Journal of Phonetics (1990) **18**, 79–92

Distinctive vowel length: duration *vs*. spectrum in Thai

722

Arthur S. Abramson and Nianqi Ren

The University of Connecticut, Department of Linguistics, Storrs, CT 06269-1145, U.S.A. and Haskins Laboratories, 270 Crown St., New Haven, CT 06511–6695, U.S.A.

Many languages are described as having a phonological distinction of length in vowels or consonants or even both. If the term is to be taken literally, we would expect to find that the underlying mechanism is control of the relative durations of the articulations for the phonemes in contrast. In languages with short and long vowels, however, it is often observed that short and long counterparts have somewhat different vowel qualities. Previous work on the vowels of Central Thai has shown that relative duration is a sufficient acoustic cue for the length distinction. Since, however, short vowels in portions of the vowel space tend to be more open, we have sought to learn how much of the perceptual burden is borne by relative duration, the major physical correlate of length, and how much by the spectrum (formant pattern), the major physical correlate of vowel quality. We have incrementally lengthened original short vowels and shortened original long vowels in minimal pairs of words embedded in sentences, thus preserving the spectral differences. The stimuli were played to native speakers for identification. The results show that the dominant cue is relative duration; however, for all the vowel pairs, the category boundary is influenced by spectral differences with, perhaps, some effect of the timing of the context as well.

1. Introduction

As with all phonological distinctions, so also with the feature of vowel length the question arises as to the likelihood that the distinction rests on a single phonetic feature. On the face of it, distinctive length seems to be little more than a matter of relative timing. Thus, in a given context, members of a "long" category are longer than their counterparts in the "short" category. For this seemingly straightforward situation, Catford, in his remarkable new textbook (1988, pp. 184–186), gives short shrift to the matter of learning to make long segments by emphasizing the need just to hold an articulatory configuration longer. We are thinking here, of course, of the prolongation of segments under voluntary control, not intrinsic length correlated, for example, with close versus open vowels or stops versus glides (Lehiste, 1970).

If the acoustic outputs of the putative length distinction could be shown to be differentiated only by relative duration, all phonologists would surely not hesitate to posit a distinctive feature of length. Indeed, the occasional, perhaps contextually

0095–4470/90/020079 + 14 $03.00/0

predictable, presence of minor concomitant acoustic features, some of them audible, might not worry these analysts. Lehiste (1970, pp. 30–33) tells us that in "quantity" languages some differences in the phonetic quality of long and short vowels can be observed, although these languages differ in how much length and quality are correlated. She illustrates the matter with plots of the first and second formants for the vowels of two languages with distinctive vowel length, Czech and Serbo–Croatian. It turns out, for example, with regard to placement in the acoustic space, that in Czech "the difference between long and short vowels is greatest for /i/ and /u/, whereas in Serbo–Croatian the greatest difference is observed in the mid vowels /e/ and /o/" (p. 32). Even in a monograph dedicated to the subject of distinctive vowel length, Durand (1946) assigns great importance to features that go together with length in a number of languages. Indeed, if vowel quality goes hand in hand with vowel length (Straka, 1959; Garnes, 1973; Catford, 1977, pp. 195–199), it may be decided that quality is paramount, or "tensity" may be invoked to "explain" the correlation between length and quality. Related to this is the general understanding that in Standard Italian and other languages with distinctively short and long *consonants,* features other than just relative duration of closure or constriction might be observed. We see, for example, that in Italian *fato* [faːto] "fate" and *fatto* [fatːo] "done", what is generally accepted as distinctive consonant length entails a reciprocal difference in vowel length. Is this an entirely redundant rhythmic feature? Are we not entitled to ask whether perceptual identification of these words is determined by two acoustic cues?

For the historical linguist, the existence at a stage of the history of a language of a feature of distinctive segment length linked to other features, even if the latter can be shown to be entailed by length, offers the possibility of diachronic sound change. If conditions are ripe, one of the concomitant features could take over the communicative function of the old primary feature, resulting in a change in the system of phonemes. For example, the Indo-European feature of distinctive vowel length is said to persist in a number of Germanic languages, among others. In the southern dialect of one of them, Swedish, experimental work (Hadding-Koch & Abramson, 1964) has shown that a difference in quality correlated with length has seemingly taken over the distinctive function in at least part of the vowel system. That is, for one pair of "short–long" vowels, acoustic cues to the distinction reside in the formant patterns rather than in relative duration, although for the other pairs tested, duration does the work. Another study (Bennett, 1968) considered the matter in a comparison of educated northern German speech and the RP accent of British English. The outlook, however, was somewhat different from that of the Swedish research and that of the present report in that Bennett did not fasten his attention upon minimal pairs allegedly distinguished by length. Instead, in both languages, he chose pairs of words differentiated not only by vowel length but also by two degrees of quality; there were short–long pairs near each other in the vowel space, i.e., with small differences in quality, and pairs farther apart in the vowel space, i.e., with greater differences in quality. Anyway, his experiments with synthetic speech yielded a general conclusion of interest here. In both languages, a pair of short and long vowels must be very close in quality for duration to be more important than spectral form (p. 76).

Let us turn now to acoustic and perceptual work on distinctive vowel length in a Southeast Asian language, Thai. For the feature of length in that language, the major acoustic differentiator seems to be relative duration (Abramson, 1962). At the same time, at least one concomitant feature can be detected. In a number of contexts, for some vowel pairs more than others, differences in quality can be heard. That is, the short member is sometimes more open than the long member. This difference in vowel quality, of course, is expected to be reflected in the formant patterns. We have sought to learn whether this other feature bears any of the communicative burden. That is, even if both acoustic features associated with this phonetic distinction are derivable from a single underlying temporal mechanism, it may still be true that listeners, having learned to avail themselves of redundancies in the signal, depend to some extent on more than one acoustic cue. Thus, the question arises as to the relative power of these putative cues. If they both carry phonetic information, are they equal, or is one of them dominant?

In addition, as implied by our remarks on historical linguistics, this work is motivated by a desire to see whether this feature of vowel quality, supposedly concomitant with a primary feature of length, has, as it were, risen in the consciousness of members of the language community to the point of being enhanced in production and established as primary.[1] Indeed, as in the Swedish case mentioned, there might be an asymmetrical condition of a mixture of features in the system of Thai vowels.

Standard Thai, which is based on the Central Thai dialect of the area around Bangkok, has a system of nine vowels that, with some contextual constraints, may be phonologically short or long, as shown here:

i	iː	ɨ	ɨː	u	uː
e	eː	ə	əː	o	oː
æ	æː	a	aː	ɔ	ɔː

Earlier work (Abramson, 1962, 1974) has shown that expressions minimally distinguished in a given context by any of these paris of vowels are well differentiated physically by relative vowel duration. In addition, experiments with gradually shortened long vowels have demonstrated that relative duration is a *sufficient* cue for perception. There is also the auditory impression of more open quality in the short members of at least some of the pairs; this is supported by formant-frequency differences.

Our hypothesis in this study was that new experiments, designed to allow for the possibility, would show that the vowel spectrum, that is, the formant pattern, which is the major physical correlate of vowel quality, is a significant secondary cue in some of the vowel pairs, especially the pairs /i iː/, /e eː/, /u uː/, and /o oː/. At the same time, relative duration should once again be shown to be the primary cue.

2. Analysis

2.1. *Linguistic material*

Five pairs of Thai words, minimally distinguished by vowel length, are shown in Thai script and phonemic transcription with English glosses in the upper portion of

[1] Another language under study, Pattani Malay, raises similar questions for its word-initial distinctive consonant length (Abramson, 1986, 1987).

WORD-PAIRS

จิบ	/cìp/	**'to sip'**	จีบ	/cì:p/	**'to flirt'**
เห็ด	/hèt/	**'mushroom'**	เหตุ	/hè:t/	**'cause'**
สด	/sòt/	**'fresh'**	โสด	/sò:t/	**'unmarried'**
ขุด	/khùt/	**'to dig'**	ขูด	/khù:t/	**'to scrape'**
ตัก	/tàk/	**'to dip up'**	ตาก	/tà:k/	**'to dry'**

SENTENCES

อย่าขุดมากเกินไป /jà: khùt mâ:k kə:n paj/
'Don't dig too much.'

อย่าขูดมากเกินไป /jà: khù:t mâ:k kə:n paj/
'Don't scrape too much.'

ไม่ทราบสดหรือเปล่า /mâj sâ:p sòt rɨ́ plà:w/
'I don't know whether it's fresh.'

ไม่ทราบโสดหรือเปล่า /mâj sâ:p sò:t rɨ́ plà:w/
'I don't know whether he's single.'

Figure 1. Pairs of Thai words, given in Thai script and phonemic transcriptions with glosses, used in the experiments. Two carrier sentences with four of the words embedded are also shown.

Fig. 1. These words, representing five of the nine vowel pairs in the Thai system, were embedded in carrier sentences, as seen in the examples at the bottom of the figure, and recorded at normal and fast rates twice by an adult native male speaker of Bangkok Thai.[2] He was a graduate student at the time. The other four vowel-pairs of the system were not included, because we could not find syntactically equivalent minimal pairs for embedding in normal sentences of the vernacular. (In the hope of achieving natural speech, for both measurements of production and testing of perception, we wished to avoid rather artificial carrier sentences.) Diphthongs were left out because of the difficulty or impossibility of segmenting them physically into vowels and glides. Of the five tones of the language, only one, the low falling tone, is represented in the figure. This was because we wanted to limit our syllable types, for ease of handling in the perceptual work, to those with final stop consonants.[3] Such syllables do not permit the occurrence of all five tones. In fact, for short and long vowels in this context, the Thai lexicon is best represented by words on the low falling tone. The speaker, as well as another Thai person and the first author, who has had long practical experience with the language, listened to

[2] Three Thai women were recorded in the same way, but their utterances were not used here.
[3] The only other available endings are nasal consonants, vowels, and glides. These were rejected in this work as awkward for later experimental manipulation.

the recordings and found them acceptable as fluent utterances of ordinary colloquial speech.

The durations of the vowels of the key-words, from the release of the first consonant to the closure of the second, were measured from waveforms. A waveform-editing computer program enabled us to magnify waveforms or any portion thereof for greater precision in marking these points. For the pair beginning in /h/, of course, there was no oral closure or constriction at the onset, so, taking the phonetic manifestation of /h/ to be simply noise-excitation of the first portion of the vowel, we marked the beginning of the [h]-turbulence as the beginning of the vowel. FFT analyses were applied to the midpoint of each vowel to measure the frequencies of the first three formants. In addition, narrow phonetic transcriptions, with reference to the IPA Cardinal Vowels, were made by the first author in two sessions separated by three weeks.

2.2. *Results*

In Table I are given the durations and formant frequencies of the vowels of the male speaker uttered in the carrier sentences at his normal conversational rate. His long vowels were on average 1.9 times longer than the short vowels. This is much as in

TABLE I. Durations and formant frequencies, each rounded to the nearest 5 ms and 5 Hz respectively, for two tokens of each vowel recorded by a male speaker for the perception experiments. Asterisks mark the tokens actually used in the test

Word	Vowel duration (ms)	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
/cìp/$_1$	30	315	2025	2740
/cìp/$_2$*	35	320	2030	2530
/cìːp/$_1$	140	250	2185	2895
/cìːp/$_2$*	105	295	2115	2745
/hèt/$_1$	135	390	1965	2500
/hèt/$_2$*	130	440	1975	2555
/hɛ̀ːt/$_1$*	195	390	2035	2565
/hɛ̀ːt/$_2$	285	370	2100	2565
/tàk/$_1$	100	880	1505	2580
/tàk/$_2$*	85	810	1580	2660
/tàːk/$_1$	135	825	1560	2705
/tàːk/$_2$*	150	850	1565	2715
/khùt/$_1$*	125	350	1110	2370
/khùt/$_2$	90	385	1100	2490
/khùːt/$_1$*	215	300	750	2570
/khùːt/$_2$	200	310	850	2490
/sòt/$_1$	65	490	1145	2665
/sòt/$_2$*	55	470	1165	2550
/sɔ̀ːt/$_1$	90	465	1050	2555
/sɔ̀ːt/$_2$*	125	455	880	2635
Av. V	85			
SD	36			
			V:/V = 1.9	
Av. V:	165			
SD	56			

earlier data (Abramson, 1962; Sittachit, 1972; Gandour, 1984; Gandour, Weinberg, Petty & Dardarananda, 1987). Within each length category, the two available tokens differ somewhat in duration, in conformity with data from larger samples (e.g. Abramson, 1962); nevertheless, there is no overlap between short and long vowels. For the front vowels in the table, the short vowels generally have higher F_1 values and lower F_2 values than their long counterparts. For the back vowels, not only is F_1 higher in the short members but so is F_2. For the pair /a aː/, on the other hand, no trend is apparent in the first two formants; all that can be said is that F_3 is higher in the long vowel.

The auditory phonetic assessment yielded an impression of slightly more open qualities for the short members of the pairs, together with slightly more rounding for long /uː/ and /oː/ than for their short counterparts. For the pair /a aː/, however, there was no consistent difference in transcription.

3. Perception

3.1. *Experiments*

As a baseline for the experiments in perception, control tests for the identification of the original words in the carrier sentences spoken at both rates were prepared. That is, all the items in Table I together with the fast versions mentioned in Section 2.1 were recorded directly from the computer files onto magnetic tape to make two randomized test orders for every word pair, with three tokens of each stimulus in each test. This means that, item by item, the subjects never knew whether to expect a short or a long vowel and whether the key word would be embedded in carrier sentence uttered at a normal or a fast rate. The resulting five tests, in two test orders each, were played through headphones to 50 undergraduates at Chulalongkorn University in Bangkok. All of them had been screened by the staff of the Department of Linguistics as native speakers of Bangkok Thai. The students, giving a total of six responses to each original production, identified the key words in Thai script.

The rest of the listening tests made use of the sentences recorded at normal speed. The utterances that were altered to make stimuli are marked with asterisks in Table I. In the circumstances, with a choice of only two tokens of each word, we considered the durations of the long and short members of each pair, their formant patterns, and the impressionistic transcriptions to choose a pair most likely to have audible differences in both length and quality for our listeners. Favoring one combination over another was a fairly subtle matter and for one pair, /tàk/*vs.*/tàːk/, perhaps arbitrary.

To make our stimuli, we lengthened the short member of each of the five pairs in steps of about 10 ms up to a duration a bit longer than that of its long counterpart. Likewise, we shortened the long member in steps of about 10 ms down to a duration a bit shorter than that of its original short counterpart. In order to avoid interfering with any formant movements present, we used spectrograms to help find a point early in the steady-state voice-excited portions of the vowel formants at which to begin either procedure. To lengthen an original short vowel, we copied one period of the waveform at this point and then repeatedly inserted it at the end of the one copied, only at zero-crossings, thus making longer and longer variants. To shorten an original long vowel, we began at the same point and removed one period after another after the first one, cutting only at zero-crossings, thus making shorter and shorter variants. No distortions detectable either by ear in the signal or by eye in the waveform were introduced. Indeed, even the fundamental-frequency pattern was not disrupted, since this early in the syllable the slight downward contour of the low falling tone—called simply "low" by others—had not yet begun or had barely begun. That is, all the stimuli sounded like natural human speech, although we were sure that a number of them in each series would be ambiguous for Thai listeners as to phonemic length category. All the variants were kept within the carrier sentence of the original word from which they were derived.

With these methods, we were able to pit possible spectral cues against duration cues. That is, we kept the original formant patterns of the vowels while altering the durations. Our hypothesis was that a lengthened short vowel should have a later crossover point in identification functions than its shortened counterpart. We supposed that spectral features of each length category would expand the range of the category over durational variants in the region that would otherwise be ambiguous. Thus, the category boundary should be pushed in the opposite direction of the starting category by spectral cues.

A listening test was prepared in three randomizations for the set of alterations of each original word together with the original word itself. A sampling of the range of stimuli was presented before each test. Each of the three test orders contained three tokens of each stimulus embedded in its carrier sentence. Written and oral instructions were given in Thai. Our 50 subjects listened through headphones and labeled each stimulus a total of nine times as one of two possible words given to them in Thai script on the answer sheets. The subjects were split into three groups for counterbalancing in the order of taking combinations of these tests and the control tests. Each group took the tests in four sessions of less than one hour each. Imperfect attendance at some of the sessions is reflected in the data reported in Section 3.2.

3.2. *Results*

We shall not display the results of the control tests, because, with a few scattered confusions at the fast rate, the subjects gave virtually 100% correct identification of all the stimuli. That is, the vowel-length distinction is still robust in Standard Thai, at least for the contexts under consideration. This means also that the words chosen as bases for the making of stimuli were themselves completely unambiguous as to phonemic length.

The perceptual results of the experimental manipulations of the short and long vowels are given in Figs 2–6. The two coded lines in each graph give us the percentage identifications of the stimuli as either short or long vowels. Since our procedures did not yield the same number of stimuli across the two conditions shown in each figure, i.e., lengthening one member and shortening the other, we extracted the 50% crossover points from cubic polynomial cures fitted to the curves for the labels as short vowels. In Fig. 2 are seen the results for the experiments with the vowels /i iː/. While in both of these tests duration is obviously a sufficient cue, the average perceptual crossover points are significantly different: $t(46) = 17.3$, $p <$

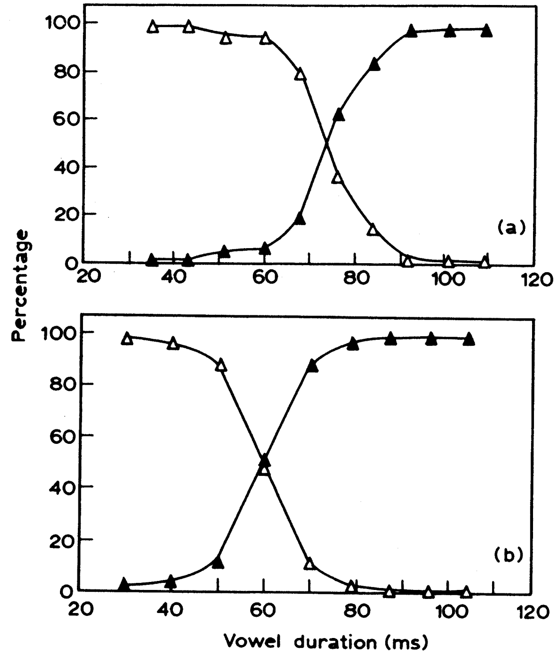

Figure 2. Percentages of labels as short (△) or long (▲) vowels of 47 subjects for variants of (a) lengthened /i/ and (b) shortened /iː/.

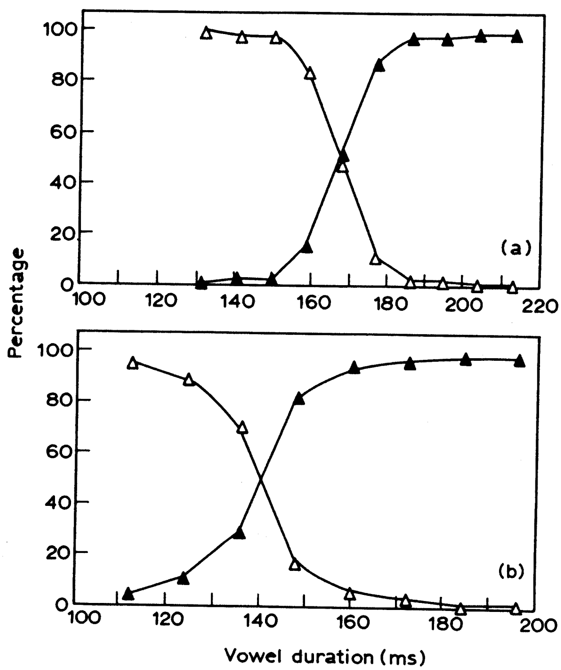

Figure 3. Percentages of labels as short (△) or long (▲) vowels of 49 subjects for (a) lengthened /e/ and (b) shortened /eː/.

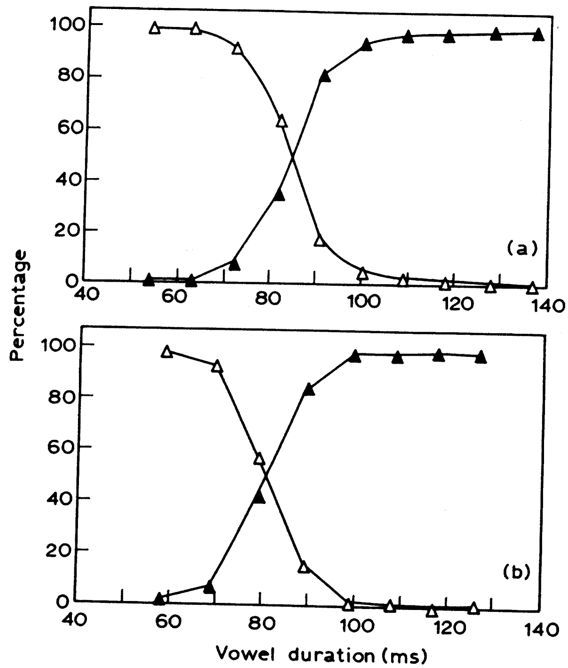

Figure 4. Percentages of labels as short (△) or long (▲) vowels of 49 subjects for (a) lengthened /o/ and (b) shortened /oː/.

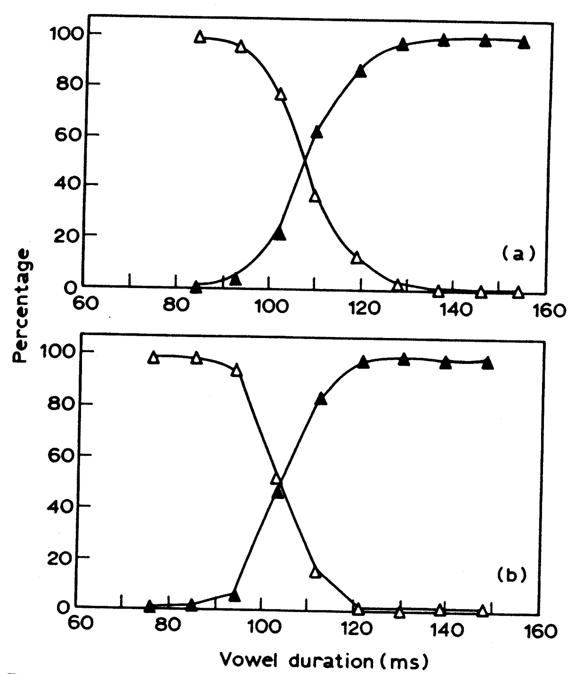

Figure 5. Percentages of labels as short (△) or long (▲) vowels of 49 subjects for (a) lengthened /a/ and (b) shortened /aː/.

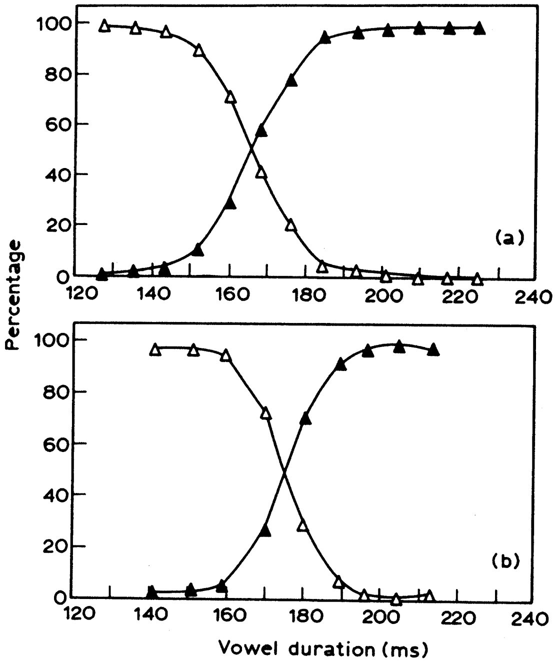

Figure 6. Percentages of labels as short (△) or long (▲) vowels of 49 subjects for (a) lengthened /u/ and (b) shortened /uː/.

0.001. For lengthened /cìp/, the crossover is at 74 ms and for shortened /cìːp/, 59 ms. That is, the zone of ambiguity is later for the original short vowel and earlier for the original long vowel. In Fig. 3, the crossover point for lengthened /hɛ̀t/ is at 167 ms; for shortened /hɛ̀ːt/ it is at 140 ms. Here again, we see the same shift; $t(48) = 24.9$, $p < 0.001$. In Fig. 4, the crossover point for lengthened /sɔ̀t/ is at 85 ms; for shortened /sɔ̀ːt/ it is at 79 ms: $t(48) = 4.9$, $p < 0.001$.

For the one pair in which the vowels seem so much the same to the ear that no difference in crossover was expected, we get the same kind of result, as seen in Fig. 5. For lengthened /tàk/ the crossover point is at 108 ms; for shortened /tàːk/ it is at 104 ms: $t(48) = 6.2$, $p < 0.001$.

Examination of the data in Fig. 6 shows that for lengthened /khùt/, the crossover point is at 166 ms; for shortened /khùːt/ it is at 175 ms: $t(48) = 8.7$, $p < 0.001$. Although the difference is significant, as in all these experiments, the direction is reversed! That is, the crossover point for the lengthened short vowel is earlier than that of the shortened long vowel.

4. Discussion and conclusion

Nearly 30 years after the first experiments (Abramson, 1962), it is evident that relative duration is still the dominant cue to distinctive vowel length in Thai. This, of course, will not surprise linguists who have seen no reason to doubt the traditional view that quantity is a feature of the Thai vowel system. The effects of speaking rate on this distinction have been investigated by Katyanee Svastikula (1986).

For all the pairs tested here, and probably for the whole vowel space, features other than length provide less powerful but pertinent cues. Our hypothesis was that a major secondary cue to the dinstinction lies in spectral differences between the long and short members of the vowel pairs. Even on this hypothesis, incrementally lengthening an original short vowel would at some point cause listeners to start hearing its long counterpart; nevertheless, the perceptual crossover zone ought to be delayed to the extent that spectral cues to "shortness" persisted. Likewise, incrementally shortening an original long vowel should induce a percept of shortness only after the spectral features of the long vowel have been overcome by sufficient decreases in duration; thus, the crossover point will be earlier in the syllable than for the lengthened short vowel. That is, in both conditions, the shift between categories will be determined mainly by relative duration, but the exact placement of the boundary will be affected by secondary cues in the spectum. This hypothesis is well supported by four out of the five pairs tested. In all of them, duration is a sufficient and powerful cue, but the placement of the boundary is significantly later for the lengthened short vowels. The pair /u uː/, however, is a puzzling exception. It is *earlier* for the series of stimuli made by lengthening the short vowel. Although this difference is statistically significant too, the direction is reversed. Once we assured ourselves that there had been no error in our records or in the manipulation of these items, we had to seek other causes. One possibility is rhythmic factors in the carrier sentences. Given a sentence that can syntactically and semantically contain either member of a pair of words minimally distinguished by the feature of vowel length, perhaps there is a tendency toward reciprocity between the duration of the vowel in the key word and the duration of the whole sentence. That is, the sentences containing short vowels will be longer than their counterparts containing long vowels. We have evidence in support of this in Table II. On the left, we see the average perceptual crossover point for the stimuli derived from each vowel and the difference between the crossovers for each word pair. On the right, we see the durations of the carrier sentences *without* the key words and the differences between them for each pair. For the first four pairs of words, as predicted and already

TABLE II. Perceptual crossover points for the pairs of source words, with differences; durations of the carrier sentences without the key words, with differences. V = short vowel; Vː = long vowel. All values are in ms

Source	Perception Crossover	V − Vː	Carrier sentence Duration	V − Vː
cìp	74		1046	
cìːp	59	15	1037	9
hèt	167		1454	
hèːt	140	27	1449	6
sòt	85		1177	
sòːt	79	6	1168	7
tàk	108		983	
tàːk	104	4	976	7
khùt	166		1300	
khùːt	175	−9	1333	−33

described, the crossover points are earlier for the original short vowels, and the carrier frames are slightly longer. For the pair in question, however, /u uː/, it is just the opposite. It must be admitted that the carrier-duration differences for the first four pairs are so small that it is psychoacoustically unlikely that they have much, if any, perceptual effect, although they are consistent with the frames of the unused items in Table I. The negative number for the carriers of the last pair, however, is relatively large. Inasmuch as it would require further experimentation to assert that this is the relevant factor for explaining the discrepancy, we can say at this time only that it is suggestive.[4]

Since relative duration is the dominant cue, it seems likely that the underlying mechanism is timing of the vowel articulation, with concomitant small differences in vocal-tract shape. The latter give rise to sufficient differences in formant pattern to yield slight but audible differences in vowel quality in association with the length feature. Thus, we seem to have here one more case of a combination of acoustic cues, differing, to be sure, in relative power, underlying a single distinctive feature. A troubling question is why a vowel-length distinction should entail any differences in the dimensions of the resonating cavities. After all, in the Introduction we pointed out that one ought to be able to hold a particular setting of the articulators for a shorter or longer time, presumably without any change in position. Here again, stimulated by observations in the literature (e.g., Durand, 1946; Lehiste, 1970, pp. 30–33), we can only speculate. If in the history of a language, a phonological distinction ostensibly based on a single phonetic feature comes into being, we generally believe that the contrasting phonemes are entered into the mental grammar with a special bond between them. That is, there is a kind of psycholinguistic reality to their status of being differentiated by a single feature rather than more than one. Thus, for example, English /b/ and /p/ would be more closely linked than either of them with /f/. Another possibility, however, is that each phoneme takes on a psychological life of its own that allows it, under mostly unexplained pressures, to accrue to itself certain bits of phonetic behavior not obviously attributable to its setting on the allegedly "distinctive" dimension. Perhaps these accrued characteristics result from not fully understood forces, tensions, and inertias of the vocal apparatus. In the particular case of distinctive vowel length, then, we might imagine that over a long period of transmitting a quantity language from generation to generation, speakers may come to produce steady-state long vowels with articulatory settings slightly different from those of the short counterparts, giving each member of a minimal pair something of a phonetic life of its own. This may happen faster in some parts of the vowel system than others. If this is happening, perhaps through natural forces governing speech events, these small differences will be audible and can provide redundant information in support of what may still be the old primary cue.

In such situations, we are probably observing a potential for a diachronic shift from length to quality (see, e.g., Hadding–Koch and Abramson, 1964) in at least part of the system and, maybe ultimately, in the whole system. Indeed, our understanding of the historic rise and fall of distinctive vowel length (de Chene, 1985) may be helped by further experimental investigations.

[4] One obvious extension of this work is to run the same experiment with the other token of /khùːt/ in Table I. Its carrier frame has a duration of 1030 ms.

Finally, for the four pairs of vowels in which the formant pattern seems to provide secondary cues to the distinction, it has to be noted that the magnitudes of the spectral differences have no bearing on the level of significance of the differing crossover points for the various pairs of vowels. On the other hand, the absolute values of the differences do differ somewhat across the pairs, for example the pair /a aː/ with its small spectral difference but its unexpectedly significant difference in crossover points, but it is hard to say much about this without a better metric for quality judgments and their correlation with formants variously distributed in auditory space.

An approach yet to be taken is to introduce spectral variants into each durational variant in our experiments. That is, for each of the stimuli used here, we ought to change the formant frequencies in small steps for further perceptual testing. Our preliminary attempt at doing this with an older computer program yielded promising tentative results with Thai subjects but also caused disturbing enough artifacts for us to discard these tests. In an extension of the present research, we plan to use a newer technique of this sort in the hope of shedding further light on the topic.

This work was supported by Grant HD-01994 from the National Institute of Child Health and Human Development to Haskins Laboratories. It is an expanded version of an oral paper given at the 63rd Annual Meeting of the Linguistic Society of America, New Orleans, Louisiana, 27–30 December 1988. The first author used much of the material, together with other work, in "The Phonetics of Distinctive Length", a talk given by invitation at the International Conference on Linguistic Approaches to Phonetics in Honor of Professor J. C. Catford, Ann Arbor, Michigan, 1–3 May 1989. He wishes to thank the National Research Council of Thailand and the Department of Linguistics, Chulalongkorn University for their hospitality and help, and The University of Connecticut for the sabbatical leave in 1988 that made obtaining the data in Thailand possible. Our patient and cheerful Thai speaker was Mr Akaravudh Osathanukrah.

References

Abramson, A. S. (1962) *The vowels and tones of standard Thai*: *Acoustical measurements and experiments* (Indiana University Research Center in Anthropology, Folklore and Linguistics, Pub. 20: *International Journal of American Linguistics*, **28**(2), Part III). Bloomington.

Abramson, A. S. (1974) Experimental phonetics in phonology : Vowel duration in Thai. *Pasaa*: *Notes and News about Language Teaching and Linguistics in Thailand*, **4**, 71–90.

Abramson, A. S. (1986) The perception of word-initial consonant length: Pattani Malay. *Journal of the International Phonetic Association*, **16**, 8–16.

Abramson, A. S. (1987) Word-initial consonant length in Pattani Malay. In *Proceedings of the XIth International Congress of Phonetic Sciences*, Vol. **6**, pp. 66–70. Tallinn, Estonia: Academy of Sciences of the Estonian S.S.R.

Bennett, D. C. (1968) Spectral form and duration as cues in the recognition of English and German vowels. *Language and Speech*, **11**, 65–85.

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

Catford, J. C. (1988) *A practical introduction to phonetics*. Oxford: Oxford University Press.

de Chene, B. E. (1985) *The historical phonology of vowel length*. New York: Garland.

Durand, M. (1946). *Voyelles longues et voyelles brèves*: *Essai sur la nature de la quantité vocalique*. Paris: Librairie C. Klincksieck.

Gandour, J. (1984) Vowel duration in Thai. *Crossroads*, **2**, 59–64.

Gandour, J., Weinberg, B., Petty, S. H. & Dardarananda, R. (1987) Vowel length in Thai alaryngeal speech . *Folia Phoniatrica*, **39**, 117–121.

Garnes, S. (1973) Phonetic evidence supporting a phonological analysis. *Journal of Phonetics*, **2**, 273–283.

Hadding-Koch, K. & Abramson, A. S. (1964) Duration versus spectrum in Swedish vowels: Some perceptual experiments. *Studia Linguistica*, **18**, 94–107.

Lehiste, I. (1970) *Suprasegmentals*. Cambridge, Mass: MIT Press.
Sittachit, K. (1972) Observations on short and long vowels in Thai. In *Tai phonetics and phonology* (J. G. Harris & R. B. Noss, editor), pp. 28–32 Bangkok: Central Institute of English Language (in Thai).
Straka, G. (1959) Durée et timbre vocalique. *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft*, **12**, 276–300.
Svastikula, K. (1986) A perceptual and acoustic study of speech rate in distinctive vowel length in Thai. Unpublished doctoral dissertation. The University of Connecticut, Storrs, Connecticut.