Consonantal Effects of Aspiration on Onset F0 in Cantonese

Xinran Ren[1], *Peggy Mok*[2]

[12]The Chinese University of Hong Kong

renxinr@gmail.com, peggymok@cuhk.edu.hk

Abstract

Consonantal effects on onset f0 are implemented differently in different languages, namely, the duration of consonantal effect, the direction of f0 change and its perceptual importance vary cross-linguistically. This study aims to investigate consonantal effects of aspiration in a tone language, Cantonese. The results showed that aspiration had a raising effect on onset f0, that is, onset f0 after aspirated stops was higher than after unaspirated stops. Besides, the aspiration-related f0 perturbations can extend to around 100ms after voicing. However, unlike f0 as a secondary cue for stop contrasts in English, when voice onset time (VOT) becomes ambiguous, f0 was not strengthened for contrast enhancement in Cantonese as well as in L2 English. This indicates that although consonantal effects in Cantonese showed phonetically similar directions and comparable duration with native English, onset f0 was not used for phonological contrast enhancement.

Index Terms: onset f0, aspiration, tone, VOT, production, Cantonese

1. Introduction

Consonantal effects on fundamental frequency at the onset of the following vowel (onset f0) are implemented differently in different languages. In English, onset f0 acts as a secondary cue to the voicing of the initial stops. It helps to decide between voiced and voiceless stops, when the primary cue VOT (voice onset time: the duration between the release of a stop and the onset of voicing) is ambiguous under certain conditions [1]. Onset f0 can also be used as a primary cue. For instance, young generations in Korean have merged the VOT difference between the lenis and the aspirated stops [2], so onset f0 becomes the only cue for distinguishing these stops.

The duration of consonant-related f0 perturbations into the vowel, and the direction of onset f0 changes (rising or falling f0) also differ among various languages. Usually, consonantal effects are more salient in non-tone languages than in tone languages [3] [4], probably because native speakers of a tone language need to inhibit such consonantal effects to use f0 for lexical tone categorization.

There are two contrasting views in the literature that account for the consonantal effects: a purely phonetic view and a more phonological view. The phonetic view believes that these effects are automatic and determined intrinsically by physiological settings for voicing production [5], so they are not controlled by the speaker. The phonological view emphasizes the speaker's active control in order to enhance perceptual distinctiveness between the stop contrasts [6] [7].

Based on the phonetic perspective, Cantonese stop contrast by aspiration should exhibit similar patterns with English voicing contrast. Although English stops are usually considered to contrast voicing, the distinction between voiced and voiceless stops is basically phonological, whereas they are phonetically realized as voiceless unaspirated and aspirated stops in syllable-initial positions, with short-lag and long-lag VOT respectively, just like Cantonese initial stops. Alternatively, consonantal effects on f0 onset in English and Cantonese should be asymmetric, because phonologically, different features [voice] and [spread glottis] are involved in stop production.

Previous studies tend to support the phonological view: English phonologically voiced and voiceless stops showed similar directions of f0 onset changes with French phonologically voiced and voiceless stops [8], although they were phonetically realized differently as voiceless unaspirated and aspirated stops in English and voiced and voiceless stops in French. Another study comparing consonantal effects in English and Spanish [9] found that the same 2 phonetic categories (voiceless unaspirated and voiced) in English and Spanish can have different effects on f0 onset in the two languages, since the stops are phonologically contrastive in Spanish but not in English.

However, while most research in this line examines languages with stops that contrast voicing, whether phonological or phonetic voicing, less is known for languages that contrast aspiration, such as Cantonese. There is no agreement yet concerning consonantal effects on onset f0 by aspiration. Some studies showed that onset f0 after aspirated stops is lower than after unaspirated stops (Hindi: [10], Mandarin: [11]), while other studies found that aspiration has a raising effect on f0 (Danish: [12]). Contradicting results were obtained even within the same languages by different studies (Thai: [3] [13]). In Cantonese, [4] showed that aspiration lowered f0 while [14] supported that aspiration raised f0.

However, [14] have several problems: it only examined f0 onset after Cantonese bilabial stops of syllables with Cantonese high-level tones, and the f0 value was calculated as an average value over the first 30ms after vowel onset.

The present study aims to investigate the above-mentioned aspects by adding more time points for f0 measurement, and it will also include 3 Cantonese level tones and different places of articulation for stops, which will possibly influence the consonantal effects on onset f0, and explore the relationship between exact VOT values and onset f0. Additionally, besides consonantal effects on onset f0 in native Cantonese, such effects in their L2 English will also be examined to understand the form of consonantal effects in interlanguage as well as to examine potential influence of English on native Cantonese patterns. English exposure is inevitable, since native Cantonese speakers in Hong Kong learn English from childhood.

In summary, this study investigates consonantal effects of aspiration on onset f0 in Cantonese in a comprehensive manner. Specifically, we ask: (1) what are the patterns like in

Cantonese, in terms of the duration of f0 perturbations after voicing as well as the direction of f0 changes? (2) How will factors of place of articulation, different tones, VOT values influence the acoustic realization of onset f0? (3) How are the consonantal effects realized in L2 English spoken by native Cantonese speakers?

2. Method

2.1. Subjects

23 college-aged native speakers of Cantonese (8 males, 15 females) were recruited at the Chinese University of Hong Kong. They started to learn English at an early age, around 2-3 years old, and their average English test score was equivalent to IELTS 6.8 (SD = 0.53). Nevertheless, they reported limited use of English outside classroom and most of them spoke English with a typical Hong Kong accent.

2.2. Reading materials and procedure

Target syllables with the stop contrasts appeared as the first syllable in $C_1V_1C_2V_2$ disyllabic non-words, with the following vowel V_1 varying among /a, i, u/. Each word was repeated 3 times.

The disyllabic non-words with aspirated and unaspirated stops in Cantonese and voiceless and voiced stops in English were further embedded in carrier phrases: “有個_井” (/jɐu kɔ _ tsɛn/, ‘there is a _ well’) in Cantonese and “It’s a _ road” in English, with the disyllabic non-words acting as an adjective of the following noun. Therefore, the stops of interest were placed in similar phonetic environment ɔ/ə_a/i/u, making Cantonese and English stops comparable in this study.

To investigate the effects of place of articulation as well as Cantonese tones in terms of tone height, stops in 3 places of articulation: bilabial, alveolar and velar, in syllables with 3 different Cantonese level tones: tone1[55], tone3[33] and tone6[22] were also included in the word list. For example, in the /a/ vowel context, 巴波-趴波 /pa55pɔ55/-/pʰa55pɔ55/, 丹波-攤波 /tan55pɔ55/-/tʰan55pɔ55/ and 篝波-溝波 /kɐu55pɔ55/-/kʰɐu55pɔ55/ are tone1 stimuli in bilabial, alveolar and velar places of articulation respectively in Cantonese, while *babber-papper*, *dabber-tabber* and *gabber-kabber* are the corresponding word pairs in English. To investigate the effects of Cantonese tones, in the bilabial place of articulation, for example, 巴/pa55/-霸/pa33/-罷/pa22/ and 趴/pʰa55/-怕/pʰa33/ are the target syllables for unaspirated and aspirated stops respectively. No /pʰa22/ word was included since Cantonese tone6 does not appear together with Cantonese aspirated stops.

Each participant was asked to read aloud all the 189 such sentences presented one at a time on the computer screen. These sentences were randomized for each participant and the order of languages was counterbalanced.

2.3. Measurements

The voicing onset was defined at the point of the zero crossing of the first glottal pulse extending through F1 and F2. VOT of the stops and f0 of the following vowel (i.e. onset f0) were obtained using Praat. The f0 values were measured at every 10ms during the first 100ms immediately after voicing onset. The decision of 100ms is in line with the previous study [4][9][15] and it also considered the overall vowel duration (at least 15.2% tokens will be left out if a period longer than 100ms is chosen). For a more accurate pitch tracking, the pitch range was set according to each speaker’s pitch floor and ceiling. Tokens with pitch tracking errors, mostly due to creaky voice, were omitted as outliers. As a result, 412 out of 46597 (0.88%) tokens were excluded. The remaining tokens were converted from Hz to semitones re 100Hz for data analysis.

For statistical analyses, SS-ANOVA was used to compare the effects of stop contrasts on onset f0 under different conditions, Pearson's correlation coefficient (r) was adopted to examine the relationship between VOT and onsetf0, and lastly, we include all the factors in mixed-effect models for a more complete picture.

3. Results and Discussion

While sections 3.1-3.6 demonstrate the patterns of consonantal effects in Cantonese and each factor influencing the realization of onset f0 with all the 11 time points or at 10ms (VOT-related data), section 3.7 includes these factors in several models using the average f0 values over the 11 time points or only 10ms (VOT-related data), so that the interactions between these factors can be examined.

3.1. Onset f0 and aspiration

Figure 1 presents the overall f0 contours for Cantonese aspirated and unaspirated tokens collapsed over place of articulation and tones: onset f0 after aspirated stops was higher than after unaspirated stops, and such f0 perturbations extended far into the vowel to almost 100ms after voicing.

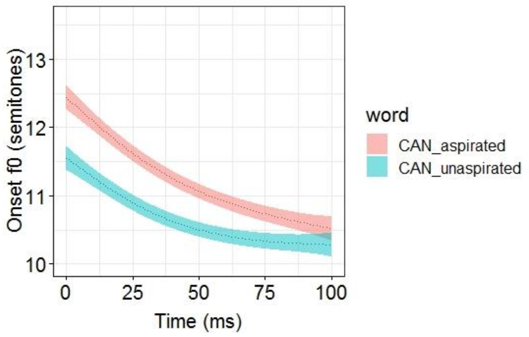

Figure 1: *Overall f0 contours after Cantonese aspirated and unaspirated stops.*

Previous studies showed that tone languages usually had a shorter consonantal effect into the following vowel [3][4], probably because f0 activities were already employed for lexical tone contrasts. However, our data observed nearly 100ms consonant-induced f0 perturbations in Cantonese, which is much longer than what is expected for tone languages.

The longer-than-expected f0 perturbations may be related to the experimental design in this study. Since we used non-words as stimuli, the speakers were likely to read the target words more carefully. The consonantal effects may be strengthened consequently.

Consonantal effects of 100ms is not likely in a tone language but is comparable to the consonantal effects in non-tone languages such as English [9] [15]. Since the speakers also learned English, it is also possible that their L2 English influenced their native Cantonese to be more English-like (see Section 3.4 for more on this point).

3.2. Onset f0 and place of articulation

In Figure 2, both the magnitude of f0 differences and the location of the crossover between the two curves indicates that the consonantal effects of aspiration appeared to be the strongest for Cantonese stops at alveolar place of articulation, while relatively weaker for stops at velar place of articulation.

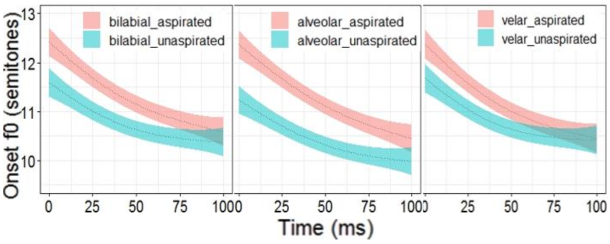

Figure 2: *F0 contours after Cantonese aspirated and unaspirated stops at bilabial, alveolar and velar places of articulation respectively.*

3.3. Onset f0 and tones

Figure 3 only displays tokens with Cantonese tone1 and tone3, since Cantonese tone6 does not appear together with Cantonese aspirated stops, thus making f0 comparison based on aspiration impossible in the tone6 condition. Figure 3 shows that the consonantal effects did not differ much on syllables with Cantonese high level tone1 and mid-level tone3, considering both the direction and the duration of onset f0 changes.

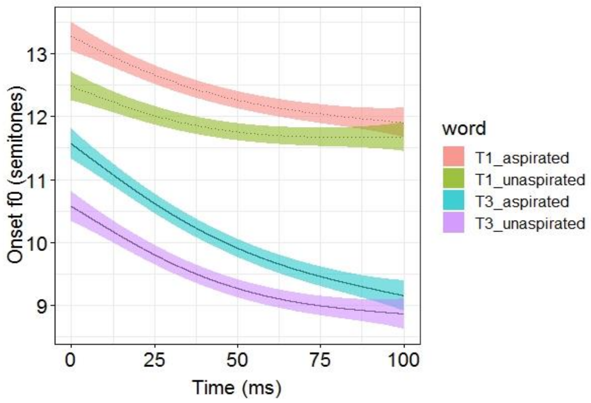

Figure 3: *F0 contours for Cantonese aspirated and unaspirated syllables with Cantonese tone1 and tone3.*

3.4. Comparing native Cantonese and L2 English

Comparing the consonantal effects of aspiration in two aspiration languages, at least phonetically, Cantonese-English exhibited similar patterns with Cantonese in that aspiration had a raising effect on the following f0: onset f0 was higher after Cantonese aspirated and English voiceless stops, whereas f0 was lowered by Cantonese unaspirated and English voiced stops.

Besides, even the duration of such f0 perturbations were similar around 100ms in both Cantonese and their L2 English.

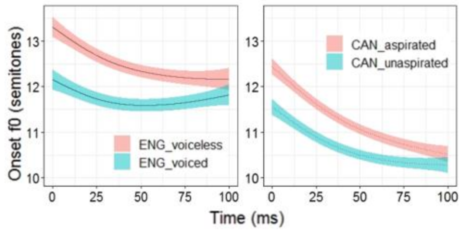

Figure 4: *Consonantal effects on onset f0 in L2 English (left) and Cantonese (right).*

This uniformity between Cantonese and L2 English may be explained in two aspects: on the one hand, the Cantonese and the English data were produced by the same group of speakers who learned English as their L2 language. Their physiological settings for speech articulation are likely to be the same in Cantonese and English production. On the other hand, similar patterns in Cantonese and L2 English can be due to a phonetic transfer between these two languages, probably backward transfer from L2 to L1. In English, many studies (e.g. [9] [15]) consistently showed that voiceless stops have a raising effect on onset f0 while voiced stops lowered onset f0, and the duration can reach nearly 100ms or longer. Given that the direction and the duration of f0 changes in Cantonese were not expected based on the literature, but closer to what is described for native English, it seems plausible that the Cantonese patterns were influenced by L2 English. Though L2-to-L1 transfer does not occur as frequently as L2 transfer to L1, it is not surprising to see backward transfer of L2 on f0, since acquiring new f0 use may not be difficult for native tone language speakers. This is also supported by the findings in [4] that native Cantonese speakers were able to use onset f0 as a cue for synthesized Cantonese aspirated and unaspirated stops, independent of the patterns in their native language.

3.5. Onset f0 and VOT

In Figure 5, f0 measurements at 10ms after voicing onset were chosen to compare with the corresponding VOT values. However, the results found no significant correlations between VOT and onset f0 within each category (Cantonese: aspirated p=0.6, r=0.02, unaspirated p=0.16, r=0.04; English: voiceless p=0.31, r=0.02, voiced p=0.16, r=0.06).

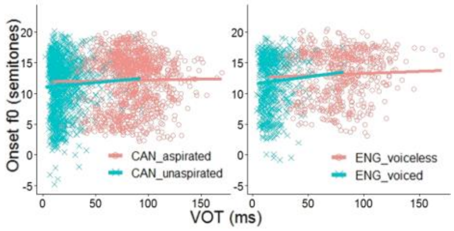

Figure 5: *Onset f0 by VOT for stop contrasts in Cantonese (left) and English (right), with regression lines fitted within each stop category.*

From a phonological view, when VOT is ambiguous, f0 will be strengthened to keep stop contrasts distinct. Accordingly, onset f0 and VOT within each stop category should be in a trading relationship: for unaspirated stops, when the VOT increases to be closer to the VOT of the aspirated category, f0 should decrease so that the relatively low f0 onset of unaspirated stops get even lower, to enhance the f0 difference for aspiration. Similarly, a negative VOT-onset f0 correlation is expected for aspirated stops as well. However, this negative correlation was not found in Cantonese and L2 English.

Therefore, although the consonant-related f0 perturbations in Cantonese were already comparable to the salient effects in non-tone language such as English, it still differs from native English in terms of the status of f0. In English, onset f0 can be a secondary cue for voicing contrast, but in Cantonese, the weighting of f0 cue for stop contrasts were ungraded due to f0 use for lexical tones and other cues, such as aspiration noise, may be stronger than f0 for the stop contrasts.

3.6. Effect of tone on VOT-onset f0 dependency

Figure 6 shows how different tones affected VOT and onset f0 relations in Cantonese (no p-value corrections for multiple comparisons): only unaspirated tone1, aspirated tone3 showed weakly positive correlations between VOT and onset f0, the correlation in the unaspirated tone3 category was even weaker.

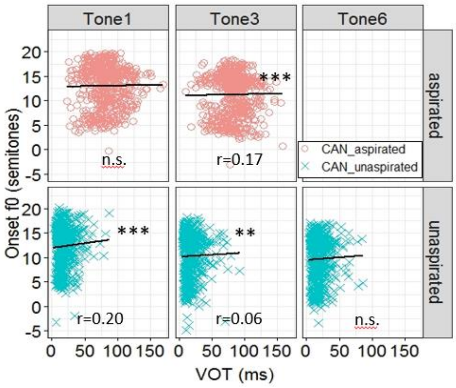

Figure 6: *Onset f0 by VOT for Cantonese aspirated and unaspirated syllables with different tones.*

A possible account could be that for the unaspirated tone1 category, longer VOT with higher f0 makes it more distinct from the aspirated tone3 category and vice versa for the aspirated tone3 category, because distinguishing different tones is a more important task than f0 differences as a cue to aspiration in Cantonese.

3.7. Interactions between aspiration and other factors

Linear mixed model analyses were adopted to include all the factors examined before (Type, Place of Articulation (POA), Tone, VOT, Language) with the focus on the interaction effects between Type (aspirated vs. unaspirated) and other factors. Only effects of interest are presented in the following tables.

The results in Table1, with Model 1 and 2 corresponding to the results in Sections 3.1-3.3 and 3.4 respectively, confirmed the raising effect of aspiration in Cantonese, which was the weakest at velar place of articulation compared with the bilabial and alveolar stops and did not differ between tone1 and tone3 conditions. Importantly, Cantonese and L2 English patterns matched in terms of both the direction and the magnitude of the consonantal effects on onset f0.

Additionally, the 2 models in Table 2, corresponding to Sections 3.5 and 3.6, showed that although for each stop category, onset f0 did not correlate with VOT values, in general, English voiceless stops exhibited a slightly more positive correlation than the voiced stops compared with Cantonese. In examining tone effects on VOT-f0 relations within Cantonese, each of the 5 stop categories showed either no significant or weakly positive correlations.

Table 1: *Mean f0: Type × Place of Articulation (POA) and Type × Tone effects in Cantonese (upper); Type effects between Cantonese and L2 English (below).*

Fixed effects	Estimate	SE	t-value
Model1: only Cantonese			
Type	0.6797	0.1236	5.498***
Type:POAbilabial	-0.1745	0.1143	-1.527
Type:POAvelar	-0.2746	0.1151	-2.386*
Type:Tone3	0.0869	0.0935	0.929
Model2: both Cantonese and L2 English			
Type	0.6751	0.1414	4.774***
Type:Language	0.2177	0.2828	0.770

Table 2: *F0 at 10ms: Type × VOT effects between Cantonese and L2 English (upper); Type × VOT, Type × VOT × Tone effects in Cantonese (below).*

Fixed effects	Estimate	SE	t-value
Model3: both Cantonese and L2 English			
Type	0.7194	0.1821	3.951***
Type:VOT	-0.0044	0.0043	-1.033
Type:VOT:Language	0.0054	0.0024	2.221*
Model4: only Cantonese			
Type	0.5198	0.2346	2.215*
Type:VOT	-0.0063	0.0045	-1.420
Type:Tone3:VOT	0.0002	0.0022	0.081
Type:Tone6:VOT	-0.0047	0.0129	-0.368

4. Conclusion

Regarding the direction and duration of aspiration-related f0 changes in Cantonese, Cantonese aspirated stops had a higher onset f0 than the unaspirated stops in all places of articulation (bilabial, alveolar, velar) with both tone1 and tone3. Such consonantal effects can extend to around 100ms after voicing, which is unexpected in Cantonese because it is comparable to non-tone languages like English. A possible explanation could be English transfer effects, but mainly phonetic transfer, because the speakers could not compensate ambiguous VOT with enhanced f0, as in languages with f0 as a secondary cue.

The divergent results in this study from [4] that found lowering consonantal effects in Cantonese only immediately following voice onset (0ms) may be due to possibly different English proficiencies, stimuli words in isolation ([4]) vs. in carrier phrases (the current study and [14] that also showed a raising effect), since running speech may change the higher vocal fold tension proposed to raise f0 in unaspirated stops [10][15] and the faster airflow rate that can explain higher f0 in aspirated stops [16]. Physiological studies are needed to provide relevant evidence. Moreover, [4] showed a trend of raising effect from 10ms, although not statistically significant.

While previous research mostly explored the phonological aspect of consonantal effects, this study offered more support for the phonetic view: Cantonese stop contrast by aspiration exhibited similar patterns with native English voicing contrast in terms of consonantal effects, given that Cantonese and English stops are both phonetically realized as aspirated and unaspirated voiceless, but they are phonologically different.

5. References

[1] A. S. Abramson and L. Lisker, "Relative power of cues: F0 shift versus voice timing", *Phonetic linguistics: Essays in honor of Peter Ladefoged*, pp. 25-33, 1985.

[2] Y. Kang, "Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study", *Journal of Phonetics*, *45*, pp. 76-90, 2014.

[3] J. Gandour, "Consonant types and tone in Siamese". *Journal of phonetics*, *2*(4), pp. 337-350, 1974.

[4] A. L. Francis, V. Ciocca, V. K. M. Wong and J. K. L. Chan, "Is fundamental frequency a cue to aspiration in initial stops?", *The Journal of the Acoustical Society of America*, *120*(5), pp. 2884-2895, 2006.

[5] J. M. Hombert, J. J. Ohala and W. G. Ewan, "Phonetic explanations for the development of tones", *Language*, pp. 37-58, 1979.

[6] P. A. Keating, "Phonetic and phonological representation of stop consonant voicing", *Language*, pp. 286-319, 1984.

[7] J. Kingston and R. L. Diehl, "Phonetic knowledge", *Language*, *70*(3), pp. 419-454, 1994.

[8] J. M**.** Hombert, "Consonant types, vowel quality, and tone", *Tone: A Linguistic Survey*. New York: Academic Press, 2014.

[9] O. Dmitrieva, F. Llanos, A. A. Shultz and A. L. Francis, "Phonological status, not voice onset time, determines the acoustic realization of onset f0 as a secondary voicing cue in Spanish and English". *Journal of Phonetics*, *49,* pp. 77-95, 2015.

[10] R. Kagaya and H. Hirose, "Fiberoptic electromyographic and acoustic analyses of Hindi stop consonants", *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics*, *9*, pp. 27-46, 1975.

[11] C. X. Xu and Y. Xu, "Effects of consonant aspiration on Mandarin tones". *Journal of the International Phonetic Association*, pp. 165-181, 2003.

[12] V. Jeel, "An investigation of the fundamental frequency of vowels after various Danish consonants, in particular stop consonants", *Annual Report of the Institute of Phonetics, University of Copenhagen*, *9*, pp. 191-211, 1975.

[13] W. G. Ewan, "Laryngeal behavior in speech", *Doctoral Dissertation, University of California, Berkeley,* 1976.

[14] E. Zee, "The effect of aspiration on the F0 of the following vowel in Cantonese", *UCLA Working Papers in Phonetics*, *49*, pp. 90-97, 1980.

[15] A. Löfqvist, T. Baer, N. S. McGarr and R. S. Story, "The cricothyroid muscle in voicing control". *The Journal of the Acoustical Society of America*, *85*(3), pp. 1314-1321, 1989.

[16] J. J. Ohala, "The physiology of tone". *Southern California occasional papers in linguistics*, *1*, pp. 1-14, 1973.