

Phonetica

Editor: K. Kohler, Kiel
Publishers: S. Karger, Basel
Separatum (Printed in Switzerland)

Phonetica 38: 126-147 (1981)

Durations and Contexts as Cues to Word-Final Cognate Opposition in English

LAWRENCE J. RAPHAEL

Herbert H. Lehman College, CUNY, Bronx, N.Y., Haskins Laboratories, New Haven, Conn., USA

Abstract. Many perceptual and acoustic studies of word-final cognate opposition have relied mainly on isolated synthetic utterances and have usually concluded that one or another of the potential cues is primary. Recent studies of natural speech segment durations in samples of connected discourse have raised questions about the efficacy of such cues. In the experiments reported here, naturally produced tokens of the word *peg* were excised from two different contexts and acoustically analyzed. One representative token from each context was then edited in several ways in order to ascertain the cue values of a number of variables. The stimuli were tested in isolation, in their original contexts and in contexts from which the word *peck* had been deleted. The results of the experiments indicate that a number of variables can be sufficient cues to cognate opposition, and that the efficacy of a particular cue depends in a large measure both on the context from which it was extracted and on the context in which it is heard. The potentially large number of cues, their temporal distribution and their variable effectiveness suggest that future research should be focussed on discovering and describing a general articulatory strategy underlying both the production and perception of cognate oppositions.

Introduction

Various temporal features of the acoustic signal have long been suggested as perceptually significant cues to the opposition of English stop and fricative cognates in word-final position. An inspection of the acoustic and perceptual studies of these features reveals that most display one or more of the following characteristics.

First, most studies have been limited to the use of isolated monosyllables or disyllables as elicited material for acoustic analysis [HOUSE and FAIRBANKS, 1953; PETERSON and LEHISTE, 1960; HOUSE, 1961; CHEN, 1970] or as stimuli for perceptual testing [DENES, 1955; NOLL, 1960; RAPHAEL, 1972; O'KANE, 1978; HOGAN and ROZSYPAL, 1980].

A second characteristic has been the suggestion that one cue is primary. In some instances the suggestion has been qualified in terms of a specific set of stimuli [NOLL, 1960; RAPHAEL, 1972], while in others it can be taken to apply generally, either through statement or implication [PARKER, 1974; O'KANE, 1978; WARDRIP-FRUIIN, 1980]. Those cues most often granted the mantle of primacy are preceding vowel duration [RAPHAEL, 1972; O'KANE, 1978], termination rate of the preceding vowel [PARKER, 1974], and the presence versus absence of voicing during consonant closure [NOLL, 1960; WARDRIP-FRUIIN, 1980].

A third characteristic has been the general use of synthetic (or hybrid-synthetic and natural) stimuli in perceptual testing [SHOLES, 1956, 1959; DENES, 1955; RAPHAEL, 1972; RAPHAEL et al., 1975]. The use of synthetic speech in these studies was motivated by at least two considerations. First, it provided the researchers with the opportunity to isolate and to specify the acoustic characteristics of independent variables and to measure their perceptual effect (albeit at the cost of eliminating other potential cues found in the more complex signal of natural speech). Second, the technology required to edit and convert natural speech signals into equally natural-sounding test stimuli was not available to many of the researchers.

To be sure, there were, among the studies cited, exceptions to the characteristics just discussed. NOLL [1960], for instance, used natural speech stimuli, while DENES [1955] and PORT [1980] have indicated the significance of the durations of both vowel and final fricative and of the ratio between them. It has not been until relatively recently, however, that studies have appeared without the characteristics mentioned above. Within the last 5 years speech samples from relatively extended texts have been subjected to acoustic analysis [KLATT, 1975, 1976; UMEDA, 1975]. Compared to the findings for isolated utterances, the results of these analyses of running speech reveal a general reduction in the durations of acoustic segments (particularly in lesser-stressed syllables), and thus in the durational differences between such segments taken from phonologically contrastive contexts. Such findings raise serious questions about the efficacy of temporal cues to cognate opposition in general, and of the vowel duration cue in particular [KLATT, 1975].

There are, too, instances of perceptual studies which have used natural speech as the basic raw material for the preparation of test stimuli [O'KANE, 1978; WARDRIP-FRUIIN, 1980; HOGAN and ROZ-SYPAL, 1980; RAPHAEL and DORMAN, 1980]. Although none of these

dealt with stimuli derived from, or tested within, a larger context, at least one [HOGAN and ROZSYPAL, 1980] provides some evidence that a multiplicity of cues, distributed in time, may be relevant to the perception of cognate opposition in word-final position. Given the findings of the acoustical studies, and assuming that the cognate oppositions can be maintained and perceived in natural speech contexts without reference to semantic constraints, it is not unreasonable to expect that as the salience of individual temporal cues diminishes with their durations, their aggregate importance to perception increases.

But what of the contexts themselves in which a member of a potential word-final cognate opposition may appear? If we can suppose at least some (minimal) cue value for any acoustic feature which regularly accompanies the manifestation of a phonological unit [LISKER, 1975; BAILEY and SUMMERFIELD, 1980], then we ought to consider the possibility that such recurring acoustic features as may be present outside the syllable in question, but within the larger context, will carry perceptual weight.

The experiments reported here were therefore designed to (1) assess the efficacy of individual acoustic cues to word-final stop cognate perception in naturally uttered syllables isolated from their original contexts, and (2) to ascertain if the contexts in which such syllables occur contain acoustic information that is perceptually relevant to the stop cognate opposition.

Experiment I

Acoustic Analyses

The stimuli used in this experiment were drawn from 1 male speaker's reading of two paragraphs which contained several occurrences of the phrases *a peck above*, *a peg above*, *a peck shorter*, and *a peg shorter*. Each of the two paragraphs was read on two separate occasions in a manner that approximated conversational speech. The words *above* and *shorter* always occurred in clause-final position, with primary stress occurring on the appropriate syllables of each word. The degree of stress on *peg* and *peck* was considerably less than that on the stressed syllables of *above* and *shorter*. Beginning with the fifth occurrence, ten consecutive tokens of each phrase were subjected to acoustic analysis, with the restriction that a true pause (silence) had to be found after any token analyzed.

Measurements were made of the following durations in the tokens of *peg* and *peck* taken from both the prevocalic and preconsonantal contexts: (1) voice onset time (VOT) of [p] as defined by LISKER and ABRAMSON [1964]; (2) vowel duration, excluding the period of aspiration following the release of the [p]; (3) duration of the terminal formant transitions of the vowel; (4) durations of the final [k] and [g] closures; (5) duration of the voicing during the final [k] and [g] closures.

With one exception, all measurements were first derived from wide-band spectrograms produced on a Kay Sonagraph, model 6061 B. In addition, the durations of all but the

Table I. Averaged durational values and standard deviations of five acoustic measures for (1) ten tokens of *peck* and *peg* in each of two contexts and for (2) the tokens of *peg* selected as the basic experimental stimuli.

	VOT of [p]	Vowel duration	Final transition Duration	Closure duration	Closure voicing duration
<i>Context: a — above</i>					
<i>peck</i> (10 tokens)	36.3 ± 3.83	69.0 ± 5.76	24.0 ± 3.74	56.9 ± 5.30	13.2 ± 4.1
<i>peg</i> (10 tokens)	32.6 ± 3.57	93.5 ± 9.68	28.0 ± 2.45	43.4 ± 6.10	42.1 ± 5.0
<i>peg</i> (basic stimulus)	36.0	101.0	30.0	40.0	40.0
<i>Context: a — shorter</i>					
<i>peck</i> (10 tokens)	34.8 ± 3.50	95.1 ± 9.06	28.0 ± 3.32	99.3 ± 3.32	11.3 ± 3.0
<i>peg</i> (10 tokens)	32.5 ± 4.72	125.2 ± 6.93	34.0 ± 4.36	82.4 ± 6.53	73.9 ± 7.0
<i>peg</i> (basic stimulus)	30.0	125.0	35.0	86.0	77.0

terminal [e] transitions were independently measured from oscillographic displays of the digitized speech samples produced by the Haskins Laboratories pulse code modulation (PCM) system. The two sets of measurements were in close agreement. The averages of those deriving from the spectrographic measurements are presented in table I.

Measuring the durations of the VC formant transitions from the wide-band spectrograms proved to be more an act of faith than of science. The commonplace that formants in running speech are in a constant state of flux was illustrated effectively by the speech samples that were subjected to analysis. Repeated measurements, including those derived from spectral sections, indicated that attempts to specify the point in time at which the VC formant transitions began were not highly reliable. Estimates of transition duration for a single token differed by as much as 34%. It thus seemed prudent to resort to a different technique, one that would at least provide auditory estimates of the effective formant transition durations.

Accordingly, a series of stimuli was prepared in which, after every utterance of *peck* and *peg* had been isolated and the closure and release of the final stop deleted from the acoustic signal, the remaining vocalic portion of each token was cut back in 5-ms steps from the termination of the vowel to its center. The stimuli prepared from each token were randomized in separate listening tests and presented via headsets to 3 trained phoneticians. The listeners were asked to say if each stimulus ended with a consonant or a vowel, and to identify the consonant if they heard one. Several tokens of an utterance of [p^he] were included in each test. It was assumed that this sort of test would reveal the effective transitional duration containing the perceptual cues for /k/ and /g/. The data for the transitions, reported in table I, were derived from the averaged durations at which the listeners first reported hearing either a CV syllable or a syllable ending in a glottal stop, rather than a CVC syllable.

The Stimuli

Following the acoustic analyses, two tokens of *peg*, one from each context, were selected as the bases for the stimuli of the principal perceptual tests of the first experiment. These tokens and their derivatives will be differentiated hereafter according to the following convention: Stimuli derived from the prevocalic context (*a — above*) will be denoted as stimulus/V, *peg*/V or *peck*/V; stimuli derived from the preconsonantal context (*a — shorter*) will be denoted as stimulus/C, *peg*/C or *peck*/C.

The criterion for selecting the basic tokens was that every potential variable has a value within 1 standard deviation (SD) of its mean in the acoustic analysis. The values of the potential variables of each of the tokens selected are shown in table I.

Table II. Durational values (ms) of the variables for each stimulus-type derived from tokens of *peg* originally occurring in two different contexts (values of the operant variable[s] for each stimulus type are italicized)

Stimulus type	VOT of [p]	Vowel duration	Final transition duration	Closure duration	Closure voicing	Release type and duration
<i>Context: a peg above</i>						
Unedited <i>peg</i>	36	101	30	40	40	[g] 12
Shortened vowel	36	72	30	40	40	[g] 12
Transition substitution	36	99	25	40	40	[g] 12
Closure voicing 1	36	101	30	40	0	[g] 12
Closure voicing 2	36	101	30	40	13	[g] 12
Closure duration	36	101	30	61	61	[g] 12
Release-burst 1	36	101	30	40	40	— 0
Release-burst 2	36	101	30	40	40	[k] 13
Shortened vowel/ transition substitution	36	70	25	40	40	[g] 12
Closure duration/ closure voicing 1	36	101	30	61	0	[g] 12
Closure duration/ closure voicing 2	36	101	30	61	14	[g] 12
<i>Context: a peg shorter</i>						
Unedited <i>peg</i>	30	125	35	86	77	[g] 15
Shortened vowel	30	93	35	86	77	[g] 15
Transition substitution	30	123	30	86	77	[g] 15
Closure voicing 1	30	125	35	86	0	[g] 15
Closure voicing 2	30	125	35	86	11	[g] 15
Closure duration	30	125	35	101	91	[g] 15
Release-burst 1	30	125	35	86	77	— 0
Release-burst 2	30	125	35	86	77	[k] 10
Shortened vowel/ transition substitution	30	90	30	86	77	[g] 15
Closure duration/ closure voicing 1	30	125	35	101	0	[g] 15
Closure Duration/ closure voicing 2	30	125	35	101	11	[g] 15

The two tokens of *peg* were then subjected to a number of editing operations which manipulated both single and combinations of two variables. All the editing tasks were performed using the Haskins Laboratories PCM system. The reader is referred to table II which presents the values of the variables for each stimulus type. Let us first consider the single-variable editing tasks, more or less following them through the syllable from left to right.

Vowel Shortening. The vowel of each of the two tokens of *peg* was shortened to approximate the average duration of the vowel in *peck* taken from the same context. Thus, 29 ms of the central portion of the vowel of *peg/V* and 35 ms of the vowel of *peg/C* were deleted. Since the cuts into, and subsequent rejoining of, the vowel had to be made at zero-crossings in order to avoid transients and other audible discontinuities, it was not possible to match exactly the durations of the shortened vowels (72 and 93 ms) to those of the average vowels in *peck* (69 and 95 ms). The durations of the vowels in the stimuli are, however, within 1 SD of the mean vowel durations found in the acoustic analyses of *peck*.

Transition Substitution. The final transitions of the vowels of *peg/V* and *peg/C* were deleted and replaced with those from tokens of the vowels of *peck/V* and *peck/C*, respectively. We should recall here that the durations derived for the transitions were based on auditory judgments, rather than on physical measurements. Because of this, and since the actual durations were not of critical importance so long as the entire transitional period was removed and replaced with that of *peck*, the two pitch periods preceding the estimated beginnings of the transitions were also deleted from the tokens of *peg*. In a like manner, two additional pitch periods preceding the VC transitions of the tokens of *peck* were excised with the transitions and inserted into the contextually appropriate tokens of *peg*. This resulted in a slight shortening of the vocalic portion of the stimuli because of differences between both the pitch period durations and the transition durations of *peck* and *peg* (table I).

Closure Voicing 1 and 2. Two versions of the closure voicing stimuli were prepared. In version 1 the entire closure period of the [g] in both tokens of *peg* was deleted and replaced with an equal duration of silence. In version 2 the closure periods of the [k]s from typical tokens of *peck* (which contained some closure voicing) replaced the [g] closures of the original tokens of *peg*. The durations of the [k] closures were shortened by 21 ms (*peck/V*) and 15 ms (*peck/C*) so that they were identical with those of the [g] closures they replaced.

Closure Duration. The durations of the [g] closures in the tokens of *peg* were increased to equal typical [k] closures of *peck*. In the case of *peg/V*, this was accomplished by reiterating pitch periods at the middle of the [g] closure so that it remained fully voiced. The durations of the pitch periods were such that the ultimate duration of the stimulus/V closure exceeded the mean closure duration for *peck* in the prevocalic context, but by less than 1 SD. The reiteration of pitch periods from the middle of the closure of *peg/C* produced a stimulus that had a closure duration within 2 ms of the mean preconsonantal [k] closure. The stimulus closure was voiced throughout 89.5% of its duration. This was quite close (within 1%) of the average duration of closure voicing found in the acoustic analyses for the ten tokens of *peg* spoken in the preconsonantal context.

The preparation of these closure duration stimuli did, in fact, necessitate the manipulation of two variables because of the (partial) coincidence of closure with voicing. I have, however, included the stimuli among the single-variable editing tasks because of the attempt to preserve the proportion of voicing found in typical [g] closures.

Release-Burst 1 and 2. Two versions of the release-burst stimuli were prepared. In version 1 the [g] bursts were deleted from the acoustic signal. In version 2 the release-bursts of the [g]s in the *peg* stimuli were replaced with those of [k]s from tokens of *peck*. The durations of the deleted [g] bursts were 12 ms (*peg/V*) and 15 ms (*peg/C*). Oscillographic traces revealed the presence of very low-amplitude periodicity in the deleted bursts. The durations of the [k] bursts used as replacements were 13 ms (*peck/V*) and 10 ms (*peck/C*). The [k] bursts were wholly aperiodic.

Let us now turn our attention to the stimuli in which two variables were manipulated.

Vowel Shortening/Transition Substitution. The preparation of these stimuli essentially combined the editing tasks described above for each of the variables individually. That is, the VC transitions from *peck* were substituted for those of *peg*, and the duration of the vowel was decreased to that of a typical vowel of *peck*.

Closure Duration/Closure Voicing 1 and 2. Two versions of this stimulus type were prepared. In both versions the closure of the [g] in *peg* was replaced by a typical [k] closure from *peck*. In version 1 any voicing present in the [k] closure was deleted and replaced by silence. In version 2 the voicing present in the [k] closure was left intact. The durations of voicing in the [k] closures that were placed into the stimuli were within 1 SD of the means found in the acoustic analyses for tokens of *peck* in same context.

The Perceptual Tests

Two perceptual tests were fashioned from the isolated stimuli just described, one for the *peg/V* stimuli and one for the *peg/C* stimuli. Five tokens of each stimulus and ten of the original, unedited utterances of *peg* were randomized in blocks of ten in each test. The inter-stimulus interval was 3 s with 10 s between blocks. There were two groups of subjects. One consisted of 5 researchers at Haskins Laboratories who had had considerable experience as subjects in perceptual tests. The other group was composed of 5 undergraduate students at Herbert H. Lehman College who had had no previous experience with tests of this type. All of the subjects had normal hearing and were free of speech disorders. Both subject groups heard the test tapes over headsets in sound-treated rooms at a comfortable listening level. They were asked to respond to each stimulus by writing either *peck* or *peg* on their response sheets.

Results and Discussions

Since no systematic differences were found between the two groups of subjects, the data for both groups have been pooled. Figure 1 displays the average and range of percentages at which the original and edited *peg* stimuli were identified as *peck*. We shall consider one stimulus type at a time, comparing the results across original contexts. The variables are characterized below in terms of their effectiveness as cues according to the percentage of *peck* responses they elicited. 'Weak' cues elicited 0-33% *peck* responses; 'moderate' cues 34-66% *peck* responses, and 'strong' cues 67-100% *peck* responses.

Unedited peg. The stimuli from both contexts were identified as *peg* more than 96% of the time. The few *peck* responses were distributed randomly across subjects.

Shortened Vowel. On average, this variable proved to be a moderately effective cue for both stimuli. We must note, however, that the range of responses was extensive. 2 subjects (1 from each group) reacted to the vowel duration in both stimuli as a strong cue; 2 others responded

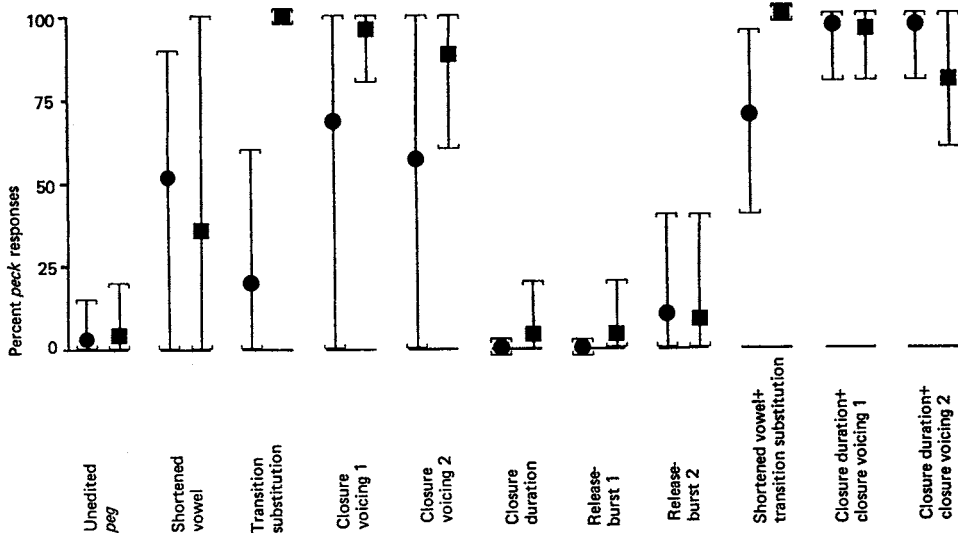


Fig. 1 Average percent and range of *peck* responses to each stimulus type heard in isolation. ● = *peg* (a — *peg* above); ■ = *peg* (a — *peg* shorter).

to it as a weak cue, providing no *peck* responses at all. The responses of the balance of the subjects were in the moderate range of cue effectiveness. It is clear from these results that the shortened vowels drawn from natural speech contexts are not sufficient cues to /k/ for most of the subjects.

Transition Substitution. We find here a striking contrast between the subjects' responses to stimuli drawn from different contexts. The formant transitions in the *peg/V* stimulus are a weak cue on average, although some subjects responded to them as a moderately effective cue. The transitions in the *peg/C* stimulus, however, are maximally effective as a cue to /k/: Each subject always identified this stimulus as *peck*. It is unlikely that the transition durations can, in themselves, underly much of the difference between the results for the two stimuli. In the first place, the shorter of the two substituted transitions occurs in the stimulus/V, which is most often judged as ending in /g/. Further, a purely durational explanation runs afoul of the results of comparing the durations of the original and substituted transitions. This comparison leaves us to explain why the 5-ms difference between transitions

35 and 30 ms in duration (stimulus/C) provides a cue that is 100% effective, whereas the same difference between transitions 30 and 25 ms in duration (stimulus/V) provides only a weak cue to /k/. Even if such an explanation were available, it would presuppose that listeners are capable of perceptually isolating transitions from the formants with which they are contiguous. Recent evidence (RAPHAEL et al., 1980] suggests, rather, that the transitions are perceptually part of the total formant duration that cues vowel length. Since the vocalic portions of the original stimuli were each shortened by only 2 ms in the editing process, it seems unlikely that the transitions, whatever their strength as cues, can be classified as durational cues per se. It may, perhaps, be more useful to look at the spectral characteristics of the vowel formants and transitions to see if they are differentially affected by the larger contexts from which they were taken.

Closure Voicing 1. In both stimuli the absence of closure voicing was, on average, a strong cue for /k/. For the stimulus/V, however, the responses ranged from 0 to 100%. 2 subjects responded to the absence of closure voicing as a weak cue, and there was a third for whom it was only moderately effective. In contrast, for the stimulus/C, all the responses were in the strong-cue range.

Closure Voicing 2. The responses to these stimuli are much like those to the closure voicing 1 stimuli, although the averages of /k/ responses are slightly lower: The average for the stimulus/V is in the moderate range of cue effectiveness, and the range of responses to the stimulus/C here extends into the moderate range of cue effectiveness. The differences are, of course, most reasonably explained by the presence of closure voicing in the closure voicing 2 stimuli. Given the relatively brief duration of this closure voicing (13 and 11 ms), the differences do suggest a considerable degree of sensitivity to the cue on the part of the listeners.

Closure Duration. This was a weak cue for both stimuli. The increase of closure duration, here accompanied by a proportional increase in closure voicing, elicited no *peck* responses to the stimulus/V and very few *peck* responses to the stimulus/C. It is interesting to note that extending the [g] closures to durations appropriate to [k] closures produced little or no perceptual change so long as the quantity of voic-

ing during the closure was proportionately appropriate to that found in productions of [g].

Release-Burst 1. The absence of a release-burst was a weak cue, eliciting no *peck* responses to the stimulus/V and very few *peck* responses to the stimulus/C.

Release-Burst 2. The substitution of [k] for [g] releases provided a weak cue to /k/. The few *peck* responses found for both stimuli were accounted for entirely by 2 subjects, and even for those subjects the cue was only moderately effective. For these stimuli, then, the presence versus absence of a release-burst appears to carry almost no information about the /k/-/g/ contrast, whereas the presence of a [k] release-burst is at best a weakly effective cue to /k/ for the subjects as a group.

Shortened Vowel/Transition Substitution. Combining the vowel duration and transition cues elicited /k/ responses at the lower end of the strong range (70%) for the stimulus/V. We might expect this result on the assumption that the individual effects of vowel shortening and transition substitution are additive when the cues are combined. The stimulus/C was judged as *peck* 100% of the time, as we would predict from the effect of the transition cue acting alone.

Closure Duration/Closure Voicing 1. This combination of variables, in which closure voicing was absent, elicited an average of almost 100% *peck* responses to both stimuli. The range of responses lies entirely within the strong-cue category. The small number of *peg* responses was provided by 2 subjects.

Closure Duration/Closure Voicing 2. The combination of these two variables, in which the original durations of [k]-closure voicing were present, proved to be a strong perceptual cue. The range of responses is somewhat greater for stimulus/C than for stimulus/V, but the number of *peg* responses is quite small and limited to 2 subjects. The results for combining the closure duration and closure voicing variables suggest the predominance of the latter cue. That is, the most salient feature of these stimuli seems to be the absence or severe limitation of the duration of the voicing during stop consonant closure, since closures of the same duration, when voiced throughout, elicit little or no *peck*

responses. We can, however, discern some effect of increasing closure duration. This effect is manifested in both the higher average and constricted range of *peck* responses to stimulus/V, when compared with the results for the uncombined closure voicing 1 and 2 variables. It appears that eliminating all or most of the voicing from the closure, while increasing closure duration, is more effective in cueing *peck* than simply limiting the duration of closure voicing.

Experiment II

The Stimuli

The second experiment was designed to assess the effects of the variables when they occur in natural speech contexts of several syllables. Accordingly, each stimulus used in experiment I was first replaced in its original context (*a* — *above* or *a* — *shorter*). In addition, a typical token of *peck* was excised from each of the contexts and replaced by each of the contextually appropriate stimuli of the first experiment. That is, e.g., the *peg/V* stimuli were substituted for a token of *peck* deleted from a *peck above*. The criterion described above for selecting tokens of *peg* as basic stimuli was applied to the selection of the tokens of *peck* that were excised. Finally, the transitions following the [g] and [k] releases in the two contextual frames of *a* — *above* were deleted and each of the stimuli/V were placed into them. Thus the *peg/V* stimuli appeared in four contexts: the original prevocalic context, both with and without the transitions following [g] release, and the prevocalic context from which *peck* had been deleted, both with and without the transitions following [k] release. The *peg/C* stimuli, on the other hand, appeared in only two contexts: the original preconsonantal context and a preconsonantal context from which *peck* had been deleted.

The editing of the contexts was performed on the Haskins Laboratories PCM system. The removal of the transitions following the [k] and [g] releases was accomplished by isolating the word *above* immediately after the preceding stop burst. This isolated segment, which sounded like /gəbav/ whether it had originally followed a [g] or [k] [LISKER, 1978a], was then cut back from the beginning in 5-ms steps until the initial [g] was no longer heard. The segment was then joined directly to the preceding [g] or [k] burst.

The Perceptual Tests

Two separate listening tests were prepared, one for each of the two contexts in which the stimuli had been placed. Five tokens of each stimulus occurred in each test. The stimuli were presented in random order in blocks of ten, with 3 s between stimuli and 10 s between blocks. Subjects, listening conditions, and response tasks were identical with those of experiment I.

Results and Discussion

The results for each context are presented separately in figures 2 and 3. The results for the isolated stimuli of the first experiment are repeated in these figures for the purpose of comparison. For each variable we will consider first the results for the stimulus/V and then those for the stimulus/C, comparing both with the results for the stimuli heard in isolation.

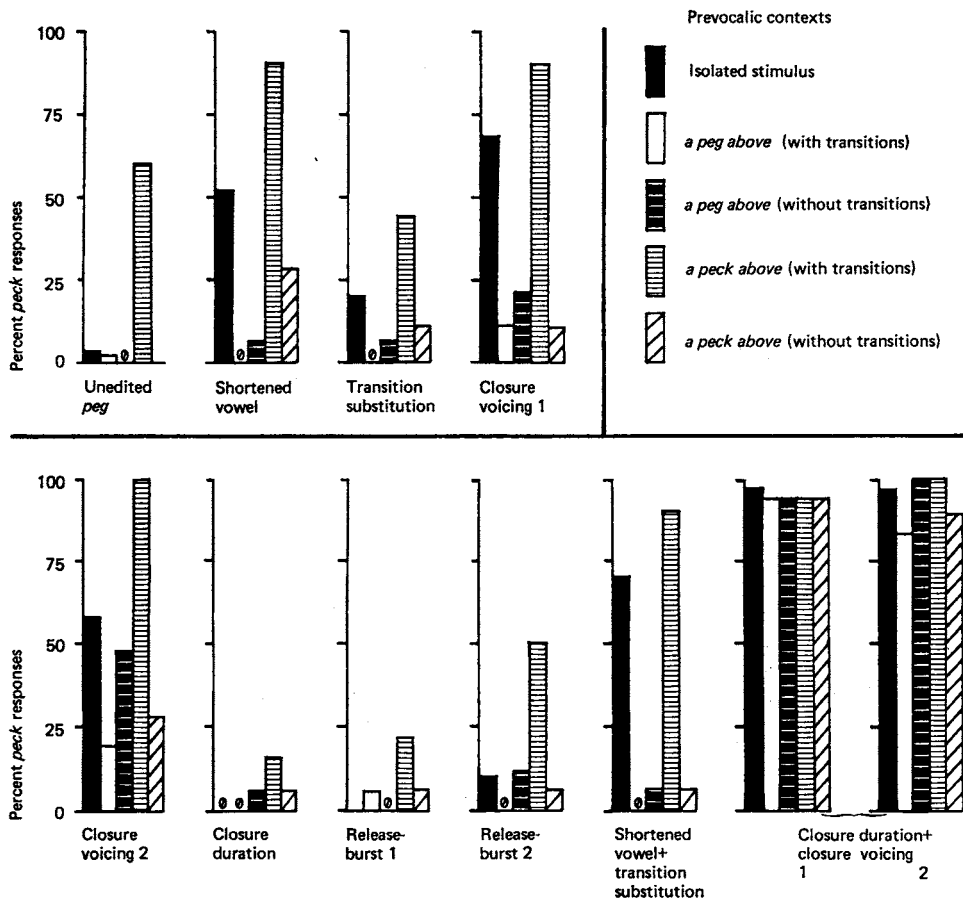


Fig. 2. Average percent of *peck* responses to each *peg*/V stimulus type heard in isolation and in prevocalic contexts with and without transitions following stop-release.

Unedited peg. The unedited stimulus/V was heard as *peg* almost 100% of the time in all contexts except for the *peck* context with the transitions intact. In that context listeners reported hearing *peck* 60% of the time, on average. 8 of the 10 subjects reported hearing *peck* more than 50% of the time, although 1 subject reported hearing *peg* 100% of the time. This result suggests that context can be of considerable importance in providing cues to the velar stop cognate opposition in final position. Moreover, the portion of the context that seems to be

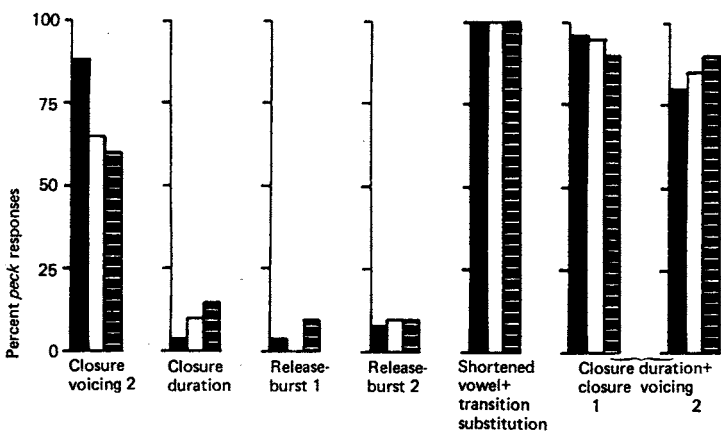
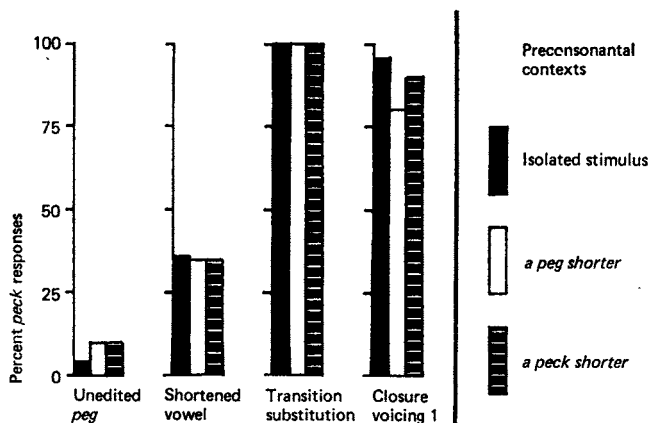


Fig. 3. Average percent of *peck* responses to each *peg/C* stimulus type heard in preconsontal contexts.

operative here is the segment containing the transitions following the release of the final consonant. In the edited context, where those transitions were deleted, we observed that subjects always identified the stimulus as *peg*. Further, although the transitions following both [g] and [k] releases cue /g/ percepts at the beginning of the isolated word *above*, listeners can generally distinguish them in terms of their effect on consonant classification when they are heard in context following the unedited tokens of *peg*.

The unedited *peg*/C stimulus was identified as *peg* 90% of the time in both contexts. We observe that both the preconsonantal *peck* context which occurs naturally without vocalic transitions following [k] release, and the prevocalic *peck* context from which the transitions had been deleted, scarcely affect perception of the unedited *peg* stimuli compared to their identification in isolation.

Shortened Vowel. In the prevocalic *peck* context with transitions the *peg*/V stimulus elicited an average of 90% *peck* responses. In the same context without transitions the stimulus elicited less than 30% *peck* responses on average, although 2 subjects reported hearing *peck* at the 60% level. The moderate cue effectiveness found for the shortened vowel in the isolated stimulus was totally overridden in both prevocalic *peg* contexts where *peck* responses averaged 6% or less.

Identification of the stimulus/C in both contexts was about the same as it was in isolation. The moderate cue value to /k/ provided by the vowel duration is thus not enhanced by its inclusion in the *peck* context nor reduced by its inclusion in the *peg* context. This result, then, differs from that for the stimulus/V in the various contexts in which it was heard.

Transition Substitution. The weak cue for /k/ provided by this variable in the isolated stimulus/V becomes moderately effective in the prevocalic *peck* context with transitions. In all other contexts it is either weakly effective or totally ineffective in eliciting *peck* responses.

The stimulus/C was always identified as *peck* in both contexts as it was in isolation. We note specifically that its presence in the preconsonantal *peg* context had no effect whatsoever on its identification as *peck*. Once again we find results which differ from those for the stimulus/V as heard in various contexts.

Closure Voicing 1. The average *peck* response of 90% for the stimulus/V in the prevocalic *peck* context with transitions is more than 20% greater than for the stimulus heard in isolation. That the [k] transitions are responsible for this result is indicated by the fact that in the transitionless prevocalic *peck* context the average *peck* response was 10%. Both prevocalic *peg* contexts largely overrode the strong cue value of the closure voicing found for the isolated stimulus/V.

The average identification of the stimulus/C in isolation as *peck* is little affected by its inclusion in either context. This result again differs from that for stimulus/V, the identification of which was markedly affected by every context into which it was placed.

Closure Voicing 2. The results for the stimulus/V are quite similar to those found for closure voicing 1, although the slightly higher average of *peck* responses in all contexts is unexpected, given the fact that this stimulus includes 13 ms of closure voicing. The stimulus/C is identified somewhat less frequently on average as *peck* when heard in either context than when heard in isolation. This result is attributable to 3 subjects who provided relatively few *peck* responses. The responses for both contexts span the full range from 0 to 100%. Once again we observe that the identification of the isolated stimulus/V was more markedly affected by being placed in both prevocalic *peg* and *peck* contexts, especially those with transitions intact.

Closure Duration. The identification of the stimulus/V as *peg* was virtually unaffected by any of the contexts. This result reinforces the conclusion that the cue value of closure duration is dependent on the duration of voicing contained in the closure. The identification of the stimulus/C as *peg* in isolation was very little affected by either of the two contexts in which it was heard.

Release-Burst 1. The identification of the stimulus/V as *peg* was little affected by its inclusion in any context. The maximum of *peck* responses (20%) occurs in the prevocalic *peck* context with transitions. The identification of the isolated stimulus/C was not changed to an appreciable degree by its inclusion in either context.

Release-Burst 2. The substitution of the [k] for the [g] burst in the stimulus/V is a moderately effective cue to /k/ only in the prevocalic *peck* context with transitions intact. In all other contexts the cue is weak, as it was for the stimulus in isolation.

The identification of the isolated stimulus/C as *peg* was essentially unchanged by its inclusion in either context. Once more we note the contrast with the results for the stimulus/V which increased from 10% average *peck* responses in isolation to 50% in the *peck* context with transitions.

Shortened Vowel/Transition Substitution. The additive effect of these variables, noted for the isolated stimulus/V, is evident only in the prevocalic *peck* context with transitions. The transitionless prevocalic *peck* context and the two *peg* contexts render the combined cue value of vowel duration and transition type virtually ineffective.

The stimulus/C was always identified as *peck* in both contexts. Given the results reported above for the variables individually, both in isolation and in context, this outcome is not surprising. We do, however, note that in both *peg* contexts the stimulus/V was almost never identified as *peck*, even though the combined variable was a strong cue to /k/ in the isolated stimulus.

Closure Duration/Closure Voicing 1. The stimulus/V was identified as *peck* 100% of the time in all contexts, as it was in isolation. We can see that the absence of voicing during the closure allows the closure duration to attain perceptual importance if we compare these results to those for closure voicing 1 where the stimulus/V with the shorter, but entirely voiceless closure is identified largely as *peg*, except in the prevocalic *peck* context with transitions. The stimulus/C was almost always identified as *peck* in both contexts, as it was in isolation. This result is virtually identical with that for the stimulus/V.

Closure Duration/Closure Voicing 2. The results for the stimulus/V are almost identical with those for closure duration/closure voicing 1, and suggest the same conclusion concerning the cue value of closure duration in the absence of closure voicing.

The 80% average rate of identification of the stimulus/C as *peck* in isolation is slightly increased in both contexts. In general, the results are much like those for the stimulus/V, with high rates of *peck* identification both in context and in isolation.

General Discussion and Conclusions

Any conclusions drawn from the results of this study must be carefully qualified in terms of its limitations. It is important to bear in mind that only a single speaker's utterances were analyzed and used as the bases for the test stimuli. Further, those utterances occurred in only two of several possible general (and of many more possible specific)

segmental contextual types. We should also remember that only one of several possible stop, fricative and affricate cognate positions has been examined. Finally, we recall that several types of prosodic and spectral features, such as fundamental frequency and formant frequency, have not been investigated here.

Bearing in mind, then, that these conclusions are based solely on the results for the stimuli used in these experiments, we may conclude first that, in syllables isolated from context, a number of variables may each provide a sufficient cue to the word-final /k/-/g/ opposition. If, for the moment, we take the measure of sufficiency to be the 'strong'-cue label as it has been used in this paper, then we can identify some cues which are generally sufficient for a particular stimulus and others which are idiosyncratically sufficient.

An example of a generally sufficient cue is the total absence of closure voicing in the stimulus/C (fig. 1). An example of an idiosyncratically sufficient cue is vowel duration, which elicited 100% *peck* responses from 2 subjects for the shortened vowel in both stimulus/V and stimulus/C.

We must emphasize here that even the description of a cue as 'generally' sufficient has to be qualified in terms of the particular token being used as a test stimulus. In this regard we note that there is only one uncombined variable that proved to be a generally sufficient cue in both isolated stimuli: closure voicing 1. It will be recalled, however, that in those stimuli all the closure voicing had been deleted and replaced by silence, and that no totally silent (voiceless) closures were found in any of the tokens of *peck* from either context in the acoustic analyses. In addition, there were subjects who never reported hearing *peck* in response to the stimulus/V of the closure voicing 1 variable. Perhaps the need to qualify (general) sufficiency in terms of a specific test stimulus is best illustrated by the transition substitution variable. The final transitions provided only a weak cue in the isolated stimulus/V, whereas the isolated stimulus/C was always heard as *peck*.

Qualifications notwithstanding, the data do indicate that closure voicing is a more consistently effective cue than others tested in these experiments. Both by itself, and in combination with the closure duration variable in isolated syllables, the closure voicing stimuli elicit the highest average percentage of *peck* responses. The effects of the prevocalic *peg* context on the potency of the closure voicing cue are, however, extensive. It remains to be seen how contexts other than those tested here affect the cue.

A second conclusion is that every variable tested had some cue value when heard in the isolated syllable and/or in context. This was true even of the release-burst and closure duration variables, although the amount of information they provided about the /k/-/g/ opposition was small. The results for the release-bursts are generally in accord with those of WANG [1959] and MALÉCOT [1958]. Closure duration, however, does carry somewhat more perceptual weight than do the bursts. We have noted earlier that the closure duration effects are dependent on the extent of closure voicing. This result is similar to that of LISKER [1978b], who noted that for word pairs such as *rapid-rabid* more than 80 ms of closure voicing precluded /p/ judgments (for a stimulus derived from *rabid*) even when closure duration was appropriate to /p/. LISKER's data also indicate that silent closures elicit more /p/ judgments as closure duration is increased. Indeed, the poststress intervocalic environment of the word-final stops in *a peck/peg above* is quite similar to that of the medial stops in *rapid* and *rabid*. There is less similarity, however, between the interstress, preconsonantal environment of *a peck/peg shorter* and the poststress environment of the stops in *rapid/rabid*. Nevertheless, the stimuli derived from the interstress preconsonantal environment also elicited more *peck* judgments when the silent (or near-silent) closure duration was appropriate to /k/ than when it was appropriate to /g/.

A final conclusion we may draw is that context may play an important role in the perceptual differentiation of /k/ and /g/ in word-final position. We should observe first that the unaltered prevocalic environment exerted a far greater influence on listeners' perceptions than did the unaltered prefricative environment. That is, the percentage of *peck* identifications for isolated syllables was affected far more radically when stimuli were heard in the context *a _____ above* than in the context *a _____ shorter*. The critical factor in the prevocalic context appears to be contained in the transitions following the release of the word-final stops. Such transitions are, of course, lacking in the prefricative context.

A particularly striking example of the effect of the [k] transitions is the unedited *peg/V* stimulus which elicited an average of 3% *peck* responses when heard in isolation and an average of 61% *peck* responses when heard in the prevocalic *peck* context with transitions intact. Indeed, 4 subjects always reported hearing *peck* in this context. When, however, the transitions were not present in the context, listeners never

reported hearing *peck* for the stimulus. Similar effects can be seen for several stimulus types, especially shortened vowel, closure voicing 1 and 2, release-burst 2 and shortened vowel/transition substitution.

The complementary effect of [g] transitions can also be observed. That is, the average of *peck* responses to any isolated stimulus/V (when it was above zero) always declined when that stimulus was placed into the *peg*/V context with transitions intact. The decline in *peck* responses was precipitous for some stimulus types such as shortened vowel (52–0.0%), closure voicing 1 (68–12%), and shortened vowel/transition substitution (70–0.0%). For other stimulus types the change was small, either because there were few *peck* identifications of the isolated stimulus (transition substitution: 20–0.0%) or because of the resistance to change of a particularly strong combination of cues (closure duration/closure voicing 1/2: 97–94%/97–83%).

Finally, for every stimulus type except release-burst 1 and closure duration/closure voicing 1, an edited version of *peg*/V, heard in the same context from which the [g] transitions were deleted, displayed a higher percentage of *peck* responses than when heard in the context with the transitions intact. These increases were generally quite modest (10% or less), although in the closure voicing 2 stimuli *peck* responses increased from an average of 17 to an average of 44%.

That the transitions should carry considerable perceptual weight for the /k/-/g/ opposition in these stimuli is not surprising, since their presence in the acoustic signal is a natural and unavoidable consequence of stop release in the prevocalic context. The sort of information carried by the transitions depends largely on the state of the laryngeal mechanism at the moment of stop release. When voicing is maintained throughout stop closure and into the stop release, as it was in the case of the token of *a peg above* used in the experiment, the transition to the following vowel will be fully voiced. When the closure and release of the stop are voiceless, the initial portion of the transitions to the following vowel will also be voiceless, as it was in the token of *a peck above* used in the experiment. More specifically, a short period of aspiration, accompanied by an attenuation of the first formant, initiated the vowel following the release of the [k]. These descriptions are, of course, much the same as those we might expect for word-initial stops that differ with regard to VOT. In the case of the word-final stops in context, however, the differences are far smaller than those

that might be found for prestressed initial stops in isolated syllables. They are small enough, in fact, that, when the word *above* is isolated from either context, it is heard as beginning with a [g], as we have noted earlier. Nevertheless, when heard in context, the presence or absence of (1) aspiration in the transitions and (2) suppression of the first formant suffice to cause the perceptual effects that have been described here.

In conclusion we should observe that the cues investigated here are arrayed variously in time throughout most of the duration of the stimulus syllables and the initial portion of the following syllables. Further, they are of diverse acoustic characteristics. These observations, taken in conjunction with the general but extremely uneven effectiveness of the variables as cues, recall the results of several recent studies which have specified multiple, temporally distributed cues to phonetic distinctions [REPP et al., 1978; DORMAN and RAPHAEL, 1980; DORMAN et al., 1980; BAILEY and SUMMERFIELD, 1980]. The weight of evidence from these studies suggests that the job of specifying the full complement of acoustic events which serve as perceptual cues to a phonetic category is as close to mounting a treadmill to oblivion as any researchers would care to find themselves. Nor do the tasks of ranking the cues in order of importance or singling out one or a subset of them as 'primary' recommend themselves as potentially fruitful labors, given the variable effects of context on the perceptual weight of individual cues.

A more productive goal for future research might be to discover and describe an articulatory explanation for the presence of any acoustic cue which is demonstrably effective in any context. Such an explanation should be the answer to the question about what a listener induces from the various, disparate and temporally distributed cues in order to make a phonetic judgment. It seems most likely that the articulatory explanation must temporally relate changes in the state of the laryngeal mechanism to the activity of the upper articulators. Because of the vagaries of consonant production in word-final position, especially for stop consonants, this temporal relationship may be somewhat complex in comparison, for instance, to that which defines VOT in initial stops. Once specified, however, it should, besides explaining the presence of acoustic cues both known and as yet undiscovered, tell us more about whether there is a general articulatory strategy which underlies both the production and perception of cognate oppositions.

References

- BAILEY, P.J.; SUMMERFIELD, Q.: Information in speech: observations on the perception of [s]-stop clusters. *J. exp. Psychol. hum. Perception Performance* 6: 536-563 (1980).
- CHEN, M.: Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22: 129-159 (1970).
- DENES, P.: Effect of duration on the perception of voicing. *J. acoust. Soc. Am.* 27: 761-764 (1955).
- DORMAN, M.F.; RAPHAEL, L.J.: Distribution of acoustic cues for stop consonant place of articulation in VCV syllables. *J. acoust. Soc. Am.* 67: 1333-1335 (1980).
- DORMAN, M.F.; RAPHAEL, L.J.; ISENBERG, D.: Acoustic cues for a fricative-affricate contrast in word-final position. *J. Phonet.* (1980).
- HOGAN, J.T.; ROZSYPAL, A.J.: Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *J. acoust. Soc. Am.* 67: 1764-1771 (1980).
- HOUSE, A.S.: On vowel duration in English. *J. acoust. Soc. Am.* 33: 1174-1178 (1961).
- HOUSE, A.S.; FAIRBANKS, G.: The influence of consonant environment upon the secondary acoustic characteristics of vowels. *J. acoust. Soc. Am.* 25: 105-113 (1953).
- KLATT, D.H.: Vowel lengthening is syntactically determined in a connected discourse. *J. Phonet.* 3: 129-140 (1975).
- KLATT, D.H.: Linguistic uses of segmented duration in English: acoustic and perceptual evidence. *J. acoust. Soc. Am.* 59: 1208-1221 (1976).
- LISKER, L.: Is it VOT or a first-formant transition detector? *J. acoust. Soc. Am.* 57: 1547-1551 (1975).
- LISKER, L.: Segment duration, voicing and the syllable; in BELL, HOOPER *Syllables and segments*, pp. 133-140 (North-Holland, Amsterdam 1978a).
- LISKER, L.: On buzzing the English /b/. *Haskins Lab. Status Rep. Speech Res.*, SR 55/56, pp. 181-188 (Haskins Laboratories, New Haven 1978b).
- LISKER, L.; ABRAMSON, A.S.: A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20: 384-422 (1964).
- MALÉCOT, A.: The role of releases in the identification of released final stops. *Language* 34: 370-380 (1958).
- NOLL, J.D.: The perceptual significance of certain acoustical correlates of consonant voicing contrasts; PhD thesis, University of Iowa (unpublished, 1960).
- O'KANE, D.: Manner of vowel termination as a perceptual cue to the voicing of post-vocalic stop consonants. *J. Phonet.* 6: 311-318 (1978).
- PARKER, F.: The coarticulation of vowels and stop consonants. *J. Phonet.* 2: 211-221 (1974).
- PETERSON, G.; LEHISTE, I.: Duration of syllabic nuclei in English. *J. acoust. Soc. Am.* 32: 693-703 (1960).
- PORT, R.F.: Linguistic timing factors in combination; in *Res. Phonet.*, Report No. 1, pp. 33-97 (Department of Linguistics, Indiana University, Bloomington, Ind. 1980).
- RAPHAEL, L.J.: Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *J. acoust. Soc. Am.* 51: 1296-1303 (1972).
- RAPHAEL, L.J.; DORMAN, M.F.: The contribution of CV transition duration to the perception of final consonant voicing in natural speech. *J. acoust. Soc. Am.* 67: Suppl. 1, p. 51 (1980).
- RAPHAEL, L.J.; DORMAN, M.F.; FREEMAN, F.; TOBIN, C.: Vowel and nasal duration as cues to voicing in word-final stop consonants: spectrographic and perceptual studies. *J. Speech Hear. Res.* 18: 389-400 (1975).

- RAPHAEL, L.J.; DORMAN, M.F.; LIBERMAN, A.M.: On defining the duration that cues voicing in final position. *Lang. Speech* 23: 297-307 (1980).
- REPP, B.; LIBERMAN, A.M.; ECCARDT, T.; PESETSKY, D.: Perceptual integration of acoustic cues for stop, fricative and affricate manner. *J. exp. Psychol. hum. Perception Performance* 4: 621-637 (1978).
- SHOLES, G.: Stop consonants in final position. *Haskins Lab. Q. Progr. Rep.* 21: Appendix 5 (1956).
- SHOLES, G.: Acoustic cues for final /s/ and /z/. *Haskins Lab. q. Progr. Rep.* 31: Appendix 3 (1959).
- UMEDA, N.: Vowel duration in American English. *J. acoust. Soc. Am.* 58: 434-445 (1975).
- WANG, W.S.-Y.: Transition and release as perceptual cues for final plosives. *J. Speech Hear. Res.* 2: 66-73 (1959).
- WARDRIP-FRUIIN, C.A.: On the status of temporal cues to phonetic categories: preceding vowel duration as a cue to voicing in final stops; PhD thesis Stanford University (unpublished, 1980).

Received: October 8, 1980; *accepted:* October 9, 1980

LAWRENCE J. RAPHAEL, Haskins Laboratories, 270 Crown Street, *New Haven, CT 06510* (USA)