)	(32.4)
/a/	243.4	177.1	249.9	178.8	251.6	184.1
	(46.7)	(30.2)	(57.9)	(43.1)	(52.9)	(44.1)

TABLE III. Closure durations and (between subject) standard deviations for test words produced in each local phonetic environment and sentential context.

	Vowel environment (/ə/)					
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	38.6	70.1	32.2	34.8	61.8	96.7
	(11.8)	(10.3)	(12.7)	(15.9)	(10.9)	(6.7)
/i/	40.3	78.2	32.0	49.6	61.6	87.5
	(12.2)	(22.2)	(11.3)	(9.1)	(11.7)	(10.5)
/a/	50.2	69.8	31.7	41.8	50.2	75.2
	(10.7)	(9.0)	(15.9)	(12.6)	(9.2)	(14.7)
Phrase-final:						
/ɪ/	63.7	98.6	68.6	72.4	88.5	103.1
	(9.4)	(14.6)	(11.0)	(20.7)	(13.8)	(17.1)
/i/	75.1	89.3	53.6	78.9	70.9	103.5
	(22.4)	(13.2)	(3.2)	(15.9)	(5.7)	(14.9)
/a/	68.7	88.8	53.8	72.3	77.6	102.8
	(10.6)	(13.3)	(9.8)	(32.6)	(7.8)	(12.4)
	Stop environment (/t/)					
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	97.7	79.5	99.5	99.5	101.1	97.0
	(36.1)	(26.0)	(33.7)	(20.1)	(8.9)	(20.8)
/i/	80.3	78.2	88.6	93.9	84.2	90.6
	(22.9)	(24.0)	(35.0)	(28.4)	(12.2)	(10.4)
/a/	79.5	74.0	77.3	134.9	81.1	84.8
	(23.0)	(17.0)	(18.1)	(51.6)	(27.5)	(12.6)
Phrase-final:						
/ɪ/	71.8	108.8	63.4	78.1	87.0	106.4
	(17.1)	(13.7)	(8.4)	(14.9)	(10.9)	(15.2)
/i/	66.7	98.1	52.8	73.4	75.4	95.8
	(13.0)	(15.7)	(6.8)	(10.1)	(7.8)	(11.6)
/a/	71.8	109.5	62.3	78.9	78.1	95.9
	(13.4)	(37.1)	(13.4)	(14.7)	(14.9)	(9.2)

ration between vowels preceding voiced and voiceless stops was greater in phrase-final position (mean difference = 68 ms) than in nonphrase-final position (mean difference = 42 ms), thus replicating, in part, previous findings (Klatt, 1976; Umeda, 1975). The relative durations of vowels preceding voiced and voiceless stops were also affected by inherent vowel duration, as evidenced by a significant interaction of inherent vowel duration and voicing [$F(2,8) = 12.80$, $p < 0.004$]. The difference between vowels preceding voiced and voiceless stops was greatest for test words containing /i/ (mean difference = 66 ms), least for test words containing /ɪ/ (mean difference = 42 ms), and intermediate for test words containing /a/ (mean difference = 55 ms). Therefore, similar to the findings of Crystal and House (1982), the shortest vowel, /ɪ/, resulted in the smallest voicing difference. Finally, voicing interacted with both sentence position and place of articulation [$F(2,8) = 8.43, p < 0.02$], primarily due to the alveolars, which exhibited the smallest difference in duration in nonphrase-final position and the largest difference in phrase-final position.

Despite these differences in relative durations of vowels, separate analyses of variance based on the significant interactions revealed that vowel duration consistently distinguished voicing at the 0.003 level of significance or beyond. Thus, even though voicing differences were modified by sentence position, inherent vowel duration, and, to a lesser extent, place of articulation, vowel duration reliably distinguished voicing categories in all instances.

2. Closure durations

Mean closure durations and standard deviations are given in Table III. Significant main effects of sentence position [$F(1,4) = 8.23, p < 0.05$], local phonetic environment [$F(1,4) = 92.09$, $p < 0.007$], place of articulation [$F(2,8) = 6.38$, $p < 0.03$], and voicing [$F(1,4) = 54.13$, $p < 0.002$] were obtained for the closure durations. The main effect of inherent vowel duration was not significant [$F(2,8) = 1.72, p > 0.05$]. Closure durations were 14 ms longer in phrase-final position (mean duration = 87 ms) than in nonphrase-final position (mean duration = 73 ms) and were 18 ms longer in the local stop environment (mean duration = 86 ms) than in the vowel environment (mean duration = 68 ms). This last result replicates, in part, results previously obtained by Umeda (1977). Closure durations for bilabials (mean duration = 86 ms) were longer than those for velars (mean duration = 77 ms), which in turn were longer than those for alveolars (mean duration = 68 ms).

Closure durations were also 18 ms longer for voiceless stops (mean duration = 86 ms) than voiced stops (mean duration = 68 ms) overall. However, voicing entered into interactions with sentence position and local phonetic environ-

TABLE IV. Consonant/vowel ratios for test words produced in each local phonetic environment and sentential context.

	/g/	/k/	/d/	/t/	/b/	/p/
	Vowel environment (/ə/)					
Nonphrase-final:						
/ɪ/	0.379	1.198	0.343	0.479	0.549	1.222
/i/	0.281	0.974	0.252	0.558	0.392	0.815
/a/	0.314	0.550	0.205	0.317	0.312	0.588
Phrase-final:						
/ɪ/	0.365	0.828	0.408	0.596	0.506	0.789
/i/	0.390	0.674	0.249	0.668	0.319	0.642
/a/	0.293	0.488	0.228	0.453	0.321	0.690
	Stop environment (/t/)					
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	1.042	1.587	1.291	1.690	1.147	1.350
/i/	0.628	1.121	0.710	1.612	0.579	0.934
/a/	0.489	0.631	0.469	1.172	0.500	0.793
Phrase-final:						
/ɪ/	0.419	1.008	0.363	0.664	0.496	0.838
/i/	0.338	0.814	0.253	0.560	0.348	0.639
/a/	0.308	0.649	0.267	0.457	0.322	0.544

ment [$F(1,4) = 8.42, p < 0.05$]; local phonetic environment, and place of articulation [$F(2,8) = 8.22, p < 0.02$]; sentence position, local phonetic environment, and place of articulation [$F(2,8) = 8.59, p < 0.02$]; and sentence position, place of articulation, and inherent vowel duration [$F(4,16) = 3.19, p < 0.05$]. Separate analyses of variance based on these interactions revealed that closure duration failed to distinguish voicing categories in approximately 53% of the total possible number of cases. In nonphrase-final position, effects of voicing at or beyond the 0.05 level of significance were obtained only for the minimal pairs /pig–pik/, /kad–kat/, /dib–dip/, and /kab–kap/ in the local vowel environment. No significant effects were obtained for test words produced in the local stop environment in nonphrase-final position. In phrase-final position, closure duration failed to distinguish voicing categories for all of the test words ending in alveolars except for the pair /kid–kit/ in the local vowel environment. In general, then, closure duration failed to distinguish voicing for most test words ending in alveolars (see also Umeda, 1977) and for most test words in nonphrase-final position, especially in the local stop environment. These results clearly demonstrate that closure duration fails to distinguish voicing consistently across sentence position, local phonetic environment, place of articulation, and inherent vowel duration, thus calling into question the universal perceptual utility of this cue to voicing of word-final stops in connected speech.

3. C/V ratios

Mean C/V ratios are shown in Table IV. Significant main effects for the C/V ratios were obtained for sentence position [$F(1,4) = 17.44, p < 0.02$], local phonetic environment [$F(1,4) = 80.97, p < 0.0003$], inherent vowel duration [$F(2,8) = 15.06, p < 0.002$], and voicing [$F(1,4) = 203.91, p < 0.0001$]. The main effect of place of articulation was not significant ($F < 1.0$). These significant effects parallel those obtained for the vowel and closure durations. Ratios were, overall, larger in nonphrase-final position than in phrase-final position, larger in the stop environment than the vowel environment, and largest for the vowel /ɪ/, intermediate for the vowel /i/, and smallest for the vowel /a/.

Voicing in terms of the C/V ratio entered into five significant interactions: (1) with inherent vowel duration [$F(2,8) = 6.16, p < 0.03$]; (2) with local phonetic environment and place of articulation [$F(2,8) = 7.21, p < 0.02$]; (3) with sentence position and inherent vowel duration [$F(4,16) = 3.33, p < 0.04$]; (4) with sentence position, local phonetic environment, and place of articulation [$F(2,8) = 6.00, p < 0.03$]; and (5) with sentence position, place of articulation, and inherent vowel duration [$F(4,16) = 3.04, p < 0.05$]. Clearly, voicing differences, as measured by the C/V ratio, were affected by each of the independent variables. To determine if the C/V ratio reliably distinguished voicing categories across the modifications imposed by each of the independent variables, separate analyses of variance based on the two four-way interactions were performed. The results of these analyses revealed that the C/V ratio failed to distinguish voicing at the 0.05 level of significance for the minimal pair /kag–kak/ in both local phonetic environments in nonphrase-final position. All other voicing distinctions were significant at the 0.05 level of significance or beyond.

4. Summary

Of the three temporal attributes examined in this investigation, vowel duration proved to to be the most reliable correlate of voicing across changes in sentence position, local phonetic environment, place of articulation, and inherent vowel duration. Although voicing for vowel duration entered into interactions with all but one of the independent variables (i.e., local phonetic environment), significantly longer vowel durations were consistently obtained for voiced than voiceless stops. Closure duration, on the other hand, fared much more poorly, significantly distinguishing voicing for less than half of the cases examined. Finally, although the C/V ratio proved more reliable than closure duration alone, it nevertheless failed to distinguish between voicing categories for /kag–kak/ in nonphrase-final position. In addition, voicing in terms of the C/V ratio did not overcome modifications of the temporal intervals imposed by the variables examined. Instead, the C/V ratio proved more susceptible to contextual effects than absolute vowel duration, entering into interactions with all of the independent variables. Thus the C/V ratio does not appear to constitute a correlate of voicing superior to that of vowel duration alone, at least in terms of the contextual effects examined in the present study. Indeed, the C/V ratio failed to distinguish voicing for all cases, whereas vowel duration alone did.

II. EXPERIMENT 2

In order to further extend and, in part, replicate the results obtained in the previous experiment, a second experiment was performed in which the local stop environment was replaced with a local fricative environment and /ə/ in the local vowel environment was replaced with /ɪ/. The effects of a local fricative environment were examined not only to extend our results, but also to overcome, in part, possible segmentation anomalies that may have arisen (especially for the alveolars) in experiment 1 in measuring postvocalic closure durations in the presence of two adjacent stop closures. Although we were careful to include in our data in experiment 1 only those test words with clearly released stops, we felt compelled to examine another obstruent in the local consonant environment that would allow more precise measurement of closure durations. In addition, we were interested in determining if closure duration would fare better as a correlate of voicing when not followed by an obstruent from the same manner class. In particular, we attempted to determine if closure durations for alveolars might more consistently signal voicing when not followed by an obstruent of the same manner *and* place of articulation. In short, experiment 2 was conducted in order to further extend and confirm our hypothesis that vowel duration alone is a more robust temporal attribute to word-final voicing of stops than either closure duration or the C/V ratio.

A. Method

1. Subjects

Two new male and two new female subjects served as volunteers. All four subjects met the same criteria as those subjects who participated in experiment 1.

2. Materials

The materials were the same as those used in experiment 1 except for the change in the sentence frames (see Table I).

3. Procedure and measurements

The procedure and method of measurement were identical to those used in experiment 1. As in experiment 1, test words with unreleased final stops were excluded from the data analysis. However, because of the use of the local fricative environment, the data from only 0.7% of the test words were excluded.

B. Results and discussion

1. Vowel durations

Mean vowel durations and standard deviations are shown in Table V. Significant main effects for vowel duration were obtained for sentence position [$F(1,3) = 17.20$, $p < 0.03$], inherent vowel duration [$F(2,6) = 59.10$, $p < 0.0001$], and voicing [$F(1,3) = 40.86$, $p < 0.008$]. Significant main effects were not obtained, however, for local phonetic environment [$F < 1.0$] or place of articulation [$F(2,6) = 3.95$, $p > 0.09$].

In contrast to experiment 1, voicing entered into only one significant interaction, namely with inherent vowel duration [$F(2,6) = 21.21$, $p < 0.002$]. The difference between vowels preceding voiced and voiceless stops was greatest for test words containing /a/ (mean difference = 84 ms), least for test words containing /ɪ/ (mean difference = 57 ms), and intermediate for test words containing /i/ (mean difference = 79 ms). Again, the shortest vowel, /ɪ/, produced the smallest voicing difference. Separate analyses of variance based on this interaction revealed that vowel duration significantly distinguished voicing at the 0.02 level of significance or beyond for each of the three vowels. Thus, as in experiment 1, vowel duration consistently distinguished voicing categories. However, unlike experiment 1, voicing differences between vowel durations were significantly modified by only one variable.

2. Closure durations

Mean closure durations and standard deviations are shown in Table VI. Only the main effects of place of articulation [$F(2,6) = 29.23$, $p < 0.008$] and inherent vowel duration [$F(2,6) = 5.94$, $p < 0.04$] were significant. Neither sentence position [$F(1,3) = 3.96$, $p > 0.05$], local phonetic environment ($F < 1.0$), nor voicing [$F(1,3) = 4.71$, $p > 0.1$] reached significance. Although the main effect of voicing was not significant, voicing entered into a significant three-way interaction with local phonetic environment and place of articulation [$F(2,6) = 8.89$, $p < 0.02$]. Separate analyses of variance based on this interaction revealed that only for test words ending in velar stops in the local vowel environment did closure duration distinguish voicing categories at the 0.05 level of significance. Thus, in approximately 83% of the cases, closure duration proved to be an unreliable correlate of voicing, supporting our earlier conclusion that closure

TABLE V. Vowel durations and (between subject) standard deviations for test words produced in each local phonetic environment and sentential context.

	/g/	/k/	/d/	/t/	/b/	/p/
Vowel environment (/ɪ/)						
Nonphrase-final:						
/ɪ/	126.2	78.9	135.3	92.7	155.9	86.4
	(17.3)	(11.6)	(41.5)	(19.3)	(49.5)	(14.7)
/i/	186.7	116.8	190.6	121.5	222.6	134.9
	(33.0)	(30.3)	(58.3)	(29.6)	(62.8)	(35.4)
/a/	229.0	152.8	231.0	164.0	208.1	139.6
	(52.9)	(27.9)	(68.8)	(27.4)	(49.5)	(13.4)
Phrase-final:						
/ɪ/	179.6	118.4	190.6	129.9	203.2	140.8
	(31.7)	(34.1)	(34.0)	(25.6)	(34.1)	(25.7)
/i/	234.0	146.7	239.9	154.4	265.6	170.4
	(9.3)	(29.1)	(33.1)	(33.4)	(31.3)	(28.3)
/a/	292.4	196.5	316.9	198.6	283.5	191.9
	(29.4)	(21.1)	(68.1)	(29.0)	(22.1)	(18.1)
Fricative environment (/s/)						
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	126.1	78.6	134.4	86.0	148.3	90.4
	(44.2)	(15.4)	(57.7)	(18.7)	(46.6)	(21.2)
/i/	172.7	105.6	184.8	115.3	188.6	147.1
	(41.7)	(32.8)	(74.8)	(34.4)	(92.1)	(27.6)
/a/	196.0	153.0	216.5	150.3	214.4	142.8
	(54.1)	(35.4)	(66.9)	(34.5)	(58.9)	(32.9)
Phrase-final:						
/ɪ/	195.1	126.0	191.9	125.0	196.5	143.5
	(45.6)	(27.0)	(41.3)	(28.0)	(19.9)	(18.6)
/i/	235.6	145.4	251.4	149.5	257.8	171.6
	(29.6)	(22.0)	(36.4)	(33.7)	(35.1)	(28.2)
/a/	281.3	188.8	303.2	193.7	285.6	184.1
	(27.9)	(11.3)	(34.0)	(28.5)	(17.9)	(22.6)

TABLE VI. Closure durations and (between subject) standard deviations for test words produced in each local phonetic environment and sentential context.

	Vowel environment (/ɪ/)					
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	60.4	88.4	55.6	79.0	81.6	115.3
	(19.2)	(31.0)	(26.8)	(53.2)	(13.1)	(30.6)
/i/	56.6	83.4	41.5	69.6	75.8	100.9
	(28.6)	(32.7)	(28.2)	(52.7)	(2.6)	(17.9)
/a/	51.5	84.2	50.7	67.8	60.5	92.7
	(16.4)	(23.0)	(33.8)	(58.9)	(23.2)	(29.2)
Phrase-final:						
/ɪ/	73.4	95.0	76.1	92.1	95.0	111.9
	(15.1)	(28.8)	(16.5)	(38.2)	(8.4)	(21.7)
/i/	67.4	95.2	61.0	92.6	78.4	100.4
	(17.5)	(23.9)	(13.0)	(34.1)	(8.6)	(16.0)
/a/	64.0	92.9	55.2	85.8	78.3	102.4
	(11.7)	(25.7)	(6.9)	(27.7)	(16.6)	(23.1)
	Fricative environment (/s/)					
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	62.8	86.1	74.9	94.0	89.3	101.8
	(24.4)	(41.9)	(23.6)	(37.6)	(10.8)	(36.8)
/i/	56.4	78.2	58.4	91.2	73.3	90.5
	(24.7)	(44.8)	(12.5)	(27.7)	(29.0)	(37.0)
/a/	59.7	74.6	61.6	88.8	76.6	89.4
	(12.9)	(32.4)	(7.7)	(22.5)	(27.0)	(39.1)
Phrase-final:						
/ɪ/	71.0	95.5	68.5	102.1	95.7	103.9
	(15.8)	(20.9)	(18.5)	(42.2)	(14.4)	(22.1)
/i/	67.2	95.1	60.4	83.5	78.5	106.1
	(12.2)	(40.6)	(13.1)	(40.7)	(11.7)	(27.2)
/a/	56.7	96.0	57.2	116.6	81.5	99.7
	(8.3)	(24.8)	(13.0)	(16.4)	(13.3)	(29.3)

duration does not constitute a consistent attribute of voicing of word-final stops in connected speech.

3. C/V ratios

Mean C/V ratios are shown in Table VII. As in experiment 1, significant main effects were obtained for inherent vowel duration [$F(2,6) = 26.70$, $p < 0.001$], and voicing [$F(1,3) = 19.47, p < 0.03$]. Also as in experiment 1, the main effect of place of articulation failed to reach significance [$F(2,6) = 4.82, p > 0.05$]. However, unlike experiment 1, the main effects of sentence position [$F(1,3) = 9.97$, $p > 0.05$] and local phonetic environment [$F(1,3) = 1.06, p > 0.3$] were not significant. The failure to observe significant main effects of environment for the closure durations or the C/V ratios in this experiment appears to have arisen because the effect of the following fricative had a smaller differential effect on closure durations compared to the effects of the following vowel. This is in contrast to the results obtained in experiment 1 in which a main effect of environment was observed for both closure durations and the C/V ratio. Differential effects of local phonetic environment were still in evidence, however, in the numerous interactions involving voicing and environment.

Turning now to the interactions involving voicing for the C/V ratios, we find a pattern of results similar to that

TABLE VII. Consonant/vowel rations for test words produced in each local phonetic environment and sentential context.

	Vowel environment (/ɪ/)					
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	0.475	1.121	0.409	0.852	0.552	1.333
/i/	0.292	0.699	0.211	0.531	0.359	0.765
/a/	0.222	0.546	0.204	0.389	0.283	0.655
Phrase-final:						
/ɪ/	0.429	0.885	0.407	0.712	0.479	0.813
/i/	0.290	0.645	0.254	0.599	0.296	0.590
/a/	0.222	0.472	0.177	0.428	0.274	0.532
	Fricative environment (/s/)					
	/g/	/k/	/d/	/t/	/b/	/p/
Nonphrase-final:						
/ɪ/	0.518	1.078	0.628	1.076	0.634	1.104
/i/	0.316	0.701	0.357	0.797	0.409	0.621
/a/	0.309	0.489	0.304	0.589	0.353	0.603
Phrase-final:						
/ɪ/	0.388	0.798	0.368	0.874	0.490	0.722
/i/	0.288	0.653	0.246	0.540	0.304	0.614
/a/	0.202	0.506	0.189	0.603	0.284	0.543

obtained in experiment 1. Voicing for the C/V ratios interacted with local phonetic environment and place of articulation [$F(2,6) = 5.79, p < 0.03$], as well as with sentence position and inherent vowel duration [$F(2,6) = 7.94, p < 0.03$]. Thus, as in experiment 1, voicing in terms of the C/V ratio was affected by each of the variables manipulated in this study.

Separate analyses of variance based on these two significant interactions revealed that the C/V ratio failed to distinguish voicing categories at the 0.05 level of significance for test words ending in alveolar stops produced in the vowel environment and for test words containing /ɪ/ produced in final position for all three places of articulation. Once again, therefore, voicing in terms of the C/V ratio was modified by contextual factors more than absolute vowel duration, and consequently failed to distinguish voicing for a number of minimal pairs.

III. GENERAL DISCUSSION

The primary goal of this investigation was to examine the extent to which three temporal attributes of phonological voicing of word-final stops consistently distinguish voicing across changes in sentential and phonetic contexts, the inherent duration of the vowel preceding the word-final stop, and the place of articulation of the word-final stop. We attempted to determine whether vowel duration, closure duration, or the C/V ratio would prove most robust, in terms of signaling voicing, under changes in the phonetic and sentential environments in which the stops were produced. Our results bear on the issue of which of these possible correlates of voicing is most likely to consistently cue the voicing distinction in perception of continuous speech.

The finding of primary importance was that the durations of vowels preceding word-final stops most consistently distinguished voicing across the various manipulations employed. Although the relative durations of vowels preceding voiced and voiceless stops varied as a function of sentence position (see Klatt, 1976, and Umeda, 1975), place of articulation, and inherent vowel duration (see Crystal and House, 1982) in experiment 1 and as a function of inherent vowel duration in experiment 2, vowel duration nevertheless proved to be a statistically reliable correlate of voicing across the phonetic and contextual variables examined. Thus, even though it is impossible to adopt a fixed criterion for judging voicing on the basis of vowel duration that is independent of vowel identity and phonetic and sentential context, it appears that, once adjustments are made for the particular environment in which the word-final stop is articulated in connected speech (cf. Raphael, 1981), vowel duration will constitute a consistent correlate of phonological voicing.

In contrast to vowel duration, which proved to distinguish voicing reliably in all cases, closure duration fared much more poorly. Indeed, closure duration failed to distinguish voicing in 53% of the cases in experiment 1 and in 83% of the cases in experiment 2. In a situation approximating continuous speech, then, differences in the closure durations of voiced and voiceless stops tend to be small and exhibit considerable variance in production, replicating, in part, the findings of Crystal and House (1982; see also Umeda, 1977). One of the primary sources of this variation in production of closure durations appears to be the identity of the segment following the word-final stop.

Given the unreliable nature of closure duration as a correlate of voicing observed in our study, it is not surprising that combining vowel and closure duration into a single abstract expression of the voicing feature proved less robust to sentential and phonetic environmental influences than simply vowel duration alone. Although the C/V ratio clearly fared better than closure duration, due to the contribution of the vowel component of the ratio, it nevertheless failed to distinguish voicing consistently in all cases. In addition, the C/V ratio is presumably a more "abstract" correlate of voicing that should not as readily succumb to contextual modifications as, for example, the absolute duration of the vowel. However, it was found that voicing as a function of the ratio was *more*, rather than less, affected by the variables manipulated. (See also Repp and Williams, 1985, for a discussion of the failure of the C/V ratio to remain constant in perception under similar contextual modifications.)

In short, the present study suggests that vowel duration may serve to be the most consistently reliable temporal cue to voicing of word-final stops of the three correlates of voicing examined here. Of course, temporal and spectral cues other than vowel duration (e.g., voicing during closure, frequency of formants at offset prior to closure) have been shown to play powerful roles in the perception of voicing (Hillenbrand *et al.*, 1984; Wardrip-Fruin, 1982; Parker, 1974; see, however, O'Kane, 1978). It would therefore be of considerable interest to establish the reliability of these cues across various sentential and phonetic environments. Nevertheless, our results suggest that vowel duration, as one cue in a possible constellation of cues, may provide a consistent, and thus, reliable correlate of voicing of word-final stops in continuous speech, once, of course, contextual modifications imposed by the articulation of words in sentences are taken into account. In this sense, then, we have addressed not the question of which of the three cues examined is primary, but the question of which attribute of voicing may most *reliably* distinguish voicing in the perception of fluent speech.

ACKNOWLEDGMENTS

This work was supported by a grant from NIH (NS-12179) to Indiana University. Both authors contributed equally to the present study. We are grateful to David Pisoni for his support and encouragement of this project and for his many helpful comments on earlier versions of this manuscript. We would also like to thank Howard Nusbaum, Robert Port, Diane Kewley-Port, and Amanda Walley for their comments, and Eileen Schwab for her statistical advice. In addition, we would like to thank Bruno H. Repp and two reviewers for their advice and suggestions.

[1]The motivation for the choice of vowels was based solely on the inherent durational differences among the three vowels reported in numerous other studies. The issue of whether the inherent durations of these vowels are primarily based on articulation or phonology was considered to be irrelevant to the present investigation.

[2]Although the choice of /t/ as one of the local phonetic environments in this experiment may have produced certain undesirable consequences for segmentation and measurement of some postvocalic closure durations, particularly for the alveolars, we chose this segment, in part, based on a previous study (Dinnsen and Charles-Luce, 1984) of Catalan that demonstrated large effects of a following voiceless stop on preceding vowel and closure durations, relative to the effects of a following vowel.

Barry, W. (1979). "Complex encoding in word-final voiced and voiceless stops," Phonetica **36**, 361–372.

Chen, M. (1970). "Vowel length variation as a function of the voicing of the consonant environment," Phonetica **22**, 129–159.

Cooper, W. E. (1975). "Syntactic control of speech timing," Ph. D. thesis (Massachusetts Institute of Technology) (unpublished).

Crystal, T. H., and House, A. S. (1982). "Segmental durations in connected speech signals: Preliminary results," J. Acoust. Soc. Am. **72**, 705–716.

Denes, P. (1955). "Effect of duration on the perception of voicing," J. Acoust. Soc. Am. **27**, 761–764.

Dinnsen, D. A., and Charles-Luce, J. (1984). "Phonological neutralization, phonetic implementation, and individual differences," J. Phonet. **12**, 49–60.

Eilers, R. E., Bull, D. H., Oller, D. K., and Lewis, D. C. (1984). "The discrimination of vowel duration by infants," J. Acoust. Soc. Am. **75**, 1213–1218.

Fitch, H. L. (1981). "Distinguishing temporal information for speaking rate from temporal information for intervocalic stop consonant voicing," Haskins Lab. Stat. Rep. Speech Res. **SR-65**, 1–32.

Hillenbrand, J., Ingrisano, D., Smith, H. L., and Flege, J. E. (1984). "Perception of the voiced–voiceless contrast in syllable-final stops," J. Acoust. Soc. Am. **76**, 18–26.

Hogan, J. T., and Rozsypal, A. J. (1980). "Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant," J. Acoust. Soc. Am. **67**, 1764–1771.

House, A. S. (1961). "On vowel duration," J. Acoust. Soc. Am. **33**, 1174–1178.

House, A. S., and Fairbanks, G. (1953). "The influence of consonantal environment upon the secondary acoustical characteristics of vowels," J. Acoust. Soc. Am. **25**, 105–113.

Klatt, D. H. (1973). "Interaction between two factors that influence vowel duration," J. Acoust. Soc. Am. **54**, 1102–1104.

Klatt, D. H. (1975). "Vowel lengthening is syntactically determined in a connected discourse," J. Phonet. **3**, 129–140.

Klatt, D. H. (1976). "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence," J. Acoust. Soc. Am. **59**, 1208–1221.

Kohler, K. J. (1979). "Dimensions in the perception of fortis and lenis plo-

sives," Phonetica **36**, 332–343.
Lisker, L. (1957). "Closure duration and the intervocalic voiced–voiceless distinction in English," Language **33**, 42–49.
Lisker, L. (1978). "*Rapid* vs. *Rabid*: A catalogue of acoustic features that may cue the distinction," Haskins Lab. Stat. Rep. Speech Res. **SR-54**, 127–132.
Lisker, L. (1981). "On generalizing the *rabid–rapid* distinction based on silent gap duration," Haskins Stat. Rep. Speech Res. **SR-65**, 251–259.
Luce, P. A., and Carrell, T. D. (1981). "Creating and editing waveforms using WAVES," Research on Speech Perception: Progress Report No. 7, Indiana University, 287–297.
Luce, P. A., and Charles-Luce, J. (1983). "Contextual effects on the consonant/vowel ratio in speech production," Research on Speech Perception: Prog. Rep. No. 9, Indiana University, 3–37.
Malécot, A. (1970). "The lenis–fortis opposition: Its physiological parameters," J. Acoust. Soc. Am. **47**, 1588–1592.
Massaro, D. W., and Cohen, M. M. (1983). "Consonant/vowel ratio: An improbable cue in speech," Percept. Psychophys. **33**, 501–505.
O'Kane, D. (1978). "Manner of vowel termination as a perceptual cue to the voicing status of postvocalic stop consonants," J. Phonet. **6**, 311–318.
Oller, D. K. (1973). "The effect of position in utterance on speech segment duration in English," J. Acoust. Soc. Am. **54**, 1235–1247.
Parker, F. (1974). "The coarticulation of vowels and stop consonants," J. Phonet. **2**, 211–221.
Peterson, G., and Lehiste, I. (1960). "Duration of syllabic nuclei in English," J. Acoust. Soc. Am. **32**, 693–703.
Port, R. F. (1978). "Effects of word-internal versus word-external tempo on the voicing boundary for medial stop closure," Haskins Lab. Stat. Rep. Speech Res. **SR-55/56**, 189–198.
Port, R. F. (1979). "Influence of tempo on stop closure duration as a cue for voicing and place," J. Phonet. **7**, 45–56.
Port, R. F. (1981a). "Linguistic timing factors in combination," J. Acoust. Soc. Am. **69**, 262–274.
Port, R. F. (1981b). "On the structure of the phonetic space with special reference to speech timing," Lingua **55**, 181–219.
Port, R. F., and Dalby, J. (1982). "Consonant/vowel ratio as a cue for voicing in English," Percept. Psychophys. **32**, 141–152.
Raphael, L. J. (1972). "Preceding vowel duration as a cue to the perception of the voicing characteristics of word-final consonants in American English," J. Acoust. Soc. Am. **51**, 1296–1303.
Raphael, L. J. (1981). "Durations and contexts as cues to word-final cognate opposition in English," Phonetica **38**, 126–147.
Repp, B. H., and Williams, D. R. (1985). "Influence of following context on perception of the voiced–voiceless in syllable-final stop consonants distinction," J. Acoust. Soc. Am. **78**, 445–457.
Umeda, N. (1975). "Vowel duration in American English," J. Acoust. Soc. Am. **58**, 434–445.
Umeda, N. (1977). "Consonant duration in American English," J. Acoust. Soc. Am. **61**, 846–858.
Wardrip-Fruin, C. (1982). "On the status of temporal cues to phonetic categories: Preceding vowel duration as a cue to voicing in final stop consonants," J. Acoust. Soc. Am. 71, 187–195.