615

Journal of the International Phonetic Association (1986) 16, 8-16

The perception of word-initial consonant length: Pattani Malay *

ARTHUR S. ABRAMSON
Dept. of Linguistics, U-145, University of Connecticut, 341 Mansfield Rd., Rm. 230, Storrs, CT 06268, U.S.A., and Haskins Laboratories, New Haven, CT 06511-6695, U.S.A.

Introduction. In experimental phonetic research, we often come across cases of multiple perceptual cues to a phonemic distinction, even though this distinction may be seen traditionally as dependent on some single phonetic feature. The question arises as to the relative power of these cues. How equally do they share the burden of communicative relevance? For example, among the several cues that emanate from the timing of the valvular action of the larynx (Lisker and Abramson (1965); Abramson (1977)), fundamental-frequency perturbations (House and Fairbanks (1953)) have been shown by some studies, apparently starting with Haggard, Ambler and Callow (1970) and Fujimura (1971), to help in the perceptual differentiation of voiced and voiceless stop consonants; however, recent work (Abramson and Lisker (1985)) suggests that this cue has very limited efficacy compared with other acoustic consequences of voice timing.

The present study is meant to combine the foregoing interest with an attempt to shed further light on the perceptual basis of LENGTH contrasts as found in many languages in which phonologically distinctive functions are borne by the relative durations of vowels (Abramson (1962); Lehiste (1970)) or of the closures and constrictions of consonants (Lehiste (1970)). Discussions of the feature of consonant length usually focus on intervocalic consonants, as in Italian and Estonian. In such cases, it is easy to demonstrate the existence of differences in consonant-closure duration and their perceptual relevance. What is uncommon is to find a language with such a distinction in word-initial, and thus potentially utterance-initial, position. In length distinctions in any context, it is not unlikely that other acoustic features will covary with the duration of the relevant span of speech. Perhaps some of these concomitant features serve as cues together with duration or, in certain circumstances, instead of duration.

Pattani Malay. Pattani Malay, the dialect of Malay spoken by some 600,000 ethnic Malays in southeastern Thailand, has a length distinction for ALL consonants in word-initial position (Chaiyanara (1983)), thus for all consonant classes of the language. The language was first called to my attention by the fieldwork of

Christopher Court and Jimmy G. Harris. Here are some word pairs with the contrast:

/ labɔ /	'to make a profit'	/ l:abɔ /	'spider'
/ makɛ /	'to eat'	/ m:akɛ /	'to be eaten'
/ siku /	'elbow'	/ s:iku /	'hand-tool'
/ bulɛ /	'moon'	/ b:ulɛ /	'many months'
/ kato? /	'to strike'	/ k:ato? /	'frog'

For all consonants with acoustic excitation of any kind during the closure or constriction, it is obvious that closure duration alone could be enough to differentiate the members of each pair in both production and perception. The first four pairs of examples are of that type, while the fifth pair, with voiceless unaspirated stops, is not. That is, in pairs of the last type, the difference in the closure durations appears only as shorter or longer medial silent gaps when the words are embedded in utterances.

For the two speakers examined in preparation for this perceptual study, the closure durations of the long consonants in both initial and medial position are on average three times longer than those of the short consonants. A more detailed presentation of duration ratios is given elsewhere (Abramson (1987)). For the voiceless stops and affricates, which are all unaspirated, measurements of duration can be made, of course, only in utterance-medial position.

If indeed the timing of the closure is the articulatory mechanism underlying the distinction, wherever the difference in closure duration is audible, it ought to be a sufficient auditory cue. In intervocalic contexts, the abrupt spectral shifts between the vocalic portions of the signal and the consonantal closure rather clearly define the acoustic span corresponding with the interval of the closure. It is in utterance-initial position that the question arises as to the sufficiency of this cue. One might expect that nasals, laterals, and fricatives with high radiation of sound during closure would be as well distinguished by listeners in initial position as in intervocalic position. Voiced stops and affricates might be an intermediate case. There is audible glottal pulsing during the closures, but its amplitude may be too low for completely reliable differentiation based on two values of voicing lead. Voiceless plosives, however, with no acoustic excitation in the closure, must surely be differentiated in initial position by other cues, possibly ones that are themselves a function of the temporal articulatory feature. These might be intensity of stop-release burst, rate of formant transitions, or fundamental-frequency perturbations. A separate study, started somewhat later, has shown that in

disyllabic words, if the first syllable begins with a long stop consonant, the amplitude of that syllable is significantly greater than that of the second syllable, but that is not so if the word begins with a short consonant (Abramson (1987)).

Experiments

Experiment 1. To establish a baseline against which to do manipulative experiments, it was first necessary to assess the perceptual robustness of the length distinction. Thirty two Pattani Malay words forming 16 short-long pairs were recorded in isolation in two random orders by two women who, although also fluent in Thai, were native speakers of Pattani Malay and lifelong speakers of it. The resulting two test orders, each one containing two tokens of each word, were played through headphones to 21 students, all native speakers, on the Pattani campus of the Prince of Songkhla University. Because this language has not been reduced to writing for popular use, it was necessary to provide the test subjects with answer sheets with a pair of possible responses written in Thai next to each item number. The possible responses were short Thai glosses for the Malay words. For each item, the subject encircled whichever of the two glosses he or she thought appropriate for the spoken word. A member of the faculty[1], also a native speaker of the language, gave the instructions on procedure and went over all the glosses to be sure that there were no misunderstandings. This method was used for all subsequent tests and seemed to cause no trouble.

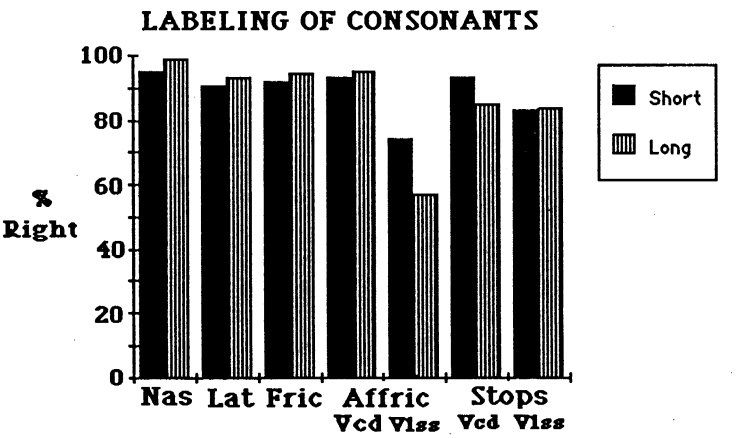

FIGURE 1. Experiment 1.

Identification of natural utterances of isolated words beginning with short and long consonants.

The results of the baseline experiment are given in Figure 1. As can be seen, the nasals, laterals, and fricatives were all identified quite well. Next, but not much lower in percentages correct, come the stops. Worst are the affricates, especially the voiceless ones, which were not labeled much better than chance. This led me to decide not to work further with the affricates in the present investigation.

Experiment 2. For this stage of my perceptual research on the distinction, I have limited my manipulations to the obvious variable of closure duration[2]. If it is indeed a major cue to the distinction, we ought to be able to bring about a shift in percept by big enough changes in closure duration. Wishing to start with consonants with very audible voicing during closure, I chose a pair of words distinguished by short and long /ɪ/: /labɔ/ 'to make a profit' versus /lːabɔ/ 'spider'. By means of wave-form editing, I shortened the initial lateral resonance of the isolated long member of the pair from its original duration of 183 msec to 63 msec in twelve 10-msec steps. Thus, the shortest variant was shorter than the original short /ɪ/ of 72 msec, which itself was not used in the experiment. I made two more stimuli by cross-splicing the lateral resonances between the two original words. That is, I removed the lateral resonances, replacing the original long one with the short one and the original short one with the long one. The resulting 15 stimuli were recorded three times each into a random order for presentation to the 21 subjects.

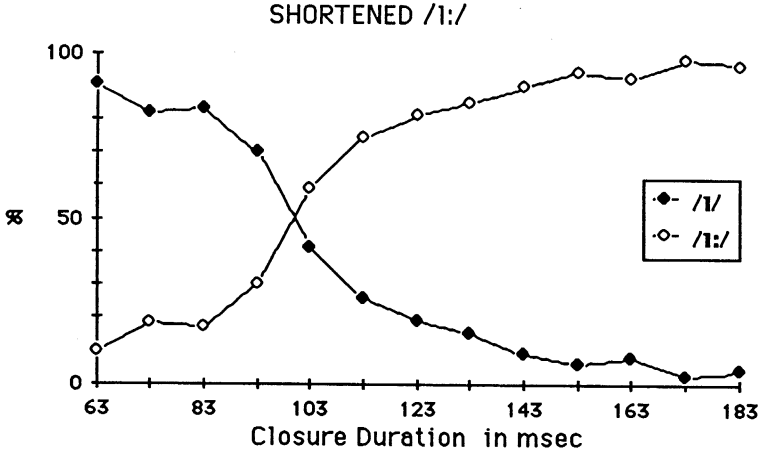

FIGURE 2. Experiment 2.

Identification functions for original long /lː/ and shortened variants. The responses to the cross-spliced stimuli are not included.

The results for the shortened / iː / series, without the cross-spliced stimuli, are given in Figure 2. The range of closure durations is shown along the abscissa, and the labeling percentages along the ordinate. The crossover zone between the two categories is centered around 100 msec. Clearly, duration is a sufficient and powerful cue to the distinction. The cross-spliced stimuli were heard virtually 100% of the time as the original words from which the resonances had been cut; this suggests that features of the constriction release had little or no cue value.

The next two experiments examined the crucial case of voiceless stops, which, of course, could be subjected to closure-duration changes by the method of wave-form editing only in intervocalic position in a neutral Pattani Malay carrier sentence suitable for the two words in question. The words chosen for both experiments were / paka / 'to use' and / pːaka / 'usable'. The plan was to shorten the long consonant and lengthen the short consonant to test, once again, the effects of variation in closure duration and, indirectly, effects of stop release.

Experiment 3. In this experiment, I shortened the closure of long / pː / from its original duration of 182 msec in fourteen 10-msec steps in a carrier sentence. The shortest variant at 42 msec was a bit shorter than the original short / p / of 47 msec, which itself was not used in the experiment. The resulting 15 stimuli were recorded three times each into a random order for presentation to the 21 subjects in the carrier sentence.

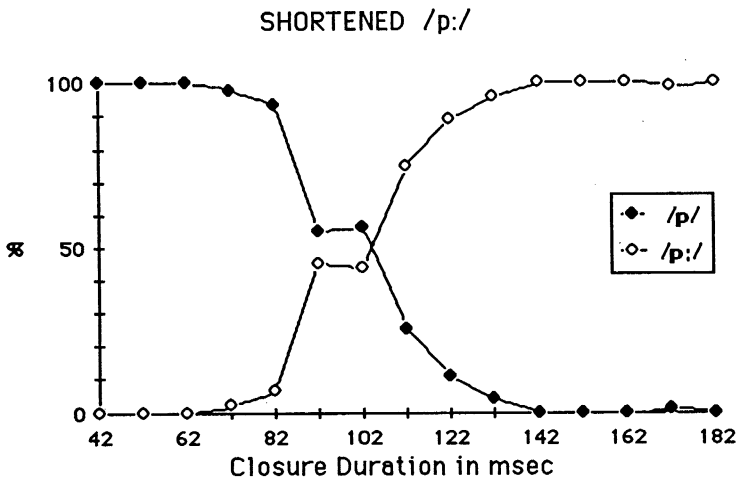

FIGURE 3. Experiment 3.
Identification functions for original long / pː / and shortened variants.

The results of Experiment 3 are given in Figure 3. The sufficiency of relative duration as a cue to the length distinction is demonstrated here for voiceless intervocalic stops. The unexplained leveling of the responses for the items at 92 and 102 msec may well be due to an artifact that remains to be uncovered. If we ignore the latter, the 50% crossover point is at 104 msec; a curve-fitting procedure, or better, a replication with new stimuli might yield a slightly earlier crossover.

Experiment 4. This experiment differed from Experiment 3 in that / paka /, the short member of the pair, was the starting point. I lengthened the closure of short / p / in the carrier sentence from its original duration of 47 msec in fourteen 10-msec steps. The longest variant was 187 msec, a bit longer than the original long / p: / of 182 msec, which itself did not appear in the experiment. This simply required using our waveform-editing program to add increments of time to the middle of the closure gap. Again, the resulting 15 stimuli were recorded three times each into a random order and played to the 21 subjects for identification as words in the carrier sentence.

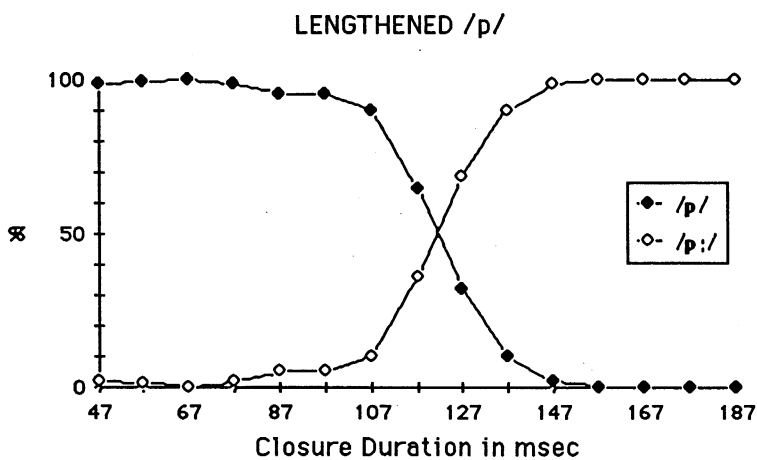

FIGURE 4. Experiment 4.
Identification functions for original short / p / and lengthened variants.

The results of Experiment 4 are given in Figure 4. Not surprisingly, the importance of closure duration is again apparent for intervocalic voiceless stops. Here, however, the perceptual crossover point is at 120 msec, considerably later than the one in Experiment 3, with or without the possible artifact in the latter. An

analysis of variance showed the large difference between the two crossover points to be highly significant [F (1, 20) = 27.48, p < 0.0001]. This implies the probable efficacy of one or more cues concomitant with that of closure duration.

Experiment 5. The goal of this last experiment was to explore the intermediate condition, that of the presence of quasiperiodic acoustic excitation during the closure but at a low amplitude. The word pair chosen was / gamɔʔ / 'approximately' versus / g:amɔʔ / 'to be shy'. I shortened the closure of long / g: / in a citation form of the word from its original duration of 197 msec in fifteen 10-msec steps. The shortest variant at 47 msec was just under the 50-msec duration of the closure of the original short / g /, which itself was not used in the experiment. The 16 resulting stimuli were recorded three times each into a random order and played to the 21 subjects for identification.

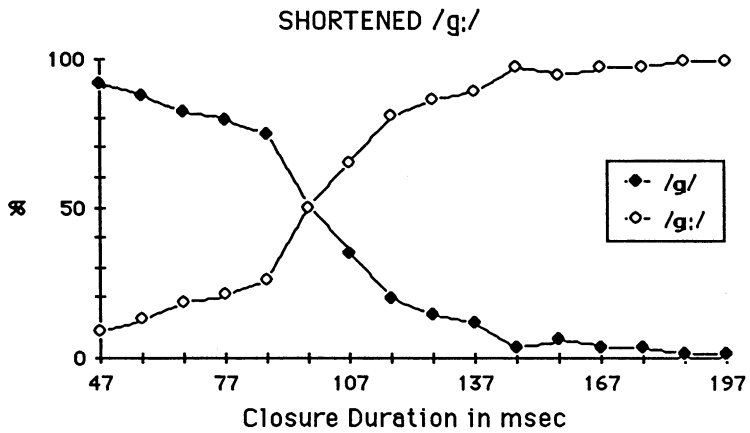

FIGURE 5. Experiment 5.
Identification functions for original long / g: / and shortened variants.

The results of Experiment 5 are given in Figure 5. The middle of the crossover zone between the two percepts is at 97 msec. Clearly, even with just low-amplitude voicing in the closures of initial stops, relative duration is a sufficient cue.

Conclusion. The word-initial consonants of Pattani Malay, except perhaps for the voiceless affricates, can be identified well as to length category in utterance-initial position. Perceptual experiments reveal the power of closure duration as a

sufficient cue in medial and, when audible, initial position for the distinction between 'short' and 'long' consonants.

The early crossover at 33% of the durational difference between utterance-initial /g/ and /g:/ seems comparable to the crossover at 31% of the difference between /ɪ/ and /ɪ:/, as contrasted with 44% in Experiment 3 and 52% in Experiment 4, both of the latter for medial voiceless stops. With so few experiments so far, we can only speculate that for utterance-initial voice-excited closures or constrictions, there is some psychoacoustic threshold below which a long consonant cannot be heard.

The results of Experiments 2 and 3 imply that for the length distinction in voiceless medial stops there is another cue, if only a secondary one, in addition to relative duration. Obviously, given the results of Experiment 1, some such feature must be at work as the PRIMARY cue in utterance-initial position. Acoustic analysis (Abramson (1987)) has shown relative amplitude to be the most promising candidate. Perceptual testing of this hypothesis is the next thing to be done.

Notes

* This is a revised and expanded version of a paper read at the 111th Meeting of the Acoustical Society of America, Cleveland, Ohio, May 12-16, 1986 (Abramson (1986)). The work was supported by NICHD Grant HD-01994 to Haskins Laboratories. The excellent help of two colleagues in Thailand, Mr. Paitoon Masmintra Chaiyanara of the Prince of Songkhla University, Pattani, and Dr. Theraphan L. Thongkum of Chulalongkorn University, Bangkok, enabled me to do the fieldwork for this study. I also wish to thank both institutions for their warm hospitality.

1. Mr. Paitoon Masmintra Chaiyanara, who is thanked in the note to the title, also went over all the words beforehand for their authenticity and for the accuracy of the glosses. He is engaged with colleagues in compiling a trilingual (Pattani Malay - Standard Malay - Thai) dictionary of the language.
2. The results of recent analytic work (Abramson (1987)) call for perceptual experiments with relative amplitude as the variable.

References

ABRAMSON, A.S. (1962). *The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments.* Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics.

ABRAMSON, A.S. (1977). Laryngeal timing in consonant distinctions. *Phonetica* **34**, 295-303.

ABRAMSON, A.S. (1986). Distinctive length in initial consonants: Pattani Malay. *Journal of the Acoustical Society of America* **79**, S27 (abstract).

ABRAMSON, A.S. (1987). Word-initial consonant length in Pattani Malay. *Proceedings of the XIth International Congress of Phonetic Sciences.* Tallinn: Academy of Sciences of the Estonian S.S.R., 68-70.

ABRAMSON, A.S. and LISKER, L. (1985). Relative power of cues: F0 shift versus voice timing. In V. Fromkin (editor), *Phonetic Linguistics: Essays in Honor of Peter Ladefoged.* New York: Academic Press, 25-33.

CHAIYANARA, P. MASMINTRA. (1983). *Dialek Melayu Patani dan Bahasa Malaysia: Satu Kajian Perbandingan dari segi Fonologi, Morfologi dan Syntaksis.* Master's thesis, University of Malaya.

FUJIMURA, O. (1971). Remarks on stop consonants: Synthesis experiments and acoustic cues. In L.L. Hammerich, R. Jakobson and E. Zwirner (editors), *Form and Substance: Phonetic and Linguistic Papers Presented to Eli Fischer-Jørgensen.* Copenhagen: Akademisk Forlag, 221-231.

HAGGARD, M.P., AMBLER, S. and CALLOW, M. (1970). Pitch as a voicing cue. *Journal of the Acoustical Society of America* **47**, 613-617.

HOUSE, A.S. and FAIRBANKS, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* **25**, 105-113.

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

LISKER, L. and ABRAMSON, A.S. (1965). Stop categorization and voice onset time. *Proceedings of the Vth International Congress of Phonetic Sciences.* Basel: S. Karger, 389-391.