Journal of Phonetics (1989) **17**, 283–298

Perception of consonant length: voiceless stops in Turkish and Bengali

Jorge Hankamer

University of California at Santa Cruz, Department of Linguistics, Santa Cruz, CA 95064, U.S.A.

Aditi Lahiri

Max-Planck-Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands

and

Jacques Koreman

University of Nijmegen, Institute of Phonetics, Erasmusplein 1, 6500 HD Nijmegen, The Netherlands

Received 23rd September 1988, and in revised form 2 June 1989

The goal of the present research was to investigate the extent to which acoustic cues other than the silent gap corresponding to closure duration might figure in the discrimination of geminate and non-geminate voiceless stops. The method, adapted from earlier studies on duration cues, was to create two sets of stimuli with the length of the silent interval varying incrementally between that of a non-geminate and that of a geminate. One set was made by artificially lengthening the silent interval of an original non-geminate in 10 ms steps up to the length of a geminate, and the other set was made by shortening an original geminate in the same manner. Starting with recorded minimal word pairs with geminate and non-geminate stops in Turkish and Bengali, sets of stimuli were constructed as described above. These stimuli were presented to native speakers of the respective languages in a word identification task, and the results were charted to see whether the identification curves were the same for the original geminates as for the original non-geminates. The result was that the curves did differ, the original geminates being identified as geminates slightly more frequently than original non-geminates at closure durations between 120 and 160 ms. The difference was statistically significant for at least some points on the curve. However, in contrast to earlier studies on consonant duration contrasts, the displacements found were less than 10 ms on the time axis, and the region of significant difference between the identification curves confined entirely to a segment of the time continuum in which no naturally occurring stimuli are found. Our conclusion is that in actual speech recognition there is no evidence that cues other than closure duration play a role in the discrimination of geminate and non-geminate stops in these languages.

0095–4470/89/040283 + 16 $03.00/0

1. Introduction

In an earlier study (Lahiri & Hankamer, 1988) of consonantal length oppositions in two unrelated languages, Bengali and Turkish,[1] we established that the silent interval corresponding to closure duration is an overriding perceptual cue for the discrimination of long (geminate) and short (non-geminate) voiceless stops in those languages. The consonant length opposition in these languages is the type generally recognized by phoneticians as a geminate *vs.* non-geminate distinction, as opposed to other possible distinctions involving length. The opposition is found only in intervocalic position, where the geminate consonant is always heterosyllabic and the simple consonant is always tautosyllabic. Several standard works insist on a distinction between this type of opposition and any opposition involving consonant length in which both long and short consonants may be tautosyllabic. Abercrombie (1967, p. 82), for example, reserves the term "long consonant" for tautosyllabic long consonants, using the term "double consonant" for articulations extending across a syllable boundary. Lehiste (1970, p. 44) draws a similar distinction, using the term "geminate" in place of Abercrombie's "double consonant".[2] The purpose of the present study is to investigate further the dominance of the closure duration cue in the geminate vs. non-geminate distinction, making use of a more delicate experimental technique. We will also examine (by measurement, but not by manipulation) acoustic factors other than the duration cue which might serve as auxiliary cues in the discrimination between geminates and non-geminates, and attempt to assess the extent to which such factors play a role in geminate *vs.* non-geminate recognition in speech perception.

There have been studies indicating that closure duration is a significant cue in the perception of consonantal length contrasts (Pickett & Decker, 1960; Obrecht, 1965; Repp, 1978, 1983). There have also, however, been studies suggesting that closure duration may play a role in the distinction of contrasts not traditionally associated with length, such as voiced *vs.* voiceless (Lisker, 1957) and fortis *vs.* lenis (Elugbe & Hombert, 1975), and that in at least some cases acoustic cues other than closure duration may play a role in the discrimination of contrasts traditionally treated as length or geminate *vs.* non-geminate oppositions. Lisker (1958), for example, found that duration differences co-occurred with differences in voiced *vs.* voiceless and tense *vs.* lax in Tamil occlusives, an opposition treated traditionally and orthographically as a geminate *vs.* non-geminate distinction. Abramson (1987*a*) found that Pattani Malay speakers can distinguish between long and short voiceless stops in utterance-initial position, where the actual duration of the closure is not reflected in an audible feature of the signal.

Such findings are perhaps not surprising. The geminate *vs.* non-geminate opposition is often associated impressionistically with features other than length, such as aspiration, affrication, "tenseness", and length of the preceding vowel (cf., for example, Catford, 1977, pp. 200–202, 210–211). Taken together with the results

[1] Henceforth, where there is no danger of confusion, we use the term "closure duration" (CD) to refer to the acoustic measure most directly corresponding to the articulatory duration of closure. In the case of voiceless stops, this will be the duration of the silent interval.

[2] Catford (1977, p. 210) repeats the above, and goes on to say that the term "geminate" is to be restricted to cases where the articulation of the consonant does not cross a morpheme boundary. In Lahiri & Hankamer (1988), we showed that phonetically there are no differences, in the two languages studied, between tautomorphemic and heteromorphemic geminates.

of the studies mentioned above, this indicates that for any consonant opposition involving length, even where the duration cue is found to be dominant, a question remains as to the presence and relative value of other acoustic properties as cues to the distinction. We propose to address this question for the case of the geminate *vs.* non-geminate opposition in voiceless stops, as characterized above and as exemplified in Turkish and Bengali.

In our earlier experiment, we tested the perceptual salience of closure duration as a cue to the geminate *vs.* non-geminate opposition in Turkish and Bengali by cross-splicing the part of the signal corresponding to closure duration between geminate and non-geminate voiceless stops. Subjects asked to identify the artificial stimuli consistently identified the cross-spliced geminates as non-geminates and *vice versa*. This experiment demonstrated that the closure duration cue overrides any other cues that may be present in the signal, when the closure duration is that of a naturally occurring geminate or non-geminate.[3]

In order to get a more delicate measure of the effect of secondary cues which might be present in the signal, we have adapted a method developed by Lisker (1957) for the study of duration cues in the voiced *vs.* voiceless opposition, and subsequently used by Elugbe & Hombert (1975) in a study of the fortis *vs.* lenis opposition and by Lisker (1958), Fukui (1978), and Abramson (1987*b*) in studies of consonant length.

The idea is to create two sets of stimuli with closure duration varying incrementally between that of a non-geminate and that of a geminate. One set is made by artificially lengthening an original non-geminate in 10 ms steps up to the length of a geminate, and the other set is made by shortening an original geminate in the same manner. Subjects are then asked to identify the stimuli. A comparison of the resulting response curves (percentage geminate responses at each duration for the two sets of stimuli) will reveal the presence of secondary cues, if there are any, affecting the identification of the stimuli as geminate or non-geminate.

Lisker (1957), employing this method, found a roughly 30 ms displacement in the crossover point between voiced *vs.* voiceless judgments of intervocalic stops in English, depending on whether the stimuli were created from original voiced stops or from original voiceless stops. Elugbe & Hombert (1975) found a roughly 20 ms displacement (approximate, measured from the graphs they present) for the fortis *vs.* lenis opposition in Ghotuo nasals. Thus, as Lisker (1957, p. 47) pointed out, the displacement of the response curves along the time axis provides a measure of the effect of cues other than duration in biasing responses to the stimuli. The greater the displacement, the more salient the other cues are in comparison to the duration cue. This method provides an indirect way of measuring the effect of secondary cues even when it is not known what they are.

Despite the fact that there have been several studies of consonant duration employing this method,[4] only two (the Marathi experiment reported in Lisker (1958)

[3] As Lisker (1958) noted, absolute closure durations vary considerably with speech rate and context (in particular, they tend to be longer in isolated words). Our samples were all words spoken in isolation, with the speakers instructed to speak at a "normal" rate of utterance. The "naturally occurring" closure durations we refer to are those found in this context.

[4] Some earlier studies on consonant duration (Obrecht, 1965; Repp, 1978, 1983) used synthesized stimuli, varying only duration, and hence would not reveal anything about secondary cues. Experiments like that of Pickett & Decker (1960), in which the duration was incremented in only one direction, similarly would not reveal the effect of secondary cues.

and the brief study of Japanese in Fukui (1978)) have dealt with clear instances of the geminate *vs.* non-geminate contrast as found in Bengali and Turkish, even though these languages probably represent the most typical case. In these languages, the consonant length opposition exists only in intervocalic position, and the geminate consonants are always heterosyllabic. Thus the opposition is not found in any contexts where a cue other than closure duration would be necessary for discrimination, as is the case in Pattani Malay. In our earlier production study, we found no measurable properties of the signal that seemed likely candidates for reliable secondary cues, in contrast to the case of the voiced *vs.* voiceless opposition studied by Lisker (1957).

As background to our own study, we describe the experiment reported in Lisker (1958), in which the Marathi pair (matə, mattə) was manipulated. Lisker incremented the CD (silent interval corresponding to closure duration) of matə in 20 ms intervals and decremented the CD of mattə similarly to create two sets of artificial stimuli. When the stimuli were played back to the Marathi speaker who produced the original words, he identified those with CD up to and including 120 ms (excluding the two shortest stimuli, at 40 and 60 ms duration, which he could not identify as words and reported as containing a voiced stop) as matə and those with CD 160 ms and greater as mattə. For the two stimuli with CD of 140 ms, he identified the one which had been created from original matə as matə, and the one that had been created from original mattə as mattə. Lisker concluded (p. 301) that "whatever the differences between the two words other than the one of closure duration, they produce only about a 20 ms shift in the boundary value between t and tt".[5]

Aside from the fact that this experiment involved only one minimal pair, the results tell somewhat less than they appear to about the magnitude of the displacement caused by the secondary cues. Lisker is no doubt correct in his main conclusion, which is that closure duration is shown to be a major cue; but the magnitude of the displacement effect could be anywhere between almost zero and 40 ms, and still be consistent with the observations. There is, after all, only one observation point in the crossover region between 120 and 160 ms. If we are interested in an accurate assessment of the magnitude of the contribution of secondary cues in the discrimination of such stimuli, it will be necessary to replicate this experiment with a greater number of observations.

In the present paper, we will show that the displacement effect of secondary cues in the geminate *vs.* non-geminate opposition in Turkish and Bengali is considerably less than the effects found for the oppositions studied by earlier investigators mentioned above. We will show that on investigating a number of minimal pairs rather than just one, the secondary effects vary considerably across different word pairs, and have no systematic correspondence with observed acoustic differences. We will argue that whatever secondary cues there are, there is no reason to believe that they have an effect on the discernment of geminates *vs.* non-geminates in actual speech, the single relevant cue being duration of the silent interval.

[5] In addition to the one native speaker of Marathi, Lisker had five native speakers of English, who knew no Marathi, judge the stimuli as to length and place of articulation. While Lisker claimed (p. 301) that these speakers' phonetic judgments were "in very close agreement with the Marathi speaker's phonemic discriminations", a close examination of the figures he presents (p. 300) indicates that they show a crossover in percept somewhere around 140–160 ms for both sets of stimuli, but they do not clearly show a shift along the duration axis differentiating the two sets.

2. Method

Earlier studies have typically been based on stimuli derived by manipulation of a single minimal pair. Given the clear dominance of the duration cue found in our earlier study, and the relative weakness of features of the signal that might serve as secondary cues, we decided to use several minimal pairs as the basis for our stimuli in both languages investigated.

In constructing the test stimuli we used six minimal word pairs in Bengali and seven in Turkish. For each language, a native speaker read the word pairs, which were recorded in a sound-proof booth on a Nagra 4.2 tape recorder. The words used in the experiment are listed in the Appendix. Since no published frequency data are available, we relied on the subjective judgments of the native speakers to ensure that for each pair the words were equally well known.

The Bengali words all contained tautomorphemic geminates, while the geminates in the Turkish words were all derived by concatenation. We restricted ourselves to coronal stops (dental and retroflex in Bengali and dental in Turkish) because in Turkish it is difficult to find minimal pairs involving geminate *vs.* non-geminate contrasts in common words at other places of articulation.

The test words were digitized on a VAX 11/750 computer with a sampling rate of 20 KHz. Since the main purpose of this experiment was to investigate the contribution of the secondary cues, we systematically measured all the acoustic details of the original test stimuli that seemed likely to turn out to be relevant. The acoustic information is given in Tables I and II. Besides the closure duration, the duration of the preceding vowel, and the VOT, we also measured the duration of the vowel offramp (VOFF) in the first syllable.[6] Debrock (1978) found that the vowel offramp was shorter for fortis consonants than when a lenis consonant followed. Perhaps the slope of the vowel offramp could potentially be more or less steep as a function of whether a single or a double consonant followed. Following Debrock (1978) and van den Broecke & van Heuven (1983), we defined this as the point halfway in time between the 90% and 10% points of the maximum energy of the vowel. This definition is commonly used in non-speech stimuli. However, as pointed out by van den Broecke & van Heuven, it is difficult to pinpoint the exact location of maximum energy in a normal speech waveform since the steady state of a vowel is never quite level. If, in the steady state of the vowel, there was equally high energy in more than one moment in time, we took the rightmost point as the point of maximum energy. All four duration measures are given in Table I.

As suggested in Abramson (1987*a*), the longer duration of the closure for geminates may lead to a higher amplitude upon the release of a long plosive. Since the release characteristics of the relevant consonants could become especially important in ambiguous situations, we also took two RMS measures. The first measure was the RMS value of the burst and the second was that of the entire second syllable (cf. Abramson, 1987*a*). Since absolute RMS values cannot be compared, both measures were normalized with respect to the first vowel. Both acoustic measures are included in Table II. In addition to the RMS values of the burst, we also took the impulse index (Fant, 1970). This measure (IMP, intensity × duration) was obtained by taking the RMS plus 20log[length] (length in number of samples) of the burst. Similar to the other RMS measures, these values were

[6] The offramp measure was suggested by one of the reviewers.

TABLE I. Closure duration, VOT, duration of preceding vowel, and vowel offramp duration for all Bengali and Turkish words (in ms). G = geminate, NG = nongeminate

	Closure		VOT		Preceding vowel		Vowel offramp	
	NG	G	NG	G	NG	G	NG	G
Bengali								
pat̪a–pat̪:a	48.20	183.65	18.90	9.95	125.90	121.35	60.90	59.65
ʃot̪i–ʃot̪:i	91.90	194.95	18.80	19.65	122.55	109.15	42.45	48.80
t̪ɔt̪o–t̪ɔt̪:o	75.20	193.65	18.55	17.65	138.65	123.70	53.70	47.30
tʃʰot̪o–tʃʰot̪:o	46.80	176.45	21.00	19.85	118.85	115.50	43.00	42.20
pɔt̪on–pɔt̪:on	70.85	179.35	19.85	19.80	138.70	129.30	53.55	43.10
bʰit̪i–bʰit̪:i	87.60	194.70	19.30	21.20	118.65	131.35	34.30	45.80
Mean	70.09	187.13	19.40	18.02	127.22	121.73	47.98	47.81
Turkish								
ata–at:a	60.90	195.90	31.50	25.60	111.10	118.60	34.40	38.10
batı–bat:ı	78.10	201.10	40.00	25.80	124.60	118.90	30.30	31.30
ete–et:e	62.70	210.00	37.70	32.90	115.30	116.20	40.60	34.90
eti–et:i	77.60	194.10	48.30	44.40	133.70	118.60	32.30	37.70
ota–ot:a	62.30	201.90	44.00	25.00	119.00	127.30	26.60	33.40
yata–yat:a	53.40	211.10	31.50	28.30	128.40	124.20	29.30	32.50
yatı–yat:ı	77.60	201.90	43.40	30.90	129.30	138.10	34.40	37.20
Mean	67.51	202.29	39.49	30.41	123.06	123.13	32.56	35.01

TABLE II. RMS values of burst, second syllable, and impulse index (all values normalized by subtracting from RMS of first vowel), and non-normalized impulse index, for all Bengali and Turkish words (in dB). G = geminate, NG = nongeminate

	Burst		Second syllable		Impulse index		Non-normalized impulse index	
	NG	G	NG	G	NG	G	NG	G
Bengali								
pat̪a–pat̪:a	16.0	13.0	6.0	−2.0	35.8	33.3	87.9	83.7
ʃot̪i–ʃot̪:i	16.0	13.0	3.0	0.0	35.0	38.4	87.8	90.1
t̪ɔt̪o–t̪ɔt̪:o	15.0	13.0	−1.0	−2.0	36.6	38.4	86.1	87.4
tʃʰot̪o–tʃʰot̪:o	14.0	12.0	1.0	−4.0	39.1	40.0	89.2	90.9
pɔt̪on–pɔt̪:on	10.0	6.0	−2.0	−4.0	42.2	45.7	90.4	93.9
bʰit̪i–bʰit̪:i	16.0	19.0	2.0	1.0	35.9	33.6	87.7	89.1
Mean	14.5	12.7	1.5	−1.8	37.4	38.2	88.2	89.2
Turkish								
ata–at:a	18.0	17.0	−4.0	−7.0	32.1	31.2	77.2	75.6
batı–bat:ı	25.0	20.0	−2.0	−4.0	26.9	27.9	75.8	76.4
ete–et:e	22.0	25.0	−3.0	−1.0	29.6	25.0	76.1	74.5
eti–et:i	23.0	22.0	−2.0	−2.0	30.6	31.1	77.3	79.2
ota–ot:a	23.0	25.0	1.0	0.0	29.4	22.9	80.0	75.2
yata–yat:a	18.0	19.0	−4.0	−3.0	30.1	30.2	79.1	77.6
yatı–yat:ı	24.0	18.0	−2.0	−3.0	28.6	31.3	77.4	81.0
Mean	21.9	20.9	−2.3	−2.9	29.6	28.5	77.6	77.1

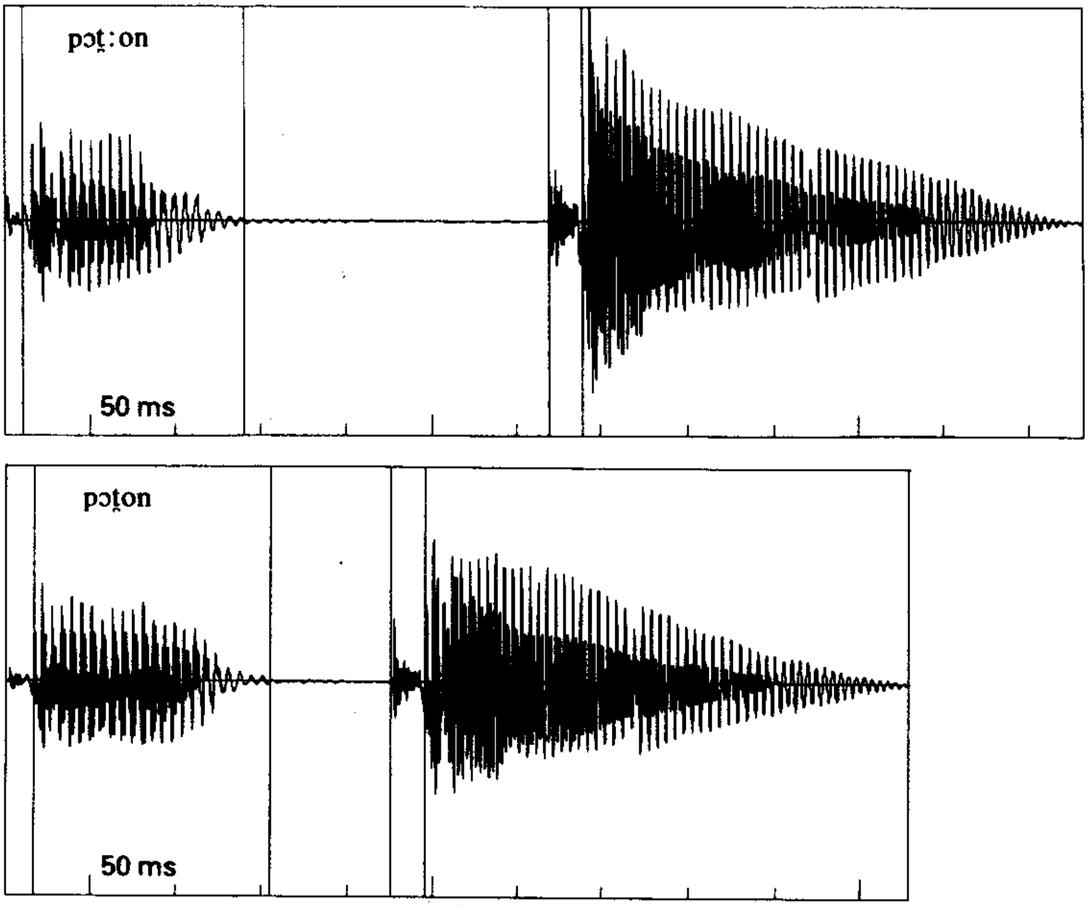

Figure 1. Sample waveforms for the geminate/non-geminate pair pɔt̪on–pɔt̪:on indicating the end points of the closures.

normalized (NIMP) with respect to the first vowel. Normalized (NIMP) and non-normalized IMP values are also listed in Table II.

Pairwise comparisons of each acoustic dimension showed significant differences for closure duration in both languages ($p < 0.001$), VOT in Turkish ($p < 0.01$), and RMS values of the second syllable ($p < 0.05$) in Bengali.

To construct the test stimuli, using a waveform editor, we replaced the silent interval corresponding to the closure of both the non-geminate and the geminate stop consonants by silence intervals of 60, 100, 110, 120, 130, 140, 150, 160, 170, 180, and 200 ms. Sample waveforms for a geminate/non-geminate pair indicating the end points of the closure durations are given in Fig. 1.[7]

We decided on these intervals on the basis of our production data (Lahiri & Hankamer, 1988) and a pilot perceptual study which showed that intervals shorter than 100 ms would always lead to non-geminate responses, and intervals longer than 180 ms would always be identified as geminates.[8] Therefore, we took equal intervals of 10 ms between these two durations for the manipulated stimuli, with 60 and 200 ms as the end points to more closely match the original production stimuli. Thus

[7] We took the silent interval representing the closure duration to begin when the vowel had completely died out. Thus any information associated with the preceding vowel would be retained in the part of the signal that we did not manipulate.

[8] In our earlier production study (Lahiri & Hankamer, 1988), involving several speakers of each language, we found considerable variation from speaker to speaker in the absolute length of the closure durations of the stops, which possibly is correlated with rate of utterance and other factors which are difficult to control in cross-speaker comparisons (cf. Lisker, 1958, p. 299). For the present experiment we created the Bengali stimuli from originals with durations shorter than the norm for Bengali, but very similar to those of our Turkish originals.

11 new stimuli were created for each word of a pair, so that including the originals we had a total of 144 test stimuli in Bengali and 168 in Turkish. For both experiments we made three separate sets of randomizations of the stimuli, ensuring that no sequences of the same word or of a word and its (non-)geminate counterpart occurred. The test items were preceded by two sets of 10 practice items which consisted of manipulated instantiations of one "non-geminate" word and one "geminate" word as well as their original counterparts.

The stimuli were then recorded on audiotape (using an Uher CR240 cassette recorder and AGFA CrO_2 Stereochrome HD60 cassettes). Three test tapes were made, each containing the practice items and one of the three randomized versions of the test items.

2.1. *Subjects*

All the subjects were native speakers of the test language. In the Bengali experiment there were seven subjects per randomization. The experiment was run in Calcutta, usually in a residential environment. For the Turkish experiment, there were seven to nine subjects per randomization; the experiment was run in Nijmegen, The Netherlands, in the local club house for Turkish employees or at the subjects' homes. The subjects listened to the tapes through Sennheiser HD222 or HD224 headphones. Their task was to write down the word they thought they heard (the experiment was free choice). The Bengali subjects used their native script, the Turkish subjects the standard Turkish orthography.

3. Results

The Bengali subjects (21 in all) responded correctly to all the original stimuli, and for the manipulated stimuli, they always wrote down an orthographically correct version of one of the words of the relevant pairs.

Some of the Turkish subjects, on the other hand, responded incorrectly to some of the original stimuli or responded with words containing a consonant other than the ones recorded (geminate and non-geminate voiceless dental). Some responded with non-standard spellings. This difference between the Bengali and Turkish subjects is probably a reflection of the fact that all the Bengali subjects were college educated, while the Turkish subjects, although literate, typically had only a few years of schooling.

The subjects' responses were scored according to whether they contained a non-geminate or a geminate consonant. Five of the Turkish subjects had less than 80% correct responses to the original stimuli, and these five were not included in the subsequent analysis. All Turkish subjects included (18 in all, six per randomization of the stimuli) scored 85% or better on responses to the original stimuli.

Figures 2 and 3 show the mean percentage geminate responses at different closure durations to Bengali and Turkish stimuli. The solid line is the curve indicating geminate responses to stimuli created from original non-geminates; the broken line is the curve indicating geminate responses to stimuli created from original geminates.

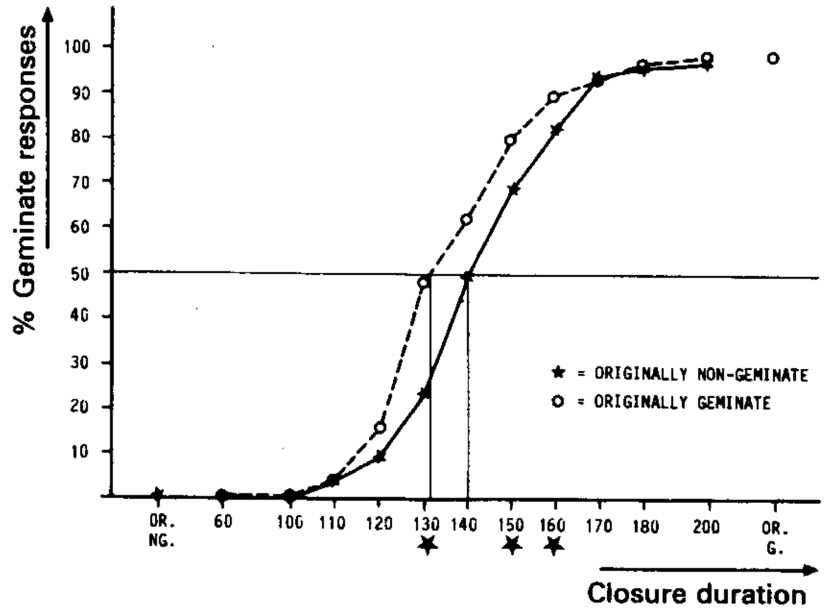

Figure 2. Mean percentage geminate responses at different closure durations to Bengali stimuli. The solid line is the curve indicating geminate responses to stimuli created from original non-geminates; the broken line is the curve indicating geminate responses to stimuli created from original geminates.

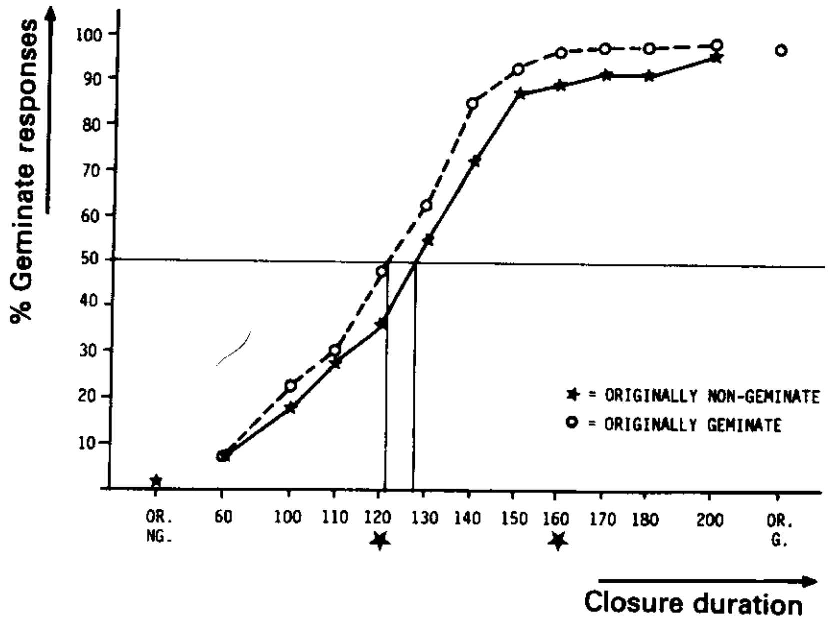

Figure 3. Mean percentage geminate responses at different closure durations to Turkish stimuli. The solid line plots responses to the stimuli created from original non-geminates and the broken line indicates responses to stimuli from original geminates.

4. Discussion

These results are superficially similar to those of previous incremental duration studies. The response curves are not identical, responses to stimuli created from original geminates crossing over from non-geminate to geminate about 8 ms earlier, in both languages, than responses to stimuli created from original non-geminates (at the 50% point, 10 ms for Bengali and 6 ms for Turkish).

The differences appear to be significant, though marginally so: a t-test on the means showed the Bengali responses to differ significantly, at the 0.05 level, at 130, 150, and 160 ms, but not at other points on the x-axis. An asterisk below a particular CD value indicates that the geminate responses at that value differ significantly depending on whether the original stimulus was a geminate or a non-geminate. The Turkish responses differ significantly (at the 0.05 level) at 120 and 160 ms. Thus significant differences are found in roughly the same range, between 120–160 ms, for both sets of stimuli. The Bengali curves are slightly displaced to the right compared to the Turkish curves (approx. 10 ms). We must conclude, following Lisker and Elugbe & Hombert, that when the closure duration cue is ambiguous, secondary cues contribute information which subjects are sensitive to in making a forced categorization.

There are, however, differences between these results and those presented in the other studies. First, the displacements between the curves in our results are smaller than those found by Lisker (1957) and Elugbe & Hombert (1975). Lisker's curves (for the contrast between voiced and voiceless stops in post-stress intervocalic position in English) are displaced by 30 ms along the duration axis, and Elugbe & Hombert's (for the fortis–lenis opposition) by about 20, while ours are less than 10 ms for both languages.

Second, in the earlier studies only one pair of words was manipulated; our results are averaged over several pairs, so we may also examine pair-wise comparisons. Figures 4 and 5 give the response curves for each word pair. As in the averaged curves, an asterisk below a particular CD value indicates that the geminate responses at that value differ significantly (at the 0.05 level) depending on whether the original stimulus was a geminate or a non-geminate. The response curves for different minimal pairs in Bengali and Turkish show considerable variation. The word pairs tʃʰo̪to/tʃʰo̪t:o and t̪o̪t̪o/t̪o̪t̪:o in Bengali, and ete/et:e, batɩ/bat:ɩ and ata/at:a in Turkish do not show significant differences at *any* CD value. For the other word pairs, no single CD value can be isolated as consistently showing a significant difference.

Let us now consider what secondary cues might be present in the signal which could possibly contribute to differentiating the consonants. Comparing the values in Tables I and II, we can see that there are few obvious differences other than closure duration between the two sets of consonants that were used as stimuli for our experiments. As we noted earlier, the only acoustic measures that were found to be significantly different in the original stimuli were VOT in Turkish and the RMS difference between preceding and following vowel in Bengali.

In our earlier study, with different subjects, we found a significant overall difference in the length of the preceding vowel in Bengali, though the difference was not significant for every subject. It appears that we can eliminate the vowel length difference as the operative secondary cue in the current experiment, because it does not correlate with differences in the response curves.

Compare the values in Table I and the matching response curves in Fig. 4 (individual curves in Bengali). The pairs bʰit̪i-bʰit̪:i and oti-ot:i have similar response curves in the critical range between 120–160 ms and we might expect matching differences in the secondary cues. For both pairs the most visible difference is in the duration of the vowel in the initial syllable. In comparing the first pair we find a 12.7 ms difference in duration, the vowel before the non-geminate being shorter. The exact opposite situation exists for the second pair. There is almost the same difference in vowel duration (13.4 ms), but here the vowel preceding the non-geminate is *longer* than that preceding its geminate counterpart.

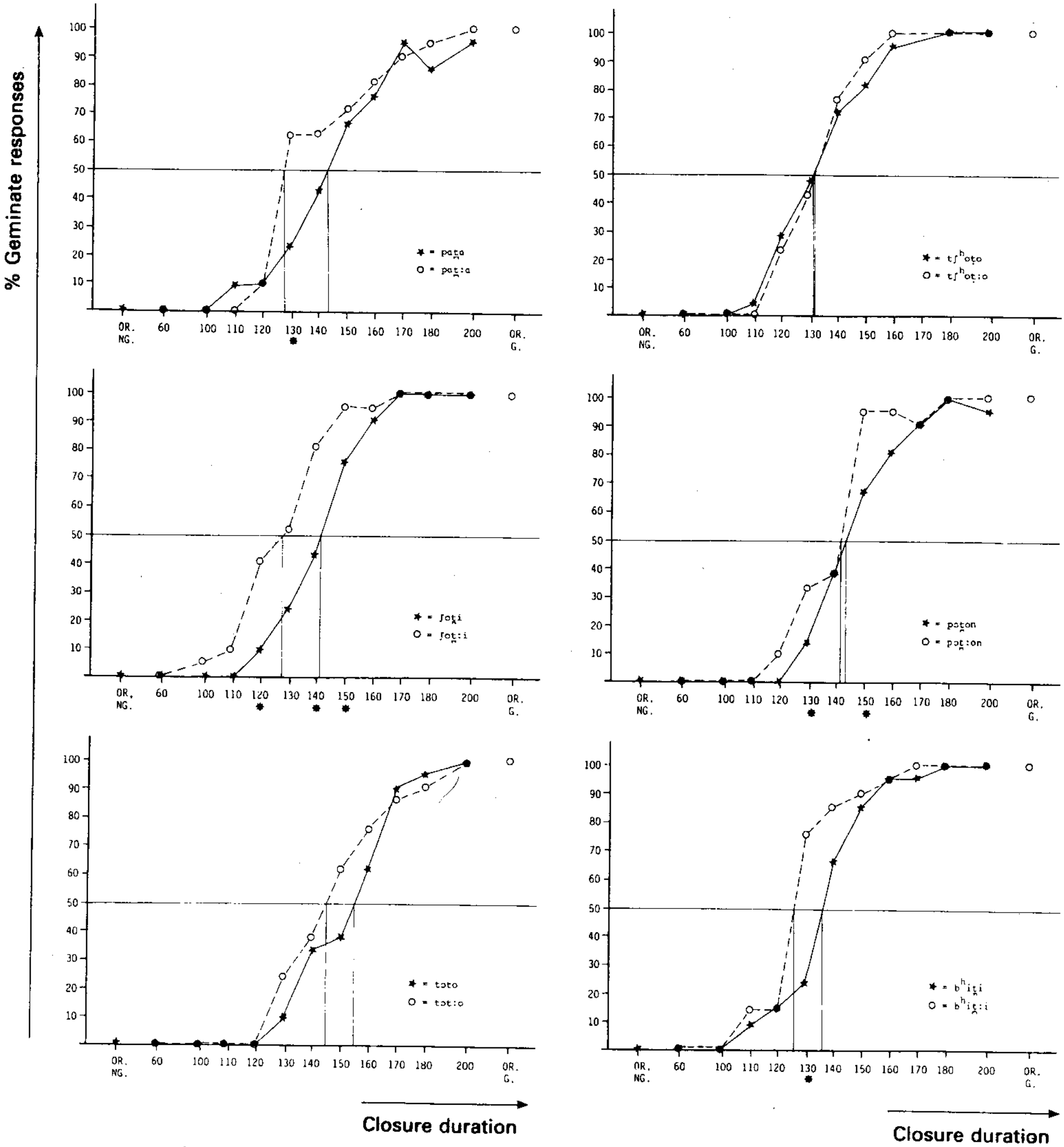

Figure 4. Percentage geminate responses to Bengali stimuli for each word pair. The solid lines plot responses to the stimuli created from original non-geminates and the broken lines indicate responses to stimuli from the corresponding geminates.

If the vowel duration cue was being used by listeners for the ambiguous stimuli, the response curves for the two word pairs should have been reversed rather than being identical. Consequently, it does not appear that it can be the vowel duration that is serving as the secondary cue.

In the data of the current experiment, the only factor besides closure duration that was found to be significant in Bengali was the RMS difference in the second

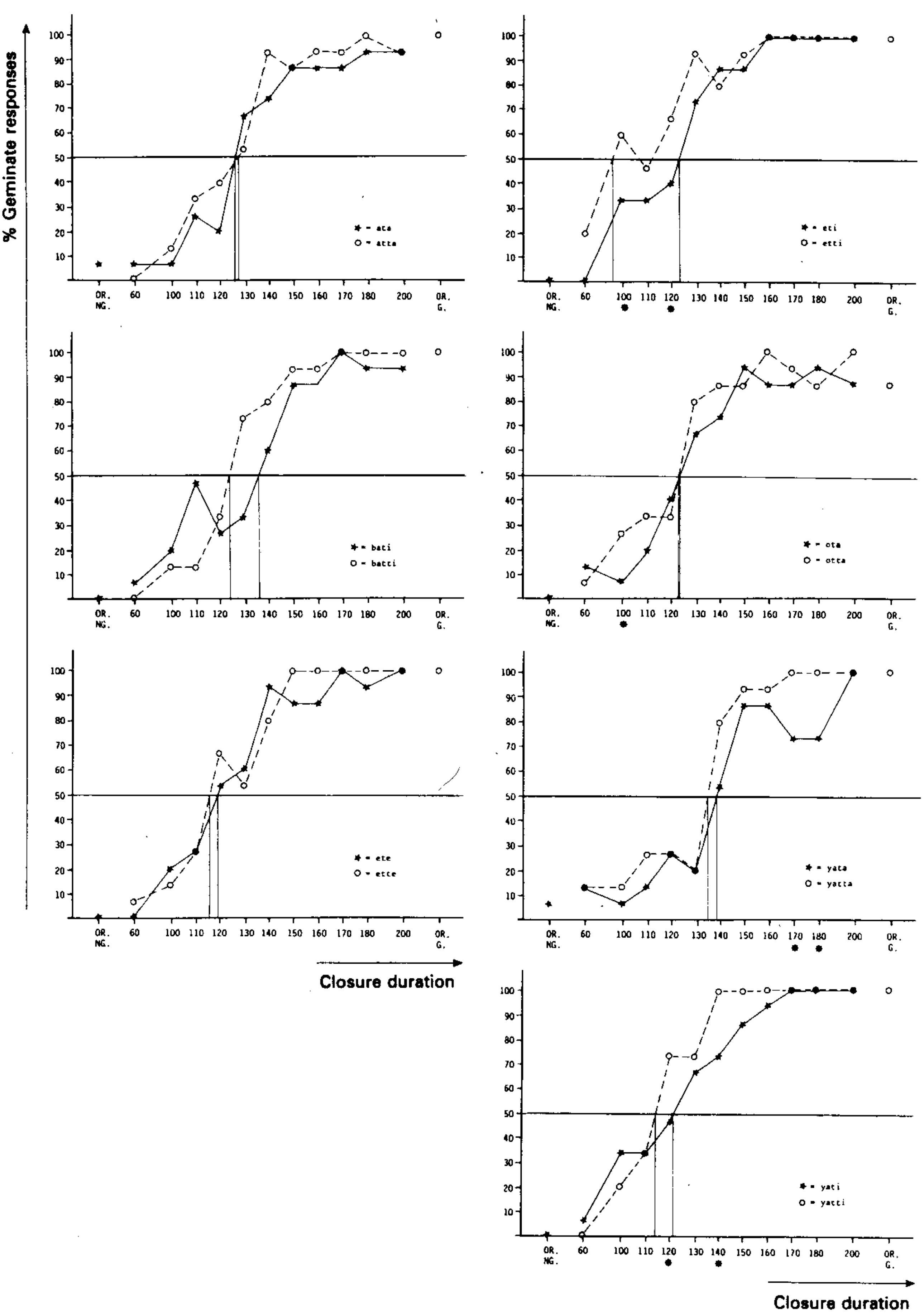

Figure 5. Percentage geminate responses to Turkish stimuli for each word pair. The solid lines plot responses to the stimuli created from original non-geminates and the broken lines indicate responses to stimuli from original geminates.

syllable (cf. Table II). Even this, however, does not appear to be a good candidate for a consistent secondary cue. The pairs bʰit̪i/bʰit̪:i, tʃʰot̪o/tʃʰot̪:o, and pat̪a/pat̪:a show increasing differences in RMS values—1.0, 5.0 and 8.0 dB respectively. The response curves, however, do not show corresponding differences. The pairs pat̪a/pat̪:a and bʰit̪i/bʰit̪:i (which show the greatest and least differences in RMS) significantly differ only at one CD value (130 ms), while the words tʃʰot̪o/tʃʰot̪:o, with an intermediate difference in RMS, have almost identical response curves. Similarly, from the differences in the response curves, it is not possible to predict the differences in RMS values due to the lack of any obvious correspondence. The pair which shows the maximum significant differences in the curves is ʃot̪i/ʃot̪:i. The 3.0 dB difference in RMS, however, falls in between the RMS difference in tʃʰot̪o/tʃʰot̪:o (5.0 dB with identical curves) and bʰit̪i/bʰit̪:i (1.0 dB with the curves differing significantly at one CD value).

The Turkish data are more variable than the Bengali, due perhaps to the differences in the subject pools mentioned in Section 3. The only measure other than CD that differs significantly is the VOT; however, there is great variation in the VOT difference from one pair to another (ranging from 3.2 to 19.0 ms). The three pairs with the most widely differing values for VOT (batı-bat:ı, ota-ot:a, yatı-yat:ı) do not show significantly different response curves compared with the other pairs. In fact, the pairs yatı-yat:ı and yata-yat:a appear to contribute equally to the general bias toward geminate responses to original geminate stimuli, even though their VOTs differ greatly (13 ms *vs.* 3 ms), and the pair with the greatest VOT difference (ota-ot:a, 19 msec) shows virtually no difference in the response curves in the range 120–160 ms, where the average curve differs most significantly.

The data were also examined to determine whether any particular acoustic dimension is statistically correlated with the percentage geminate responses. For each pair, the difference in the eight acoustic measures (see Tables I and II) was correlated with the difference in percentage geminate responses at the points where there were overall significant differences (130, 150 and 160 ms for Bengali, and 120 and 160 ms for Turkish). The correlations do not show a systematic pattern. In Bengali, at 150 ms the NIMP measure correlates significantly with the geminate responses ($p = 0.017$, $r = 0.903$), while at 160 ms, the VOFF measure shows significant correlation ($p = 0.013$, $r = 0.905$). No other correlations were significant. In Turkish, none of the correlations were significant. This corroborates the conclusions we have drawn from inspection of the individual response curves: nothing that we were able to measure in the acoustic properties of the stimuli systematically correlates with the differences in the response curves, and the nature of the hidden cue or cues remains a mystery.

5. Conclusion

Our experiments have shown that the responses of both Bengali and Turkish subjects are biased by secondary features of the acoustic signal when the closure duration cue is in the ambiguous region between 120–160 ms. Examination of the properties of the stimuli, however, failed to reveal any feature which could systematically be counted on to provide a secondary cue. Hence we do not know the source of the bias, nor even whether it is due to a single feature in either language. Possibly the bias is due to a combination of cues, each by itself too subtle for our measurements to detect.

We can, of course, rule out lexical bias of the kind found by Ganong (1980), since both members of our minimal pairs were real words; similarly, it seems safe to assume that there is no significant contamination by frequency effects, given the number of pairs examined and the similar results for the two unrelated languages.

It is of course not surprising that there should be secondary features associated with closure duration in the production of geminate and non-geminate stops. What is surprising, in the light of earlier studies such as Lisker (1957), is the comparatively small effect in perception produced by these secondary cues, along with their variability and subtlety. Our failure to isolate any single cue, for either language, which might account for the perceptual bias is a distinctly negative result; we have, however, learned some things which can be taken as positive results. First, that only closure duration, of the several features of the signal that we measured, serves as a reliable cue to the geminate *vs.* non-geminate distinction, across speakers and word pairs in these two languages. Second, that the region in which the bias is evident is confined to the range 120–160 ms, a region in which no naturally occurring stops exist in the sample utterances used in the experiment.

In interpreting the results of our study, we address two questions: first, what is the significance of the much smaller displacement along the time axis that we found compared with those of the Lisker (1957), Elugbe & Hombert (1975), and Abramson (1987*b*) studies? and second, what do the experiments tell us about cues actually used in perception of the geminate *vs.* non-geminate contrast in normal speech? The answer to the first question, we believe, is that there is a relation between the phonological status of a duration distinction and its phonetic manifestation. The opposition we have studied is one intrinsically linked to the syllable structure of the language: geminate consonants only appear as a final consonant of one syllable joined with an initial consonant of another syllable. The Lisker (1957) and Elugbe & Hombert (1975) studies established that duration provided a significant cue in certain oppositions which, while they certainly involve duration as a relevant cue, are phonologically not the same type as the geminate *vs.* non-geminate opposition found in Turkish and Bengali. We cannot tell, from the descriptions presented, whether the oppositions there ever involve heterosyllabic consonants, but they certainly appear to be oppositions that can be found in utterance initial position. Perhaps it is the case that the difference in effect of secondary cues correlates with the possibility of contrast in other than intervocalic position.

Consider, for example, the difference between our results and those of Abramson, who found a displacement of between 20 and 30 ms in Pattani Malay. Pattani Malay has a long *vs.* short opposition in word-initial, and hence potentially utterance-initial, position, where for voiceless stops closure duration could not be a cue at all. For any language which has the opposition for voiceless stops in other than intervocalic position, it is clear that it cannot be distinguished in such a position by closure duration alone, even if that is a strong cue in some environments. As Abramson (1987*b*) himself points out (p. 149), languages like Pattani Malay where a consonant length opposition is found in word-initial position are relatively rare. Languages like Turkish and Bengali, with the distribution of geminates such that closure duration alone could always be a sufficient cue, represent the common case.

To the second question, the answer is that the silent gap of the closure is by far the most salient and perhaps the only dependable cue. Our results show that subjects can be sensitive to acoustic cues present in artificially created stimuli, where the CD cue is rendered uninformative. They do not tell us anything about whether these secondary cues are relevant in the categorization of naturally occurring consonants in real speech, since the effects of the secondary cues are found only in a narrow region within the no-man's-land on the CD-continuum where no naturally occurring stops exist.[9]

We know now that there must be cues other than closure duration present in the naturally occurring samples which bias the percept when the duration cue is in the ambiguous region. The extent of the bias, however, is smaller than that found in studies of other contrasts involving duration and perceptible only in the aggregate sample, not reliably in the responses to stimuli from individual word pairs. If we had studied one such pair, taken at random, we might have found the bias and we might not. We conclude that, whatever the secondary cues are, their effect is surprisingly small even when the duration cue is maximally ambiguous, and the variations in the response curves to individual pairs indicates that they are probably not very reliable cues. There is no reason to believe that they have any value at all in the perception of real speech, where the dominant duration cue clearly determines the category of the stop.

There is, consequently, no reason not to believe that the geminate *vs.* non-geminate opposition in languages like Turkish and Bengali, where the distribution of geminates is restricted to intervocalic position, is systematically a contrast in nothing but length, the acoustic measure for which is the duration of closure.

The waveform editing system at the Max-Planck Speech Laboratory was developed by Henning Reetz.

This research was partly supported for Jorge Hankamer by the Institute for Turkish Studies, the Santa Cruz Syntax Research Center, and the Max-Planck-Institute for Psycholinguistics.

We would like to thank John Ohala for his comments and suggestions on an earlier study, which led directly to the investigation presented here, and Herbert Schriefers for comments on a draft of the paper which led to significant improvements. A version of this paper was presented at LSA 1987. We would like to thank the participants for helpful comments. We owe thanks also to two anonymous reviewers, whose comments led to substantial improvements in the paper.

References

Abercrombie, D. (1967) *Elements of general phonetics*, Edinburgh: Edinburgh University Press.

Abramson, A. (1987*a*) Word-initial consonant length in Pattani Malay. In *Proceedings XIth international congress of phonetic sciences*, vol. 6, pp. 68–70.

Abramson, A. (1987*b*) The perception of word-initial consonant length: Pattani Malay, SR-91, Haskins Laboratories, New Haven, pp. 149–156.

Alderson, A. D. & Iz, F. (1959) *The concise Oxford Turkish dictionary*. Oxford: Oxford University Press.

[9] A small caveat is necessary here. No naturally occurring stops are found in the transition region in the sample stimuli used for the experiment. In other data recorded during our earlier study, there were instances of non-geminate stops with closure durations in what we are here calling the ambiguous region. These instances, however, were from the speech of a speaker (of Bengali) whose stops, whether geminate or non-geminate, were systematically longer than those of other speakers in our sample. We know [illegible]thing about how listeners adjust for differences in tempo or in individual speaker variation.

van den Broecke, M. P. R. & van Heuven, V. J. (1983) Effect and artifact in the auditory discrimination of rise and decay time: speech and nonspeech, *Perception and Psychophysics*, **33**(4), 305–313.

Catford, J. C. (1977) *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.

Debrock, M. (1978) An acoustic correlate of the force of articulation, *Journal of Phonetics*, **5**, 61–80.

Dev, A. T. (1973) *Student's favorite dictionary*, Calcutta: Dev Sahitya Kutir.

Elugbe, B. & Hombert, J. M. (1975) Nasals in Ghuotuo: /Lenis/ or [Short]? in *Nasalfest* (C. A. Ferguson, L. M. Hyman & J. Ohala, editors), Stanford University, California.

Fant, G. (1970) *Acoustic theory of speech production* (second printing). The Hague: Moulton.

Fukui, S. (1978) Perception for the Japanese stop consonants with reduced and extended durations, *Bulletin of the Phonetic Society of Japan*, no. 59, 9–12.

Ganong, W. F. (1980) Phonetic categorization in auditory word perception, *Journal of Experimental Psychology: Human Perception and Performance*, **6**(1), 110–125.

Lahiri, A. & Hankamer, J. (1988) The timing of geminate consonants, *Journal of Phonetics*, **16**, 327–338.

Lehiste, I. (1970) *Suprasegmentals*. Cambridge, MA: MIT Press.

Lisker, L. (1957) Closure duration and the intervocalic voiced-voiceless distinction in English, *Language*, **33**(1), 42–49.

Lisker, L. (1958) The Tamil occlusives: short vs. long or voiced vs. voiceless?, *Indian Linguistics, Turner Jubilee Volume*, I, 294–301.

Obrecht, D. H. (1965) Three experiments in the perception of geminate consonants in Arabic, *Language and Speech*, **8**, 31–41.

Pickett, J. M. and Decker, L. R. (1960) Time factors in perception of a double consonant, *Language and Speech*, **3**, 11–17.

Repp, B. (1978) Perceptual integration and differentiation of spectral cues for intervocalic stop consonants, *Perception and Psychophysics*, **24**(5), 471–485.

Repp, B. (1983) Bidirectional contrast effects in the perception of VC-CV sequences, *Perception and Psychophysics*, **33**(2), 147–155.

Appendix. Bengali and Turkish words[10]

Non-geminate		Geminate	
Bengali			
pat̪a	leaf	pat̪:a	whereabouts
ʃot̪i	virtuous wife	ʃot̪:i	truth
t̪ɔt̪o	so much	t̪ɔt̪:o	fundamental truth
tʃʰoṭo	small, petty	tʃʰoṭ:o	very tiny
pɔt̪on	downfall	pɔt̪:on	settlement, foundation
bʰit̪i	fear	bʰit̪:i	base, root
Turkish			
ata	horse (dat)	at:a	horse (loc)
batı	west	bat:i	sink (past)
ete	meat (dat)	et:e	meat (loc)
eti	meat (acc)	et:i	do (past)
ota	grass (dat)	ot:a	grass (loc)
yata	yacht (dat)	yat:a	yacht (loc)
yati	yacht (acc)	yat:ı	lie down (past)

[10] For reference, the reader may consult the following dictionaries: for Bengali, Dev (1973); for Turkish, Alderson & Iz (1959).