STABILITY OF ACOUSTIC CUES OF THE THREE-WAY VOICING CONTRAST IN THAI MOTHERS' STOP PRODUCTION

Chutamanee Onsuwan[1, 2], Juthatip Duangmal[1], Nawasri Chonmahatrakul[1], Naoto Yamane[3], Hyun Kyung Hwang[3], Reiko Mazuka[3,4]

[1]MARCS-CILS NokHook BabyLab and [2]Department of English and Linguistics, Faculty of Liberal Arts, Thammasat University, Thailand
[3]Laboratory for Language Development, RIKEN Center for Brain Science, Japan
[4] Department of Psychology & Neuroscience, Duke University, U.S.
consuwan@tu.ac.th; juthatip.d@gmail.com; nawasri@hotmail.com; yamanen@brain.riken.jp; hyunkyung.hwang@riken.jp; reiko.mazuka@riken.jp

ABSTRACT

Recent findings have shown age-related and possibly sex-related shift in stops' VOT values in some Asian languages (e.g., Japanese and Korean). This study examines VOT modifications of Thai stop contrasts (/b/, /p/, and /ph/) in infant directed speech (IDS). Ten Thai mother-infant dyads were recruited; infants were from two age groups: 4–6 and 8–10 months. The mothers produced repetitions of target syllables in carrier sentence while interacting with their infant and later with experimenter (ADS). Durational measures were taken from the stops and the following vowel. Durational differences in IDS vs. ADS of same mothers within each group (younger vs. older infants) were compared. Vocalic differences (longer in IDS) were significant in all factors. Overall, VOT differences in IDS vs. ADS were non-significant, but slightly wider interquartile range and more instances of varied phonetic realizations were observed in IDS, such as "bleed" in [b] and fricated stop in [ph].

Keywords: infant directed speech, mothers, stops, Thai, voice onset time (VOT)

1. INTRODUCTION

1.1. Changes in progress for stop voicing contrasts in Japanese and Korean

Among many acoustic cues for stop voicing contrasts (i.e., F0, vowel duration), VOT values have been the most widely studied and have played a significant role in differentiation of voicing contrasts [1, 7, 15, 20]. Recently, a number of studies in languages such as Japanese and Korean have reported changes in progress with regards to stop contrasts and shifts in VOT values. In several dialects of Japanese (voiced vs. voiceless contrast), there are regional and generational shifts with voiced stop from voiced to "devoiced" in young speakers, especially greater among female speakers [23, 24]. In Korean (Seoul dialect) (tense (or fortis), lax (or lenis), and aspirated stops), in phrase-initial position there is a VOT contrast reduction between lax and aspirated stops while F0 has become a more important cue. Again, it appears that the shift is more advanced among young female talkers [2, 21].

1.2. Modifications and phonetic variability in IDS

Substantial body of research has reached general agreement that speech directed to infants (and to young children) is modified in many ways and differs considerably from speech directed to adults [9, 13, 14]. In current literature, the speech modifications have manifested in the form of hyper and hypo-articulation, phonetic cue enhancement, and some degree of phonetic/phonological variation. It has been shown that prosodic modifications [10] and enhancement of phonetic cues in vowel quality [12, 19] are common in IDS. Fewer studies have looked at segmental aspects of IDS. Sundberg and Lacerda (1999) investigated VOT production in IDS in Swedish and reported some degree of non-canonical phonetic realizations (e.g., fricated stops) in IDS as well as ADS [22]. Interestingly, Foulkes et al. (2005) examined the use of variant (t) in Tyneside (England) and found that patterns of phonetic variants of (t) were present in child directed speech (CDS), but differed from those in ADS [11].

1.3. Variation of VOTs in IDS

Voicing contrasts in terms of VOT values in IDS (vs. ADS) have been documented in a few studies with somewhat mixed and inconclusive results. For English, Malsheen (1980) reported a significant overlap reduction in a subset of data [17] while Baran et al. (1977) found overall adult-like VOTs in IDS [3]. On the other hand, Swedish voiced and voiceless stops showed significantly shorter VOTs in IDS than in ADS [22]. Moreover, recent findings from Nepali stops (a four-way contrast) showed hypo-articulated trends in IDS in voiced and voiceless stops [4].

2. THAI STOPS

Thai three-way stop contrasts (voiced: /b/ /d/, voiceless unaspirated: /p/ /t/ /k/, and voiceless aspirated: /ph/ /th/ /kh/) have been a focus of many pioneering research and extended studies on VOTs [7, 15, 20]. However, little is known as to whether VOTs in Thai stops vary as a function of phonetic/ phonological context, speech register, and sociolinguistic variable.

2.1. VOT and contrastive vowel length

Stability of VOTs in Thai voiceless stops across phonemic vowel length (long vs. short) context has been reported in [18]. Despite the fact that Thai long vowels are generally twice as long the short counterparts, phonemic vowel length did not show any significant influence on preceding VOT.

2.2. Variation of (kh)

Chalermsanyakorn (2000) investigated variation of (kh) which exhibits two variants: [kh] (standard form) and [kxh] (fricated stop). The latter appeared more frequently in stressed position, when followed by /i(i)/, and more prevalent among younger speakers (sex was not a research variable) [6].

The present study is part of a larger cross-linguistic investigation of stops in Asian languages (Japanese, Korean, and Thai). Given the gaps in previous research, the current study was set out to explore VOT modifications of Thai stop contrasts: /b/, /p/, and /ph/ among young Thai females in IDS vs. ADS. Moreover, non-canonical instances of the stops were examined and categorized.

3. METHOD

3.1. Experimental setup

3.1.1. Speech materials

A few sets of monosyllabic pseudowords were constructed focusing on initial stops and lexical tones. For this study, [bee], [pee], [phee], all carry mid tone, were analyzed. Each target word was produced in a stressed and non-final position within a carrier sentence. [nîi ___khà lûuk/ ták].

3.1.2. Thai mother-infant dyads

Ten Thai mother-infant dyads were recruited and participated in an infant perception study prior to recording session. The mothers were ranging in age from 24–38 years old (M = 32.5 years). The infants were of two age groups: 4–6 months (2 girls, 3 boys) and 8–10 months (2 girls, 3 boys).

3.1.3. Recording procedure

In a sound-attenuated room, each mother wore a head-mounted microphone and was recorded (on TASCAM DR-100MKIII Linear PCM Recorder) while sitting and interacting, using a set of attractive toys labeled with the target words (identical toys and labels for all mothers), with their own infant (IDS) and then later with an experimenter (ADS). The experimenter was present at all times, the infant was only in the room during the IDS recording. The mothers were asked to naturally produce a series of carrier sentences. Each contains a target word and is repeated at least five times. They were allowed to practice before the real recording began. Speaking rate was not directly controlled since the mothers were encouraged to talk comfortably at their normal speaking rate. It should be noted that the mothers' IDS was generally slower than their ADS. A more natural way to elicit IDS data is possible, but here our intention was to maintain similar phonetic contexts across IDS and ADS recordings.

3.2. Acoustic analysis

Durational measures for initial stop's VOT and the following vocalic portion of 300 speech tokens (3 targets × 5 repetitions × 2 registers (IDS vs. ADS) × 10 mothers) were taken manually using Praat computer software [5]. VOT measurements followed guidelines proposed in [1, 8].

Second and (sometimes) third coders took part in analysis of non-canonical instances. A number of tokens (19 out of 300; 6.3%) were excluded from our quantitative analysis because of one or more of the following factors:

- Disfluency (1% in [b] ADS). Pause or filled pause was inserted before the target word.
- Lack of pre-voicing in [b]. See section 5.2.1.
- Fricated stop in [ph]. See section 5.2.2.
- Incomplete closure in [ph] See section 5.2.2.
- Voicing in [ph]. See section 5.2.2. .

4. RESULTS

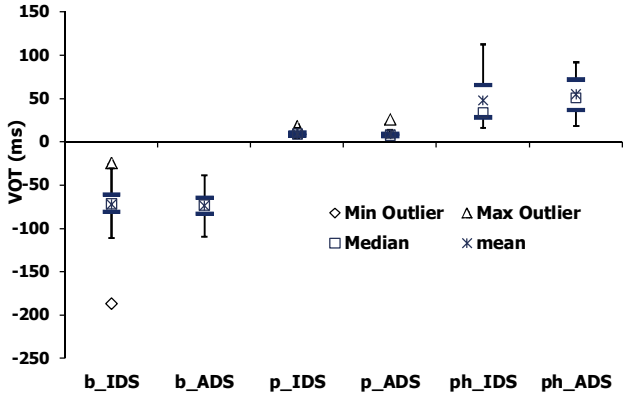
A three-way repeated measures of Analysis of Variance was conducted for VOT, and another for vowel duration where Register (IDS vs. ADS) and (Voicing) Type (/b/ vs. /p/ vs. /ph/) as within subject factors and Infant Age (4–6 months vs. 8–10 months) as a between subject factor.

Durational measures of initial stop's VOT and the following vowel taken from [bee], [pee], and [phee] in IDS vs. ADS (averaged across two infant age groups) are shown in Table 1.

Table 1: Durational measures (ms) of initial stops and following vowels taken from [bee], [pee], and [phee] in IDS vs. ADS of Thai mothers (averaged across infant age groups). ± indicates SD.

			IDS	ADS
/b/	**voicing lead (ms)**	average	−72.6 (±16)	−74.2 (±16)
		range	162.1	71.5
IDS=43 tokens		min–max	−187– −24.9	−110.1– −38.6
ADS=46 tokens		median	−72.07	−73.24
		Q1	−81.40	−83.54
		Q3	−61.21	−64.80
		IQR	20.2	18.7
	following V (ms)	average	228.2 (±72)	192.2 (±31)
		range	318.7	175.7
		min–max	119.6–438.3	115–290.7
/p/	**voicing lag (ms)**	average	8.7 (±2)	8.2 (±2)
		range	13.8	20.7
IDS=50 tokens		min–max	3.7–17.5	4.9–25.6
ADS=50 tokens		median	8.9	7.2
		Q1	6.7	6.5
		Q3	10.3	9.2
		IQR	3.6	2.6
	following V (ms)	average	213 (±60)	179 (±29)
		range	260.0	147.4
		min–max	120.2–380.3	108.3–255.7
/ph/	**voicing lag (ms)**	average	47.5 (±21)	54.4 (±18)
		range	96.6	73.9
IDS=43 tokens		min–max	15.6–112.2	17.7–91.6
ADS=49 tokens		median	33.5	50.4
		Q1	28.0	36.6
		Q3	65.3	71.6
		IQR	37.2	35.0
	following V (ms)	average	224.1 (±65)	175.66 (±25)
		range	268.9	156.6
		min–max	97.2–366.2	102.9–259.5

Figure 1: Variability of VOT values in /b/, /p/, and /ph/ in IDS vs. ADS of Thai mothers (averaged across infant age groups). The unfilled squares indicate the median. The lines above and below the median show Q3 and Q1. The whiskers cover ±1.5×IQR, and the unfilled triangles and diamonds indicate outliers.



4.1. VOT

Main effect of Type was significant [F(2, 16)=220, p<.01], but not Register [F(1, 8) <1, p=.93] or Infant Age [F(1, 8)<1, p=.46]. Neither Type × Infant Age [F(2,16)<1, p=.44] nor Type × Register [F(2,16)=1.2, p=.33] was significant. But, the three-way interaction Infant Age × Type × Register was significant [F(2,16)=7.58, p <.05]. To further analyze the effect of Type and Register, two Infant Age groups were analyzed separately. For 4–6 mo. group, the main effect of Type was significant [F(2,8)=70.71, p <.05], but not Register [F(1,4) <1, p=.98] or Type × Register interaction [F(1, 8) <1, p=.41]. Pairwise comparison among means showed that all 3 pairs were significantly different from each other (p<.05). For 8–10 mo. group, the main effect of Type was significant [F(2,8)=296.73, p <.05], while Register [F(1,4)=1.03, p=.37] was not. Type × Register was significant [F(2,8)=17.91, p<.01]. In a simple main effect analysis, ADS showed significantly longer VOT than IDS only in /ph/ (p <0.05).

In Table and Figure 1, VOT values differ significantly upon different voicing categories; separated VOT values for /b/, /p/, and /ph/ are clearly shown. Even though VOT differences in IDS vs. ADS are non-significant, slightly wider interquartile range and wider range in distribution are noted for /b/, /p/, and /ph/.

4.2. Comparison of VOT values

In general, our VOT values fall within the ranges that have been previously reported as shown in Table 2. However, in [15], /b/ exhibited somewhat bigger and /ph/ smaller VOT values.

Table 2: VOT values (average (and ±SD or range) in ms.) reported in previous studies.

study	recording method	/b/	/p/	/ph/
this study	spoken word in carrier sentence	−73.4 (±19.5)	8.5(±3.1)	51.2 (±23.6)
Lisker & Abramson (1964) [15]	connected speech; conversation	−35 (−50–−20)	8 (0–15)	37 (25–45)
L-Thongkum et al. (2011) [16]	read word in isolation	−86	14	75
Shimizu (1996) [20]	read word in carrier sentence	−104 (±18.5)	5 (±1.6)	73 (±22.2)

4.3. Vowel duration: /ee/

The vowel duration data was analyzed in the same way. Unlike VOT, however, none of Infant Age [F(1, 8)<1, p=.71], Type [F(2, 16)=1.90, p=.18], Type × Infant Age [F(2,16)=2.03, p=.16], Infant Age × Type × Register [F(2,16)<1, p=.95] was significant and

only the main effect of Register was significant [$F(1,8)=6.22$, $p<.05$]; vowel duration of IDS was significantly longer than that of ADS. The ranges of vowel duration in IDS were larger than ADS (Table 1), consistent with previous studies that found IDS segments to be more variable than those in ADS.

5. DISCUSSIONS

The results of the present study showed that today's Thai mothers produce stops with VOTs that are remarkably similar to those reported over 50 years ago by Lisker and Abramson [15]. The VOTs differed reliably for the three-way stop contrasts. The VOTs, however, did not differ significantly between IDS and ADS. It was not the case that Thai mothers produced stop contrasts in IDS and ADS tokens indistinguishably. Instead, mothers showed modifications in terms of vowel duration. Interestingly, in the course of our analysis, varied phonetic realizations of the stops were observed in IDS and ADS, some of which are discussed here.

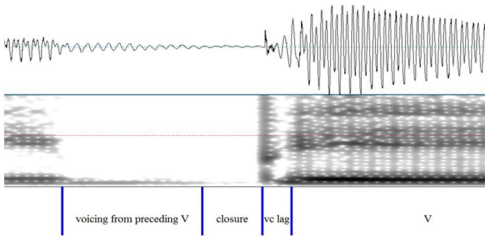
5.1. Phonetic realizations

In agreement with Sundberg and Lacerda (1999) and Foulkes et al. (2005), our IDS data contains instances of phonetic (non-canonical) realizations of [b] and [ph] which appear (based on our observation) in relatively greater degree than in ADS. These instances were excluded from the quantitative analysis as mentioned in 3.2, but are categorized and listed here.

5.2.1. Phonetic realizations of [b]

- Lack of pre-voicing in [b]. Voicing continued after the preceding segment [ii], followed by brief closure and voicing lag (positive VOT). (7% in IDS; 3% in ADS; 5 mothers). This type of realization was quite common in English voiced stops and was referred to as "bleed" in [8].

Figure 2: Example of lack of pre-voicing ("bleed") in [b].

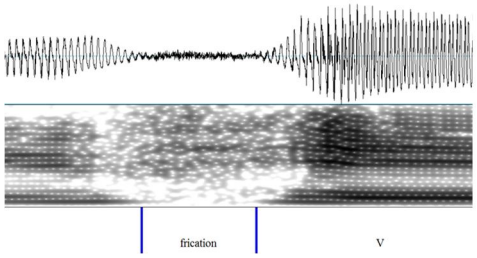


- Relatively clearer and sharper release burst in ADS than IDS in [b]. This pattern has been observed only in some mothers.

5.2.2. Phonetic realizations of [ph]

- Fricated stop in [ph] (4% in IDS; 1% in ADS; 3 mothers). This finding is interesting since it was previously reported that the fricated stop [kxh] was a variant of Thai (kh) (Section 2.2). Although this is still speculative, frication in this case might be triggered by phonetic context or could be an "emerging" variant of (ph).

Figure 3: Example of fricated stop in [ph].



- Incomplete closure in [ph] (2% in IDS; 2 mothers).
- Voicing in [ph]. There was unbroken voicing from preceding vowel [ii] with no closure in [ph] (1% in IDS; 1 mother).

6. CONCLUSION AND FUTURE WORK

Voicing contrasts in terms of VOT values are preserved in Thai young females both in IDS and ADS. In fact, our data showed clearly separated VOT values for /b/, /p/, and /ph/ with some small overlapping between /p/ and /ph/ in the two registers. Crucially, unlike vocalic portions (here long vowels) which were greatly hyper-articulated in IDS, stops' VOTs remained relatively unaffected. Despite stability of VOTs as acoustic cues for voicing contrasts in Thai stops, our data clearly suggested that phonetic (non-canonical) realizations appeared in IDS no less frequently than in ADS.

Areas for future research will include an extended study to look at VOT modifications in /d/ /t/ /th/ /k/, and /kh/. As alveolar and velar stops generally have larger VOT values than the bilabials (i.e., less crowded voicing system), some interesting pattern might emerge. Lack of pre-voicing in [b] found in this data set could be attributable to preceding vowel context and further analysis of [b] in other phonetic contexts should be pursued. Finally, the question of whether the fricated stop in [ph] is simply a phonetic realization or an "emerging" (sociophonetic) variant is also worth investigating.

7. ACKNOWLEDGEMENT

The research was funded in part by JSPS Grant-in-Aid for Scientific Research S (16H06319) and (#4903-17H06382/17H06383) to Reiko Mazuka.

8. REFERENCES

[1] Abramson, A. S., & Whalen, D. H. 2017. Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of phonetics*, 63, 75–86.

[2] Bang, H. Y., Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T. J. 2018. The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics*, 66, 120–144.

[3] Baran, J. A., Laufer, M. Z., & Daniloff, R. 1977. Phonological contrastivity in conversation: A comparative study of voice onset time. *Journal of Phonetics*, 5, 339-350.

[4] Benders, T., Pokharel, S., & Demuth, K. 2018. Hypo--articulation in infant-directed speech: The case of the Nepali four-way stop voicing contrast. *The 43rd Annual Boston University Conference on Language Development.* Boston: USA.

[5] Boersma, P., & Weenink, D. 2011. Praat: Doing Phonetics by Computer [Computer Program] Version. 5.3. 21. 2012.

[6] Chalermsanyakorn, S. 2000. *Variation of (kh) in Bangkok Thai by Phonetic Environments and Speakers' Age (master's thesis)*. Chulalongkorn University, Bangkok.

[7] Cho, T., & Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of phonetics*, 27(2), 207–229.

[8] Davidson, L. 2010. Variation in stop releases in American English spontaneous speech. *J. Acoust. Soc. Am.*, 128(4), 2458–2458.

[9] Davis, B. & Lindblom, B. 2001. Phonetic variability in baby talk and development of vowel categories. In: F. Lacerda, C. von Hofsten & M. Heimann (eds.), *Emerging Cognitive Abilities in Early Infancy*. Mahwah, NJ: Lawrence Erlbaum, 135-171.

[10] Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16(3), 477-501.

[11] Foulkes, P., Docherty, G., & Watt, D. 2005. Phonological variation in child-directed speech. *Language*, 88(1), 177-206.

[12] Kalashnikova, M., Carignan, C., & Burnham, D. 2017. The origins of babytalk: smiling, teaching or social convergence?. *Royal Society open science*, 4(8), 170306.

[13] Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., ... & Lacerda, F. 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277(5326), 684-686.

[14] Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. 2014. Infant-directed speech enhances temporal rhythmic structure in the envelope. In: *Fifteenth Annual Conference of the International Speech Communication Association (Interspeech)*.

[15] Lisker, L., & Abramson, A. S. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3), 384–422.

[16] L-Thongkum, T., Inthajamornrak, C., Sawanakunanon, Y., Sombatnan, N.B., Awirutthiyothin, T., Petcharat, N.,…Jiamjuankhao, T. 2011. *Thai Sounds: An Acoustic Studies.* Bangkok: Chulalongkorn University Press. (in Thai)

[17] Malsheen, B. J. 1980. Two hypotheses for phonetic clarification in the speech of mothers to children. *Child phonology*, 2, 173–184.

[18] Onsuwan, C. 2005. *Temporal Relations between Consonants and Vowels in Thai Syllables* (Doctoral dissertation). University of Michigan.

[19] Ratner, N. B. (1984). Patterns of vowel modification in mother–child speech. *Journal of child language*, 11(3), 557-578.

[20] Shimizu, K. 1996. *A cross-language study of voicing contrasts of stop consonants in Asian languages*. Tokyo: Seibido.

[21] Silva, D. J. 2006. Variation in voice onset time for Korean stops: A case for recent sound change. *Korean Linguistics*, 13(1), 1-16.

[22] Sundberg, U., & Lacerda, F. 1999. Voice onset time in speech to infants and adults. *Phonetica*, 56(3-4), 186-199.

[23] Takada, M. 2012. Regional and generational variation of VOT in Japanese. *The First International Conference on Asian Geolinguistics*, 273-282.

[24] Takada, M., Kong, E., Yoneyama, K., & Beckman, M. E. 2015. Loss of prevoicing in modern Japanese/g, d, b/. *Proc. 18th ICPhS Glasglow.*