Language exposure modulates the role of tone in perception and long-term memory: Evidence from Cantonese native and heritage speakers

RACHEL SOO [a] & PHILIP J. MONAHAN [a, b]

[a] *Centre for French and Linguistics, University of Toronto Scarborough, Canada,* [b] *Department of Linguistics, University of Toronto, Canada**

1 Cantonese heritage speakers

Given their second language (L2) dominance over their first (L1), HERITAGE SPEAKERS (HSs) are unique outliers to the traditional scenario where the language learned at birth is the one speakers continue to be dominant in into adulthood. HSs are typically children of immigrant families who learn their parents' language first, the HERITAGE LANGUAGE (HL), but later acquire greater proficiency in the majority language of the wider community as an L2 (Valdés 2000, Montrul 2011). Though HSs are a growing population in multilingual societies, little is known about their proficiency in the HL. The present paper examines the perception and production of lexical tone in Cantonese HSs.

While early language experience has been shown to give HSs an advantage over adult L2 learners of the HL (Oh et. al. 2003), HSs also underperform relative to NATIVE SPEAKERS (NS) due to category assimilation between the HL and L2 (MacKay et. al. 2001). To date, there is no account for HS proficiency for a linguistic feature of the HL that does not exist in the L2. As such, does the lack of an assimilatory target for the HL feature in the L2 help or hinder HSs in maintaining this particular feature? This is the case for HSs of tone languages, such as Cantonese, who must maintain lexical tone in the HL, when it is not used in a more dominant L2, such as English.

* We would like to thank Diana Archangeli and Stephen Matthews from Hong Kong University and Alexei Kochetov from the University of Toronto for facilities support. This paper benefitted greatly from the comments made by audiences at the 2016 Southern California Undergraduate Linguistics Conference, the 2016 Cornell Undergraduate Linguistics Conference, and the 43rd Annual Meeting of the Berkeley Linguistics Society. We are especially grateful to Janessa Tam, Michelle Chui, Alex Chu and Jessica Yeung for their help in data collection and Connie Ting, Fion Chu and members of the UTSC Linguistics Laboratory for ongoing support. This research was funded by SSHRC Insight Development Grant (#430-2015-00647) awarded to Philip J. Monahan and a University of Toronto Scarborough Student Conference Travel Grant awarded to Rachel Soo.

2 Tone acquisition versus tone perception

Cantonese is a tone language that makes use of six tones on open syllables to differentiate words (Figure 1A). Perceptual sensitivity to tone is acquired early, around 4 months (Yeung et. al. 2013). This is well before HSs become dominant in the L2 (Kondo-Brown 2006). Therefore, we might expect tone perception to be robust even after HSs lose dominance in their L1. Conversely, tone is the least robust cue in spoken word recognition compared with consonants and vowels even for NSs (Tsang & Hoosain 1979, Taft & Chen 1992, Cutler & Chen, 1997, Weiner & Turnbull 2015). This conflict between the age at which tone is acquired and its relative salience in word recognition poses an interesting problem for speech perception and sheds light on how much more there is to learn about tone language HSs and HS phonology more generally. To the best of our knowledge, the current paper is the first to address this niche with quantitative data by investigating tone sensitivity during spoken-word recognition in Cantonese HSs. Since the more dominant L2 does not readily utilize tone, we hypothesize that Cantonese HSs may use tone less reliably in word recognition relative to NSs for whom tone is already a weak cue.

3 Tone production and AX discrimination

A tone production task and AX discrimination task were conducted to collect primary data on how Cantonese HSs differ from Cantonese NSs.

3.1 Participants

21 HS (17 female, mean age = 22.4) and 21 NS (12 female, mean age = 23.3)[1] were recruited from the University of Toronto and Hong Kong University, respectively. All participants completed a *Bilingual Language Profile* (BLP) questionnaire, which provides a relative weighting of language dominance (Birdsong et. al. 2012)[2].

[1] All participants conformed to specific eligibility requirements, reported normal speech, hearing, and vision, were compensated for participation and provided informed written consent prior to taking part in the experiment.

[2] A negative score indicates Cantonese dominance, while a positive score indicates English dominance. All NSs scored negatively (mean = -97.91) and all HSs scored positively (mean = 69.31).

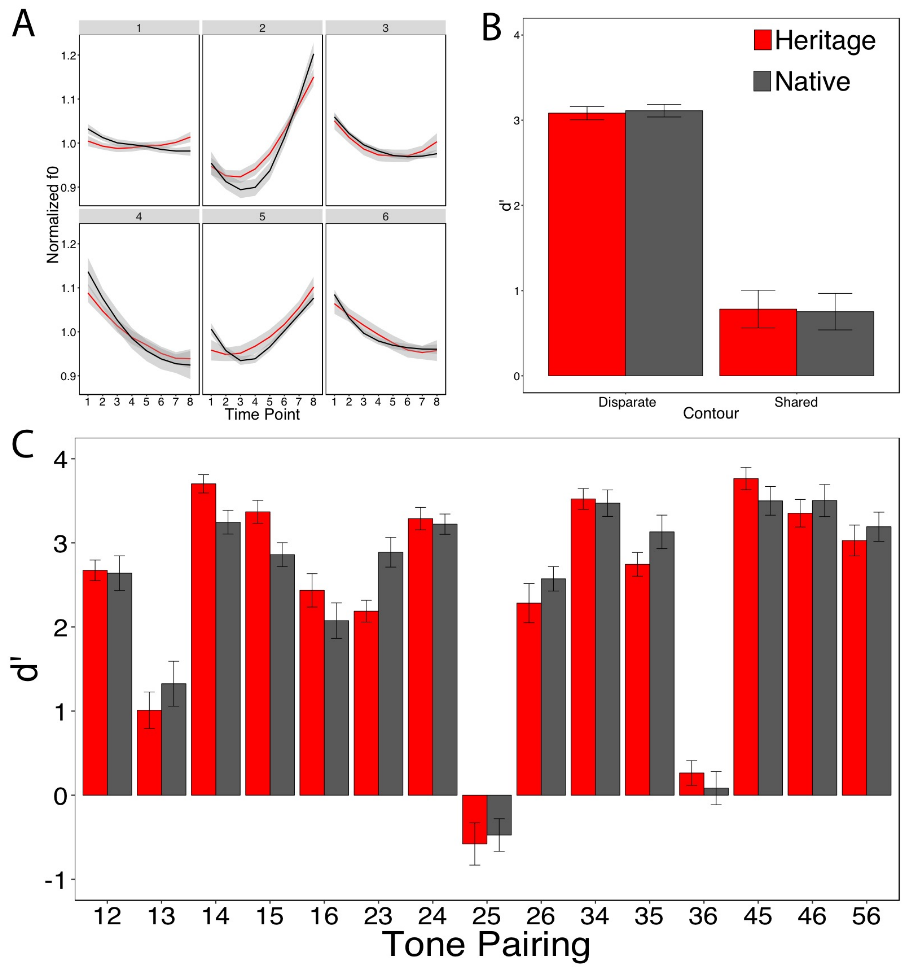
3.2 Production methods and results

Participants were recorded producing three repetitions of six Cantonese tone minimal pairs in /ɑ i o/ vowel conditions[3]. All instructions were provided in Cantonese for NSs. HSs were instructed in English and provided with a Romanization and English translation alongside the Chinese character during elicitation. Recordings were made with an *Audio Technica 2050* multi-pattern condenser microphone on a *Zoom H4N* digital voice recorder with a sampling rate of 44.1 kHz and 16-bit depth.

Individual word tokens were segmented in *Praat* (Boersma & Weenink 2015). Pitch (Hz) was extracted at 8 equally spaced points throughout the duration of the word and then normalized[4]. Figure 1A presents the pitch contours for all six Cantonese tones. Overall, HSs produced tones similar to NSs. The mean contours for each group overlap and match the flat, rising, and falling slopes as appropriate for the respective tones.

[3] 3 vowel conditions × 6 tones × 3 repetitions = 54 tokens per speaker.
[4] Each token's average pitch across the contour was subtracted from each of the eight individual points to eliminate inter-speaker variation (Khouw & Ciocca 2007).

Figure 1: Panel A: Heritage speaker (red) and native speaker (black) tone contours in the production experiment. Panel B: d' scores categorized by tone contours (shared or disparate) in the AX discrimination task. Panel C: d' scores separated by tone pairings in the AX discrimination task. All error bars and shading reflect standard error of the mean.



3.3 AX discrimination methods and results

On each trial in the AX discrimination task, participants listened to a pair of the Cantonese syllable /ji/ with either the same or different tone produced by a male and female NS. The inter-stimulus interval (ISI) was 500 ms. Participants made same-different judgements via button press. The inter-trial interval (ITI) was 1000 ms.

D-prime (d′) scores (Macmillan & Creelman 2005) were calculated for each of the 15 possible tone pairings. By-participant reaction times greater than or less than 2.5 standard deviations of the individual subject means were removed (3.1% of the total data). As with the production experiment, HSs perform similar to NSs (Figure 1C). More specifically, the two groups discriminate tone pairings with disparate contours well (e.g., flat-falling pair 1-4) and those with shared contours poorly (e.g., rising-rising pair 2-5; Figure 1B). There was a main effect of Contour ($F(1,40) = 551.53$, $p < 0.001$) but no effect of Group or a Contour × Group interaction ($F < 1$).

While it is somewhat surprising that HSs perform so native-like in a perception task, we suggest that the similarities between the two groups can be explained by the demands of the task itself. Previous studies have shown that at short ISIs, speakers only need to retain phonetic information in short-term memory and tap into phonetic-perceptual level processing (Werker & Logan 1985). Therefore, the success of HSs in approximating NS performance can be attributed to the fact that both groups have access to a phonetic-perceptual level of processing for tones in an AX discrimination task with short ISIs. If it is true that HSs do not robustly encode tone on a more abstract-phonological level, we expect them to perform worse than NSs in tasks that require long-term memory or phonological levels of processing.

4 MDRP

To test this hypothesis, we performed a MEDIUM-DISTANCE REPETITION PRIMING (MDRP) experiment, which can reveal differences between highly-proficient groups of bilinguals (Pallier et al. 2001).

4.1 Methods

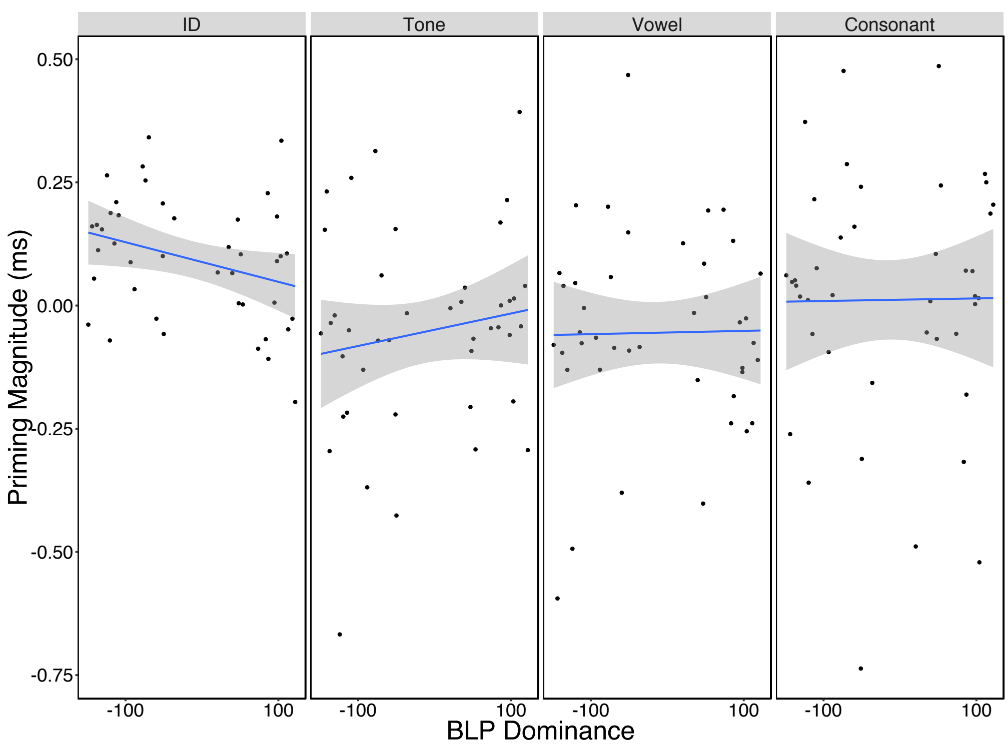
The same 21 HSs and 21 NSs listened to monosyllabic Cantonese words (n = 98) and non-words (n = 98) in isolation and made a lexical decision response on each trial. Items were either repeated (identity pairs: *be2-be2*) or followed by their corresponding minimal pair (minimal pairs: *be2-be5*) differing in a single Cue (Consonant, Vowel, Tone) 8 to 20 trials later. Words were matched for token frequency by condition using *PyCantonese* (Lee 2015).

Since stimuli pairs are separated across a medium distance (8-20 trials), listeners must rely on long-term memory and tap into abstract levels of representation (Pallier et al. 2001, Sumner & Samuel 2009). In this regard, if HSs do not have robust tonal representations at a more abstract level of representation, they will not experience identity priming across medium distances and will not appropriately distinguish minimal pairs.

4.2 Analysis and results

Figure 2 provides priming magnitudes against BLP dominance scores. Priming magnitude was calculated for each pair by subtracting the reaction time (RT) to the second item (*be2-**be5***) from the RT to the first item (***be2**-be5*). A positive priming magnitude indicates that the second item is primed by the first.

Figure 2: BLP dominance scores plotted against priming magnitude in identity priming conditions (collapsed across Cue) and minimal pair priming conditions (separated by Cue).



There is a negative correlation between language dominance and priming magnitude for the identity pairs ($t(40) = -2.07$, $p < 0.05$, $r = -0.31$); namely the more English dominant the speaker, the less priming they show across medium distances. This lack of priming suggests that speakers for whom English is the more dominant language are unable to hold tonal information in memory long enough to be primed by the time that the second identity pair is presented. Similarly, there is a non-significant but positive trend in the tone minimal pair case ($r = 0.16$): the more English dominant the speaker, the more priming there is across medium distances, indicating that tonal representations are weak and easily confusable. Since minimal pair priming does not extend across medium distances, participants are likely mistaking the second minimal pair for being a repetition of the first, as if it were identity priming. These findings point toward the lack of strong tonal representations for English dominant speakers at an abstract-phonological level.

5 Conclusions

The goals of the present study were two-fold: a) to provide primary data on Cantonese HS production and perception of tone, and b) to investigate tone sensitivity during spoken-word recognition. We compared Cantonese HSs with Cantonese NSs in a production and AX experiment and showed that HSs perform similarly to NSs. We argued that this is due to the nature of the AX discrimination task, which does not rely on long-term phonological representations. Using a task that taps into such representations (i.e., MDRP), we found data in favour of a fossilized tone-system due to minimal exposure to the tone L1 (Cantonese), as the tone-less L2 (English) becomes more dominant. Ultimately, it appears that differing degrees of language exposure affect the levels at which tone representations can be accessed and stored during perception and spoken-word recognition. HSs who have less exposure are consequently only able to access linguistic cues at a phonetic-perceptual level and not at an abstract-phonological level.

REFERENCES

BIRDSONG, DAVID; LIBBY GERTKEN; and MARK AMENGUAL. 2012. Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. *COERLL*, University of Texas at Austin. Online: https://sites.la.utexas.edu/bilingual/.

BOERSMA, PAUL, and DAVID WEENINK. 2015. Praat: Doing phonetics by computer, version 5.4.19.Online: http://www/praat/org/.

CUTLER, ANNE, and HSUAN-CHIH CHEN. 1997. Lexical tone in Cantonese spoken-word processing. *Perception and Psychophysics* 59. 165-179.

KHOUW, EDWARD, and VALTER CIOCCA. 2007. Perceptual correlates of Cantonese tones. *Journal of Phonetics* 35. 104–117.

KONDO-BROWN, KIMI. 2006. *Heritage language development: Focus on East Asian immigrants*. Vol. 32. Amsterdam, the Netherlands: John Benjamins Publishing.

LEE, JACKSON. 2015. PyCantonese: Cantonese linguistic research in the age of big data. Talk presented at the Childhood Bilingualism Research Centre, Hong Kong.

MACKAY, IAN R. A.; JAMES EMIL FLEGE; THORSTEN PISKE; and CARLO SCHIRRU. 2001.Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America* 110. 516-528.

MACMILLAN, NEIL A.; and C. DOUGLAS CREELMAN. (2005). *Detection Theory: A User's Guide*. Mahwah, NJ: Lawrence Erlbaum Assoc.

MONTRUL, SILVINA. 2011. Introduction. *Studies in Second Language Acquisition* 33. 155–161.

OH, JANET S.; SUN-AH JUN; LEAH M. KNIGHTLY; and TERRY KIT-FONG AU. 2003. Holding on to childhood language memory. *Cognition* 86. B53-B64.

PALLIER, CHRISTOPHE; ANGELS COLOMÉ; and NÚRIA SEBASTIÁN-GALLÉS. 2001. The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science* 12. 445-449.

SUMNER, MEGAN, and ARTHUR G. SAMUEL. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60. 487–501.

TAFT, MARCUS, and HSUAN-CHIH CHEN. 1992. Judging homophony in Chinese: The influence of tones. In Hsuan-Chih and Ovid J. L. Tzeng (eds.), *Language processing in Chinese*, 151-172. Amsterdam, the Netherlands: Elsevier.

TSANG, KEUNG, and RUMJAHN HOOSAIN. 1979. Segmental phonemes and tonal phonemes in comprehension of Cantonese. *Psychologia* 22. 222-224.

VALDÉS, GUADALUPE. 2000. Teaching heritage languages: An introduction for Slavic-teaching professionals. In Olga Kagan and Benjamin Rifkin (eds.), *The learning and teaching of slavic languages and cultures: Toward the 21st century*, 375-403. Bloomington, IN: Slavica.

WEINER, SETH J., and RORY TURNBULL. 2016. Constraints of tones, vowels and consonants on lexical selection in Mandarin Chinese. *Language and Speech* 59, 59-82.

WERKER, JANET F., and JOHN S. LOGAN. 1985. Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37. 35-44.

YEUNG, H. HENNY; KE HENG CHEN; and JANET F. WERKER. 2013. When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language* 68. 123-139.