CONTACT-INDUCED TONOGENESIS IN HONG KONG ENGLISH

Ping Hei Yeung

Georgetown University
py101@georgetown.edu

ABSTRACT

Tonogenesis has been commonly described as a language-internal phonetic process. Yet, recent literature on contact languages in Africa suggests that it may also occur when languages with different word-prosodic systems come into contact. To test whether language contact is indeed a possible trigger of tonogenesis, this study compares the pitch production of monosyllabic content and function words by 39 speakers of Hong Kong English (HKE), a contact variety with tonal Cantonese substrate, with 30 speakers of American English (AE). Dynamic F0 was measured using Praat and statistically evaluated with Generalized Additive Mixed Modelling (GAMM). HKE speakers were found to produce content words with a significantly higher F0 than function words, while AE speakers did not show any significant difference in F0. The results indicate that HKE is a tone language with two contrastive lexical tones. Thus, apart from language-internal phonetic factors, language contact may also motivate tonogenesis.

Keywords: Tone, language contact, tonogenesis, sound change, English

1. INTRODUCTION

Tonogenesis, the phonologization of lexical tones in a previously non-tonal language, is often described as a language-internal process motivated by laryngeal articulation of consonants, e.g. Kammu [1]–[3], vowel height, e.g. U [2], [4], and F0 correlates of stress, e.g. Swedish [5], [6]. Recently, studies on contact languages in Sub-Saharan Africa and the Caribbean suggest that tonogenesis may also occur as a result of contact between languages with different word-prosodic systems, especially when the substrates are tonal Niger-Congo languages and the superstrates are non-tonal Indo-European languages. Such examples include Central African French [7], Equatorial Guinean Spanish [8], Ghanaian English [9], Nigerian English [9], Pichi [10], and Saramaccan [11]. Even though their superstrates are non-tonal Indo-European languages, they have all become tonal with a {H, L} two-tone system, which can be illustrated by the minimal pair *wood* (H tone) vs. *would* (L tone) in Nigerian English [8], [9].

The reinterpretation of F0 correlates of non-tonal L2 stress systems by L1 tone language speakers has been proposed to be the source of tonogenesis in these languages. Since stressed syllables in many non-tonal languages like English receive a H* pitch accent, tone language speakers would treat the pitch accent as lexical and assign the stressed syllables a H tone. Conversely, since unstressed syllables do not carry any pitch accent, they would tend to have a lower pitch than stressed syllables, and therefore tone language speakers would assign unstressed syllables a L tone [7], [8].

Thus, tonal minimal pairs in these contact languages often originate from distinctions between content words and function words. In Pichi, an English-based creole, the minimal pair *bay* 'by' (L tone) vs. *báy* 'buy' (H tone) differ not only in their tone but also their syntactic category: the function word *bay* receives a L tone while the content word *báy* receives a H tone [10]. Function words in English often have two variants: a weak form and a strong form. The unstressed weak form does not receive any pitch accents while the strong form does. Since the unstressed variants of function words have a much higher frequency than the stressed variants [12], the former would be more salient to L2 speakers, favoring them to analyze those words as having a L tone.

A similar observation has been made for Hong Kong English (HKE), a contact variety of English with tonal Cantonese substrate. Due to the presence of lexical pitch contours in polysyllabic content words, Yiu [13] and Wee [14] classified HKE as a tone language with a {H, Ø} two-tone system. However, it remains unclear whether the tones in HKE are contrastive as previous studies did not provide any examples of minimal pairs or compare pitch production with non-tonal varieties of English. Thus, the objective of this study is to examine whether tonogenesis has occurred in HKE by comparing pitch production of monosyllabic content and function words between speakers of HKE and American English (AE). Given the evidence from other contact languages, it is hypothesized that speakers of HKE would make a tonal distinction between content and function words by using a higher pitch for the former and a lower pitch for the latter, while speakers of AE would not.

2. METHODOLOGY

2.1. Participants

39 speakers of HKE aged between 18-58 (22 females, 17 males) and 30 speakers of AE aged between 18-30 (15 females, 15 males) were recruited. Participants in the HKE group were all born and raised in Hong Kong, and they were fluent bilinguals in both English and Cantonese. Participants in the AE group on the other hand were all born and raised in the United States, and they spoke AE as their native language. Participants reported no history of speech or hearing impairments.

2.2. Materials and procedure

The HKE data were drawn from a larger sociolinguistic study that includes sociolinguistic interviews, a word list reading task, and a minimal pair reading task. The AE group performed the same two reading tasks but they did not take part in the sociolinguistic interviews. For the purpose of this study, only the speech data from the word list reading task were analyzed. In this task, both groups were instructed to read aloud a list of 155 words, among which 9 pairs of segmentally identical content words and function words were the targets, which are shown in Table 1. By using content word/function word homophones, the segmental effects on F0 were controlled. These words were chosen also because they were reported to be tonal minimal pairs in tonal contact varieties of English [9].

Table 1: Syntactic category of the 18 target words. Words on the same row are segmentally identical in both varieties.

Function word	Content word homophone
be	bee
but	butt
or	ore
an	Ann
their	there
by	buy
for	four
to	two
would	wood

Data collection took place in quiet public spaces or the participants' home in Hong Kong for the HKE group and a sound-attenuated booth at Georgetown University for the AE group. Words were presented one at a time in a pseudo-random order on a laptop computer and they were embedded in the carrier phrase “say ___ again” to keep the phonological environment constant. The target words were placed at the middle of the carrier phrase to avoid the effect of boundary tones. The target words were primed to be the focus of the phrase such that they receive a H* pitch accent in AE regardless of their syntactic category [15], [16]. Participants were instructed to read aloud each word along with the carrier phrase three times, resulting in a total of 54 target tokens per speaker. Audio was recorded at a 44.1 kHz sampling rate and 16-bit sample depth using condenser microphones.

2.3. Data analysis

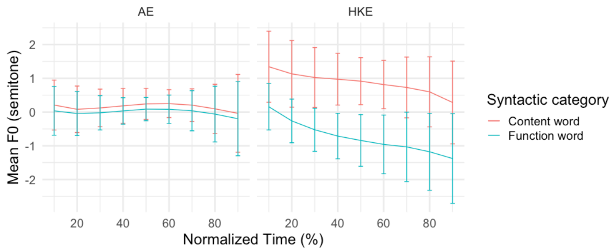
The audio recordings were segmented using Montreal Forced Aligner [17], which were then manually reviewed and corrected. Dynamic F0 was measured using Praat [18] at every 10% interval between 10% and 90% of the sonorant portion of the target words. The tokens were then coded for their syntactic category, adjacent segments, duration, number of repetitions, and English variety. Duration was defined as the duration of the sonorant portion of the word, while adjacent segment was defined as the onset and coda consonants of the word. After that, the F0 values of each token were converted into semitones and z-score normalized by speaker to account for human pitch perception and individual differences in voice pitch.

The normalized F0 was analyzed using Generalized Additive Mixed Modelling (GAMM) [19]. GAMM was fit using the *mgcv::bam()* function [19], [20] and evaluated using functions from the *itsadug* and *gratia* packages [21], [22]. The dependent variable was the normalized F0. The independent variable was the interaction of English variety and syntactic category. The reference level of this interaction term was content word by AE speakers. The interaction was included as a parametric effect and as a smooth in the model. The smooth included time, which is the normalized time point of the measurements. A smoothing term models the non-linear F0 values over time. The parametric effects of duration and repetition, as well as their fixed interactions with time were included to model their influence on F0. A difference smooth accounted for the non-linear effect of adjacent segments. Speaker was included as a random intercept and slope interacting with syntactic category to model the variable effect of syntactic category on individual speakers. Word was included as a random intercept and slope interacting with English variety to model variety-specific differences in the articulation of individual words. To account for the relationship between measurements taken at consecutive time points, an autoregressive error term was included. An AR1 model was incorporated to account for autocorrelation of residuals. The model was calculated using a scaled-t distribution to correct for non-normality of the model residuals as indicated by *mgcv::gam.check()*.

3. RESULTS

The F0 values of content and function words in the two varieties relative to the mean F0 before z-score normalization are given in Figure 1. At the 10%, 50%, and 90% intervals, AE speakers had a mean F0 of 0.208 (SD±0.74), 0.242 (SD±0.461), and -0.038 (SD±1.154) semitones for content words and 0.036 (SD±0.723), 0.087 (SD±0.346), and -0.198 (SD±1.099) semitones for function words. HKE speakers on the other hand had a mean F0 of 1.344 (SD±1.053), 0.917 (SD±0.697), and 0.283 (SD±1.227) semitones for content words and 0.157 (SD±0.687), -0.842 (SD±0.766), and -1.378 (SD±1.329) semitones for function words. Content words had a higher F0 than function words by 0.172, 0.155, and 0.16 semitones in AE and 1.187, 1.759, and 1.661 semitones in HKE.

Figure 1: Mean F0 (semitone) of content and function words by English variety.



In regard to the results of GAMMs, Figure 2 provides the predicted F0 trajectory by each syntactic category for speakers of AE and HKE, while Figure 3 demonstrates the significance of the F0 difference between the two syntactic categories for each variety. HKE speakers produced the content words with a significantly higher F0 throughout the entire sonorant duration than the function words. As shown in Figure 2, the 95% confidence interval of the content words did not overlap with that of the function words. Likewise, Figure 3 indicates a significant difference in F0 throughout the whole sonorant duration. Then, as shown in Figure 3, the degree of F0 difference between content and function words changed throughout the sonorant duration of the target words. The difference increased from the beginning to around the 60% point of sonorant duration, and then the difference decreased from that point onwards. The shape of the F0 trajectory also differed between the two syntactic categories. While a falling F0 contour was observed across both categories, function words were characterized by a sharper drop in F0 than content words.

Figure 2: Fitted GAMM smooths with 95% confidence interval for F0 (z-score) of content and function words by English variety.

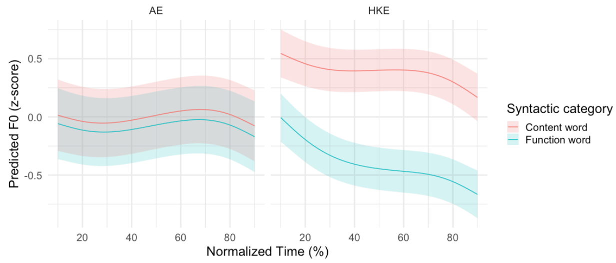
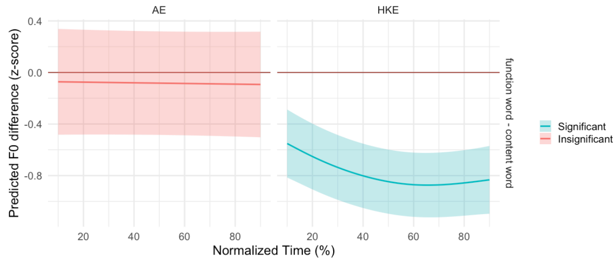


Figure 3: Difference plot for function and content words by English variety. Blue line indicates significant difference while red line indicates insignificant difference.



Conversely, the F0 of the AE group stayed relatively constant throughout the sonorant duration regardless of syntactic category. As shown in Figure 3, although the function words had a slightly lower F0 than the content words, the difference was not significant. It is also shown in Figure 2 by the complete overlap of the 95% confidence intervals of the predicted F0 contours of content and function words. Therefore, there was no significant effect of syntactic category on the dynamic F0 of AE speakers. Furthermore, the confidence intervals for each syntactic category in the AE group were wider than those of the HKE group, which indicates that AE speakers produced the target words with greater variability in F0. Thus, speakers of HKE and AE differed in their F0 production of monosyllabic content and function words. Throughout the sonorant duration, HKE speakers produced the content words with a significantly higher F0 than function words, while no effect of syntactic category was found for AE speakers.

4. DISCUSSION

The results of the current study support the hypothesis that HKE is a tone language with two lexical tones. Unlike speakers of AE who maintain the same pitch contour regardless of syntactic category, speakers of HKE produced the content words with a significantly higher pitch than the function word homophones. This difference is apparent in the production of the two phrases "*Say four again*" and "*Say for again*", which are shown in Figures 4 and 5. The HKE speaker in Figure 4 produced two distinct pitch contours: a higher pitch for "*four*" and a lower pitch for "*for*". The target words in the two phrases are specified for their lexical tone, with the content word *four* being assigned a higher tone and the function word *for* a lower tone. On the other hand, the AE speaker in Figure 5 produced the two sentences with the same pitch contour, corroborating existing literature on English prosody [15], [16]. Since the target words were the focus of the carrier phrase, they became prosodically prominent regardless of their syntactic category. The function word *for* appeared in its stressed form, making it both segmentally and suprasegmentally indistinguishable from its content word homophone *four*.

Figure 4: F0 contour of "*Say four again*" (upper) and "*Say for again*" (lower) by a speaker of HKE.

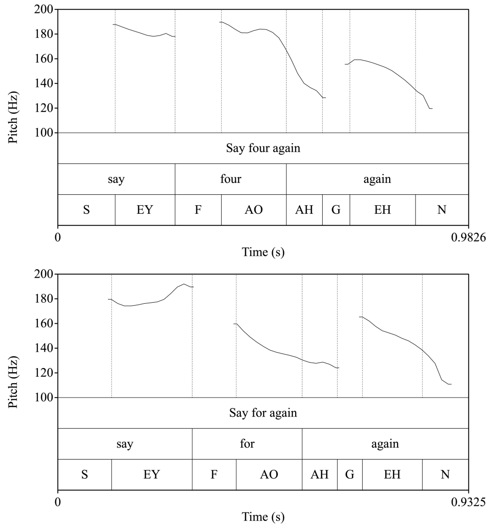
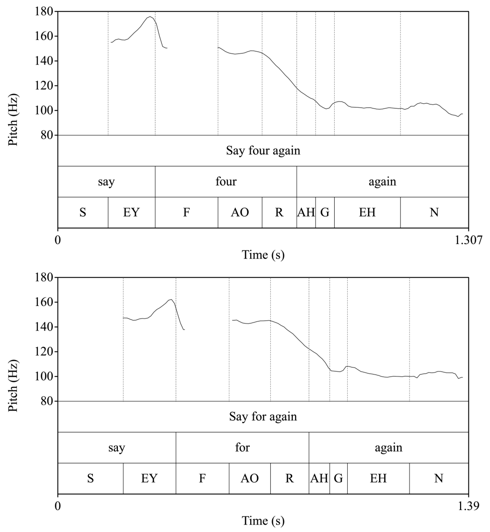


Figure 5: F0 contour of "*Say four again*" (upper) and "*Say for again*" (lower) by a speaker of AE.



However, the HKE tones observed in the current study differ in their trajectory from those reported in Yiu [13] and Wee [14]. Instead of level tones with a stable pitch throughout the duration, the ones found in monosyllabic content words and function words involve a falling pitch contour, which is shown in both the predicted F0 trajectory by GAMM in Figure 2 and the raw F0 trajectory in Figures 1 and 4. Such a difference might not be phonological but instead a result of tonal coarticulation in the carrier phrase. Following Wee [14], the preceding word in the carrier phrase *say* would have a H surface tone and the following word *again* would have a M-H surface tone. When the tone in the target word differs from those of the adjacent tone-bearing units, there would be transitions in pitch target that create pitch contours. More in-depth analysis is required to determine whether the tones assigned to the monosyllabic content and function words are phonologically the same as the ones found in previous studies.

HKE and contact languages in Sub-Saharan Africa and the Caribbean share the same tonal inventory and the same strategy in tone assignment: they all have two lexical tones differing in their pitch levels, and they all assign a higher tone to content words and a lower tone to function words [9]. This could be caused by similarities in both the substrates' tonal inventories and the superstrate's intonation system. The various Niger-Congo languages spoken in Sub-Saharan Africa and Cantonese spoken in Hong Kong all have level tones distinguished by their pitch level relative to each other [23], [24]. On top of that, the superstrates like English, French, and Spanish all assign pitch accents to specific syllables of a word, and their function words are all generally unstressed by default [7], [15], [16], [25], [26]. Both factors might have facilitated the formation of the {H, L} two-tone systems found in these tonal contact languages.

5. CONCLUSION

With the production of tonal minimal pairs in HKE as evidence, this study demonstrates that language contact between tonal substrates and non-tonal superstrates is indeed a possible trigger of tonogenesis. Although some North Germanic languages like Swedish also underwent tonogenesis that stemmed from the reinterpretation of F0 correlates of stress [5], [6], the case of HKE is unique as it is a contact-induced change rather than a language-internal change. Given that tonogenesis has occurred repeatedly in various contact languages from HKE to Central African French, the notion of lexical tone being a marked phonological feature that readily disappears during language contact [27], [28] is challenged.

6. REFERENCES

[1] S. Premsrirat, 'Khmu dialects: a case of register complex and tonogenesis', in *Cross-linguistic studies of tonal phenomena: historical*

development, phonetics of tone, and descriptive studies, Tokyo, 2003.
[2] J. O. Svantesson, 'Tonogenetic mechanisms in northern Mon-Khmer', *Phonetica*, vol. 46, no. 1–3, pp. 60–79, 1989, doi: 10.1159/000261829.
[3] J. O. Svantesson and D. House, 'Tone production, tone perception and Kammu tonogenesis', *Phonology*, vol. 23, no. 2, pp. 309–333, Aug. 2006, doi: 10.1017/S0952675706000923.
[4] J. Kingston, 'Tonogenesis', in *The Blackwell Companion to Phonology*, M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice, Eds., John Wiley & Sons, Ltd, 2011, pp. 1–30. doi: 10.1002/9781444335262.wbctp0097.
[5] E. Gårding, *The Scandinavian word accents*. Lund, Sweden: Gleerup, 1977.
[6] T. Riad, 'The Origin of Scandinavian Tone Accents', *Diachronica*, vol. 15, no. 1, pp. 63–98, Jan. 1998, doi: 10.1075/dia.15.1.04ria.
[7] B. Guri, 'Prosodie et contact de langues: le cas du système tonal du français centrafricain', PhD Thesis, Université de Nanterre - Paris X, 2012. [Online]. Available: https://tel.archives-ouvertes.fr/tel-00789349
[8] G. B. Steien and K. Yakpo, 'Romancing with tone: On the outcomes of prosodic contact', *Language*, vol. 96, no. 1, pp. 1–41, 2020, doi: 10.1353/lan.2020.0000.
[9] C. Gussenhoven, 'On the intonation of tonal varieties of English', in *The Oxford Handbook of World Englishes*, M. Filppula, J. Klemola, and D. Sharma, Eds., Oxford University Press, 2014.
[10] K. Yakpo, *A grammar of Pichi*. Language Science Press, 2018. [Online]. Available: https://langsci-press.org/catalog/view/85/91/1370-1
[11] J. Good, 'Tone and accent in Saramaccan: charting a deep split in the phonology of a language', *Lingua*, vol. 114, no. 5, pp. 575–619, May 2004, doi: 10.1016/S0024-3841(03)00062-7.
[12] A. Bell, D. Jurafsky, E. Fosler-Lussier, C. Girand, M. Gregory, and D. Gildea, 'Effects of disfluencies, predictability, and utterance position on word form variation in English conversation', *J. Acoust. Soc. Am.*, vol. 113, no. 2, p. 1001, Jan. 2003, doi: 10.1121/1.1534836.
[13] S. Yiu, 'Aspects of tone in Cantonese English', The University of Hong Kong, Hong Kong, 2014. doi: 10.5353/th_b5481868.
[14] L.-H. Wee, 'Tone assignment in Hong Kong English', *Language*, vol. 92, no. 2, pp. e67–e87, 2016, doi: 10.1353/lan.2016.0039.
[15] D. R. Ladd, *Intonational phonology*. Cambridge University Press, 2008.
[16] E. Selkirk, 'The prosodic structure of function words', in *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc, 1996, pp. 187–213.
[17] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, 'Montreal Forced Aligner'. 2017. [Online]. Available: http://montrealcorpustools.github.io/Montreal-Forced-Aligner/
[18] P. Boersma and D. Weenink, 'Praat: doing phonetics by computer [Computer program]'. 2022. [Online]. Available: http://www.praat.org/
[19] S. N. Wood, *Generalized Additive Models: An Introduction with R*, 2nd ed. New York: Chapman and Hall, 2017. doi: 10.1201/9781315370279.
[20] S. N. Wood, 'mgcv: Mixed GAM computation vehicle with automatic smoothness estimation'. 2022. [Online]. Available: https://CRAN.R-project.org/package=mgcv
[21] J. van Rij, M. Wieling, R. H. Baayen, and H. van Rijn, 'itsadug: Interpreting time series and autocorrelated data using GAMMs'. 2022. [Online]. Available: https://CRAN.Rproject.org/package=itsadug
[22] G. L. Simpson, 'gratia: Graceful ggplot-based graphics and other functions for GAMs fitted using mgcv'. 2022. [Online]. Available: https://gavinsimpson. github.io/gratia/
[23] S. Matthews and V. Yip, *Cantonese: A Comprehensive Grammar*. Routledge, 2013.
[24] D. Odden, 'Tone: African languages', in *The handbook of phonological theory*, Cambridge: Blackwell Publishers Ltd., 1995, pp. 444–475.
[25] J. I. Hualde and P. Prieto, 'Intonational variation in Spanish: European and American varieties', in *Intonation in Romance*, S. Frota and P. Prieto, Eds., Oxford University Press, 2015, pp. 350–391. doi: 10.1093/acprof:oso/9780199685332.003.0010.
[26] I. Roca, 'Saturation of Parameter Settings in Spanish Stress', *Phonology*, vol. 22, no. 3, pp. 345–394, 2005.
[27] P. Trudgill, 'Contact and Sociolinguistic Typology', in *The Handbook of Language Contact*, R. Hickey, Ed., John Wiley & Sons, Ltd, 2010, pp. 299–319.
[28] J. H. McWhorter, 'Identifying the Creole Prototype: Vindicating a Typological Class', *Language*, vol. 74, no. 4, pp. 788–818, 1998, doi: 10.2307/417003.