1126

Original Paper

Phonetica 1999;56:44–55

Received: December 15, 1998
Accepted: March 22, 1999

Perceiving Final Voiceless Stops without Release: Effects of Preceding Monophthongs versus Nonmonophthongs

Leigh Lisker

Haskins Laboratories, New Haven, Conn., USA; University of Pennsylvania, Philadelphia, Pa., USA

Abstract

Prepausal postvocalic stops in English are reported to occur both with and without audible release bursts, more or less randomly, and this difference is said to be without distinctive function. However, there is evidence that an English final stop, absent its release, may be of reduced intelligibility, particularly as to its place of articulation. Without audible release a final stop's place is conveyed mainly, perhaps entirely, by frequency shifts in the vowel formants. Among the vowels of English some are diphthongal, a property also signaled by formant shifts. The question arises: Is the intelligibility of an unreleased stop significantly affected by the phonetic nature of the vowel? Perceptual testing of appropriately chosen nonsense monosyllables ending in [p˺ t˺ k˺] indicates that these stops are generally somewhat less intelligible after diphthongs. However, not all three stops are affected equally, [k˺] perception being especially reduced in this context.

Introduction

Phonetic description is usually focused on properties of a speech signal that are thought to serve a message-differentiating function, most particularly those by which words are distinguished. Some less 'functional' properties do of course also get a certain amount of attention, e.g. English stop aspiration, usually called a predictable accompaniment of voicelessness under certain conditions, and hence not a distinctive feature of the language, does nevertheless play some role in a listener's perception. Another nondistinctive stop attribute, that of audible release in final position, has been less often assessed for any communicative function. The status of these two differences, [±aspirated] and [±released], is of course somewhat different in English: [+aspirated] and [−aspirated] voiceless stops have markedly different and nonoverlapping distributions within words, while [+released] and [−released] final stops occur more or less sporadically in very much the same contexts. Thus, to produce [pʰ] instead of [p] will violate a phonological rule of the language, whereas making a choice

KARGER
Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 1999 S. Karger AG, Basel
0031–8388/99/0562–0055
$17.50/0
Accessible online at:
http://BioMedNet.com/karger

Leigh Lisker
127 Fitzwater Street
Philadelphia, PA 19147 (USA)
Tel. +1 (215) 925 8241
E-Mail llisker@babel.ling.upenn.edu

between a prepausal [p] and [p'] is perhaps no more than a 'stylistic' matter. (Gimson [1962, p. 151] tells us that 'The non-release of final plosives is a feature of colloquial RP. Careful speakers, however, tend to release such plosives audibly and those who, in ordinary conversational style, use the unexploded variety will often use an audible release in more formal circumstances.') A possible exception, where the [±release] difference perhaps oversteps the stylistic to the point of being phonologically crucial, might be the case of final stops following /s/ (and other fricatives), if in fact it is true that 'in words like "task" the identification of the final stop must evidently be attributed to the spectral properties of the stop burst; since the stop is not adjacent to any vowel, there can be no transition cue' [Halle et al., 1957, p. 107].

That English stops may or may not be released finally reminds us once again of the rather exceptional nature of this class of consonant. English stops in general show a greater degree of phonetic variability than do many other phonetic segment types in the language: prevocalically p t k range from voiceless unaspirated to heavily aspirated, while b d g range from fully voiced (i.e. 'prevoiced') to voiceless unaspirated, and the six stops occur both with and without release when utterance-final. Moreover it would seem that releases may run a gamut from barely audible to heavily aspirated. Now in general the acoustic properties that signal the place of articulation of a stop consonant are distributed over a stretch of signal that often straddles the interval of oral closure, a preceding 'implosive' transition and the release plus explosive transition. Since stops in utterance-initial position are in general easily identified as to place (and rare indeed would be the language that limited stops to noninitial positions), we can readily believe that the acoustic information provided by a closing transition is not indispensable to its intelligibility. (There do not seem to be languages in which a word-initial stop must in utterance-initial position be preceded by an enunciative bit of 'sound stuff' that might carry such a transition.) The role of the interval just following closure is somewhat less clear. There are languages which do not allow stops either word- or utterance-finally (e.g. Italian, Tamil, Japanese), while in some other languages that do allow final stops they are regularly produced with audible release, i.e. a kind of vestigial vowel (e.g. French, Polish, German). Both these language types might suggest that post-closure properties are essential for place perception. However, there exist languages in which final stops occur regularly without audible release (e.g. Thai, Cambodian, Korean), and from such languages one may plausibly infer that releasing a final stop is perceptually unmotivated, and that the closing transition suffices as the carrier of place information. (It may be noted that in languages in which final stop contrasts are neutralized it seems to be more often voicing rather than place distinctions that are dispensed with.) Oddly enough, in no language that allows prepausal stop does the phonetically perfectly possible and readily controllable [±release] difference appear to serve any phonological function. From their wide-ranging survey of stops in the world's languages Henton et al. [1992] come to the same conclusion.

Earlier Work on English Final Unreleased Stops

The fact that speakers of English fluctuate in their management of final stops, producing them sometimes with and sometimes without audible release, so that the linguistic significance of this difference is commonly dismissed as nil, does not mean either that the occurrence of release makes no contribution to stop intelligibility or that

speakers' behavior is absolutely random in respect to this feature. A number of studies, all carried out in the fifties, assessed the perceptual function of prepausal release, and while these studies and the findings they report show some degree of variability, they do agree in concluding that stops deprived of audible releases suffer a considerable loss of intelligibility. Thus Householder [1956] obtained a mean score correct for [pˀ tˀ kˀ], in nonsense syllables that included these stops as well as glottal stop after 14 vowels, of roughly 52.5%, while Halle et al. [1957], examining stops after five vowels, determined a mean correct score of about 58% for the same stops. (Householder [1956] noted differences in the degree of intelligibility of final unreleased stops produced by different talkers.) Malécot [1959] looked at unreleased stops after the two vowels /ɔ/ and /ɛ/, and his subjects did somewhat better, achieving an overall percent score of about 76% correct in identifying unreleased p t k, while for these stops when artificially unreleased (i.e. 'dereleased') their score was virtually the same (78.7%). In a study that looked at unreleased stops following just the vowel /ɪ/ Wang [1959] found that phonetically naive listeners correctly identified words ending with the three voiceless stops minus releases at better than 90%. All three studies agreed in concluding that velars without releases tend to be less accurately identified than either alveolars or labials, though this does not hold true following some vowels, /ɪ/ in particular. The two studies in which a range of vocalic contexts was considered, those of Householder [1956] and Halle et al. [1957], were also agreed in finding the lower intelligibility of velar place to be especially notable after 'tense' vowels. However, in a brief review of these papers Delattre [1958] ventured to conclude from their data that the perceptibility of unreleased velar stops depends not on the 'tensity' of a preceding vowel, but rather on whether it is lip-spread or lip-rounded, being lower following a vowel of the latter type. Moreover, unlike Householder [1956] and Halle et al. [1957], who offered no phonetic explanation for the observed vowel-dependent effect, Delattre [1958] explained it as a consequence of a difference in the transitional 'loci' for velars that is conditioned by the rounding status of an adjacent vowel.

During the same period in which Householder [1956], Halle et al. [1957], Malécot [1959], and Wang [1959] presented their findings, the present author was engaged in a somewhat similar study [some of its findings were presented in the then house-organ of the Haskins Laboratories; Lisker, 1957, 1958], where the question addressed was whether the cue value of a closing transition for a final postvocalic stop consonant is affected when the vowel is at the same time undergoing the formant frequency modulation that can signal a diphthongal glide. To test for a difference in the intelligibility of final unreleased stops following monophthongal as against diphthongized vowels, a set of nonsense monosyllables was devised in which each of the 13 nuclei i ɪ e ɛ æ ɑ ɔ o ʊ u ɔj aj aw was paired with each of the three codas p t k. For each rhyme six onsets were chosen so that no onset + rhyme combination made a meaningful form in English. This set of syllables was recorded by a single speaker of American English long resident in the New York City area. He was told to pronunce each form with the stress and intonation appropriate to that of an isolated one-word declarative sentence. Several tokens of each syllable were recorded, and in all of them the final stops were produced with audible releases. For each syllable type a single token was chosen so as to make up a stimulus set of near uniform loudness and intonation shape. This set was used in a preliminary test, from which it turned out that each syllable coda, upon presentation in randomized order, was identified as the intended stop by the author and two colleagues. These syllables were then subjected to waveform editing in which the final releases

Table 1. Confusion matrices representing 1958 data

		p	t	k	NC				p	t	k	NC
Monophthongs												
ɪ	p'	*100*	0	0	0		æ	p'	*88*	10	0	2
	t'	2	*91*	7	0			t'	0	*90*	5	5
	k'	0	0	*100*	0			k'	3	7	*90*	0
ɛ	p'	*60*	40	0	0	0	ɔ	p'	*75*	4	21	0
	t'	0	*97*	3	0	0		t'	8	*88*	4	0
	k'	0	0	*100*	0	0		k'	5	7	*88*	0
ɑ	p'	*94*	6	0	0		ʊ	p'	*60*	9	31	0
	t'	22	*56*	22	0			t'	0	*93*	7	0
	k'	3	11	*86*	0			k'	3	11	*86*	0
Semidiphthongs												
u	p'	*83*	0	17	0		o	p'	*88*	7	5	0
	t'	0	*100*	0	0			t'	0	*98*	0	2
	k'	53	20	*20*	7			k'	67	14	*19*	0
i	p'	*96*	0	0	4		e	p'	*86*	14	0	0
	t'	40	*48*	2	10			t'	42	*50*	5	3
	k'	44	27	*25*	4			k'	25	47	*28*	0
Diphthongs												
aj	p'	*88*	11	1	0		aw	p'	*54*	24	14	8
	t'	2	*92*	6	0			t'	11	*78*	7	4
	k'	15	15	*27*	0			k'	39	33	*14*	14
ɔj	p'	*63*	15	3	19							
	t'	35	*50*	0	15							
	k'	13	52	*10*	25							

For each combination of vowel and final stop a total of 96 responses were obtained from 8 listeners. Boldface figures are percent correct judgments.
NC = No consonant.

were removed. For each syllable nucleus a 36-item test was composed in which each of the truncated 18 syllables (6 onsets × 3 codas) occurred twice in a random order.

Each of the 13 tests prepared (one for each nucleus) was presented, via loudspeaker at a 'comfortable' level in a sound-isolated room, to 8 educated but phonetically untrained native speakers of American English (New York City region). Interstimulus intervals were fixed at 3 s. Listeners were asked to decide, for each stimulus syllable, whether it terminated in p, t, k, or perhaps in no consonant at all. Distributions of their judgments over the four response categories for the 13 vowels are tabulated as percentages in table 1, where each value represents means of 96 responses for each rhyme (6 onsets × 2 presentations × 8 listeners). In line with the hypothesis motivating the experiment the syllable nuclei are grouped tentatively into three categories: monophthongs, semidiphthongs (i.e. nuclei of variable degree of diphthongization, perhaps particularly variable before voiceless codas), and diphthongs. (As must be the case, at least in my opinion, the decision as to whether a vowel in a consonantal context is monophthongal or diphthongized is a matter for the auditory judgments of trained phoneticians. In this case, judgments by 3 such listeners were obtained.)

Table 2. Vowels arranged in ascending order of percent correct judgments of final p' t' k' means over all three stops taken together and separately

p' t' k'		p'		t'		k'	
ɔj	41	aw	54	i	48	ɔj	10
aw	49	ɛ	60	e	50	aw	14
e	55	ʊ	60	ɔj	50	o	19
i	56	ɔj	63	ɑ	56	u	20
o	68	ɔ	75	aw	78	i	25
u	68	u	83	ɔ	88	aj	27
aj	69	e	86	æ	90	e	28
ʊ	79	aj	88	ɪ	91	ʊ	83
ɑ	79	o	88	aj	92	ɑ	86
ɔ	83	æ	88	ʊ	93	ɔ	88
ɛ	86	ɑ	94	ɛ	97	æ	90
æ	89	i	96	o	98	ɛ	100
ɪ	97	ɪ	100	u	100	ɪ	100

N = 96 judgments from each of 8 listeners.

When the 13 nuclei are sorted according to the overall percentage correct identifications obtained for the three final stops taken together (table 2, left column), we find that they vary over a wide range – from 41% for ɔj to 97% for ɪ. The mean overall percent correct identification, about 70%, is a considerably higher score than either of those reported by Householder [1956] and Halle et al. [1957]. It may be noted that the nuclei fall into two rather than three nonoverlapping groups, the monophthongs ɪ ɛ æ ɔ ɑ ʊ, all with mean cross-place values above 75%, and the nonmonophthongs (semidiphthongs and diphthongs considered together) e i u o aj aw ɔj, all with values of less than 70%. These relations are shown more clearly in figure 1, where percent correct responses are grouped according to these two vowel type contexts. While in figure 1A we see that the stops taken together seem clearly more intelligible after monophthongs, this relationship is not manifested for all three stops, since for p' intelligibility is virtually the same after the two kinds of vowels. Thus the overall advantage of the monophthongal context reflects the case of t' and k', more particularly the latter. The magnitudes of these differences in mean intelligibility are about 12% for t' and as much as 70% for k'. Since back in the fifties it was not yet the fashion in phonetic research to subject findings to statistical testing, the individual response data underlying the displays in figures 1 and 2 were not saved, and hence are no longer available for any more exact measure of statistical significance. (It is to be noted that no statistical testing was reported in any of the published papers of the fifties cited above.)

Aside from all these experiments aimed at determining the perceptual significance of final stop release, there were sporadic observations of release behavior on the part of English speakers that indicated that the occurrence of releases might not be entirely random. Thus from Trager and Smith [1951, p. 31], one gathers that prepausal stops, at least the voiceless ones, are commonly unreleased, release being a feature of 'emphatic' speech. Wang [1959, p. 71], found that 'before terminal junctures in informal American English voiced plosives are released less frequently than are voiceless plosives'. On the other hand, Parker and Walsh [1981] found the voiced stops to be more often released than the voiceless, and that, moreover, stop release 'is a function of

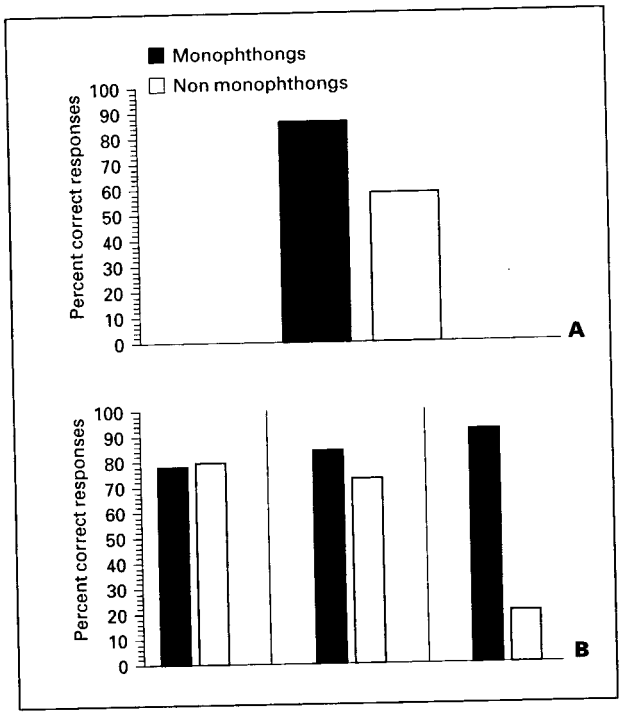

Fig. 1. Data from table 2 grouped according to contextual vowel types. **A** Percentage correct responses summed over the three stops. **B** Data for each stop separately: p' (left), t' (center), and k' (right).

the tenseness of the preceding vowel' (p. 355), the 'tense' vowels being i e ɔ o u, which are dubbed 'tense' rather than 'lax' simply because they are free to occur word-finally in stressed syllables. Their study was limited to alveolar stops, but they apparently felt that its finding held true for all three places of stop articulation. Crystal and House [1988] considered stops in the speech of subjects reading connected texts, and reported that prepausal stops were 'complete' (i.e. accompanied by releases detectable in spectrograms) in only 48% of 163 tokens examined. A much more ambitious survey by Byrd [1993] considered some 1,130 sentence-final stops occurring in the TIMIT data base [a description of TIMIT is provided by Keating et al., 1994], and found about 40% to be produced without audible releases, with no significant difference between voiced and voiceless ones except in the case of the velars, with the 130 /k/s observed being 85% released. There was, moreover, a sex-based difference in the pronunciation of stops – female speakers more frequently produced prepausal stops with audible releases.

Two Recent Experiments

Released vs. Unreleased Stops

In the very recent past two further experiments have been run to measure once again the intelligibility of final stops without releases in postvocalic contexts. The main purpose of the first was to learn how well the results obtained in the fifties would be

replicated for a set of stimuli recorded by a new speaker and tested on a new group of subjects. A set of 84 nonsense syllables was recorded, all of them terminating in a voiceless stop following each of 14 syllabic nuclei: the seven monophthongs ɪ ɛ æ ɑ ʌ ʊ ɔ and the seven nonmonophthongs i e u o aj aw ɔj. In half of the syllables tested the final stops were produced with clearly audible releases, in half they were produced without release bursts. (The stimuli with released stops were included primarily in order to ensure that 'optimal' stimuli – those with releases being presumably such – would achieve intelligibility scores very close to 100%.) Unlike in the experiment 40 years ago, this time testing was computer-facilitated (with the use of the software program PsyScope), each session involving just a single listener, who was required to make a forced choice among 'p', 't', and 'k' in responding to each stimulus (heard in pairwise presentation), but was free to take as much time as he/she wanted in making a response. Tapping a key (p, t, or k) on the computer keyboard triggered presentation of the next stimulus in the PsyScope-generated randomization. Stimuli were presented by headphones at a comfortable listening level, each listener being tested individually. Nine listeners served as test subjects, and they provided a total of 45 judgments for each stimulus rhyme.

As should have been expected, if the speaker who provided the stimuli performed as required, the recording and reproducing systems were of satisfactory audio fidelity, and the test environment was also suitably noise-free, the responses to the stimuli terminating in released stops were very nearly 100% 'correct', i.e. in conformity with those intended by the speaker providing the stimuli. (Percentage correct responses varied over the 9 listeners from a high of 100% to a low of about 98.4%, with a mean overall of 99.6%.) For the stimuli ending in the unreleased stops, the distribution of p t k judgments recorded (table 3) was roughly similar to that shown in table 2, from which we may conclude that intelligibility again varied both with the vowels and the stops. The overall intelligibility of this new set of unreleased stop stimuli seems to be markedly higher than any of those reported in any of the earlier studies. In fact, a rough comparison indicates that overall percent correct responses were better than 25% higher. Whether this difference is to be attributed to differences between the 2 speakers who provided the stimuli of the two tests, to differences in the fidelity of recording and/or reproduction, to differences in the backgrounds of the listeners tested (unlikely, I believe), or perhaps to differences in the way in which the tests were administered, it is now impossible to determine. It might most reasonably be supposed that the high intelligibility scores recorded by all 9 subjects were facilitated by the fact that they were tested individually, that stimuli were presented over high-quality headphones, and that subjects were free to take as long as they wanted in making each response choice.

When the percentage correct responses to the 42 syllable nuclei (14 vowels × 3 stops) are grouped with respect to release type, vowel type, and stop place (fig. 2), we find, first of all, that the released stops over both vowels and stops have mean intelligibility scores of over 90%, while for the unreleased stops the intelligibility pattern is similar to that found in the earlier study. First of all, monophthongs overall elicited more correct responses than did the diphthongized vowels. In the second place, there was a clear stop-place effect, although there is not the sharp division, in the case of /k/, between monophthongs and nonmonophthongs that was found in the study of the fifties. On the other hand, p and t judgments seem to pattern somewhat more like those for /k/ with respect to vowel type. An analysis of variance (repeated measures) of the data for the released and the unreleased stops showed that the difference between

Table 3. Vowels arranged according to percent correct judgments of final p t k, both released and unreleased

p t k		p		t		k	
Released							
ɔj	96	aw	98	æ	98	ɔj	87
o	99	o	98	aw	98	o	98
aw	99	aj	100	o	100	aj	98
aj	99	ɛ	100	aj	100	ɛ	98
ɛ	99	æ	100	ɛ	100	u	98
æ	99	u	100	u	100	æ	100
u	100	i	100	i	100	i	100
i	100	ɔj	100	ɔj	100	aw	100
e	100	e	100	e	100	e	100
ʊ	100	ʊ	100	ʊ	100	ʊ	100
ʌ	100	ʌ	100	ʌ	100	ʌ	100
ɪ	100	ɪ	100	ɪ	100	ɪ	100
ɔ	100	ɔ	100	ɔ	100	ɔ	100
ɑ	100	ɑ	100	ɑ	100	ɑ	100
Unreleased							
ɔj	67	e	70	aj	80	ɔj	32
aw	74	aw	77	i	82	aw	45
i	81	ɔj	84	ɔj	84	o	54
o	82	o	93	e	92	u	54
u	83	i	93	ɛ	93	i	67
e	84	u	95	ɪ	95	aj	79
aj	86	ʊ	95	ɔ	98	e	84
ʊ	95	aj	100	o	98	ʊ	91
æ	97	ɑ	100	u	99	æ	91
ʌ	98	ɔ	100	aw	100	ʌ	95
ɛ	98	ʌ	100	ɑ	100	ɑ	98
ɪ	98	æ	100	ʊ	100	ɔ	98
ɔ	99	ɛ	100	ʌ	100	ɛ	100
ɑ	99	ɪ	100	æ	100	ɪ	100

N = 45 judgments from 9 listeners.

the two kinds of stops was highly significant [$F(1, 8) = 96.5$, $p < 0.001$, and that both vowel types and stop place were just as important: for vowels $F(1, 8) = 101.6$, $p < 0.001$; for stops $F(2, 16) = 27.6$, $p < 0.001$]. Moreover, interactions among the three factors were all significant at levels better than of $p = 0.001$. From post-hoc paired t testing of the mean percent correct scores for each of the stops as these vary with vowel type it appears that while mean percent correct identifications for all three stops are lower after nonmonophthongs, vowel-dependent differences are only significant in the case of the unreleased stops, and there the differences are statistically of somewhat variable significance: most significant in the case of k' ($p < 0.001$), still highly significant for p' ($p = 0.004$), and just skirting the boundary of significance for t' ($p = 0.052$).

Unreleased vs. 'Dereleased' Stops

The second experiment was designed to learn whether the intelligibility of a final stop minus an audible release varies depending on whether the stop was originally produced without release or whether an original release signal was removed by editing.

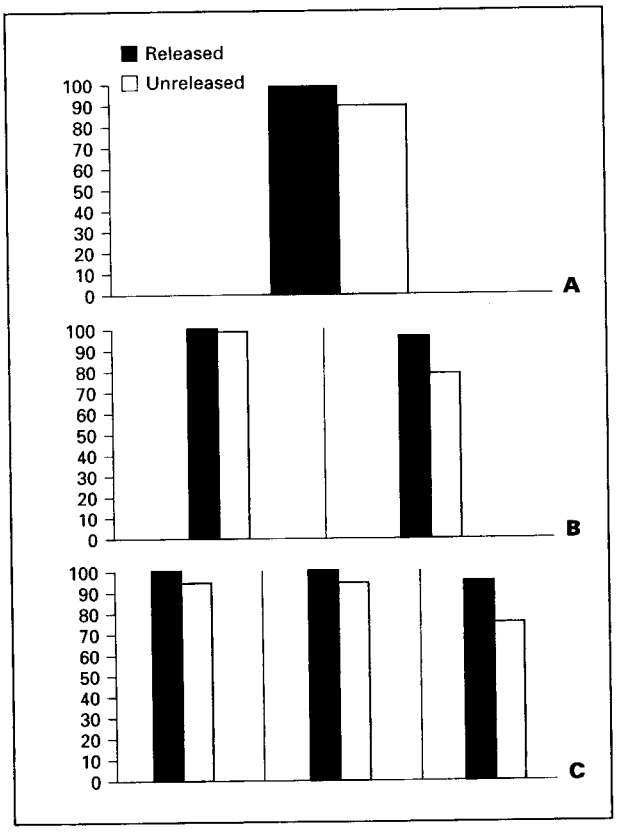

Fig. 2. Data from table 3 arranged to show mean percent correct judgments for released and unreleased stops. **A** Means summed over all vowels and stops. **B** Means according to vowel type: monophthongs (left panel) and nonmonophthongs (right panel). **C** Means for the individual stops: p' (left panel), t' (center panel), and k' (right panel).

It seemed plausible to suppose that there might be a difference in the nature of the closing gestures for the two varieties of stop, such that the transitions to the unreleased closures would be of enhanced intelligibility. To be sure, Malécot [1958] reported a negative finding on this score, but his data were restricted to stops after just the two monophthongal vowels, ɛ and ɔ. On the other hand, the possibility of such an articulatory difference has just been suggested to explain the greater ability of Thai listeners as against American in their ability to identify the regularly final unreleased stops of Thai [Abramson, A.S.; Tingsabadh, K.: Thai final stops: cross-language perception, submitted]. To test for the presence of a difference over a greater variety of vowels the two kinds of final stops (unreleased and 'dereleased') were presented in nonsense syllables after the eight monophthongs ɪ ɛ æ ʌ ɔ ʊ ɑ ɚ and the seven diphthongized vowels i e u o ɔj aj aw. A total of 61 responses for each of the 90 rhymes (2 nonrelease types × 15 vowels × 3 stops) was recorded from 15 subjects, 6 new ones in addition to the 9 who had served earlier. The overall percentage correct responses as they varied over the 15 vowels (table 4) yields a picture that is again quite similar to those conveyed by tables 2 and 3: intelligibility summed over the three stops is greater following the monophthongs, and this holds true for both unreleased and dereleased stops. An analysis of variance examining the effects of closure type (unreleased and dereleased) vs. vowel

Table 4. Vowels arranged according to percent correct judgments of final unreleased and dereleased p t k: means over all three stops taken together and separately

p˺ t˺ k˺ p˺		t˺		k˺			
Unreleased							
aw	54	aw	51	e	54	aw	10
u	74	ɔj	85	aj	72	u	31
ɔj	78	ɚ	85	i	78	ɔj	46
i	79	e	89	ɔj	80	i	53
e	80	ɔ	90	ɚ	92	e	56
o	80	u	92	ɪ	95	o	72
aj	86	o	93	ʌ	97	aj	79
ɚ	90	ɛ	95	ʊ	97	ɚ	92
ɔ	95	i	97	o	98	ɔ	95
ʊ	97	æ	97	ɔ	98	ʊ	95
æ	97	aj	98	ɛ	98	æ	95
ɛ	98	ɪ	100	ɑ	100	ɛ	100
ɪ	98	ʌ	100	æ	100	ɪ	100
ʌ	99	ʊ	100	u	100	ʌ	100
ɑ	100	ɑ	100	aw	100	ɑ	100
Dereleased							
u	64	aw	77	e	48	u	15
aw	69	u	77	aj	74	aw	31
ɔj	76	ɔj	80	i	82	o	52
e	78	ɛ	90	ɪ	92	ɔj	52
o	81	o	92	ɚ	93	aj	75
aj	83	ɚ	95	ɔ	95	ɚ	77
i	87	ʌ	97	ɔj	95	i	79
ɚ	89	ɔ	97	ɛ	98	ʊ	82
ʊ	93	ʊ	97	ʌ	98	e	89
ɛ	96	æ	98	aw	98	ɑ	95
ʌ	96	e	98	æ	100	ɔ	97
ɔ	96	aj	98	ʊ	100	ɪ	98
ɪ	97	ɪ	100	ɑ	100	ʌ	98
ɑ	98	i	100	o	100	ɛ	100
æ	99	ɑ	100	u	100	æ	100

N = 61 judgments from 15 subjects.

type vs. stop place revealed that while differences in both stop category and vowel type were significant, as was true for the data of the previous test [for vowel type $F(1, 14) = 257.4$, $p < 0.001$; for stops $F(2, 28) = 43.1$, $p < 0.001$], the overall difference between unreleased and dereleased stops was not [$F(1, 14) = 0.056$, $p = 0.82$]. An examination of this unreleased-dereleased difference for each vowel type and each stop showed very minor differences in percent correct means, differences that in no case reached statistically significant values. Thus it appears that either the speaker who provided the test stimuli performed essentially the same closing gesture before released and unreleased closures, or, if he did not, then our English-speaking listeners did no better with one type than the other (fig. 3).

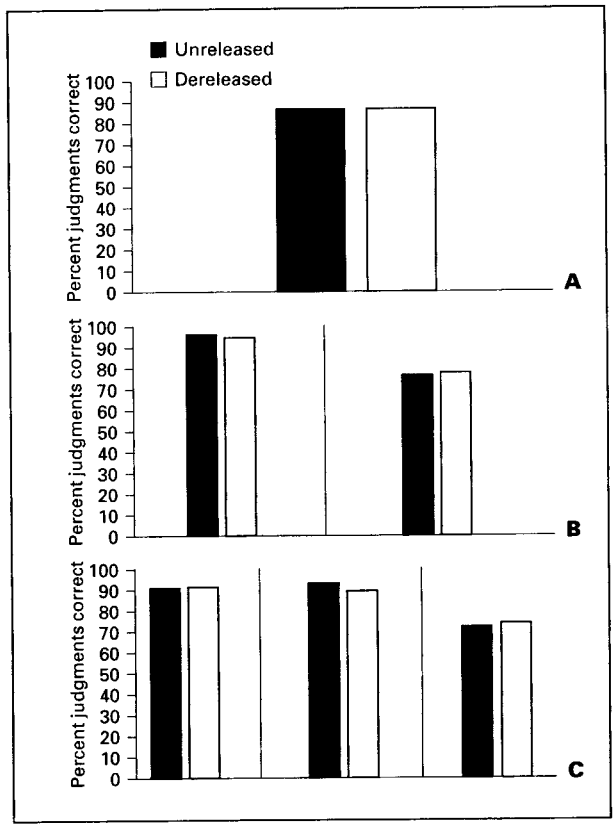

Fig. 3. Data from table 3, showing differences in mean percent correct judgments of unreleased and dereleased p' t' k'. **A** Means over both vowel types and three stops. **B** Means for monophthongs (left panel) vs. nonmonophthongs (right panel). **C** Means for p' (left panel), t' (center panel), and k' (right panel).

Discussion

The data derived from these last two experiments can be said to agree roughly with those gathered in the fifties, bearing out the hypothesis that motivated the original study. Postvocalic English unreleased final stops tend to be of lower intelligibility, at least for English-speaking listeners, than are those with releases, and this lowered intelligibility is manifested particularly in the case of stops that are preceded by nonmonophthongs. However, this effect does not appear equally for the voiceless stops at all three places of articulation: while p' and t' now and again suffer a certain loss after nonmonophthongs, it is k' that is especially and regularly vulnerable. Furthermore, there is considerable variation in stop intelligibility within each class of vowel contexts. Thus p' suffers little or no loss after i aj, while t' after u o aw is as intelligible as it is after any monophthong. This might suggest that where diphthongal and stop transitional shifts are apt to be very different in direction, the listener is best able to resolve the formant movements into components that reflect the articulatory gestures that gave rise to them. So the generally lowered intelligibility of final stops without release bursts cannot be entirely attributed to diphthongization of a preceding vowel – there must be other factors involved. Of course, we must also bear in mind that the data just presented are derived from single tokens of the vowel-stop combinations examined, and we can-

not be sure as to how much cross-token variability might turn up with much more extensive testing. (Thus the relatively low intelligibility scores for pˀ after the monophthongs ɛ and ʊ in tables 1 and 2 were nowhere near matched in the data of either table 3 or table 4, and possibly reflect accidental 'glitches' either in the test stimuli or in the testing environment.) Lastly, we should not restrict ourselves to phonetic explanations: perhaps the lower intelligibility of pˀ and kˀ after aw and ɔj is not so much a reflection of difficulty in disentangling consonantal and vocalic components in the diphthongal context as it is of a phonotactic fact of English – that the rimes awp awk ɔjp and ɔjk are not found (not 'permitted?') in the language.

Finally, by way of conclusion, the following question might be posed: Granted that English speakers occasionally indulge themselves in final unreleased stops, with presumably little effect on their overall intelligibility in running speech (a top-down effect?), why do there exist languages for which permissible final stops are *mandatorily* unreleased? The behavior of an English speaker who, in speaking 'deliberately', releases all final stops, can be plausibly explained on the ground that the speaker 'knows' that releases enhance intelligibility. Can the fact that a Thai speaker, no matter how deliberate the style, will refrain from releasing final stops possibly be explained on the same basis? And if it seems difficult to explain the existence of final unreleased stops as a norm by supposing that they are more intelligible than released stops, it seems just as far-fetched to presume that in some language-independent sense (whatever that might be) they are somehow 'easier' to produce.

Acknowledgments

The most recent research reported above was supported by NIH grant DC-02717 to Haskins Laboratories. Thanks are due to the many members of the Laboratories who served as willing and attentive subjects in the tests asministered. For their patient instruction in the understanding and management of the several software programs used in the gathering and statistical evaluation of the data reported, I would in addition like to express my gratitude to the following colleagues: Arthur Abramson, Steven Frost, Einar Menci, Leonard Katz, and Douglas Whalen.

References

Byrd, D.: 54,000 American Stops. UCLA Working Papers Phonet. *83:* 1–19 (1993).
Crystal, T.H.; House, A.S.: Segmental durations in connected-speech signals: current results. J. acoust. Soc. Am. *83:* 1553–1585 (1998).
Delattre, P.: Unreleased velar plosives after back-rounded vowels. J. acoust. Soc. Am. *30:* 581–582 (1958).
Gimson, A.C.: An introduction to the pronunciation of English (Arnold, London 1962).
Halle, M.; Hughes, G.W.; Radley, J.-P.A.: Acoustic properties of stop consonants. J. acoust. Soc. Am. *29:* 107–116 (1957).
Henton, C.; Ladefoged, P.; Maddieson, I.: Stops in the world's languages. Phonetica *49:* 49–65 (1992).
Householder, F.W. Jr.: Unreleased PTK in American English; in Halle, Lunt, McLean, van Schooneveld, For Roman Jakobson, pp. 235–244 (Mouton, The Hague 1956).
Keating, P.A.; Byrd, D.; Flemming, E.; Todaka, Y.: Phonetic analyses of word and segment variation using the TIMIT corpus of American English. Speech Commun. *14:* 131–142 (1994).
Lisker, L.: The effect of glides on following implosive transitions. Haskins Laboratories, 24th Quarterly Progress Report, App. 12 (1957).
Lisker, L.: The effect of glides on terminal implosive transitions, App. 4 (1958).
Malécot, A.: The role of releases in the identification of released final stops: a series of tape-cutting experiments. Language *34:* 370–380 (1958).
Parker, F.; Walsh, T.: Voicing cues as a function of the tense/lax distinction in vowels. J. Phonet. *9:* 353–358 (1981).
Trager, G.L.; Smith, H.L. Jr.: An outline of English structure (Battenburg Press, Norman 1951).
Wang, W.S.-Y.: Transition and release as perceptual cues for final plosives. J. Speech Hear. Res. *2:* 66–73 (1959).