Influence of bilingualism or caregiver input? Variation in VOT in simultaneous bilingual preschoolers in Singapore

Jasper H. Sim

National Institute of Education, Nanyang Technological University, Singapore
jasper.sim@nie.edu.sg

ABSTRACT

This exploratory study examines variation in voice onset time (VOT) in 12 English-Mandarin and English-Malay simultaneous bilingual preschoolers born and raised in Singapore. These children were not only different in their second L1, but they and their caregivers were also different in their language dominance, as are typical of speakers in such sociolinguistic contexts. This study seeks to explore effects of bilingualism and caregiver input on VOT development. The findings revealed differences in English VOT between Chinese and Malay children, and between Malay children—all of whom highly English dominant—that reflect the same variation previously found in the child-directed speech of Singaporean caregivers. The results also revealed category assimilation of Malay voiceless unaspirated and English fortis stops in some Malay children that could be attributed to the phonetic overlap in their input. Some unexplained individual variation and differences further highlight the complexity of bilingual phonological acquisition in such contexts.

Keywords: quality of input, language contact, multi-dialectal, English-Malay, English-Mandarin.

1. INTRODUCTION

In many bilingual contexts, the phonetic input that children receive is not homogeneous. Caregivers who are late-L2 speakers may provide phonetic input that is inconsistent, and these caregivers, as well as those experiencing L1 attrition, may also exhibit phonetic details that differ from their monolingual peers in their child-directed speech (CDS) [1]–[3].

Not all differential production is due to acquiring an L2 late. In multicultural and multilingual contexts, between-speaker variation can arise from, *inter alia*, caregivers speaking a different other L1 or heritage language; even those speaking the same languages may differ in their language dominance (e.g. in terms of amount of use of their 2L1s), thereby exhibiting accent differences [4], [5]. Moreover, in communities that have experienced long-term language contact and/or shift, features that had once emerged from effects of individual bilingualism may be transmitted to and retained by later generations of L1 speakers [5]–[7]. In other words, what appear as effects of bilingualism may in fact be learnt through the input.

Children are sensitive to sub-phonemic details in the input, and its fine-grained variation can moderate language outcomes [8], [9]. Stoehr and colleagues [2] examined the production of VOT by Dutch-German preschoolers who acquired the heritage language German from their mothers who spoke L1 German, and the majority language, Dutch, from L1-speaking fathers and L2-speaking mothers. They found that individual variation in the children's VOT was associated with the variation in both their mother's non-native L2 and attrited L1. Sim and Post [9] also found that Singaporean mothers who released English coda stops to a lesser degree also had children who tended to not release their stops, and the same was true for mothers who released their stops to a higher degree. Importantly, these studies show that variable language outcomes in bilinguals may be attributed to input properties and not (solely) due to effects of bilingualism such as cross-linguistic influence (CLI) that involves assimilation of or interactions between the two phonological systems [10]–[13].

This present study explores the variation in the implementation of stop voicing contrast in English-Mandarin and English-Malay simultaneous bilingual preschoolers. Singaporean Mandarin employs a two-way distinction between unaspirated and aspirated /p, t, k/ stops, and Malay employs a two-way distinction between truly voiced /b, d, g/ stops and unaspirated voiceless /p, t, k/ stops. In a recent work [14], it was revealed that while Singaporean Malay and Chinese early bilingual mothers were similar in their use of VOT for stop voicing contrast in their English CDS towards preschoolers (i.e. longer positive VOT for fortis, and shorter VOT/lead VOT for lenis stops), they were dissimilar in where in the VOT continuum the contrasts were made: Malay mothers employed shorter positive VOT but longer lead VOT, whereas Chinese mothers produced longer positive VOT but shorter lead VOT. Mothers who were less English dominant produced even shorter VOT in their fortis stops, thereby exhibiting smaller phonetic contrasts. The findings suggest that Malay children, especially those raised by Malay-dominant caregivers, are exposed to English stops that phonetically overlap with the same stops in Malay to a large degree.

This study builds on this work by exploring the effects of individual bilingualism and caregiver input on VOT development. It aims to examine (i) whether and how English VOT patterns between Singaporean Malay and Chinese children who are equally highly English dominant are different, and (ii) whether and how English and Malay VOT patterns are different between Malay-dominant and English-dominant Malay children.

2. METHODOLOGY

2.1. Participants

Participants (10 boys, 2 girls) were nine Malay and three Chinese Singaporean preschoolers between the ages of 44 and 76 months (M_{age} = 60.8) who were raised by early bilingual caregivers. The children were exposed to local varieties of English and Malay or Mandarin (for the Chinese) by the age of three. The language dominance of their mothers was measured using the Bilingual Language Profile (BLP) [15]; possible scores ranged from –218 (Malay or Mandarin dominant) to +218 (English dominant). The amount of use of English by the children, the proxy for language dominance for this study, was calculated from an accumulated measurement of the estimated amount and proportion of time for which their 2L1s were used with significant people in their immediate ecosystem, in self-interaction and in their exposure to media, as reported by their caregivers.

The three Chinese children (**CHI**) were all highly English dominant ($M_{\%\text{Eng use}}$ = 88, *SD* = 5.8), and were raised by highly English-dominant mothers (M_{BLP} = 119.7, *SD* = 50.1). The Malay children were divided into three groups: six Malay children were highly English dominant ($M_{\%\text{Eng use}}$ = 80.3, *SD* = 7.5), three of whom (**MLY-L.ENG**) were raised by mothers who were *less* English dominant (M_{BLP} = 47.1, *SD* = 2.0), while three others (**MLY-M.ENG**) were raised by mothers who were *more* English dominant (M_{BLP} = 76.0, *SD* = 22.3). The final three Malay children (**MLY-DOM**) were more Malay dominant ($M_{\%\text{Eng use}}$ = 54.7, *SD* = 7.1) and were raised by Malay-dominant mothers (M_{BLP} = -9.25, *SD* = 21.1).

2.2. Materials and recording procedure

Target words with close/high vowels (Table 1) were elicited through a picture card naming task primarily conducted by one of the caregivers (typically the mother). Each word was elicited twice. The Malay target words were only elicited from the Malay children, after a short interaction in Malay between caregiver and child.

The recording took place in a quiet room in the respective homes of the participants. The child's responses were recorded using a NAGRA ARES-MII recorder, through a lapel microphone that was pinned to the collar of the child. Back-up recording was also performed by the author using a Zoom H5 recorder. Both recorded at a sampling rate of 44.1 kHz at 16 bit.

Stop	English	Malay
p	pea (swimming) pool	pipi *cheek* putih *white*
t	tea, 't' two	tiga *three* tujuh *seven*
k	key cooking	kitab *book* kuda *horse*
b	bee (peek-a-)boo	bibir *lips* buka *open*
d	'd' doing	diri *stand* duduk *sit*
g	geese good	gigi *teeth* gula-gula *candies*

Table 1: Stimuli

2.3. Acoustic analysis

VOT boundaries for all word-initial stops were placed manually based on acoustic cues in the waveforms spectrograms on Praat [16]. VOT was defined as the time interval between the release burst as signalled by a sharp peak in waveform energy and the onset of voicing, taken to be the nearest zero-crossing of the onset of periodicity in the waveform. If voicing occurred during the closure, VOT was measured from the onset of glottal pulses in the waveform and/or visual cues in the spectrogram that indicated voicing onset up to the first release burst. 44 tokens could not be reliably measured due to noise and overlapping speech and were excluded. Faster speaking rate is associated with shorter VOT; vowel duration was used in this study as a proxy for local speech rate [17].

2.4. Statistical analysis

Stops with positive VOT (+VOT) were analysed separately from prevoiced stops (–VOT) because of the bimodal distribution. Linear regression analyses were conducted using R. Model selection was based on parsimony. In all models, the random effect structure included random intercepts for subject and word only if they significantly improved model fit. To evaluate the contribution of each predictor, pairwise model comparisons between a full model and a more restricted model that excluded the predictor under consideration were performed using likelihood ratio tests. Continuous predictors were z-standardised. Categorical predictors were weighted effect coded. Outliers in the raw measurements were detected using the interquartile range method. After checking the normality of residuals, influential outliers were removed using Cook's distance, before the final, optimised models were obtained and reported below.

3. RESULTS

The VOT means of English fortis and lenis stops (with +VOT and –VOT) and Malay voiceless (VL) and voiced (V) stops are presented in Table 2.

Stops	Chi	Mly-dom	Mly-L.Eng	Mly-M.Eng
Eng Fortis *SD (n)*	**120.6** 63.3 (41)	**62.1** 34.2 (44)	**48.8** 50.0 (41)	**46.8** 36.5 (46)
Eng Lenis (+) *SD (n)*	**27.1** 24.1 (19)	**35.7** 15.2 (19)	**27.4** 17.0 (21)	**30.4** 15.6 (10)
Eng Lenis (–) *SD (n)*	**–193.5** 97.4 (16)	**–133.5** 41.9 (16)	**–79.2** 64.0 (11)	**–104.5** 41.5 (23)
Mly VL *SD (n)*	—	**38.6** 28.2 (38)	**31.2** 18.8 (29)	**36.3** 24.3 (37)
Mly V *SD (n)*	—	**–102.9** 22.6 (11)	**–71.1** 45.8 (9)	**–80.5** 37.2 (22)

Table 2: VOT means (ms) of Eng. and Mly stops.

3.1. Stops with +VOT

Figure 1 shows the distribution of +VOT of English fortis/lenis and Malay voiceless stops for each child.

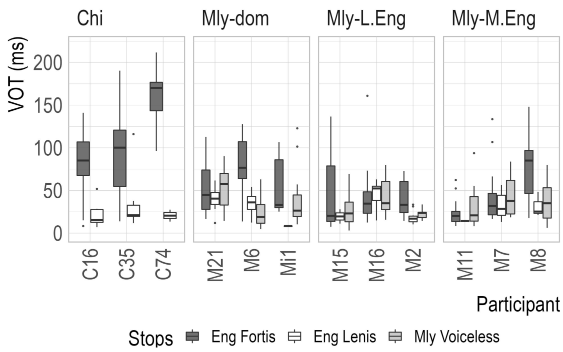

Figure 1: +VOT (ms) of English fortis/lenis and Malay voiceless stops of individual child by group.

3.1.1 Phonemic contrast of English stops with +VOT

The +VOT of English fortis and lenis stops produced by all children were first analysed. The full model included as fixed effects (first in list = reference level for all models henceforth) vowel duration, place of articulation (POA; velar, alveolar, bilabial), group (MLY-M.ENG, CHI, MLY-DOM, MLY-L.ENG), and type (fortis, lenis), and the two-way interaction between group and type. In the reduced model (*obs.* = 223, influential outliers removed = 18, marginal R^2 = .62, conditional R^2 = .75), all were significant predictors. Only for fortis stops, Chinese children produced longer VOT than Malay-dominant Malays, b = 69.64, t(9.05) = 6.32, $p < .001$, and even more so than English-dominant Malays (MLY-L.ENG: b = 87.32, t(9.33) = 7.86, $p < .001$; MLY-M.ENG: b = 84.69, t(8.99) = 7.70, $p < .001$). Marginal differences between the Malay groups were not significant.

In addition, Chinese children exhibited phonemic contrast between fortis and lenis stops by producing significantly longer +VOT for fortis stops than lenis, b = 97.6, t(43.4) = 11.81, $p < .001$. The contrast was smaller in MLY-DOM Malays, b = 27.7, t(37.8) = 3.48, p = .001, marginally in MLY-M.ENG Malays, b = 18.7, t(64.1) = 2.04, p = .04, and was not significant in MLY-L.ENG Malays, b = 11.7, t(36.6) = 1.47, p = .15.

In sum, after controlling for vowel duration and POA, Chinese children produced much longer +VOT for English /p, t, k/ than all Malays, thereby exhibiting greatest phonetic contrast between English fortis and lenis stops. Compared to English-dominant Malays, MLY-DOM Malays produced greater fortis-lenis contrasts; the fortis-lenis contrasts produced by MLY-L.ENG Malays were marginal and not significant.

3.1.2 English and Malay VL stops with +VOT

Malay children's English fortis and lenis stops with +VOT and their Malay /p, t, k/ stops were compared. Malay voiceless stops are normally unaspirated, as are English lenis stops. The full model included vowel duration, POA, group, type (Eng. fortis, Malay voiceless, Eng. lenis) and the two-way interaction between group and type. In the reduced model (*obs.* = 268, influential outliers removed = 17, marginal R^2 = .41, conditional R^2 = .54), the main effects of vowel duration, POA and type were significant predictors. Compared to the weighted grand mean, only English lenis stops had significantly shorter +VOT (β = -0.50, b = -13.26, t = -3.47, $p < .001$). The +VOT of English fortis stops was not significantly longer than Malay voiceless unaspirated stops (b = 3.48, t = 0.70, p = .77).

In sum, after controlling for vowel duration and POA, their Malay voiceless unaspirated stops were more similar to English fortis stops in +VOT than to their English lenis stops. That there was no significant interaction between type and group suggests that there were no significant differences between groups.

3.2. Stops with –VOT

3.2.1 Likelihood of prevoicing

English lenis and Malay voiced stops were more likely short-lag than prevoiced. The Chinese children prevoiced 16/35 (46%) English lenis stops, 16/35 (46%) by MLY-DOM, 11/32 (34%) by MLY-L.ENG, but 23/33 (70%) by MLY-M.ENG. For Malay voiced stops, MLY-DOM prevoiced 11/31 (35%) stops, 9/31 (29%) by MLY-L.ENG, and 22/33 (67%) by MLY-M.ENG. As can be seen in Figure 2 below, some children did not (consistently) produce prevoiced stops.

3.2.2 Length of –VOT

Figure 2 shows the distribution of –VOT of English lenis and Malay voiced stops for each child.

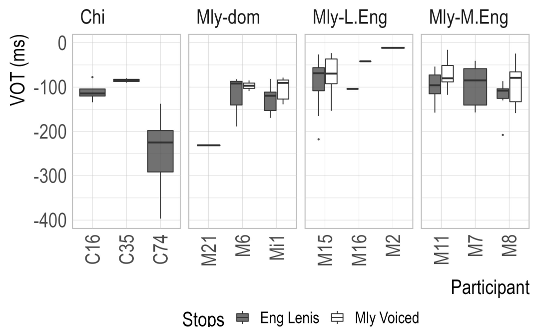

Figure 2: –VOT (ms) of English lenis and Malay voiced stops of individual child by group.

The regression analysis for –VOT length included only data of Malay children who produced at least five prevoiced stops: M6, Mi1, M15, M11 and M8. In the full model that included vowel duration and the two-way interaction between language and group as predictors (*obs.* = 81, R^2 = .18), only vowel duration significantly improved model fit. In other words, there was no significant difference between English lenis and Malay voiced stops in their –VOT for the five children producing prevoiced stops.

4. DISCUSSION

Variation in VOT in 12 English-Mandarin and English-Malay simultaneous bilingual preschoolers raised in Singapore was examined. They were not only different in their second L1, but they and their caregivers were also different in their language dominance, as is typical of speakers in such sociolinguistic contexts. This is an exploratory study; the findings should be interpreted with caution due to the small sample size and number of words per child that were examined.

The analysis on the +VOT of English stops of all children revealed that, for fortis stops only, Chinese children produced longer +VOT than all the Malays, thereby also exhibiting the largest phonetic contrast between English fortis and lenis stops with +VOT. Amongst the Malay children, the fortis-lenis contrast was greater for MLY-DOM children than for English-dominant MLY-M.ENG Malays. The contrast was marginal and not significant for MLY-L.ENG children.

Effects of bilingualism such as CLI and language dominance bear less explanatory power for these findings, as we should expect small differences between CHI, MLY-L.ENG, and MLY-H.ENG children, who all used English highly frequently, on an average of 83% of the time [13]. These differences instead suggest an influence of caregiver input, as they reflect the same variation found in the VOT of Singaporean caregivers [14]: as mentioned, in their CDS, Chinese mothers produced fortis stops with longer +VOT than all Malay mothers (as did CHI children). Compared to their English-dominant peers, Malay-dominant mothers produced shorter +VOT in their fortis stops, thereby exhibiting smaller contrasts (as did MLY-L.ENG children). The results, along with those of previous work, e.g. [2], [9], [18], indicate that an explanatory model of phonological acquisition needs to consider qualitative properties of caregiver input.

That MLY-DOM children produced greater fortis-lenis phonetic contrasts than their English-dominant peers was unexpected, since previous studies showed that Malay-dominant educated Singaporean adults had much shorter +VOT for English /p, t, k/ than their English-dominant peers, due to potential CLI or input of caregiver who were L2 English speakers [5], [14]. Interestingly, the mothers of MLY-DOM children were all Malay language teachers. Whether the greater contrast between English fortis and lenis stops in these children was due to the influence of caregiver input (i.e. their mothers also exhibiting greater English contrasts) or due to richer Malay input (and therefore enhanced phonemic category dissimilation between their 2L1s [19]) remains to be studied.

In the analysis that compared Malay children's Malay voiceless unaspirated stops with their English stops produced with +VOT, it was revealed that, across groups, the VOT of Malay voiceless unaspirated stops was not significantly different from the VOT of English fortis stops. Instead of being a result of CLI, since this was observed across groups, it could be attributed to the considerable phonetic overlap in the input between Malay voiceless unaspirated stops and English fortis stops in the short-lag region that could have led to category assimilation [11], [19]. Note, however, that based on visual inspection of the raw VOT in Figure 1, some children did distinguish these stops by producing longer English fortis stops, although such individual variation was averaged within groups.

Finally, it remains unclear if the lack of consistent prevoicing of Malay stops, if produced at all, is due to developmental reasons [13] or a result of mothers' reduced prevoicing [7]. One question that cannot be answered in this study is why MLY-M.ENG children produced more prevoiced English and Malay stops than others overall, when it is the MLY-DOM children who should presumably have been exposed to more prevoiced stops in their input by their Malay-dominant mothers.

5. REFERENCES

[1] M. S. Fish, A. García-Sierra, N. Ramírez-Esparza, and P. K. Kuhl, "Infant-directed speech in English and Spanish: Assessments of monolingual and bilingual caregiver VOT," *Journal of Phonetics*, vol. 63, pp. 19–34, Jul. 2017, doi: 10.1016/j.wocn.2017.04.003.

[2] A. Stoehr, T. Benders, J. G. van Hell, and P. Fikkert, "Bilingual preschoolers' speech is associated with non-native maternal language input," *Language Learning and Development*, vol. 15, no. 1, pp. 75–100, 2019, doi: 10.1080/15475441.2018.1533473.

[3] G. Khattab, "Acquisition of Lebanese Arabic and Yorkshire English /l/ by bilingual and monolingual children: A comparative spectrographic study," in *Current Issues in Linguistic Theory*, vol. 319, Z. M. Hassan and B. Heselwood, Eds. Amsterdam: John Benjamins Publishing Company, 2011, pp. 325–354. doi: 10.1075/cilt.319.15kha.

[4] M. Amengual, "Type of early bilingualism and its effect on the acoustic realization of allophonic variants: Early sequential and simultaneous bilinguals," *International Journal of Bilingualism*, vol. 23, no. 5, pp. 954–970, Oct. 2019, doi: 10.1177/1367006917741364.

[5] J. H. Sim, "'But you don't sound Malay!': Language dominance and variation in the accents of English-Malay bilinguals in Singapore," *English World-Wide*, vol. 40, no. 1, pp. 79–108, 2019, doi: 10.1075/eww.00023.sim.

[6] S. Kirkham, "Ethnicity and phonetic variation in Sheffield English liquids," *Journal of the International Phonetic Association*, vol. 47, no. 1, pp. 17–35, Apr. 2017, doi: 10.1017/S0025100316000268.

[7] R. Mayr and A. Siddika, "Inter-generational transmission in a minority language setting: Stop consonant production by Bangladeshi heritage children and adults," *International Journal of Bilingualism*, vol. 22, no. 3, pp. 255–284, Jun. 2018, doi: 10.1177/1367006916672590.

[8] A. Cristià, "Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category," *J Acoust Soc Am*, vol. 129, no. 5, pp. 3271–3280, May 2011, doi: 10.1121/1.3562562.

[9] J. H. Sim and B. Post, "Variation in quality of maternal input and development of coda stops in English-speaking children in Singapore," *Journal of Child Language*, pp. 1–26, Aug. 2021, doi: 10.1017/S0305000921000593.

[10] J. E. Flege, C. Schirru, and I. R. A. MacKay, "Interaction between the native and second language phonetic subsystems," *Speech Communication*, vol. 40, no. 4, pp. 467–491, Jun. 2003, doi: 10.1016/S0167-6393(02)00128-0.

[11] J. A. Barlow, P. E. Branson, and I. S. B. Nip, "Phonetic equivalence in the acquisition of /l/ by Spanish–English bilingual children," *Bilingualism: Language and Cognition*, vol. 16, no. 1, pp. 68–85, Jan. 2013, doi: 10.1017/S1366728912000235.

[12] S. Kirkham and K. M. McCarthy, "Acquiring allophonic structure and phonetic detail in a bilingual community: The production of laterals by Sylheti-English bilingual children," *International Journal of Bilingualism*, vol. 25, no. 3, pp. 531–547, Jun. 2021, doi: 10.1177/1367006920947180.

[13] M. Kehoe and K. Kannathasan, "Development of voice onset time in monolingual and bilingual French-speaking children," *Lingua*, vol. 252, p. 102937, Mar. 2021, doi: 10.1016/j.lingua.2020.102937.

[14] J. H. Sim, "Variation in VOT in English child-directed speech of English-Mandarin and English-Malay early bilinguals in Singapore," Under review.

[15] D. Birdsong, L. M. Gertken, and M. Amengual, "Bilingual Language Profile: An easy-to-use instrument to assess bilingualism." COERLL, University of Texas at Austin, 2012. [Online]. Available: https://sites.la.utexas.edu/bilingual/

[16] P. Boersma and D. Weenink, "Praat: doing phonetics by computer." 2022. [Online]. Available: http://www.praat.org/

[17] R. H. Kessinger and S. E. Blumstein, "Effects of speaking rate on voice-onset time in Thai, French, and English," *Journal of Phonetics*, vol. 25, no. 2, pp. 143–168, Apr. 1997, doi: 10.1006/jpho.1996.0039.

[18] M. Ramon-Casas, S. Cortés, A. Benet, C. Lleó, and L. Bosch, "Connecting perception and production in early Catalan–Spanish bilingual children: language dominance and quality of input effects," *J. Child Lang.*, pp. 1–22, Dec. 2021, doi: 10.1017/S0305000921000787.

[19] J. Flege, "Language contact in bilingualism: Phonetic system interactions," *Laboratory Phonology*, vol. 9, Jan. 2007.