The Acquisition and Dissolution of the Voicing Contrast in Thai

Jack Gandour and Soranee Holasuit Petty
Purdue University

Rochana Dardarananda, Sumalee Dechongkit, and Sunee Mukngoen Mahidol University

This paper reports data on the acquisition and dissolution of the voicing contrast in Thai word-initial stops, using voice onset time (VOT) as a phonetic measure of this phonological contrast. Thai 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 ph/) and alveolar (/d t th/) places of articulation and a two way opposition (voiceless unaspirated, voiceless aspirated) at the velar (/k kh/) place of articulation.

VOT is defined as the temporal relation between the onset of glottal pulsing and the release of the initial stop consonant (Lisker & Abramson, 1964). In relation to the moment of stop release, voicing detected before the release is called voicing lead, while voicing starting after the release is called voicing lag. Normal adult speakers of Thai produce distinct, nonoverlapping distributions of VOT for homorganic word-initial stops: voiced unaspirated stops in the long lead region of the VOT continuum, voiceless unaspirated stops in the zero or short lag region, and voiceless aspirated stops in the long lag region (Lisker & Abramson, 1964; Gandour, 1985). In contrast, Gandour & Dardarananda (1984) reported that VOT productions do not fall into discrete nonoverlapping regions for a group of three Broca's aphasics. No previous VOT acquisition studies of Thai word-initial stop consonants are available in the literature. However. Sarawit (1976) and Tuaycharoen (1979), in their longitudinal diary studies of individual children, a Thai-English bilingual and a Thai monolingual, respectively, have reported that voiceless unaspirated stops appear to be the first acquired, voiceless aspirated stops the next acquired, and voiced unaspirated stops the last acquired.

The aim of this paper is to determine the extent to which there exists a correspondence between adult dissolution and child acquisition of voicing contrasts in word-initial stops in Thai. Specifically, we will evaluate the degree to which the order of acquisition of voicing contrasts is similar to the order of dissolution, the degree to which the effects of number and type of voicing categories on the acquisition of stop consonants are similar to those on the dissolution of stop consonants, and the degree to which the VOT distribution pattern of adult aphasics is similar to that of normally-developing children.

Any attempt, however, to establish a relationship between the nature of the processes of production in development and those in pathology must be made with extreme caution. The adult aphasic is an individual who had a fully acquired and normally organized linguistic system and who then sustained brain damage that affected its normal processing. In contrast, the child's linguistic system is at best only partially developed or incompletely organized. Blumstein's (1978) cautionary remarks on speech perception in pathology and ontogeny also apply to speech production: "Consequently, any correspondence between the performance of aphasics and infants [children] on perception [production] tasks does not necessarily imply that the aphasic represents an ontogenetically earlier stage of language processing or that the perceptual [production] processes in pathology and development are the same. Nevertheless, such correspondences may reveal important evidence suggestive of the types of processes that seem to underlie speech perception [production] in man" (p.4).

Method

Subjects

Four groups of non-brain-damaged and one group of brain-damaged subjects participated in this study (see Table 1). The normal groups comprised seven young adults, seven 7-year-old children, seven 5-year-old children, and seven 3-year-old children. The brain-damaged group consisted of two Broca's aphasics. The patients were classified on the basis of results of a Thai diagnostic aphasic examination (Gandour 1982) and neurological tests. Both patients had unilateral lesions in the left frontal region. All subjects were native speakers or learners of Bangkok Thai.

Table 1.
Characteristics of Individual Subjects

Group ^a	Subject	Age ^b	Sex	Handedness	Etiology	Time Since Onset
A A A A A A	1 2 3 4 5 6 7	25;0 29;3 20;6 26;0 24;4 27;2 24;5	F F F M M F			
7 7 7 7 7 7	1 2 3 4 5 6 7	7;6 7;2 7;8 7;0 7;0 6;9 6;10	F F M F F F M			

5 5 5 5 5 5 5	1 2 3 4 5 6 7	4;9 4;8 5;0 5;3 5;1	M F F M F			
3 3 3 3 3 3	1 2 3 4 5 6 7	2;11 2;8 2;9 3;1 3;2 3;2 3;1	M F M F F			
B B	1 c 1 d	49;6 56;0	F M	Right Right	Stroke Stroke	3;0 0;3

A = adult group; 7 = 7-year-old group; 5 = 5-year-old group; 3 = 3-year-old group; B = Broca group.

Speech Materials

The speech sample consisted of a near-minimal triplet of the three way contrast in VOT for word-initial bilabial stops, a near-minimal triplet of the three way contrast in VOT for word-initial alveolar stops, and a minimal pair for the two way contrast in VOT for word-initial velar stops (see Table 2). Tone was held constant for each set of homorganic stops: bilabials, falling tone; alveolars, mid tone; and velars, low tone.

Table 2.
Test Stimuli

	Bilabial	Alveolar	Velar	
Voiced Unaspirated	/baan/'house'	/daaw/'star'		
Voiceless Unaspirated	/paa/'aunt'	/taa/'eye'	/kaj/'chicken'	
Voiceless Aspirated	/pʰaa/'cloth'	/thian/'candle'	/kʰaj/ˈeggˈ	

bin year; months

corresponds to B2 in Gandour & Dardarananda (1984).

dcorresponds to B3 in Gandour & Dardarananda (1984).

Recording Procedure

High quality recordings of multiple productions of the eight words spoken in isolation were made in a soundproof booth at the Speech and Hearing Clinic of Ramathibodi Hospital. Recordings were made on a Nagra IV-S tape recorder with a Sony ECM-150 microphone. Children sat comfortably on the floor of the booth. They were recorded with the microphone attached to the front of their clothing approximately 15 cm from the lips. Each adult speaker was seated and wore a custom-made headband designed to hold the microphone 15 cm from the lips. All recordings were completed in a single session.

Children and aphasics were cued with pictures of the referents of the eight stimulus words. They were instructed to name the referent shown in the picture. Normal adult speakers were cued with index cards containing the written Thai words. They were instructed to say each word in isolation. The pictures or index cards were presented in a different random order to each subject.

Measurement Procedure

Subject's productions were analyzed using a Voiceprint 700 sound spectrograph. VOT values were measured to the nearest 5 msec from conventional wideband spectrograms along with overall amplitude traces. For a more detailed description of spectrographic measurements, see Gandour (1985). The total number of observations subjected to statistical analysis was 64 for each non-brain-damaged subject; 85 and 97 for brain-damaged subjects Broca 1 and Broca 2, respectively.

Results and Discussion

Normal Adults

Figure 1 shows the frequency distributions of VOT values associated with homorganic word-initial stops for the group of normal adults. No overlap in VOT distribution is seen between any of the homorganic stop consonants. The values for all stop categories cluster into discrete regions along the VOT continuum. VOT values for /b/ and /d/ lie in the long voicing lead region; for /ph/, /th/, and /kh/ in the long voicing lag region; and for /p/, /t/, and /k/ at zero or in the short voicing lag region. These findings for our seven normal adults are in essential agreement with Lisker & Abramson's (1964) VOT data for three normal speakers and Gandour's (1985) VOT data for five normal speakers.

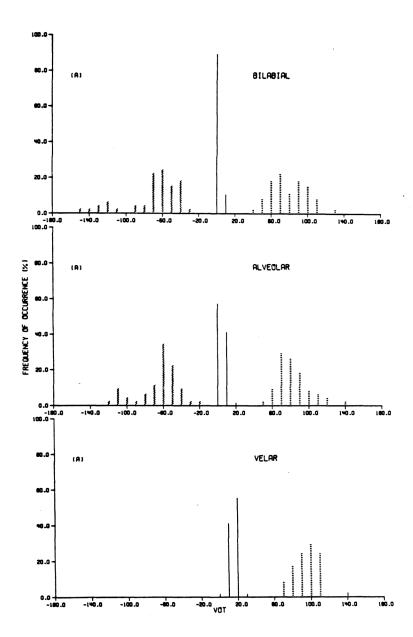


Figure 1. Distribution of VOT productions for bilabial, alveolar, and velar stop consonants in word-initial position for a group of seven normal Thai adults. The abscissa of each graph represents VOT values in 10-msec intervals and the ordinate the percentage of productions falling within the specified VOT interval. Voiced unaspirated stops, /b/ and /d/, are indicated by the slashed lines; voiceless unaspirated stops, /p/, /t/ and /k/, by the solid lines; and voiceless aspirated stops, /ph/, /th/, and /kh/, by the striped lines.

Children vs. Normal Adults

Figures 2, 3, and 4 show the frequency distributions of VOT values associated with homorganic word-initial stops for the groups of 3-year-old, 5-year-old, and 7-year-old children, respectively. Considerable overlap in VOT distribution between voiced and voiceless unaspirated stop consonants is observed in the groups of 3-year-old and 5-year-old children. Over half of VOT values associated with /b/ and /d/ lie in the short voicing lead region or at zero for 3-year-old and 5-year-old children. Minimal overlap is observed between voiceless unaspirated and voiceless aspirated stops at all three places of articulation. Only sporadic occurrences of overlap in VOT distribution between homorganic stop consonants are seen in the group of 7-year-old children. VOT values fall into separate regions along the VOT continuum for all stop categories.

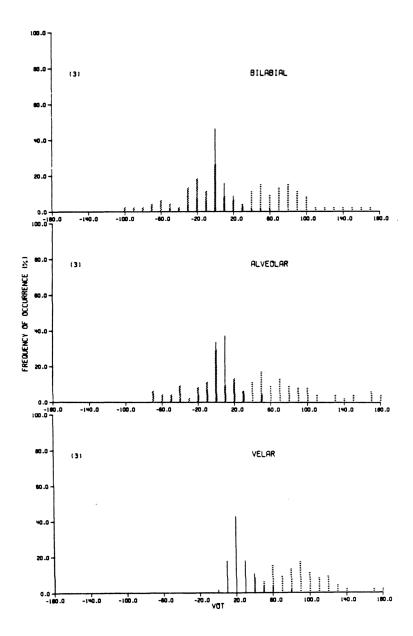


Figure 2. Distributions of VOT productions for bilabial, alveolar, and velar stop consonants in word-initial position for a group of seven 3-year-old children.

See also caption for Figure 1.

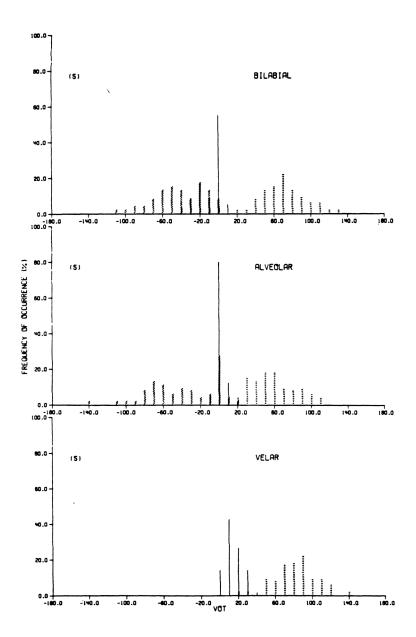


Figure 3. Distribution of VOT productions for bilabial, alveolar, and velar stop consonants in word-initial position for a group of seven 5-year-old children.

See also caption for Figure 1.

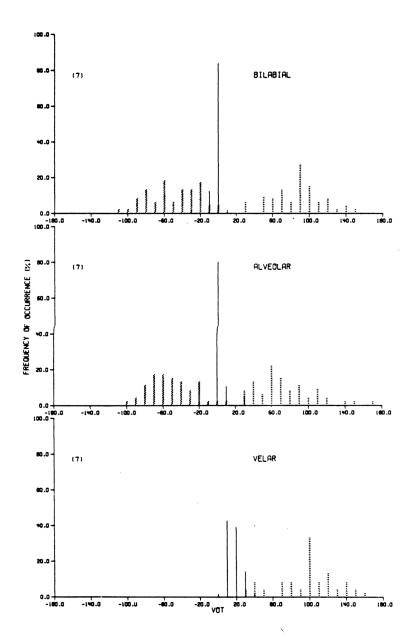


Figure 4. Distribution of VOT productions for bilabial, alveolar, and velar stop consonants in word-initial position for a group of seven 7-year-old children.

See also caption for Figure 1.

Results of a three factor (age x voicing category x place of articulation) analysis of variance with repeated masures on two factors showed that mean VOT differed significantly as a function of age (F=3.38; df 3.24; p < .0345), voicing category within place of articulation (F=359.86; df 5,120; p < .0001), and age x voicing category interaction (F=4.69; df 15,120; p < .0001). Post hoc comparisons revealed that differences in mean VOT as a function of age group were significant for /b/ and /d/ only: adults vs. 3-year-olds (p < .01); adults vs. 5-year-olds (p < .05); 7-year-olds vs. 3-year-olds (p < .01); 5-year-olds vs. 3-year-olds (p < .05). Adults, 7-year-olds, 5-year-olds, and 3-year-olds showed mean VOT values of -65.4, -48.2, -39.8, and -17.3 msec, respectively, for /b/; -63.3, -49.5, -36.3, and -6.0 msec, respectively, for /d/. No other comparisons between age groups in the mean VOT of a stop consonant were significant. Within each age group, post hoc comparisons were all significant (p < .01) except for /b/ vs. /p/ and /d/ vs. /t/ in the group of 3-year-olds. These findings are consistent with those reported in VOT acquisition studies of word-initial stops in other languages that have a voiced unaspirated category characterized by long voicing lead. Macken and Barton (1980) found that Spanish children did not produce the voiced stop category with voicing lead until after age four. Allen (1985) reported that lead voicing was nearly absent for phonemically voiced stops in French children aged 1:9 to 2:8.

Aphasics vs. Normal Adults

Figures 5 and 6 show the frequency distributions of VOT values associated with homorganic word-initial stops for two Broca's aphasics. Considerable overlap in VOT distribution, especially between voiceless unaspirated and voiceless aspirated stops, is observed at the velar place of articulation for Broca 1 and at all three places of articulation for Broca 2. the majority of Broca 1's VOT values associated with $/k^h/(53.9\%)$ are less than +70 msec, the minimum value observed for the voiceless aspirated velar stop in the group of normal adults. The majority of Broca 2's VOT values associated with $/p^h/$ (64%), $/t^h/$ (56%), and $/k^h/$ (67%) are less than +45, +55, and +70 msec, respectively, the minimum values observed for the three voiceless aspirated stops in the group of normal adults. Thus, both Broca's aphasics exhibited a tendency to produce voiceless aspirated stops with a shorter voicing lag than that of normal adult speakers. Although some overlap in VOT distribution is seen between voiced unaspirated and voiceless unaspiratd stops, VOT productions associated with both /b/ and /d/ predominately lie in the long voicing lead region of the VOT continuum. For /b/, 56% and 50% of VOT values are -50 msec or longer in voicing lead in the productions of Broca l and Broca 2, respectively; for /d/, 75% and 60%, respectively. 1's VOT distributions associated with voiceless unaspirated stops are not markedly different from those of normal adults, with the majority of VOT values lying at zero or in the short voicing lag region. contrast, Broca 2's VOT productions of voiceless unaspirated stops are considerably longer than those of normal adults at all three places of articulation. The majority of Broca 2's VOT values associated with /p/ (62%), /t/ (100%), and /k/ (56%) are greter than +10, +20, and +30 msec, respectively, the maximum values observed for the three voiceless unaspirated stops in the group of normal adults.

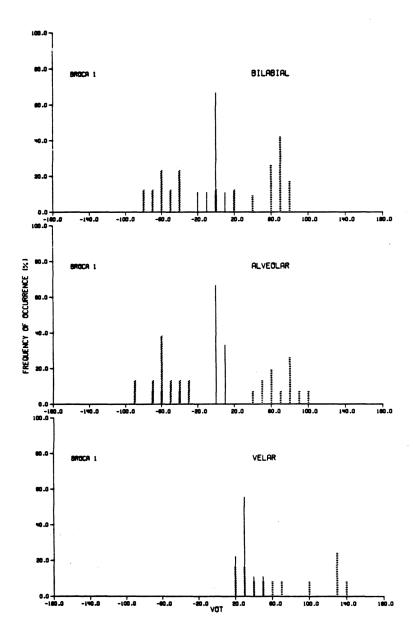


Figure 5. Distribution of VOT productions for bilabial, alveolar, and velar stop consonants for Broca's aphasic (Broca 1). See also caption for Figure 1.

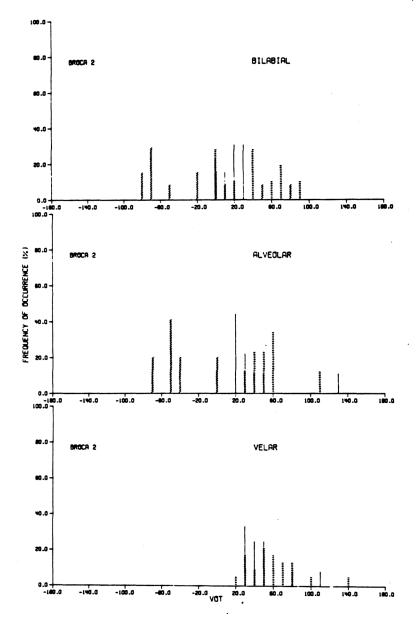


Figure 6. Distribution of VOT productions for bilabial, alveolar, and velar stop consonants for another Broca's aphasic (Broca 2). See also caption for Figure 1.

In an earlier investigation of VOT production in aphasia (Gandour & Dardarananda, 1984; cf. Figures 7-9), these same two Broca's aphasics exhibited substantial overlap in the frequency distribution of VOT values at all three places of articulation. In that earlier study, both Broca's aphasics exhibited shorter-than-normal voicing lag for voiceless aspirated stops and shorter-than-normal voicing lead for voiced unaspirated stops. This pattern of VOT production resulted in a marked compression of the VOT continuum. In agreement with Gandour & Dardarananda's (1984) findings, both Broca 1 and Broca 2 similarly produced shorter-than-normal voicing lag for voiceless aspirated In disagreement with their findings, neither Broca 1 nor Broca 2 tended to produce voiced stops with a shorter-than-normal voicing lead. Although the earlier study used a different set of Thai words and an imitation task, it is not readily apparent how to account for the discrepancy in VOT productions associated with voiced stops in terms of those methodological differences. We should point out, however, that the deficit pattern in VOT production for a given stop category may vary from individual to individual of the same aphasic syndrome. For example, Itoh, Sasanuma, Tatsumi, Murakami, Fukusako, & Suzuki (1982) reported that for voiced stops, three out of four Japanese anterior aphasics tended to produce shorter voicing lead than normals, but one tended to produce longer voicing lead. More research is clearly warranted to determine the extent to which the performance of an individual aphasic may vary on VOT production as a function of stimuli, stop category, and task.

The two Broca aphasics in the present study evidenced significant deficits in VOT production. All previous investigations of VOT production in aphasia in English (Blumstein, Cooper, Zurif, & Caramazza, 1977; Blumstein, Cooper, Goodglass, Statlender, & Gottlieb, 1980; Freeman, Sands, & Harris, 1978; Kent & Rosenbek, 1983) and Japanese (Itoh et al., 1982) have reported similar findings for Broca's aphasics or apraxic patients. However, the extent of overlap and compression of the voicing categories along the VOT continuum in these anterior patients varies from individual to individual within each language. In the present study, Broca 2 exhibited a more severe impairment in VOT production than Broca 1.

In agreement with all previous interpretations of VOT production deficits in Broca's aphasics or patients with apraxia of speech (Blumstein et al., 1977, 1980; Freeman et al., 1978; Itoh et al., 1982; Kent & Rosenbek, 1983), we also conclude that the underlying basis of the disorder is primarily phonetic rather than phonological. The aberrant VOT distribution patterns indicate that these anterior patients have difficulty in timing vocal fold vibration relative to the moment of release of the oral articulator. That is, the motor programming for the synchronization of the laryngeal and supralaryngeal articulatory gestures is impaired.

Children vs. Aphasics

A comparison of the VOT distribution patterns of the two Broca's aphasics (Figures 5-6), especially Broca 2, to those of the group of 3-year-old children (Figure 2) indicates that the overlap of VOT values associated with opposing stop categories is more extensive in the group of Broca's aphasics. Not only do aphasics and 3-year-olds

differ in the extent of overlap, but also in the particular stop categories that are principally involved in the overlap of VOT values. In the group of 3-year-olds, considerable overlap is observed between voiced unaspirated and voiceless unaspirated stops (i.e., /b/ vs. /p/, /d/ vs. /t/), whereas in the group of Broca's aphasics overlap is much more extensive between voiceless unaspirated and voiceless aspirated stops (i.e., /p/ vs /ph/, /t/ vs /th/, /k/ vs. /kh/). Overlap in VOT production in the group of 3-year-olds stems primarily from a shorter-than-normal voicing lead for voiced stops. In the group of Broca's aphasics, overlap in VOT production is caused by a shorter-than-normal voicing lag for voiceless aspirated stops as well as a slightly longer-than-normal voicing lag for voiceless unaspirated stops.

As measured by differences in mean VOT, 3-year-old Thai children were already producing adult-like VOT productions of voiceless aspirated stops at all three places of articultation. Consequently, we were unable to track the developmental changes in VOT over time. In their VOT acquisition study of English-learning children, Barton & Macken (1980) suggested that children first overshoot adult VOT values of voiceless aspirated stops and then only gradually draw them back towards adult values. Assuming that language breakdown is a mirror image of language development, we might predict that Broca's aphasics would similarly overshoot VOT values associated with long lag stops. Instead, they tend to undershoot VOT values of long lag stops, which leads to a shorter-than-normal voicing lag. If Barton & Macken's (1980) interpretation of their English data is supported by VOT acquisition data from other languages that have long lag stops, then it would appear that children and aphasics are dissimilar with respect to VOT productions associated with voiceless aspirated stops. In the absence of transition data on Thai voiceless aspirated stops in the present study, we of course have no way to resolve this issue conclusively.

In the case of both children and aphasics, the effects of the number and type of opposing stop categories on VOT production are inextricably confounded with each other. For the two way contrast between /k/ and $/k^h/$, 3-year-olds exhibit minimal overlap, whereas the two Broca's aphasics exhibit extensive overlap. The same pattern of VOT distribution between voiceless unaspirated and voiceless aspirated stops for the group of 3-year-olds and the Broca 2 aphasic can also be observed in the three way contrasts at the bilabial and alveolar places of articulation. Prima facie, 3-year-olds find two way contrasts to be less difficult than three way contrasts. voiced unaspirated stop category is missing at the velar place of articulation. For the group of 3-year-olds, overlap occurred primarily between voiced and voiceless unaspirated stops. basis of overlap in VOT distribution, Broca 1 had more difficulty with two way than three way contrasts. Broca 1's outcome is opposite to that expected if the number of opposing stop categories is a primary influence on VOT production. Thus, in the case of both children and aphasics, it appears that the type of stop category, instead of the number of categories, primarily determines the pattern of VOT distribution.

Previous diary studies of two Thai-learning children (Sarawit. 1976; Tuaycharoen, 1979) reported that voiceless unaspirated stops are the first acquired. The findings of the present study indicate that the voiced stops are the last acquired. Collectively, the findings from the earlier Thai diary studies, from the present study, as well as from stop acquisition studies in other languages indicate that stop voicing categories in word-initial position are acquired, in order. voiceless unaspirated (zero or short lag), voiceless aspirated (long lag), and voiced unaspirated (long lead). Kewley-Port & Preston (1974) and Cooper (1977) hypothesized, on the basis of articulatory and aerodynamic considerations, that zero or short lag stops are easier to produce than either long lag stops or stops with voicing lead. No claims were made about the relative complexity of stops with long voicing lag vs. stops with long voicing lead. The later development of adult-like productions of Thai voiced stops suggests that lead voicing is more difficult to produce than zero or short lag voicing. Moreover, the late emergence of Thai voiced stops as opposed to voiceless aspirated stops suggests that long lead voicing is inherently more difficult to produce than long lag voicing.

Data from Thai aphasics is not as clearcut as regards the relative complexity of long lead and long lag stops. In the present study, both Broca's aphasics apparently had more difficulty in producing long lag stops than long lead stops. However, in Gandour & Dardarananda (1984), these same two aphasics exhibited shorter-than-normal voicing lead as well as shorter-than-normal voicing lag. Taken together, these findings indicate that more research is required before we can reach any firm conclusions about the relative complexity of long lead vs. long lag stops based on VOT production in language breakdown.

In conclusion, our VOT production data do not suggest a simple mirror-image relationship between language development and language breakdown. To the contrary, differences observed in the VOT distribution patterns of children and aphasics suggest that the parameters modifying articulatory competence are not identical in each. The phonological and phonetic systems of the child reflect different organizations from those of the adult aphasic. Thus, dissolution of language cannot be described simply in terms of a relapse to an earlier stage of acquisition.

Acknowledgments

Research was supported in part by National Science Foundation Grant No. BNS-7905854 and National Institute of Neurological and Communicative Disorders and Stroke Grant No. NS-18116 to the first author. The data were collected during May-December 1980 while the first author was affiliated with the Division of Communication Disorders, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand. The hospitality and cooperation of Poonpit Amatyakul, Director of the Speech and Hearing Clinic, Ramathibodi Hospital, is gratefully acknowledged. Requests for reprints should be sent to Jack Gandour, Department of Audiology and Speech Sciences, Heavilon Hall, Purdue University, West Lafayette, IN 47907.

References

- Allen, G.D. (1985). How the young French child avoids the pre-voicing problem for word-initial voiced stops. <u>Journal of Child Language</u>, 12, 37-46.
- Barton, D., & Macken, M.A. (1980). An instrumental analysis of the voicing contrast in word-initial stops in the speech of four-year-old English-speaking children. <u>Language and Speech</u>, 23, 159-169.
- Blumstein, S.E. (1978). The perception of speech in pathology and ontogeny. In A. Caramazza & E. B. Zurif (Eds.), <u>Language acquisiton and language breakdown</u> (pp. 3-23). Baltimore: Johns Hopkins University Press.
- Blumstein, S., Cooper, W.E., Goodglass, H., Statlender, S., & Gottlieb, J. (1980). Production deficits in aphasia: A voice-onset time analysis. <u>Brain and Language</u>, 9, 153-170.
- Blumstein, S., Cooper, W.E., Zurif, E.B., & Caramazza, A. (1977). The perception and production of voice-onset time in aphasia.

 Neuropsychologia, 15, 371-383.
- Cooper, W.E. (1977). The development of speech timing. In S.J. Segalowity & F.A. Gruber (Eds.), <u>Language development and neurological theory</u>, (pp. 357-373). New York: Academic Press.
- Freeman, F.J., Sands, E.S., & Harris, K.S. (1978). Temporal coordination of phonation and articulation in a case of verbal apraxia: A voice onset time study. Brain and Language, 6, 106-111.
- Gandour, J. (1982). A diagnostic aphasia examination for Thai.

 <u>Linguistics of the Tibeto-Burman Area, 6</u> (2), 65-76.
- 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. & Dardarananda, R. (1984). Voice onset time in aphasia: Thai. II. production. Brain and Language, 23, 177-205.
- Itoh, M., Sasanuma, S., Tatsumi, I.F., Murakami, S., Fukusako, Y., & Suyuki, T. (1982). Voice onset time characteristics in apraxia of speech. Brain and Language, 17, 193-210.
- Kent, R.D., & Rosenbeck, J.C. (1983). Acoustic patterns of apraxia of speech. <u>Journal of Speech and Hearing Research</u>, 26, 231-249.
- Kewley-Port, D. & Preston, M.S. (1984). Early apical stop production: a voice onset time analysis. <u>Journal of Phonetics</u>, 2, 195-210.
- Lisker, L. & Abramson, A.S. (1964). A crosslanguage study of voicing in initial stops. Word, 20, 384-422.

- Macken, M.A. & Barton, D. (1980). The acquisition of the voicing contrast in Spanish: a phonetic and phonological study of word-initial stop consonants. <u>Journal of Child Language</u>, 7, 433-458.
- Sarawit, M. (1976). Assimilation in child language acquisition. In T. W. Gething, J. G. Harris, & P. Kullavanijaya (Eds.), <u>Tailinguistics in honor of Fang-Kuei Li</u> (pp. 205-213). Bangkok: Chulalongkorn University 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.