LINGUISTICS OF THE TIBETO-BURMAN AREA
VOLUME 12:1 - SPRING 1989

A Case Study of Abnormal Phonological Development in Thai

Jack Gandour
Soranee Holasuit Petty
Purdue University

Rochana Dardarananda
Mahidol University

Clinical data are especially instructive to linguistics when they show that certain components or functions of the language system may be separately disrupted. In those cases where all functions are equally disturbed, we gain no insight into the hierarchical nature of language organization or mechanisms underlying different language components. However, in those cases where only some elements or functions are disturbed, we can sometimes bring evidence to bear on issues regarding the relative complexity of linguistic units and interaction of different language components. One potential source of clinical data for linguistics comes from the effects of delayed language development.

Up to the present, the bulk of published research on delayed language development had focused on English-speaking subjects. There is a critical need for information about the effects of delayed language development in other languages. Without such information, we cannot determine those features of delayed language development that are common across languages from those that are specific to particular languages. To our knowledge, there are no published data available on delayed language development in Thai, the national language of Thailand. The present study represents an initial effort to fill this information gap.

Because of the wide range of intersubject variability among language-disordered populations (Leonard 1983), it may not always be wise to represent their behavior by reporting only a group mean. Unlike the case of research with normal subjects, one cannot consider the within group variance of the sample of speech/language-disordered subjects as consisting primarily of random error variance but must seriously consider the possibility that a large part of this variance is caused by theoretically important individual differences (cf. Caramazza & Martin 1983). A single subject design eliminates this particular problem. In a single subject design, on the other hand, it is essential that multiple measures across a variety of tasks are taken in order to draw valid conclusions. The present case study of a Thai-speaking subject (LI) is offered as an example of the kind of detail that is required for drawing valid conclusions about selected speech production capabilities of an individual subject.

Because of the clinico-perceptual impression of disturbances in certain aspects of LI's timing and fundamental frequency (F_0) control, this paper reports on an acoustic-perceptual investigation of important linguistic distinctions in the Thai language associated with timing and F_0. We selected different aspects of timing control that are associated with linguistic units that occur at different levels of linguistic representation.

Thai is a language that has a maximum of three voicing categories of stops in word-initial position: a three-way opposition (voiced unaspirated, voiceless unaspirated, voiceless aspirated) at the bilabial (/b, p, p^h/) and alveolar (/d, t, t^h/) places of articulation and a two-way opposition (voiceless unaspirated, voiceless aspirated) at the velar (/k, k^h) place of articulation. Previous studies of the production of voice onset time (VOT) in normal adult speakers of Thai have shown distinct nonoverlapping distribution of VOT for homorganic word-initial stops in isolated words (Gandour 1985; Lisker & Abramson 1964). VOT values for /b, d/ lie in the long voicing lead region, for /p^h, t^h, k^h/ in the long voicing lag region, and for /p, t, k/ at zero or in the short voicing lag region. VOT associated with word-initial stops provides a measure of consonant timing at the syllable level.

In Thai, differences in the duration of short and long vowels are used phonemically (e.g., /hàt/ 'to practice' vs. /hàat/ 'beach'; /bòt/ 'to grind' vs. /bòot/ 'temple'; /khun/ 'you' vs. /khuun/ 'to multiply'). Although there are concomitant differences in vowel quality for the mid and high vowels, the overriding importance of duration has been amply demonstrated in perception experiments with Thai normal adults (Abramson 1962). Previous studies of the production of vowel length in normal adult speakers of Thai have shown that short and long vowel productions fall into distinct, nonoverlapping regions along the duration continuum (Abramson 1962, 1974). Vowel duration associated with a quantity opposition is a measure of vowel timing at the syllable level.

Rhythm has traditionally been defined in terms of the timing of syllables and the intervals between them. Pike (1945) divided languages into two rhythmic categories: languages in which all syllables are equally spaced in time (syllable-timed) or languages in which only stressed syllables are equally spaced in time (stress-timed). Luangthongkum (1977) concluded that rhythm in Thai has characteristics of both syllable-timed and stress-timed languages. Rhythmic patterns associated with syllables, phrases, and sentences provide a measure of timing control at the extrasyllabic or discourse level.

Thai is a tone language. In a tone language, every syllable or word may have a distinctive pitch pattern as part of its phonemic structure. Thai has five contrastive pitch patterns, or lexical tones, traditionally labeled mid, low, falling, high, and rising (Gandour 1975). The primary acoustic correlate of tone is generally considered to be F_o states and movements. In prepausal position in Thai, the mid tone can be described as mid level with a final drop, low tone as low falling, falling tone as high falling, high tone as high rising, and rising tone as low rising (see Abramson 1962, 1976 for acoustical description of Thai tones). F_o contours associated with the five Thai lexical tones provide a measure of F_o control.

Accordingly, in this paper we report on the findings of four acoustic-perceptual experimental studies of LI's production of VOT of homorganic word-initial stops (Study 1), vowel duration of phonemic short and long vowels (Study 2), durational patterns of syllables, phrases, and sentences in connected speech (Study 3), and F_o patterns of lexical tones (Study 4).

Case Report

LI was a 16-year-old male boy, who was seen at Ramathibodi Hospital in the summer of 1980, at which time the data for this study were collected. LI was a monolingual native speaker of Central Thai. At the time of testing for this study, LI had been receiving speech therapy at Ramathibodi Hospital biweekly for about a year and a half. When LI was first seen at age 15, audiological testing demonstrated LI's hearing to be normal. His ENT exam was normal. His neurological evaluation was normal. The neurologist's diagnosis was "developmental articulation disorder" with a recommendation for speech therapy. According to the speech pathologist, LI exhibited a "severe articulation disorder". Almost all consonants were produced as glottal stops. Vowels were distorted, including both monophthongs and diphthongs. His tonal production was aberrant. With therapy, his consonant production improved dramatically; some vowels were still distorted, especially diphthongs; tone production improved slightly. By self report, LI didn't start speaking until he was about 3 years old. Other than a delay in language development, LI's medical history was normal. His birth was without complications. He did not suffer from any prolonged childhood diseases. At the time of testing, LI had already completed nine years of formal education.

Study 1: Voicing Contrast in Word-Initial Stops

Perception

Method

Subjects. Data were obtained from LI for the present study. For comparison purposes, data for one normal, male adult subject were taken from a previously published report (Gandour & Dardarananda 1982).

Stimuli. A total of 37 stimuli for each place of articulation were synthesized by means of a computer-controlled parallel resonance synthesizer at Haskins Laboratories (Lisker & Abramson 1970, pp. 563-564). The continuum ranged in VOT from -150 to +150 msec: from -150 to -10 in 10 msec steps, from -10 to +50 in 5-msec steps, and from +50 to +150 in 10-msec steps.

Each stimulus consisted of three steady-state formants of a vowel of the type [ ]. Bilabial, alveolar, and velar stops were made by adding appropriate release bursts and formant transitions to the beginning. For voicing lead, only low-frequency harmonies of the buzz source were used; for voicing lag, the interval between burst and onset of pulsing was excited by hiss alone, the first formant was suppressed. Each stimulus variant had a fundamental frequency of 114 Hz with a falling contour near the end of the vowel.

For each place of articulation, two test tapes were constructed, each consisting of a separate random presentation of two occurrences of each stimulus.

Listening Procedure. Each subject was tested individually in a quiet room. The test tapes were played on a Nagra IV-S tape recorder and the signal was presented through TDH-39 headphones at a comfortable listening level.

Subjects were instructed to identify each stimulus by pointing to a card displaying the Thai printed letter and picture of its alphabet keyword: /b/ 'leaf', /p/ 'fish', /p^h/ 'tray', /d/ 'child', /t/ 'turtle', /t^h/ 'soldier', /k/ 'chicken', /k^h/ 'water buffalo'.

Results and Discussion

Identification functions for the VOT continua for the normal subject and LI are shown in Figure 1. The curves give the distribution of a subject's identification of the synthetic speech stimuli as functions of VOT values. As shown in Figure 1, the labeling curves for the normal subject reflect three distinct perceptual categories for the bilabials and alveolars, and two for the velars. His performance replicates the findings of Lisker & Abramson (1970) both in terms of location of the VOT boundaries and steepness of the identification functions. LI's labeling curves, on the other hand, reflect only two distinct perceptual categories regardless of place of articulation. For the bilabials, there is no sharp perceptual boundary between /b/ and /p/; for the alveolars, there is no sharp perceptual boundary between /d/ and /t/. His labeling curves for /p^h/, /t^h/, and /kh/ are similar to those for normals. With the exception of four VOT values in the voicing lead region of the continuum, LI's labeling curve for /k/ also mirrors that of normals. Thus, LI's deficit in VOT perception appears to be restricted primarily to the voicing

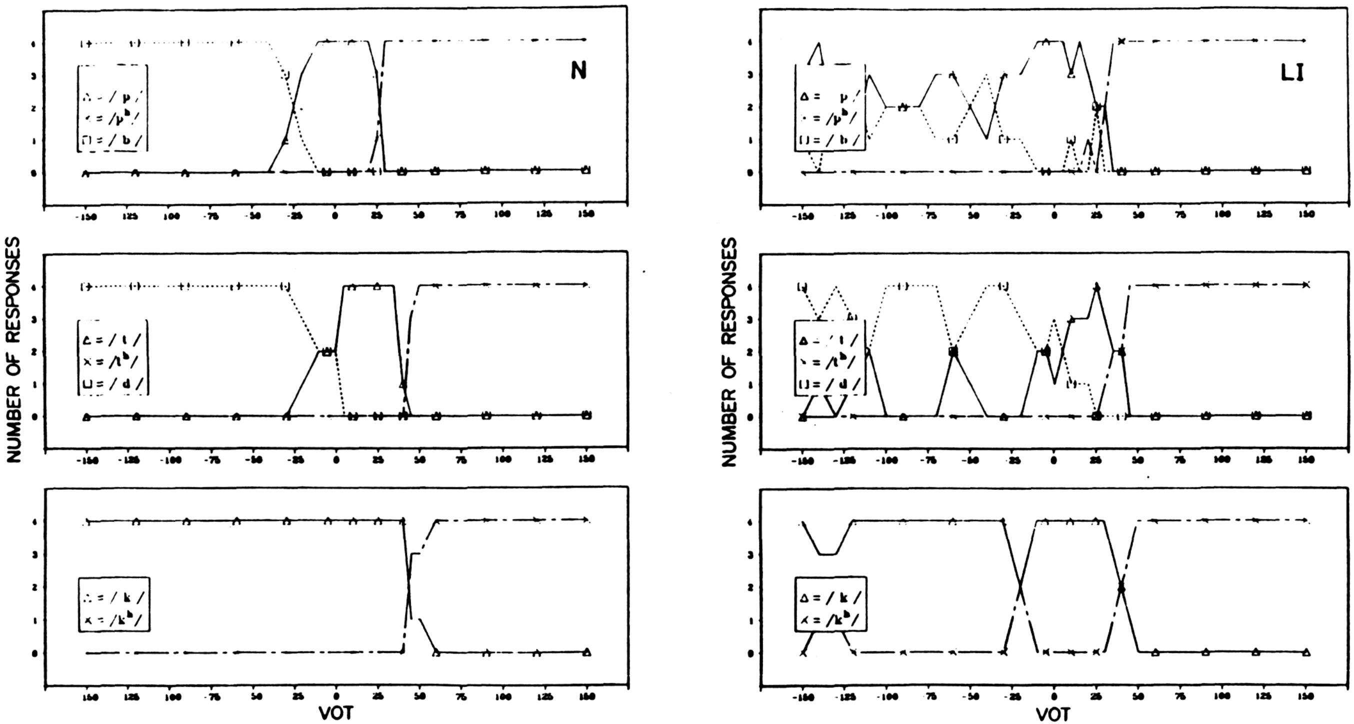


Fig. 1. Identification functions for VOT continua at three places of articulation for a normal Thai adult (N: adapted from Gandour & Dardarananda 1982) and LI.

lead region of the continuum.

Production

Method

Subjects. Data were obtained from LI for the present study. For comparison purposes, data were taken from previously published reports on the perception (Gandour, Weinberg, Petty & Dardarananda 1987a) and production (Gandour 1985) of VOT by five normal adult speakers.

Speech Materials. Test stimuli were eight monosyllabic Thai words. These eight words formed a minimal triplet for the three-category distinction in VOT at each of the bilabial (/baan/ 'to bloom'; /paan/ 'birthmark'; /p^haan/ 'ignorant') and alveolar (/dam/ 'black'; /tam/ 'to pound'; /t^ham/ 'to do') places of articulation and a minimal pair at the velar (kan/ 'to prevent'; /k^han/ 'to itch') place of articulation. All words began with an initial stop consonant followed by a low back unrounded vowel and a final nasal consonant. Tone (mid) was held constant across all eight words.

Recording Procedure. LI was instructed to produce the words in a carrier sentence at a comfortable speaking tempo. A total of 80 utterances (8 words x 8-11 tokens) were collected. For a more detailed description of the recording procedure, see Gandour 1985.

Listening Procedure. A closed-set identification procedure was used in the listening tests. Under this procedure, a trial consisted of the presentation of a single word. Eight to eleven productions of each of the eight words were assembled into three blocks of trials. In block 1, tokens of /baan/ 'to bloom', /paan/ 'birthmark' and /p^haan/ 'ignorant' were presented; in block 2, tokens of /dam/ 'black', /tam/ 'to pound/, and /t^ham/ 'to do'; and in block 3, tokens of /kan/ 'to prevent' and /k^han/ 'to itch'. The productions of words within each block were presented in random order. The response interval for each trial was 3 seconds.

Three native, adult speakers of Thai participated in the listening tests. Each listener was provided with a set of response sheets which contained a closed set of stimulus words for each trial. Response sheets were typed in Thai orthography. Listeners were instructed to identify each stimulus by circling the word on their response sheets corresponding to what they heard. A block of three practice trials were presented before proceeding with the actual experiment. The test tapes were played on a Nagra III tape recorder, and the signal was presented through TDH-39 headphones at a comfortable listening level (i.e., about 70 dB SPL).

Measurement Procedure. Conventional wide-band spectrograms with amplitude traces were made of 80 utterances produced by LI (Kay Digital Sona-Graph model 7800). The measurement procedure was the same as that described in Gandour 1985.

Results and Discussion

The percentages of correct identification of the word-initial stops produced by normal speakers and LI are summarized in Table 1. Normal speakers' productions of all eight stops were correctly identified at a perfect or near-perfect level of accuracy. LI's productions of /p^h, t, t^h, k^h/ were also correctly identified at a perfect or near perfect level of accuracy. Although better than chance, his productions of /p/ were perceived at a level considerably lower than that of normal speakers. His intended productions of /b, d/, however, failed to achieve even chance levels of identification. None of his intended productions of /k/ were perceived correctly.

TABLE 1. Identification of Thai word-initial stops produced by normal speakers and a Thai child with abnormal phonological development

Speakers	/b/	/p/	/pʰ/	/d/	/t/	Stops /tʰ/	/k/	/kʰ/
Normal[a]	96	96	100	100	100	100	100	100
LI	20	73	100	21	90	100	0	100

Note: Values are expressed in percentages of correct identification.
[a]Normal = 5 normal Thai adults from Gandour et al. 1987a.

The frequency distributions of VOT values associated with the production of bilabial, alveolar, and velar stops by normal speakers and LI are illustrated in Figure 2. For the group of normal speakers, no overlap in VOT values is found between any of the voicing categories at each of the three places of articulation. The values for all stop consonants cluster into well defined regions along the VOT continuum. The VOT values for /b, d/ all lie in the long voicing lead region, for /pʰ, tʰ, kʰ/ in the long voicing lag region, and for /p, t, k/ at zero or in the short voicing lag region. These findings for our five normal adults are in agreement with Lisker & Abramson's (1964) VOT data.

LI's VOT distributions differ from those of normals at all three places of articulation. Few VOT values lie in the long voicing lead region (< -60 ms) for either /b/ or /d/. VOT values for /b, d/ lie primarily at zero or in the short voicing lead region. The majority of values for /p, t/ lie at zero or in the short voicing lead region. Thus, there is substantial overlap between voiced unaspirated and voiceless unaspirated stops at both bilabial and alveolar places of articulation. VOT values for /k/ lie primarily in the long voicing lag region (> +60 ms). LI's distributions of VOT values for /pʰ, tʰ, kʰ/ are similar to those of normals. Because his VOT values for /k/ are appreciably longer than normal, there is considerable overlap between /k/ and /kʰ/.

Descriptive statistics for the distribution of VOT productions for normal speakers and LI are given in Table 2. A mean VOT for LI was considered to be significantly different from the mean VOT for normal adults if LI's value was equal to or greater than each group mean plus two standard deviations or equal to or less than each group mean minus two standard deviations (cf. Kent, Netsell & Abbs 1979, p. 633). Only four of LI's eight mean VOT values fell within two standard deviations of the mean VOT values for normal adults: /pʰ, tʰ, kʰ/ and /t/.

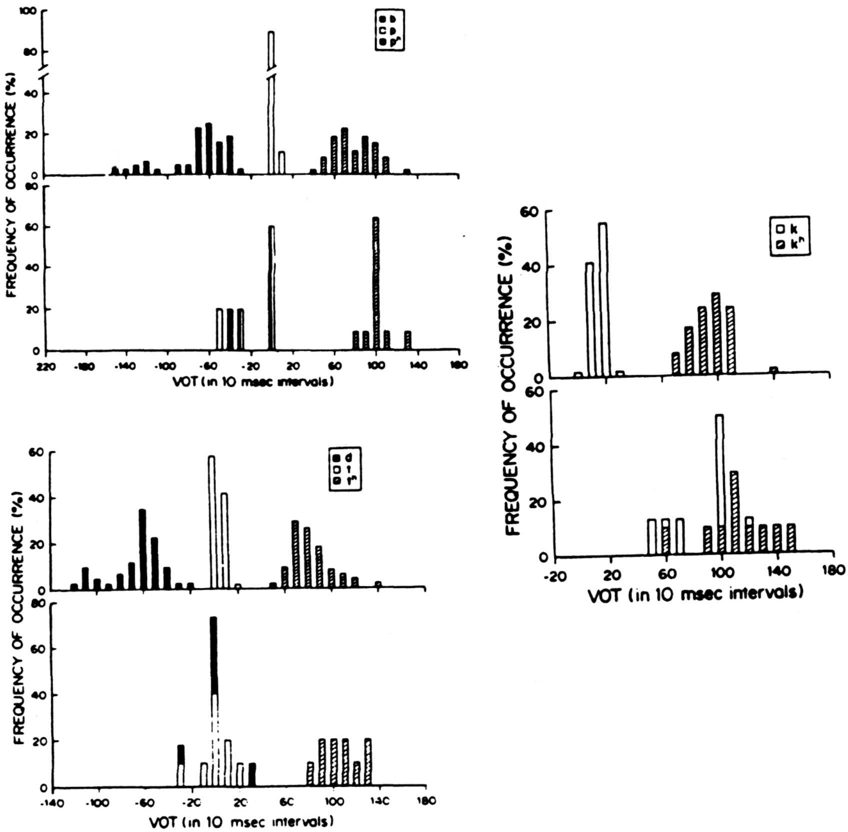


Fig. 2. Distribution of VOT productions for bilabial, alveolar, and velar stop consonants in word-initial position for a group of 5 normal Thai adults (upper panel: adapted from Gandour 1985) and LI (lower panel). The abscissa of each graph represents VOT values in 10-ms intervals and the ordinate the percentage of utterances falling within the specified VOT interval.

TABLE 2. VOT comparison to normal adults

	/b/		/p/		$/p^h/$	
Speakers	*M*	*SD*	*M*	*SD*	*M*	*SD*
Normal[a]	-120	11	6	3	85	11
LI	-26	25	-30	19	99	11
	/d/		/t/		$/t^h/$	
Speakers	*M*	*SD*	*M*	*SD*	*M*	*SD*
Normal	-100	7	7	4	87	13
LI	-6	13	3	16	104	16
			/k/		$/k^h/$	
Speakers			*M*	*SD*	*M*	*SD*
Normal			21	4	99	11
LI			88	21	111	25

Note: Values are expressed in terms of ms.
[a] Normal = 5 normal Thai adults from Gandour 1985.

Results of a two-factor (voicing x place of articulation) analysis of variance with repeated measures showed that LI's mean VOT differed significantly as a function of voicing category within place of articulation, $F(5,72) = 108.83$, $p < .05$. Post hoc Newman-Keul's comparisons indicated significant ($p < .05$) differences in mean VOT between /b, p/ and $/p^h/$, and between /d, t/ and $/t^h/$. Differences between /b/ and /p/, /d/ and /t/, and /k/ and $/k^h/$ were not significant.

These findings combined suggest that LI is capable of signaling no more than two voicing categories along the VOT continuum. Because of his difficulty in producing stops with long voicing lead, voiced unaspirated and voiceless aspirated stops do not occupy discrete regions along the continuum. LI's difficulty with VOT, however, was not restricted to the long voicing lead region. Unlike normal speakers, the majority of his VOT values for /p/ lie in the short voicing lead region, for /k/ in the long voicing lag region. In view of his productions of /p/ and /t/, we might have expected LI to be able to signal the two-way voicing contrast between /k/ and $/k^h/$ successfully. The fact that he was unable to do so points to a specific problem in VOT production associated with /k/. Perhaps it is the short voicing lag that is typically associated with productions of /k/ by normal speakers. What remains consistent across all three places of articulation is LI's preservation of voiceless aspirated stops with long voicing lag, i.e. $/p^h, t^h, k^h/$.

A comparison of LI's VOT production data to those of normally-developing Thai children (Gandour, Petty, Dardarananda, Dechongkit & Mukngoen 1986) reveals that LI's production of word-initial stops, with the exception of /k/, is at a stage of development similar to that of Thai 3-year-olds. For a group of seven 3-year-old children, Gandour et al. (1986) reported considerable overlap in VOT distribution between voiced unaspirated and voiceless unaspirated stops. Over half of VOT values associated with /b/ and /d/ lie in the short voicing lead region or at zero. There was minimal overlap between voiceless unaspirated and voiceless aspirated stops at all three places of articulation. By the age of 3, Thai children are already producing adult-like VOT productions of voiceless aspirated stops, i.e. /p^h, t^h, k^h/. The late emergence of Thai voiced unaspirated stops as opposed to voiceless aspirated stops suggests that long lead voicing is inherently more difficult to produce than long lag voicing. These differences in order of acquisition as a function of voicing category would lead us to predict that a Thai child with a specific phonological impairment will exhibit more difficulty with voiced unaspirated stops. This prediction is borne out in LI's data.

With regard to abnormal phonological development, it is of interest to determine the extent to which there is a relation between production and perception abilities. The presence of an impairment in both speech perception and speech production suggests a central disorder in the phonological component, and therefore, an association between production and perception abilities. In contrast, the presence of an impairment in either speech production or speech perception separately suggests a peripheral disorder in the articulatory phonetic encoding mechanism or in the acoustic phonetic decoding mechanism, and therefore, a dissociation between production and perception abilities. LI's data on VOT perception and production are inconclusive with respect to this question. His performance on both perception and production tasks was abnormal in the voicing lead region of the continuum for both bilabials and alveolars, thus suggesting an association between perception and production. However, LI's performance on perception and production of velars suggests a dissociation between perception and production. He was able to label the VOT stimuli in a normal-like fashion, yet unable to produce nonoverlapping distributions of VOT for /k/ and /k^h/.

Study 2: Vowel Length

Method

Subjects. Data were obtained from LI for the present study. For comparison purposes, data were taken from a previously published report on the perception and production of the vowel length contrast by a group of normal adult speakers (Gandour, Weinberg, Petty & Dardarananda 1987b).

Speech Materials. Test stimuli were six monosyllabic Thai words. These six words formed three minimal pairs for the contrast in vowel length: (1) /bàt/ 'card' vs. /bàat/ "Thai monetary unit'; (2) /pàt/ 'to brush off' vs. /pàat/ 'to make smooth and level', and (3) /p^hàt/ 'to stir fry' vs. p^hàat/ 'briefly'. Vowel quality, tone, and final consonant were held constant across pairs. In pair 1, the word-initial bilabial stop was voiced unaspirated; in pair 2, voiceless unaspirated, and in pair 3, voiceless aspirated.

Recording Procedure. All six words were printed in large Thai letters on 3 x 5 cards for presentation to the subject. The subject was instructed to produce the words in the carrier frame /nîi _____/ 'This (is) _____' at a comfortable speaking rate. His productions were tape-recorded using a Sony ECM-150 microphone and a Nagra IV-S tape recorder in a soundproof booth in a single session. A total of 60 utterances (6 words x 10 repetitions) were recorded.

Listening Procedure. LI's productions of each of the six words were spliced out from the original recording, and reassembled into a randomly-ordered set of 60 trials. Each trial consisted of the presentation of a single word. The test tape

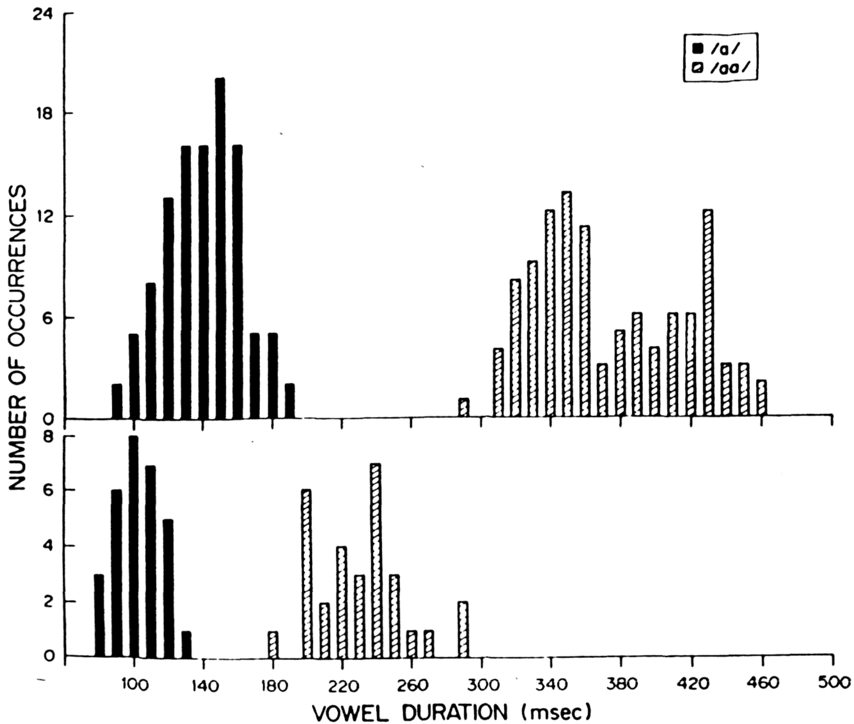


Fig. 3. Distribution of vowel duration for short and long vowel phonemes in monosyllabic words produced by a group of 5 normal adult speakers (upper panel; adapted from Gandour 1987b) and LI (lower panel). The abscissa represents vowel duration values falling within 10 ms steps; the ordinate represents the number of occurrences falling within the specified duration interval.

was played on a Nagra IV-S tape recorder, and the signal was presented through TDH-39 headphones at a comfortable listening level. Both the listeners and task (closed-set identification) were the same as those in Study 1.

Measurement Procedure. Conventional wide-band spectrograms with amplitude traces were made of 60 utterances produced by LI using a Kay Digital Sona-Graph model 7800. The measurement procedure was the same as that described in Gandour et al. 1987b.

Results and Discussion

All vowels produced by 4 normal speakers were perceived as intended by 10 native Thai listeners (Gandour et al. 1987b). Similarly, LI's productions of short and long vowels were perceived at a near-perfect level of accuracy (95%). No errors were made in the identification of LI's short vowel productions.

The distribution of short and long vowel productions for the group of normal adults and LI are shown in Figure 3. Short and long vowel productions fall into distinct, nonoverlapping regions along the duration continuum for normal speakers. Although there was a compression of the duration continuum toward the short end, LI similarly demonstrated no overlap between short and long vowel productions.

The means and standard deviations of the duration of Thai short and long vowels are summarized in Table 3. LI's short and long vowel productions, respectively, were on average 28% and 39% shorter than those of normals. Yet, the mean duration of LI's short and long vowel productions were significantly different, $t(58) = 23.22$, $p < .0001$. LI's duration ratio of the means of long to short vowels was 2.3, the group of normal speakers' was 2.6. Although LI's vowel durations were considerably shorter than those for normal speakers, he was still able to control relative timing differences associated with the vowel length contrast.

TABLE 3. Vowel duration for short and long vowels in monosyllabic words

Speakers	Short M	Short SD	Long M	Long SD
Normal[a]	142	22	375	43
LI	102	14	230	27

Note: Values are expressed in ms.
[a]Normal = 4 normal Thai adults from Gandour et al. 1987b.

Study 3: Rhythm in Connected Speech

Method

Stimuli. Three readings of a passage were obtained from LI. Data based on one reading of the passage by each of 5 normal speakers were taken from Luangthongkum's (1977) study. The passage itself was a Thai adaptation of the story of a parrot as told by Blanche DuBois in Tennessee Williams' play A Streetcar Named Desire. The majority of words in the reading passage were monosyllabic. The Thai adaptation was carefully designed by Luangthongkum for the purpose of making syllable duration measurements. The Thai reading passage and an English translation are presented in Figure 4.

เรื่องนกเอี้ยง

เอ ... นึกดูก่อน เรื่องที่เคยเล่า อ้า ... นึกออกแล้ว เอาเรื่องนกเอี้ยงดีกว่านะ มีนกเอี้ยงตัวหนึ่ง มันด่าเก่งเป็นไฟเลย รู้คำหยาบมากมายเหลือเกิน แล้วก็วิธีเดียวที่จะให้มันเงียบได้ นั่นก็คือ ต้องเอาผ้าคลุมกรงเสีย มันจะได้คิดว่ามืดแล้ว ได้เวลานอนเสียที ทีนี้นะ รุ่งเช้าวันหนึ่งยายสาวแก่เจ้าของนกนั่นก็เปิดเอาผ้าที่คลุมกรงออกอย่างเคย รู้ไหมใครเดินผ่านมาพอดี บาทหลวงจ้ะ ยายเจ้าของแกรีบวิ่งกลับมาคลุมกรงเจ้านกอย่างเร็วไว แล้วก็วิ่งไปที่ประตูเชิญบาทหลวงเข้ามาในบ้าน นกมันก็นิ่งเงียบ และแล้ว ... ยายสาวแก่ก็ถามพ่อบาทหลวงว่า ให้ใส่น้ำตาลกี่ช้อนในกาแฟ ไอ้เจ้านกมันก็ตะโกนโพล่งออกมาว่า "แหม ... ให้ตายเถอะน่ะ ไอ้กลางวันนี่มันสั้นเป็นบ้าเลยเว้ย!"

The Story of a Myna Bird

Hm ... let me think ... the story that I have ever told. Ah! I can remember now ... the story of a myna. There was a myna who was very keen in using bad language. He knew a lot of dirty words. The only one way that could make him quiet was to cover his cage with a piece of cloth. This would make him think that it was dark and that it was the time to go to bed. Well, one early morning, the spinster who was the bird's owner came to uncover the cage as usual. Do you know who was passing by? A Roman Catholic priest! The owner of the bird rushed back to cover the cage as quickly as she could. Then, she ran to the door and invited the priest to come in. The bird was quiet. Finally, the spinster asked the priest, "How much sugar do you want in your coffee? The bird, then, shouted out loudly, "Damn it! How come the day is so short!"

Fig. 4. Thai adaptation of the story of a myna bird and English translation.

__Measurement Procedure__. Recordings of the three readings of the passage by LI were played into a Pitch Instruments F_0-intensity meter. The audio, F_0, and intensity signals were played into three different channels of a Honeywell 1508 oscillograph operating at a chart speech of 100 mm/s.

Identification of syllable boundaries was determined from inspection of the audio, F_0, and intensity-meter signals. For the majority of adjacent syllables in the passage either the offset of the first syllable or the onset of the second syllable was voiceless, or they were separated by a syntactic pause. In those cases, syllable onset was operationally defined to occur when there was a sharp increase in intensity from the baseline, syllable offset when there was a sharp decrease in intensity to the baseline. In the few remaining cases, adjacent syllables had different tones. Syllable onset and offset were determined on the basis of shifts in direction of F_0 associated with the different tones. A pause was operationally defined to occur when the speech waveform approached baseline for at least 200 ms. The portion of the waveform tracing between two pauses was designated as a phrase. In this context, phrases were phonetically-structured entities that did not necessarily coincide with syntactic phrases. Syllable and pause durations were measured by the first author to the nearest centisecond using a transparent millimeter rule. All syllable and pause durations were reanalyzed on a separate occasion for an estimate of intrajudge reliability (r = .99). The total number of syllables subjected to duration measurements differed between readings because of the speaker's omissions of words in each of his three readings of the paragraph or the occurrence of unmeasurable portions of the oscillographic records.

Identification of stressed syllables in the readings of the passage by LI was performed by the second author, a native speaker of Thai with a postgraduate degree in linguistics. She listened to the tape recordings of the reading passage and identified which syllables were perceived to be stressed. Each recording of the paragraph was presented in two separate sessions spaced approximately one week apart. Her stress judgments were identical for the two presentations of each recording of the paragraph. Stressed syllables in the readings of the passage by the five normal speakers were taken from Luangthongkum 1977.

For comparison purposes, duration measurements of syllables and pauses in the readings of the passage by five normal speakers were taken directly from Luangthongkum 1977. Her measurements were identically determined from inspection of oscillographic records to the nearest centisecond.

__Results and Discussion__

Mean durations of syllables and pauses based on three readings of the passage by LI and one reading each of the passage by Luangthongkum's (1977) five normal speakers are summarized in Table 4. LI took a longer time to read the passage aloud than normal speakers. Although LI's total reading time was based on seven more syllables in comparison to normal speakers, his total reading time was, on the average, 69 seconds longer. As measured in syllables per second, LI and normal speakers spoke at a rate of 1.5 and 3.3 syllables/seconds, respectively.

TABLE 4. Duration of pauses and syllables

Measures	LI *M*	LI *SD*	Normal *M*	Normal *SD*
	Speakers			
Total duration of passage(s)	123.9	7.5	54.5	2.5
Total number of pauses	113.7	2.1	23.2	1.3
Total duration of pauses(s)	72.8	4.2	14.1	1.6
Percent pause duration	58.7	0.4	25.7	1.8
Total number of syllables	183.3	0.6	177.0	1.3
Total duration of syllable(s)	51.2	3.4	40.4	1.2
Mean syllable duration(cs)	27.9	2.0	23.0	1.0
Syllables per second	1.5	0.1	3.3	0.1
Pause-to-syllable duration ratio	1.4	0.1	0.3	0.1

Note: Normal = 5 normal Thai adults from Luangthongkum 1977.

All pause time measures for LI were larger than for normal speakers (see Table 4). LI exhibited a larger number of pauses, a greater amount of total pause time, a larger pause-to-syllable duration ratio (total pause time divided by total phrase time), and a greater percentage of pause time. Although LI's total duration of syllables and mean syllable duration were only slightly longer than those for normal speakers, his syllables per second were less than half those of normal speakers. These findings indicate that LI paused more frequently than normal speakers, and produced syllables at a slower rate.

Phrase timing measures for LI and normal speakers are presented in Table 5. The total number of phrases produced by LI was almost five times greater than the total number produced by the normal speakers. In terms of mean phrase duration, maximum number of syllables per phrase, and mean number of syllables per phrases, LI clearly produced shorter phrases.

TABLE 5. Phrase timing

Measures	LI *M*	LI *SD*	Normal *M*	Normal *SD*
	Speakers			
Total number of phrases	114.3	1.5	24.2	1.3
Mean phrase duration (cs)	25.7	0.6	134.0	8.0
Maximum syllables per phrase	4.3	0.6	19.0	0.0
Mean syllables per phrase	1.6	0.1	7.3	0.4

Note: See note to Table 4.

The time between the onset of one stressed syllable and the onset of the next stressed syllable was designated as an interstress interval. A one-syllable interstress interval indicates that two adjacent syllables are stressed; a two-syllable interstress interval indicates that an unstressed syllable occurs between two stressed syllables; and so on. The overwhelming majority of LI's interstress intervals were one-syllable (Table 6). He rarely produced two-syllable, and never produced three- or four-syllable interstress intervals. Although one-syllable interstress intervals accounted for more than half of normal speakers' productions, two-syllable intervals occurred frequently, three- and four-syllable intervals occasionally.

Table 6. Occurrence of Interstress Intervals

	Speakers			
	LI		Normal	
Interstress Interval	M	SD	M	SD
One-syllable	98.1	0.3	56.3	1.4
Two-syllable	1.9	0.3	35.3	1.2
Three-syllable	-	-	3.9	0.5
Four-syllable	-	-	2.9	0.8

Note: Values are expressed in percentages. See also note to Table 4.

A comparison of syllable duration as a function of stress and position is presented in Table 7. Mean percentage of stressed syllables was 98% for LI, 63% for normal speakers. The stressed-unstressed mean syllable duration ratios for LI and the normal speakers were 1.9 and 2.0, respectively. Because of the scarcity of unstressed syllables in his speech, the fact that LI's duration ratio is almost identical to that of normals should not be overinterpreted. LI's inability to produce appropriate temporal patterns becomes more apparent when sentence position is taken into account.The mean duration of LI's nonprepausal syllables was about 32% greater than for normal speakers. The prepausal-nonprepausal duration ratio was 0.97 for LI, 1.7 for the normal speakers. Further, LI produced a greater percentage (62%) of prepausal syllables than the normal speakers (13%).

TABLE 7. Effect of Stress and Position on Syllable Duration

	Speakers			
	LI		Normal	
Measures	M	SD	M	SD
Total number of stressed syllables	180.7	0.6	112.0	2.3
Total number of unstressed syllables	3.3	0,6	65.0	2.3
Mean stressed syllable duration(cs)	28.2	2.0	28.0	0.2
Mean unstressed syllable duration(cs)	14.9	1.5	14.0	0.1
Total number of prepausal syllables	113.7	2.1	23.2	1.3
Total number of nonprepausal syllables	69.7	2.5	154.2	2.5
Mean prepausal syllable duration(cs)	26.6	3.3	35.7	1.9
Mean nonprepausal syllable duration(cs)	27.4	1.8	20.8	0.6

Note: See note to Table 4.

The findings of the present study clearly indicate that LI exhibited an impairment in timing in connected speech. By comparison to normal adults, LI's temporal patterns in syllables and phrases were abnormal. His slower overall speaking rate, shorter phrases, uniform syllable durations regardless of sentence position, and overwhelming predominance of one-syllable interstress intervals all contributed to the dysrhythmic quality of LI's speech. These findings are not restricted to oral reading. The dysrhythmic quality of LI's oral reading directly reflects his spontaneous speaking abilities. LI tended to produce speech on a syllable-by-syllable basis. In so doing, an economy in timing patterns was achieved insomuch as temporal dependencies among syllables were minimized. Such timing patterns appear to reflect a simplification of motor programming of larger-sized linguistic units.

Study 4: Tone

Perception

Method

Subjects. Data were obtained from LI for the present study. For comparison purposes, data for 21 normal adult subjects were taken from a previously published report (Gandour, Carney, Nimitbunnasarn & Amatyakul 1984).

Stimuli. Two sets of tonal stimuli, natural and synthetic, were used in this experiment. The first set consisted of natural speech tokens of the citation forms of /k^haa/ 'to be stuck', /k^hàa/ 'a kind of spice', /k^hâa/ 'to kill', /k^háa/ 'to engage in trade' and /k^hǎa/ 'leg' produced by an adult male native speaker of Thai (see Figure 5). The second set consisted of synthetic versions of the same words produced on the Haskins Laboratories' formant synthesizer in which only the F_0

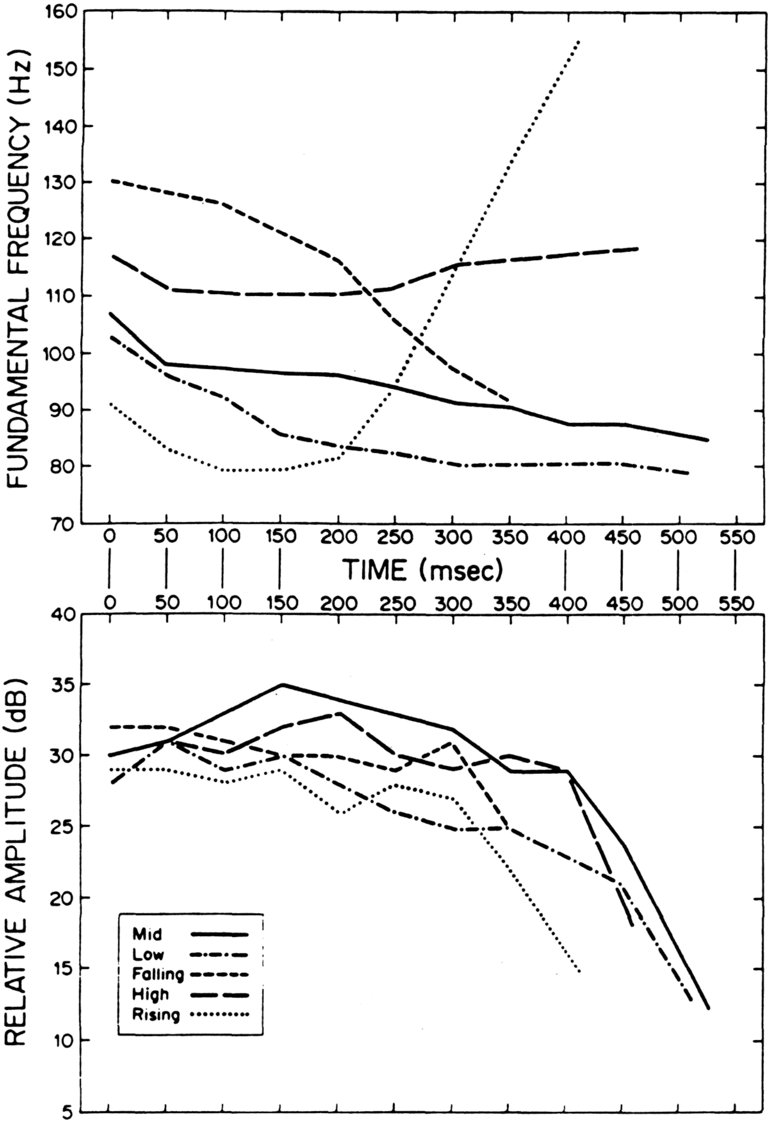


Fig. 5. Fundamental frequency and amplitude contours used in the natural speech stimulus set.

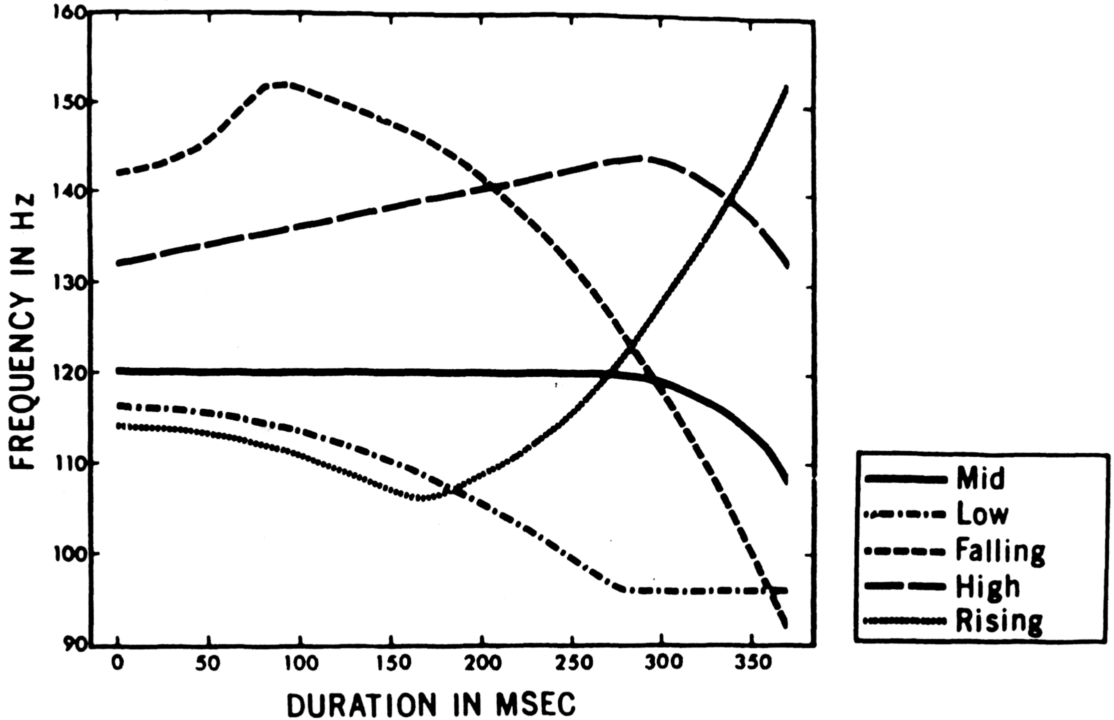


Fig. 6. Fundamental frequency contours used in the synthetic speech stimulus set (adapted from Abramson 1975, by permission of Department of Linguistics, Chulalongkorn University, Bangkok, Thailand).

parameter was varied (see Figure 6). Each stimulus had a flat amplitude except for a slight rise at the beginning and a slight fall at the end, and a total duration of 375 ms (see Abramson 1975, p. 5, for details of speech synthesis procedure). The synthetic stimuli enabled us to assess the perceptual efficacy of F_0 alone in the signaling of Thai tones.

Two test tapes were constructed for each stimulus set. Each test tape contained 25 trials, 5 repetitions of each stimulus tone. For both stimulus sets, tonal stimuli were presented in two different random orders, one order of presentation per test tape.

Listening Procedure. Each subject was tested individually in a large sound-treated booth. The test tapes were played on a Revox tape recorder, and the signal was presented through TDH-39 headphones at a comfortable listening level. Subjects were instructed to identify each stimulus presented auditorily by circling the appropriate word on their answer sheets.

Results and Discussion

For the normal listeners, overall mean percentages of correct identification of the five natural speech and the five synthetic speech tones were 98% and 94%, respectively (Gandour et al. 1984); for LI, 80% and 98%, respectively. A comparison between normal listeners and LI of the identifiability of individual tones from the natural speech and synthetic speech stimulus sets is shown in Figure 7. For normal listeners, mean percentages of correct identification were 90% or above for all tones from either stimulus set except for the mid tone from the synthetic speech stimulus set; for LI, 90% or above for all tones from either stimulus set except for the mid and falling tones from the natural speech stimulus set. In the case of the natural speech tones, 88% and 80% of the total number of normal listeners' and LI's confusion errors, respectively, involved the mid or low tones. In the case of the synthetic speech tones, all of the normal listeners' and LI's confusion errors involved the mid or low tones. These combined findings indicate that LI's perception of Thai tones was comparable to that of normal listeners both quantitatively and qualitatively.

Production

Method

Subjects. Data were obtained from LI for the present study. For purposes of comparison, data on normal speakers' production of Thai tones were taken from a previously completed investigation (Gandour, Petty, & Dardarananda 1988).

Speech Materials. A set of five, monosyllabic Thai words, minimally distinguished by tone, was used in this investigation: /k^{h}aa/ 'to be stuck', /k^{h}àa/ 'galangal, a rhizome', /k^{h}âa/ 'to kill', /k^{h}áa/ 'to engage in trade/, /k^{h}ǎa/ 'leg/'. The same set of words was used in Gandour et al. 1988).

Recording Procedure. All five words were printed in large Thai letters on 3 x 5 cards. A minimum of 50 cards (5 words x 10 repetitions) were presented in a pseudorandom order to LI. He was instructed to produce the words in isolation. To avoid start and end effects in the production of these utterances, extra cards were placed at the top and bottom of the deck.

LI was tested in a sound-proof booth in a single session. His utterances were recorded using a Sony ECM-150 microphone and a Nagra IV-S tape recorder. LI was seated and wore a custom-made headband designed to hold the microphone 15 cm from his lips. Prior to actual recording, LI was screened for his ability to read the

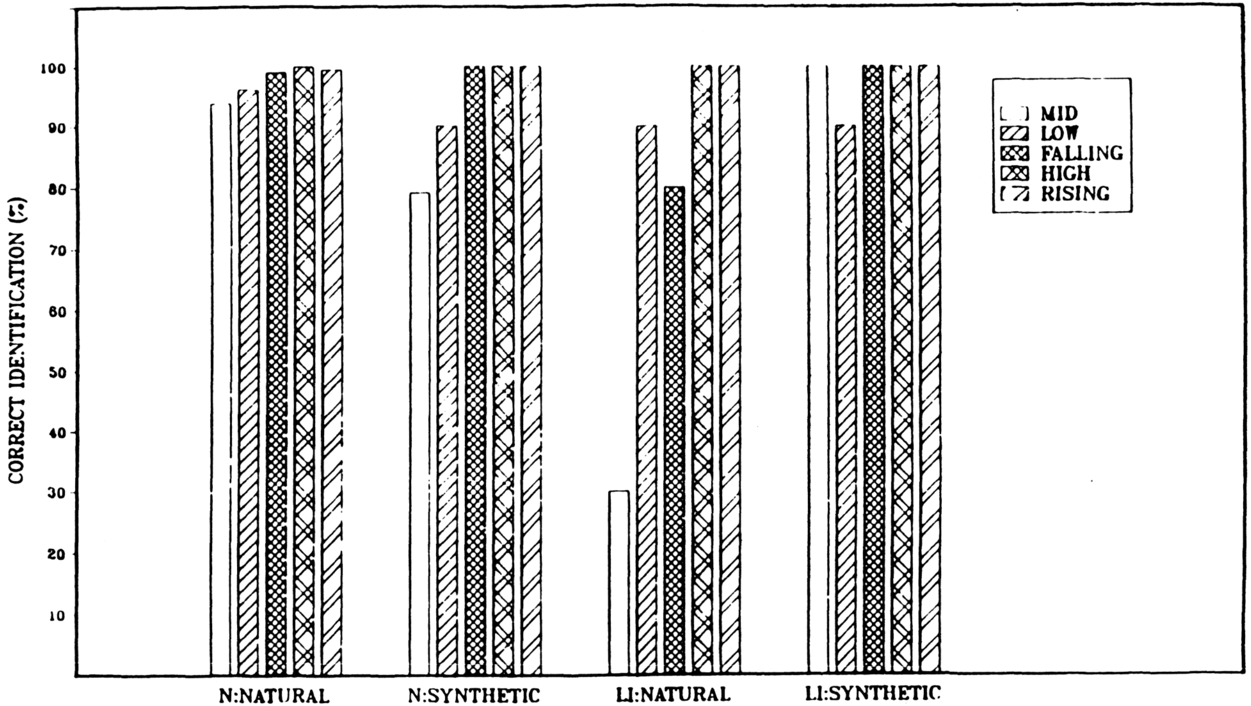


Fig. 7. Mean % of correct identification of each tone from the natural speech and synthetic speech stimulus sets for a group of normal Thai adults (N: adapted from Gandour et al. 1984) and LI.

cue cards. He was asked to point to the card upon hearing a native Thai speaker produce the word. This pretest consisted of a pseudorandom presentation of two occurrences of each word. LI responded correctly in all ten trials.

Listening Procedure. A closed-set identification procedure was used in the listening tests. Under this procedure, a trial consisted of the presentation of a single word. Productions of the five words were presented in a different random order for LI and the normal speakers. The response interval for each trial was 3.5 seconds.

Three native Thai-speaking adults participated in the listening tests. Each listener was provided with a set of response sheets containing the closed set of stimulus words for each trial. Response sheets were typed in Thai orthography. Listeners were instructed to identify each stimulus by circling the word on their response sheets corresponding to what they heard. The test tapes were played on a Nagra III tape recorder, and the signal was presented through TDH-39 headphones at a comfortable listening level (i.e., about 70 dB SPL).

Measurement Procedure. All utterances were low-pass filtered at 4 kHz and digitized at a 10 kHz sampling rate for speech analysis using ILS (Interactive Laboratory System, Signal Technology, Inc.) programs implemented on a Data General Nova 4 computer. F_0 contours were extracted using a cepstrum method of analysis. Syllable duration measurements were made to the nearest 10 ms applying cursors to the waveform and F_0-intensity curves displayed simultaneously on a Tektronix 4010 graphics terminal.

In addition, a measure was obtained of intraspeaker variability in F_0 contours for each of the five Thai tones produced by LI and one normal, male speaker. To compare the point-to-point variability of F_0 contours for repetitions of each tone from each speaker, an ensemble-averaging approach was taken (cf., Atkinson 1976, pp. 441-442). All repetitions of each tone made up an ensemble of F_0 contours for that tone by that speaker. Because all repetitions of each tone were not of equal duration, the F_0 contours were normalized for duration on a percentage scale. The ensemble mean and standard deviation were determined at each 10% interval. These ensemble-averaged statistics made it possible to evaluate changes in ensemble characteristics of intraspeaker contours for each tone as a function of time.

LI's intended productions of the rising tone were characterized by a creaky voice quality, sometimes interrupted by a glottal stop. Because of the resultant aperiodicity in the acoustic signal, F_0 measurements could not be performed reliably for his rising tone.

Results and Discussion

Mean percentages of correct identification of the five Thai tones produced by the group of normal speakers and LI are presented in Figure 8. In agreement with previous data on the perception of Thai tones (Abramson 1962, 1975, 1976; Gandour et al. 1984), Gandour et al.'s (1988) group of five normal speakers similarly achieved a near-perfect level of accuracy in signaling tonal contrasts. Listeners correctly identified all five tones produced by normal speakers at levels ranging from 95% to 100%. LI, on the other hand, was unable to signal any of the five tones at a level of proficiency comparable to that of normal speakers. Although somewhat better than chance, LI's productions of the mid and low tones achieved levels of correct identification of only 43% and 47%, respectively. His intended productions of the falling tone achieved only chance level of identification, of the high and rising tones below chance level. These findings reveal LI's tone system to be severely impoverished from the standpoint of production. However, LI's tonal perception is intact. The disparity in LI's performance on the perception and production tasks indicates that LI's tonal deficit is restricted to production only. This

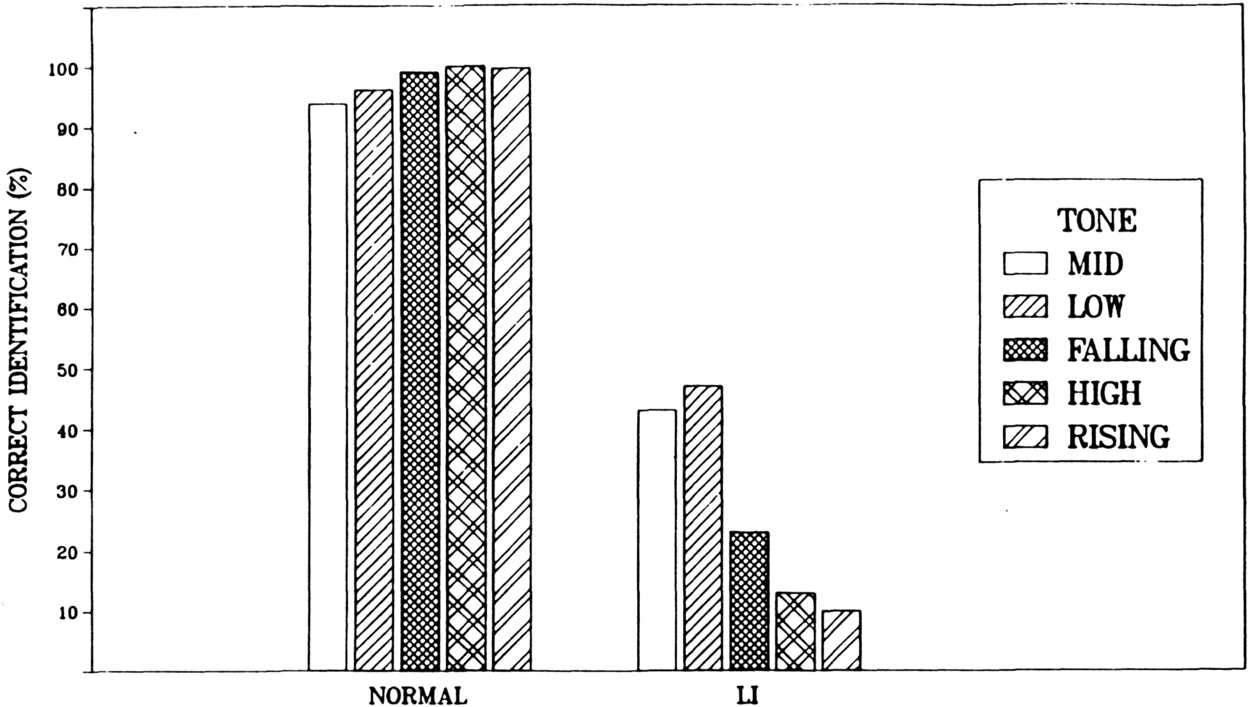


Fig. 8. Mean % of correct identification of each Thai tone produced by a group of normal speakers (adapted from Gandour et al. 1988) and LI.

dissociation between production and perception leads us to speculate that LI's problem is neuromotor underlyingly instead of psycholinguistic.

The listener responses to each of the five Thai tones produced by the group of normal speakers and LI are given in Table 8. Confusion between normal speakers' mid and low tones accounted for 86% of their total number of errors. In contrast, only 17% of listeners' identification errors of LI's tonal productions were confusions between his mid and low tones. Almost all (96%) confusion errors of LI's tonal productions involved substitution of either the mid or low tone. The mid tone was the predominant perceptual response irrespective of tonal category. Thus, LI's performance differed from that of normal speakers in kind as well as in degree.

TABLE 8. Confusion matrices of responses to Thai tones produced by normal speakers and LI

Normal (n = 5)[a]						
	Responses					
Stimuli	Mid	Low	Falling	High	Rising	
Mid	457	23				480
Low	15	475				490
Falling			490			490
High				500		500
Rising				6	484	490
	472	498	490	506	484	2450

LI						
	Responses					
Stimuli	Mid	Low	Falling	High	Rising	
Mid	13	13	1	2	1	30
Low	16	14				30
Falling	22	1	7			30
High	25		1	4		30
Rising	11	16			3	30
	87	44	9	6	4	150

Note: The total number of responses to each of the five phonemic tones is displayed in the form of a confusion matrix. Stimuli are listed by row, responses to the stimuli are entered in the cells along the main diagonal of the matrix, incorrect responses in the off-diagonal cells.
[a]Data taken from Gandour et al. (1988).

Average F_o contours normalized for time as a percentage of total duration of each of the five tones produced by one of the normal, male speakers and LI are displayed in Figure 9. For the normal speaker, F_o contours associated with the mid and low tones fall steadily throughout, and with the falling, high, and rising tones change in slope approximately halfway through their duration. The falling tone is relatively level and then falls sharply; the high and rising tones fall slightly and then rise. Average F_o contours for the other four normal speakers were similar to those shown in Figure 9. In terms of overall shape, LI's average F_o contours associated with the mid, low, falling, and high tones are reasonably similar to those of normal speakers. However, LI's mid and low tone productions do exhibit a considerably steeper slope than those of normals. In terms of relative position in the tone space, LI's F_o contours for the falling and high tones are below that for the mid tone. Just the opposite is true in the normal tone space. Despite the similarity in shape of LI's falling and high tonal contours to those of normal speakers, the combined acoustic-perceptual findings clearly show that shape alone is insufficient as a perceptual cue to Thai tones. Relative height of the tones in the tone space is also an important perceptual cue. The fact that listeners perceived practically all of LI's tonal productions to be either mid or low is accounted for by LI's restricted tone space. LI's productions of the mid, low, falling, high, and presumably rising tones throughout their duration occupy a region of the tone space normally occupied by the mid and low tones only.

The findings from this study converge with previous data on the chronological order in which Thai tones are acquired. In her longitudinal study of a normally-developing Thai child between the ages of 3 and 18 months, Tuaycharoen (1979) reported that her subject began using the mid and low tones lexically at the age of 0;11. At this age, both mid and low tones were produced correctly and consistently; mid and low tones were substituted for the falling, high, and rising tones. At 13 months, the child began to use the rising tone on lexical items. From the end of the 15th month to 18 months, the falling and high tones began to appear on lexical items. However, they were not produced consistently. Although no specific data are presented, Tuaycharoen stated that her subject did not master all the tones until the age of 23 months.

Interestingly, LI's impoverished tone system mirrors that of an early stage in the acquisition of Thai tones. By the age of 11 months, Thai children have acquired the mid and low tones, but not the falling ,high, and rising tones (Tuaycharoen, 1979). It appears not to be accidental that LI was able to produce the mid and low tones at a level of proficiency considerably higher than chance. The two tones that are acquired earlier in normal tone acquisition are precisely those tones that appear in LI's tone system.

The point-to-point variability in F_o contours of each of the five tones produced by one normal, male speaker and LI is displayed in Figure 10. The ensemble mean can be considered as the target F_o contour for a particular tone, the ensemble standard deviation as a measure of the random perturbations that result in deviations from the target contour. A series of four separate t-tests (collective level of significance = .04, alpha = .01) revealed that LI's ensemble-averaged standard deviations were significantly larger than those for the normal speaker on the mid and low tones only. The fact that LI exhibited increased variability in his

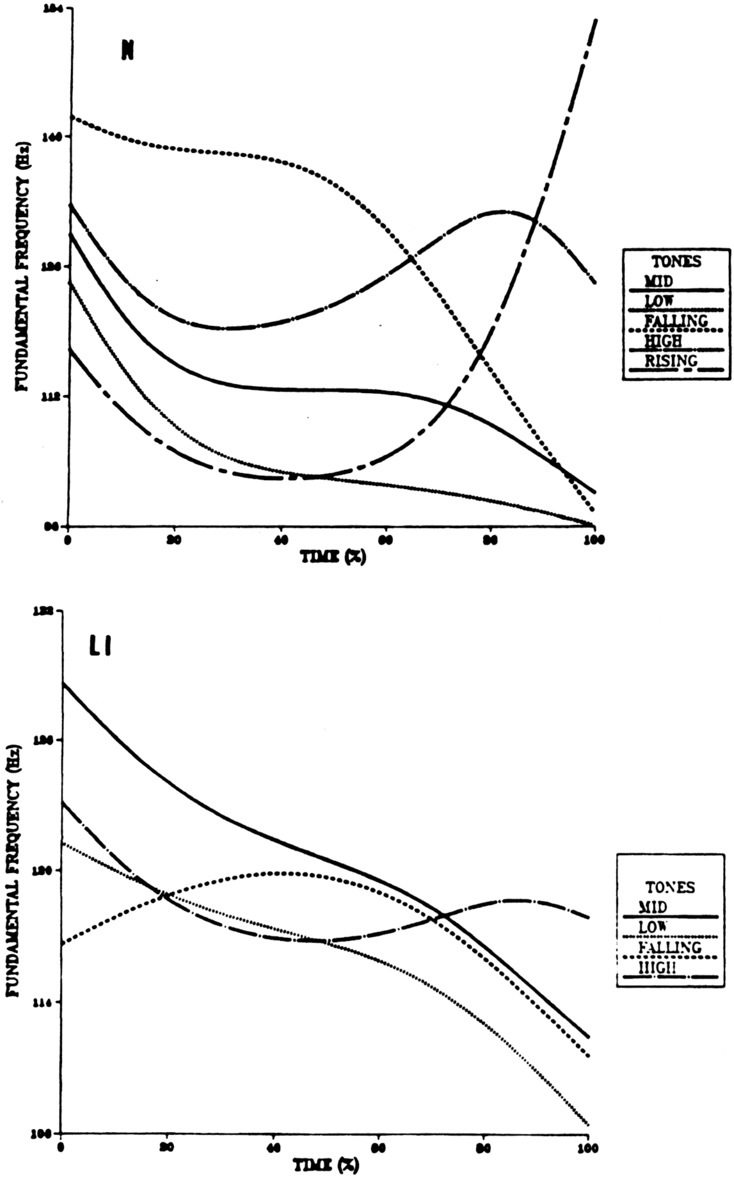
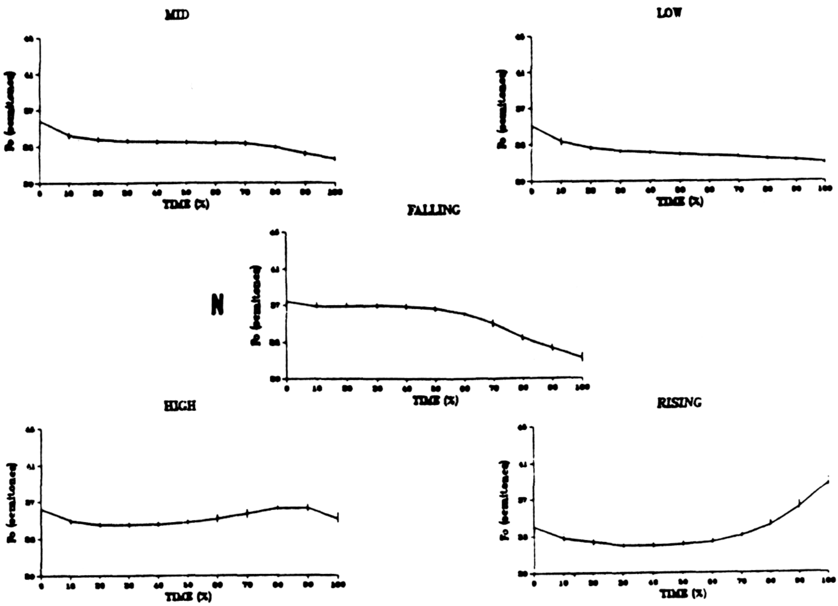
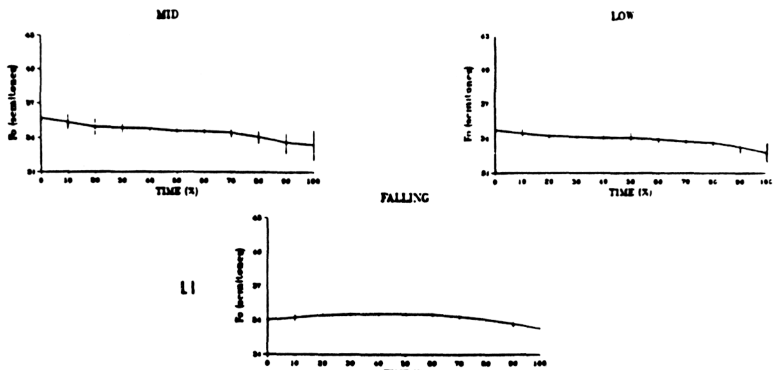


Fig. 9. Average F_o contours in Hz normalized for time on a % scale of Thai tones on citation forms of monosyllabic words for a normal male adult speaker (N: adapted from Gandour et al. 1988) and LI.





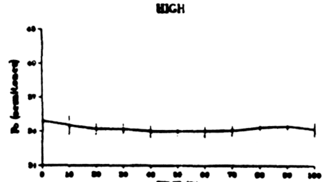


Fig. 10. Variability of F_o contours of the Thai tones in semitones normalized for time on a % scale for monosyllabic words produced in isolation by a normal male adult speaker (N) and LI.

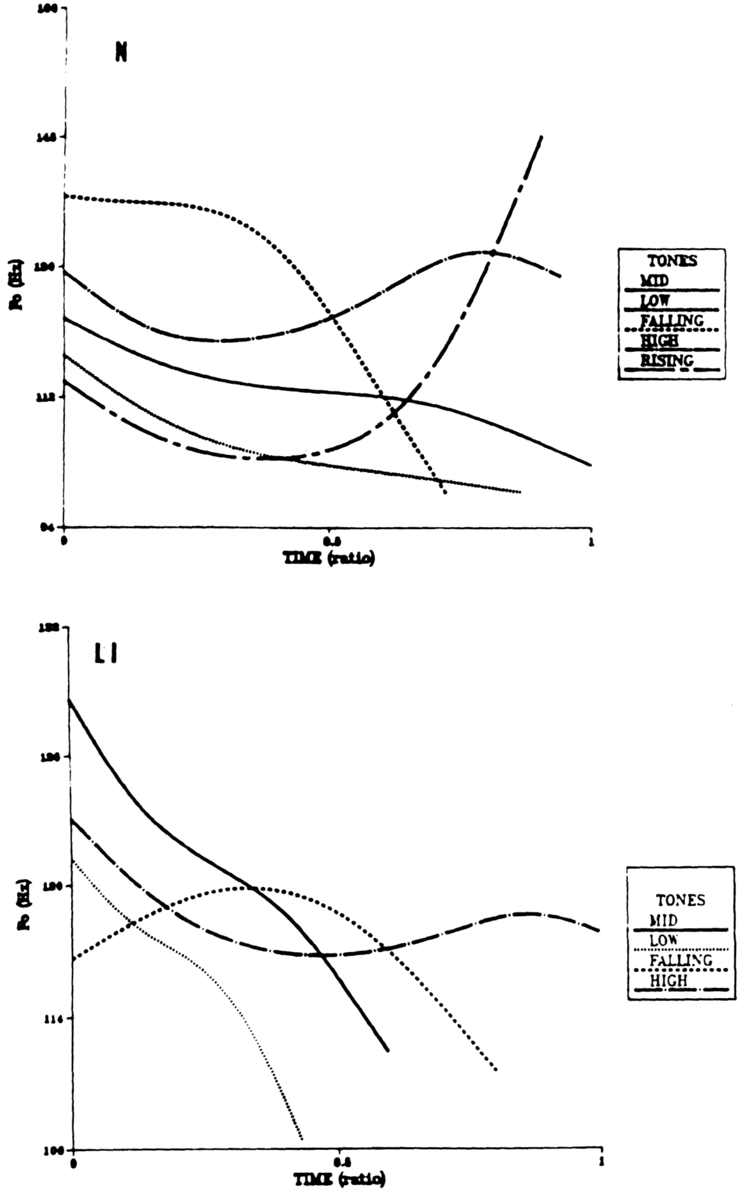


Fig. 11. Average F_0 contours in Hz normalized for time on a ratio scale for Thai tones on citation forms of monosyllabic words for a normal male adult speaker (N: adapted from Gandour et al. 1988) and LI.

intended productions of the mid and low tones, but not the high and falling tones, when compared to the normal speaker, emphasizes that his difficulty was not simply a phonetic one. He apparently was consistent in his production of F_0 contours associated with the high and falling tones. His problem was that his F_0 contours were not acceptable exemplars of those two particular tones.

In Figure 11, average F_0 contours of the five tones normalized for time as a ratio to the tone with the longest average duration are shown for one of the normal, male speakers and LI. For this normal speaker, as well as the other four not shown here, the mid tone was longer than the low tone, and the falling tone was the shortest of the five. In contrast, LI's low tone was the shortest. Whereas the normal's mid and low tones are both longer than falling, just the reverse is true for LI. By an absolute scale of measurement, LI's average tonal durations in ms (mid=164; low=118; falling=220; high=275) were all considerably shorter than those for the normal speaker (mid=605; low=523; falling=438; high=570). It is possible that abnormal durations of LI's F_0 contours on either absolute or relative measurement scales were responsible in part for listeners' identification errors.

General Discussion

The combined results from the three studies on timing indicate that LI's impairment in speech timing is not pervasive in the sense that all phonetic manifestations of speech timing are not equally impaired. LI's timing deficit in the production of VOT did not include the voiceless aspirated stops. Although his vowel durations were abnormally short, LI managed to maintain the relative contrast in vowel length in isolated monosyllables. His speech timing in connected speech, however, was grossly aberrant. LI's speech output was filled with lengthy pauses; his syllables were produced with equal stress. As a result, temporal patterns in Thai conditioned by stress or sentence position were severely disrupted. It is also possible that his timing difficulties in larger-sized linguistic units stem in part from his residual articulatory problems.

The results from the study on F_0 indicate that LI's tone space is aberrant. All five Thai tones were impaired to varying degrees. Only two of LI's tones, mid and low, were perceived at above chance levels of identification, but at levels substantially lower than those of normal speakers. Earlier research on the acquisition of tone (see Clumeck 1980; Li & Thompson 1978 for reviews) would lead us to expect tone to be fairly robust even in the case of delayed language development. LI's tonal difficulties were somewhat unexpected. It is true that LI exhibited a severe articulatory disorder at the time of diagnosis. Perhaps, in severe cases of delayed language development, tones are as vulnerable as consonants and vowels. To our knowledge, there are no earlier published reports on the speech of phonologically-impaired children who are native speakers of Thai, or for that matter, any other tone language. This is an empirical question that clearly warrants further research.

Footnote

Research supported in part by NIH Grant No. NS-24539 and NSF Grant No. BNS-7905854. The data were collected under the authority of the National Research Council of Thailand during the period from May to December 1980 while the author was affiliated with the Division of Communication Disorders, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand. The hospitality of Poonpit Amatyakul, Director, and his staff of the Speech and Hearing Clinic at

Ramathibodi Hospital is gratefully acknowledged. Reprint requests should be sent to Jack Gandour, Department of Audiology and Speech Sciences, Heavilon Hall, Purdue University, West Lafayette, Indiana, U.S.A. 47907

References

Abramson, A.S. 1962. The vowels and tones of Standard Thai: Acoustical measurements and experiments (Publication No. 20). Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics.

Abramson, A.S. 1974. Experimental phonetics in phonology: Vowel duration in Thai. Passa, 4(1), 71-90.

Abramson, A.S. 1975. The tones of Central Thai: Some perceptual experiments. In J.G. Harris & J. R. Chamberlain (Eds.), Studies in Tai linguistics in honor of William J. Gedney (pp. 1-16). Bangkok: Office of State Universities.

Abramson, A.S. 1976. Thai tones as a reference system. In T.W. Gething, J.G. Harris, & P. Kullavanijaya (Eds.), Studies in Tai linguistics in honor of Fang-Kuei Li (pp. 1-12). Bangkok: Chulalongkorn University Press.

Atkinson, J.E. 1976. Inter- and intraspeaker variability in fundamental voice frequency. Journal of the Acoustical Society of America, 60, 440-445.

Caramazza, A. & Martin, R.C. 1983. Theoretical and methodological issues in the study of aphasia. In J.B. Hellisge (Ed.), Cerebral hemisphere asymmetry: Method, theory, and application (pp. 18-45). New York: Praeger.

Clumeck, H. 1980. The acquisition of tone. In G.H. Yeni-Komshian, J.S. Kavanaugh & C.A. Ferguson (Eds.), Child Phonology, Volume 1: Production (pp. 257-275). New York: Academic Press.

Gandour, J. 1975. On the representation of tone in Thai. In J.G. Harris & J.R. Chamberlain (Eds.), Studies in Tai linguistics in honor of William J. Gedney (pp. 170-195). Bangkok: Office of State Universities.

Gandour, J. 1985. A voice onset time analysis of word-initial stops in Thai. Linguistics of the Tibeto-Burman Area, 8(2), 68-80.

Gandour, J., Carney, A., Nimitbunnasarn, C. & Amatyakul, P. 1984. Tonal confusions in Thai patients with sensorineural hearing loss. Journal of Speech and Hearing Research, 27, 89-97.

Gandour, J. & Dardarananda, R. 1982. Voice onset time in aphasia: Thai. I. perception. Brain and Language, 17, 24-33.

Gandour, J., Petty, S. & Dardarananda, R. (1988). Perception and production of tone in aphasia. Brain and Language, 35, 201-240.

Gandour, J., Petty, S., Dardarananda, R., Dechongkit, S. & Mukngoen, S. 1986. The acquisition and dissolution of the voicing contrast in Thai. Linguistics of the Tibeto-Burman Area, 9(2), 36-52.

Gandour, J., Weinberg, B., Petty, S. & Dardarananda, R. 1987a. Voice onset time in Thai alaryngeal speech. Journal of Speech and Hearing Disorders, 52, 288-294.

Gandour, J., Weinberg, B., Petty, S. & Dardarananda, R. 1987b. Vowel length in Thai alaryngeal speech. Folia Phoniatrica, 39, 117-121.

Kent, R.D., Netsell, R. & Abbs, J.H. 1979. Acoustic characteristics of dysarthria associated with cerebellar disease. Journal of Speech and Hearing Research, 22, 627-648.

Leonard, L.B. 1983. Discussion: Part II: Defining the boundaries of language disorders in children. In J. Miller, D.E. Yoder, & R. Schiefelbusch (Eds.), Contemporary issues in language intervention. ASHA Reports, 12 107-112.

Li, C.N. & Thompson, S.A. 1978. The acquisition of tone. In V.A. Fromkin (Ed.), Tone: A linguistic survey (pp. 271-284). New York: Academic Press.

Lisker, L. & Abramson, A.S. 1964. A crosslanguage study of voicing in initial stops. Word, 20, 384-422.
Lisker, L. & Abramson, A.S. 1970. The voicing dimension: Some experiments in comparative phonetics. In Proceedings of the Sixth International Congress of Phonetic Sciences (pp. 563-567). Prague: Academia Publishing House of the Czechoslovak Academy of Sciences.
Luangthongkum, T. 1977. Rhythm in standard Thai. Unpublished doctoral dissertation, University of Edinburgh.
Pike, K.L. 1945. The intonation of American English. Ann Arbor: University of Michigan Press.
Tuaycharoen, P. 1979. An account of speech development of a Thai child: from babbling to speech. In T.L. Thongkum, V. Panupong, P. Kullavanijaya, & M.R.K. Tingsabadh (Eds.), Studies in Tai and Mon-Khmer phonetics and phonology in honour of Eugenie J.A. Henderson (pp. 261-277). Bangkok: Chulalongkorn University Press.