Cite as:
Tse, Holman. 2012. Lexical Tone vs. F0 Effects on VOT in Cantonese. Unpublished Manuscript. University of Pittsburgh. Pittsburgh, PA

Lexical Tone vs. F0 Effects on VOT in Cantonese

Holman Tse
University of Pittsburgh
Department of Linguistics
2816 Cathedral of Learning
Pittsburgh, PA 15217
USA
hbt3@pitt.edu

Abstract

This study investigates the effects of lexical tone on the Voice Onset Time (VOT) of prevocalic stops in Cantonese. It builds on literature showing how various factors affect VOT without a resultant loss in phonological contrast and also ties in a body of research on tone and consonant interaction. The specific research questions addressed are (1) Does tone have an effect on VOT in Cantonese?; (2) If so, what kind of an effect does it have?; and (3) Is this effect purely an automatic articulatory consequence of F0 modulation or is this effect also mediated by lexical tone and hence a secondary cue that contributes towards maintaining phonological contrasts between different tonal categories? To address these questions, the speech of 6 native speakers (5 male and 1 female) of Hong Kong Cantonese was examined. A total of 80 tokens of words contrasting in tone and aspiration were analyzed for each speaker for a grand total of 480 tokens. Results from an ANOVA test showed that there is a statistically significant effect ($p < 0.001$) of tone category on VOT. A post-hoc analysis revealed a two-way split between words with a low-falling (21) or mid-rising (25) tone, which have higher VOT, and words with either a mid-level (33) or high-level tone (55), which have lower VOT. A Pearson's Correlation Test also showed statistical significance ($p = 0.01$) with an inverse relationship between VOT and onset F0. An analysis of individual speakers, however, showed that this correlation was statistically significant for only 3 of the 6 subjects while ANOVA tests on each individual speaker showed statistical significance for all 5 male subjects but not for the female subject. The more consistent results on the ANOVA tests on individual speakers suggest that VOT differences can serve as a secondary cue used to distinguish between different tonal categories. This study has implications for developing a better understanding of how phonological contrasts are implemented in tonal languages.

1. Introduction

The study presented in this paper addresses whether or not lexical tone has an effect on the Voice Onset Time (VOT) of prevocalic stops in Cantonese, a Sinitic (Chinese) language with a two-way contrast between voiceless aspirated (long-lag) and voiceless unaspirated (short-lag) stops. In addressing this question, this paper brings together two major research areas focused on the phonetic implementation of phonological categories. The first is research on factors that affect VOT without a loss in contrast between stop categories while the second is research on tone (phonetically implemented as F0) and consonant interaction. The findings from this study have implications for developing a better understanding of the phonetics-phonology interface.

Voice Onset Time (VOT), which is defined as the time interval between the release of the stop and the onset of voicing for the following vowel, was first introduced by Lisker & Abramson (1964) as an acoustic cue to distinguish between different stop categories. This cross-linguistic study of 11 languages found that in spite of variation in the precise VOT values for different stop phonemes in these languages, VOT values tend to cluster into a maximum of three groupings. These three groupings correspond to the stop phonemes in each of the languages examined. This includes voiced stops, which all have negative VOT values, voiceless unaspirated stops, which all have positive VOT values slightly above zero (short-lag), and voiceless aspirated stops, which have positive VOT values in a range that is distinctly above the range found for voiceless unaspirated stops (long-lag). This clustering of VOT values suggests that VOT can serve as an acoustic cue for distinguishing between different stop categories.

Subsequent research has shown that the range of VOT values corresponding to each of these clusters is not stable. Various factors can significantly increase or decrease VOT values while still maintaining phonological contrasts between different stop categories. Factors that have been examined include place of articulation (Cho & Ladefoged, 1999), prosodic position (Cho, 2003; Van Dam, 2003), vowel height (Klatt, 1975), and rate of speech (Kessinger & Blumstein, 1997). Perception of VOT boundaries has also been investigated and has shown how listeners compensate for speaking rate effects by recognizing intrinsic timing ratios (Boucher, 2002; Summerfield, 1981). This body of literature raises the question of what else can affect VOT without a resultant loss in phonological contrast. Lexical tone is the factor investigated in this paper. The question of what can synchronically affect VOT may also be related to the diachronic question of tone and consonant interaction. For example, as will be discussed in Section 2, there has been a body of literature on tone and consonant interaction based on the observation that there is a relationship between voicing contrasts and the diachronic development of tonal contrasts. While a strong correspondence has been observed between low tone (and low F0) and voicing (or low VOT) for tonal languages with a voicing contrast, research on an analogically similar correspondence between tone (and F0) and VOT for languages that have contrasting voiceless stops (ex: short-lag vs. long-lag) has shown mixed results.

The paper specifically addresses the following research questions:

(1) Does tone have an effect on VOT in Cantonese?

(2) If so, what kind of an effect does it have? (Ex: Shorter VOT with low tone?)

(3) Is this effect purely an articulatory consequence of F0 modulation or is this effect also mediated by lexical tone and hence a secondary cue that contributes towards maintaining phonological contrasts between different tonal categories?

Once these questions are addressed, this paper addresses the extent to which the findings are language specific and whether or not cross-linguistic generalizations can be made. As the question of tone affecting VOT fundamentally involves the issue of tone and consonant interaction, the literature review (Section 2) that follows will begin with a discussion of this topic. Section 3 reviews research on the various factors affecting VOT. Relevant phonological information about Cantonese is introduced in Section 4. The methodology for this study is presented in Section 5. This is followed by a presentation of the results in Section 6 and a discussion of the findings in Section 7. This paper concludes in Section 8.

2. Diachronic Studies of Tone and Consonant Interaction

Much research on tone and consonant interaction has focused on diachronic accounts of how tonal distinctions developed based on voicing distinctions in consonants. The fact that voiced stops in tone languages that have them often appear in syllables with low tones is a correspondence that has long been observed. Haudricourt's (1954) pioneering study of Vietnamese tonogenesis offered a phonetically motivated explanation to this observation. By comparing the modern reflexes of several Southeast Asian languages and historical records of earlier stages of these languages, he argued that the Vietnamese tone system emerged from a former voicing opposition that turned into an opposition between high tone and low tone with voiceless aspirated stops developing into high tone and voiced stops developing into low tone.

Subsequent studies including Hombert (1975, 1978) provided instrumental data that supported Haudricourt's claim by showing that prevocalic voiceless stops have a raising effect on the F0 onset of the following vowel while voiced stops have a lowering effect. This differential effect on F0 onset based on voicing contrasts is, thus, what leads to the widespread correspondence between low tone and voiced stops and between high tone and voiceless stops. Modern Chinese dialects have been shown to have developed their current tone and stop inventories in this way. The variation found in the modern reflexes of current dialects reflects different stages of this process. For example, the Songjiang Wu dialect retains a 3-way stop contrast that includes voiced stops as well as voiceless aspirated and voiceless unaspirated stops. Voiced stops occur only in low tones as shown in the examples below:

[-voice]	Tonal Contour	Gloss	[+voice]	Tonal Contour	Gloss
ti	52	‘low’	di	31	‘lift’
ti	44	‘bottom’	di	22	‘younger brother’
ti	35	‘emperor’	di	13	‘field’

Table 1: Tone and Voicing Correspondences in Songjiang Wu (From Yip, 2002, p. 7)

Some dialects such as Cantonese lack voiced stops and can be considered more advanced in the tonogenesis process. The current tonal inventory is derived from a three-tone system in Middle Chinese that split into a six-tone system. As the chart below shows, all words beginning with voiced stops in Middle Chinese correspond to low tones in Modern Cantonese. While in some dialects, the reflexes retained the voiced stops, in Cantonese these voiced stops became either voiceless aspirated (long-lag) or voiceless unaspirated (short-lag) stops while retaining their low register tone.

	Middle Chinese Tone A	**Middle Chinese Tone B**	**Middle Chinese Tone C**
HIGH REGISTER TONES (Yin) voiceless aspirated and unaspirated stops (unchanged)	55/53 both aspirated and unaspirated	35 both aspirated and unaspirated	33 both aspirated and unaspirated
LOW REGISTER TONES (Yang) *voiced stops > voiceless aspirated and unaspirated stops	21 voiceless aspirated	23 voiceless aspirated	22 voiceless unaspirated

Table 2: The Modern Cantonese Tone System for Syllables with Stop Onsets[1]

Hombert et al. (1979) discuss two phonetically motivated hypotheses to account for this effect. According to one hypothesis based on aerodynamics, the production of voiced stops involves a decrease in pressure drop and this consequently leads to lower F0 at the onset of the

[1] Information for this table was adapted from (Haudricourt, 1972).

following vowel. The aerodynamic effect of voiceless stop production, however, results in higher F0. The second hypothesis is that the F0 differences are due to differences in vocal fold tension.

Whatever proves to be the underlying mechanism, what is clear is that there is a widespread correspondence between voicing and tone. This pattern has led Hyman & Schuh (1974) to make the generalization that "consonants affect tone, but tone does not affect consonants." Thus, while any diachronic tone change rule can also be a synchronic phonological rule, they claim that the reverse cannot be true. If this generalization were correct, it would suggest that tonal effects on VOT do not exist. Maddieson (1974, 1976, 1977), however, disagrees with this generalization and cites numerous examples that he claims are exceptions that need to be reinterpreted. One example is Jingpho, which exhibits a voicing alternation between [yàk] ('difficult') and [yàggai] ('it is difficult'). This alternation does not occur in words with high tone. For example, there is no voicing alternation between [cát] ('tight') and [cáttai] ('it is tight'). Maddieson (1974) argues that this is a counterexample to Hyman & Schuh's (1974) generalization because it clearly shows a case of low tone triggering a change on the following consonant. Hyman (1976) responds by saying that the Jingpho case is not an exception because it involves breathiness and it happens to be the case that all Jingpho low tones are accompanied by breathiness. He also reclarifies his original claim by restating it as "consonants affect pitch, but pitch does not affect consonants" (Hyman, 1976, p. 94). This reformulated statement excludes cases such as the Jingpho example, which he believes involve segmental changes triggered by breathiness rather than by pitch.

Hyman also clarifies his original statement by saying that "what we are interested in is cases where a tone possibly affects the <u>segmental</u> [emphasis in original] characteristics of a consonant" (1976, p. 92). Thus, low-level phonetic changes such as VOT differences induced by tonal differences, as is the case for this paper, would be excluded from Hyman & Schuh's (1974) generalization. Nevertheless, Maddieson (1977) pushes the issue of directionality of effect even further by citing examples from instrumental studies that show tone affecting consonants without changing segmental characteristics. For example, Maddieson makes the generalization that "higher tones condition longer consonants" (1977, p. 103). This claim is based on several sources including his work on Thai which showed that the voiced bilabial stop /b/ has a longer duration when preceding the high and falling tones than when preceding the low, mid, and rising tones. He also cites Zee (1977) which shows that the duration of /s/ in /si:/ in Taiwanese is longer when preceding high tones. Gandour & Maddieson (1976) provide further evidence showing that tones can affect consonants. This study involved measurements of larynx movement for a Thai speaker and showed that when producing low tones, the larynx moved more downward than when producing high tones. This physiological effect shows that the following tone can affect a preceding consonant. Though Maddieson (1974, 1976, 1977) presents many examples that he believes show tone affecting consonants, he still agrees with Hyman (1976) that clear cases of consonants affecting tone rather than the reverse are far more common.

3. VOT in synchronic studies

To reiterate, the main focus of this paper is on low-level phonetic effects that do not change the stop category. This builds on a large body of literature on the phonetics-phonology interface by studying how different factors can affect phonetic production without having an effect on the phonological characteristics of individual sounds. Experimental studies involving VOT measurements have formed a major part of this body of literature.

Lisker & Abramson's (1964) study of VOT provided evidence showing how phonetic categories do not necessarily map one-to-one with phonological categories. Based on the clustering of VOT values into different groups, Lisker and Abramson established three universal phonetic categories: short-lag (voiceless unaspirated with short positive VOT), long-lag (voiceless aspirated with long positive VOT), and pre-voiced (negative VOT). These three phonetic categories correspond to different categories depending on the language. For languages with three stop categories such as Eastern Armenian and Thai, the mapping is relatively straightforward. Pre-voiced stops are phonologically voiced; Short lag stops are phonologically voiceless unaspirated; Long-lag stops are phonologically voiceless aspirated stops.

Korean is also a language with three stop categories. Unlike Eastern Armenian and Thai, however, the three stops are produced at different points along the VOT continuum. The Korean phonological categories include one described as tense, long, and glottalized (with positive VOT values near zero), another described as lax and slightly aspirated (with VOT slightly higher than tense stops), and a third described as lax and heavily aspirated (with VOT higher than the other two categories). All three categories are voiceless and, thus, have positive VOT. Though Korean differs from other three-category languages in not having its three stop categories more spread out along the VOT continuum, it still makes use of different ranges of VOT values to distinguish among these three different categories.

Languages that have two categories of stops are more variable in how they make use of VOT values. Some languages such as Dutch, Spanish, Hungarian, and Tamil have a contrast between phonetically pre-voiced and phonetically short-lag stops. Others like English and Cantonese, however, have a contrast between phonetically short-lag and phonetically long-lag stops. Keating (1984) illustrates the discrepancy between phonetic and phonological categories by contrasting English and Polish, which are both two-category languages. Although both of these languages are described as having voiced and voiceless stops, the phonetic realization of this distinction is different. In Polish, voiced stops are phonetically pre-voiced while voiceless stops have low positive VOT. In English on the other hand, voiceless stops have much higher VOT (long-lag) while voiced stops are only slightly pre-voiced. Keating argues that a polarization principle operates in both of these two languages such that in English, the short-lag stop is slightly pre-voiced while the short-lag stop in Polish is slightly aspirated. This polarization principle maximizes the phonetic differences in language specific stop categories. Yet, in spite of the different phonetic sounds that are actually produced in these languages, the universal feature [±voice] is all that is needed in phonological representations of these languages even if, as is the case for English, the [+voice] category has both voiced and voiceless allophones.

While the previous section makes clear that VOT values cluster together in an organized way to maintain voicing and aspiration contrasts, the picture gets more complicated when we examine some of the wide range of factors that can shift VOT values in different directions without losing these contrasts. One of these factors, place of articulation, has been widely attested cross-linguistically. Even Lisker & Abramson (1964) noticed a pattern with the place of articulation in their data but they did not try to account for why place of articulation may affect VOT. In a study of 18 languages, Cho & Ladefoged (1999) show that the further back the place of articulation, the longer the VOT for voiceless stops especially for velars and uvulars compared to coronals and labials although this generalization holds less true for the difference between labials and alveolars and the difference between velars and uvulars. Cho & Ladefoged (1999) suggest that some of the deviations from this place of articulation generalization may be due to language-specific effects interacting with universal phonetic constraints. For instance, some of the languages they studied contrast plain oral and ejective stops. In this case, shifting VOT values can be seen as a way of maximizing contrasts between different stop categories.

Another factor shown to affect VOT values is prosodic position. In English, for example, stress enhances VOT contrasts such that in stressed syllables pre-vocalic voiceless stops have higher VOT while underlying 'voiced' stops are more likely to be pre-voiced than in other prosodic positions (Lisker & Abramson, 1967). Increasing the VOT differences between the two stop categories can, thus, be seen as a way of using stress to emphasize this distinction. Van Dam (2003) shows other stress-induced patterns in English. First of all, he shows that underlying voiceless stop onsets in stressed syllables have higher VOT than onset stops in syllables following stress. This study also shows that onset stops in a position not adjacent to stressed syllables have intermediate VOT values. These prosodic effects are language-specific as Dutch has been shown to display the opposite effect with underlying voiceless stops having shorter VOT in stressed contexts (Cho, 2003). Cho attributes the difference between English and Dutch to the different phoneme inventories of the two languages. Yet, the same principle of prosodic strengthening operates in both languages. While Dutch has a contrast between voiced and short-lag voiceless stops, the two stop categories in English are both voiceless. The difference can be seen in the fact that in Dutch, the voiceless stop involves glottal adduction while in English the voiceless stop involves glottal abduction. Prosodic strengthening in Dutch would mean more adduction and hence lower VOT while prosodic strengthening in English would mean more abduction and hence higher VOT. In all of these cases, the prosodic effects on VOT values can be seen as a way of maintaining the phonemic contrasts that exist in a particular language by enhancing the differences.

A third variable that has been shown to affect VOT values is rate of speech. Kessinger & Blumstein (1997) examine the effects of speaking rate in three different languages with different stop inventory systems: Thai (pre-voiced, short-lag, and long-lag), French (pre-voiced and short-lag), and English (short-lag and long-lag). As rate of speech increases, VOT values for both long-lag and pre-voiced stops in the languages that have them approach zero. A slower speaking rate results with the inverse of these effects and, thus, more negative VOT for pre-voiced stops and more positive VOT for long-lag stops. Interestingly, the VOT for short-lag stops remains relatively stable at all rates of speech. There was also no overlap found between pre-voiced and short-lag and only minimal overlap found between short-lag and long-lag at faster rates of speech. The short-lag category, which is found in all three languages studied, seems to act as a "phonetic anchor" to help maintain stop contrasts by keeping the VOT values of one category in place. While Kessinger & Blumstein (1997) investigated speech production, Summerfield (1981) and Boucher (2002) investigate speech perception of VOT boundaries. Both of these studies show that listeners learn to compensate for speaking rate effects by recognizing intrinsic timing ratios. These studies show that in spite of variation in absolute VOT values across speaking rates, stable boundaries that separate stop categories can also be found when intrinsic timing relations such as the ratio of VOT/syllable duration are examined.

Finally, several studies have also shown tone to have an effect on VOT. For example, Pearce (2005) examined Kera, a Chadic language spoken in Africa. Like many other tonal languages, Kera appears to have developed a high/low tone contrast through a voicing distinction. Though diachronically speaking, this is clearly a case of consonants affecting tone, Pearce (2005) argues that this is not the case when looking at the language synchronically. The evidence she presents comes from recordings, pitch tracks, and acoustic measurements. In particular, she shows that what have previously been described as voiced stops in Kera are actually voiceless and have short-lag positive VOT values. These stops correspond with low tone while stops with longer VOT values correspond with high tone. Though there is a general correspondence between longer VOT and high tone, the evidence presented does not show a bi-modal split of VOT values based on a voicing contrast. The gradient VOT values, Pearce suggests, indicate that voicing is no longer a contrastive feature in Kera stops. Nevertheless, VOT remains as a secondary phonetic cue that enhances the distinction between high, mid-level, and low tones. Pearce concludes that tone is underlying rather than voicing and that it is the tone (or more precisely the F0 onset) that affects the duration of the preceding consonant rather than the consonant that affects the tone. Pearce (2009) further develops this analysis of Kera with perceptual data confirming that Kera speakers rely more on F0 than on VOT as a perceptual cue while the reverse is true for English speakers.

Herrera (2003) also examined VOT in her study of pre-nasalized stops in Santa Clara Mazatec, a three-category language with long-lag voiceless, short-lag voiceless, and pre-nasalized voiced stops. What is notable about this study is that the pre-nasalized stops were shown to have properties of both voiced and voiceless aspirated stops. First of all, like voiceless aspirated stops the VOT values of these stops were shown to be relatively high. In fact, the average VOT values surpass the range normally found for voiceless aspirated stops in languages that have them (94-126 ms for alveolars and 103-218 ms for velars). In general, the high level tone correlated with longer average VOT for pre-nasalized velar stops than did the low level tone though this correlation does not appear to be statistically significant. For the pre-nasalized alveolar stops, however, only one of the 3 speakers produced longer VOT for high-level tone words while the other two speakers had similar VOT values for both high and low level tones. Secondly, the pre-nasalized stops were shown to have a lowering effect of the F0 of the following vowel much like typical voiced stops. With an uneven effect of tone on VOT but a much clearer effect of the pre-nasalized stops in affecting the F0 of the following vowel, Herrera concludes that in Mazatec, it is consonants that affect tone rather than tone that affects consonants contrary to the results reported for Kera (Pearce, 2005, 2009).

For Shanghainese King & Schiefer (1990), which is a three-category language (long-lag voiceless, short-lag voiceless, and voiced) like Mazatec, high tone was also found to co-occur with longer VOT. While Herrera (2003) examined only pre-nasalized stops, however, King & Schiefer's (1990) study of Shanghainese looked only at short-lag and long-lag stops. Still, the results for these two languages as well as for Kera show the same effect of longer VOT with high tones.

For Mandarin and Sixian Hakka, which are both two-category languages with stops that contrast in aspiration, the effect of tone on VOT seems to be quite different. Liu et al. (2008) found that in Mandarin it is the mid falling-rising and the mid-rising tone that have longer VOT than the high-level and high-falling tones. This seems to be the exact opposite correlation found in Kera, Mazatec, and Shanghainese. Chen et al. (2009) also obtained the same results for Mandarin but found that this correlation held true only if the data included nonce words. When the nonce words were excluded, there was no clear pattern found. The same study also looked at Hakka and found significantly shorter VOT for what are identified as Tone 4 (mid-falling) and Tone 8 (high level). Unlike the effect found for the other cited studies, the effect for Hakka seems to be one of tone duration rather than of tone height. Both Tone 4 and Tone 8 are identified as short tones that occur in syllables that end with a stop.

While there seems to be consensus on the correspondence between voicing distinctions and F0 differences, the research on voiceless unaspirated vs. voiceless aspirated stops have shown mixed results. This may explain why the VOT and tone interaction studies cited above do not show consistent results either. Zee (1980), for example, shows that in Cantonese, voiceless aspirated stops raise the F0 onset of the following vowel. Xu & Xu (2003), on the other hand, show the opposite effects for Mandarin. Hombert & Ladefoged (1977) show no significant difference between the effects of voiceless aspirated vs. unaspirated stops on the following F0 in English and French, which are both non-tonal languages. Yet, even different studies of the same language have shown contradicting results. For example, while Gandour (1974) shows that in Thai the F0 onset is lower following voiceless aspirated stops than following voiceless unaspirated stops, Ewan (1976) shows higher F0 following aspirated stops. Bradshaw (1999) has gone as far as to argue that only voiced segments are ever involved in consonant-tone interactions. Yet, Lai (2004) speculates that the lack of agreement in voiceless consonant-tone interaction studies is due to the fact that they did not take into account factors such as place of articulation and following vowels. In designing her study on the effects of aspiration on F0 in Taiwanese, Lai (2004) took into account these factors and also took measurements of F0 onset, F0 at regular intervals, end F0, tonal duration, and VOT. Results from this study showed that aspiration has a raising effect on F0 onset and that this raising effect disappears around the midpoint of the tonal duration. Though VOT was not the main focus of this paper, Lai (2004) does mention that results from a correlation test show that longer VOT values tend to co-occur with lower F0 onsets. Her explanation of this is that while aspiration basically raises F0 due to faster airflow, the effect decreases over time as aspiration gets longer and as airflow decreases with time. Thus, longer VOT translates into a disappearance of the F0 raising effect that would otherwise occur with aspiration. Any study that looks at the possible effect of tone on VOT should also take into account the effect of aspiration on F0. Effects on F0 may simply be an automatic articulatory response, which should not vary between different languages. Xu (2005), however, would suggest that such effects are local and do not change the underlying F0 targets associated with tonal categories. Tone as an abstract phonological unit would thus be underlying and would be what would trigger any low-level changes in preceding consonants in the syllable in which they are found.

4. Cantonese Tone and Stop Inventory

Following the classification proposed by Lisker & Abramson (1964), Cantonese is a two-category tonal language that contrasts between short-lag and long-lag voiceless stops. The stop inventory includes pairs of bilabial, alveolar, and velar stops. One pair of alveolar affricates involving contrastive aspiration is also part of the consonantal inventory. All of these sounds are represented in IPA in Table 3.

	Bilabial	**Alveolar**	**Velar**	**Affricates**
Aspirated (long-lag)	pʰ	tʰ	kʰ	t͡sʰ
Unaspirated (short-lag)	p	t	k	t͡s

Table 3: Aspiration Contrasts in Cantonese

The tonal inventory of the language includes six contrasting tones. In Table 4 and in the rest of the paper, the Chao (1930) tone numbering system will be used. This system indicates relative pitch using a scale of 1-5 with 5 being the highest and 1 being the lowest. ‘q’ represents a checked tone, which occurs only in closed syllables. The checked tones 5q, 3q, and 2q occur in complementary distribution with the 55, 33, and 22 tones respectively and are essentially shortened versions of this latter group of tones. Though synchronically the checked tones are not separate tones, the traditional tonal classification scheme considers them separate to indicate the diachronic relationship of these tones.

	Sample Word CV(C) + Chao tone # (including alternate representations)	Tone Description	Chinese Character	Gloss
Yin (High Register)	si 55/53[2], 44[3]	High-level	詩	'poem'
	si 25, 35	Mid(high)-rising	史	'history'
	si 33, 44	Mid-level	試	'to try'
	sik 5q	High	識	'to know'
	sɛk 3q, 4q	Mid	錫	'to kiss'
Yang (Low Register)	si 21, 11, 22	Low-falling	時	'time'
	si 23, 13, 24	Low-rising	市	'city'
	si 22, 33	Low-level	事	'matter'
	sik 2q	Low	食	'to eat'

Table 4: Cantonese Tonal Inventory Chart (Traditional Classification) [4]

Though there are words pronounced as /si/ with all six contrastive tones, this is not the case for every possible CV(C) sequence in Cantonese. For instance, there are no words with the 22 tone that begin with aspirated stops. There are also no words beginning with unaspirated stops with the 21 and 23 tones. The few exceptions include onomatopoeic words and words formed through lexical reduplication. These inventory gaps can easily be explained by looking at how the current six tone system developed from a three tone system as shown in Table 1. Voiced stops became aspirated stops for two of the three Middle Chinese tone categories and unaspirated stops for the third tone category. This results in 5 possible tones for words starting with aspirated stops and 4 possible tones with words starting with unaspirated stops. The most important implication of this is that direct comparisons between aspirated and unaspirated stops can be

[2] The 55 and 53 tones have merged for most Cantonese speakers.

[3] Commas (,) that appear in this column are used to separate different representations of each tone category that have appeared in the literature. The Chao number representation for Cantonese tones that I adopt in the rest of this paper includes the first entry that appears for each category in this chart. Thus, they are 55, 25, 33, 21, 23, and 22 for the non-checked tones.

[4] Information in this table comes from several sources including Matthews & Yip (1994); Yip (2002); Yue-Hashimoto (1972).

made for a maximum of only 4 tones. It is also worth noting that Cantonese has been described as a sandhi-poor language with the few tonal changes that do occur being better described as "tone change" due to morphological processes rather than "tone sandhi" due to phonological processes (M. Y. Chen, 2000; Matthews & Yip, 1994). For this reason, tone sandhi is not considered a factor in this study.

Figure 1 is a graphic representation of the tonal contour inventory of one of the subjects. The chart includes only the four tone categories examined in this study. It is generally representative of all the subjects who participated in this study and suggests that the tones should be described as 55, 33, 25, and 21.

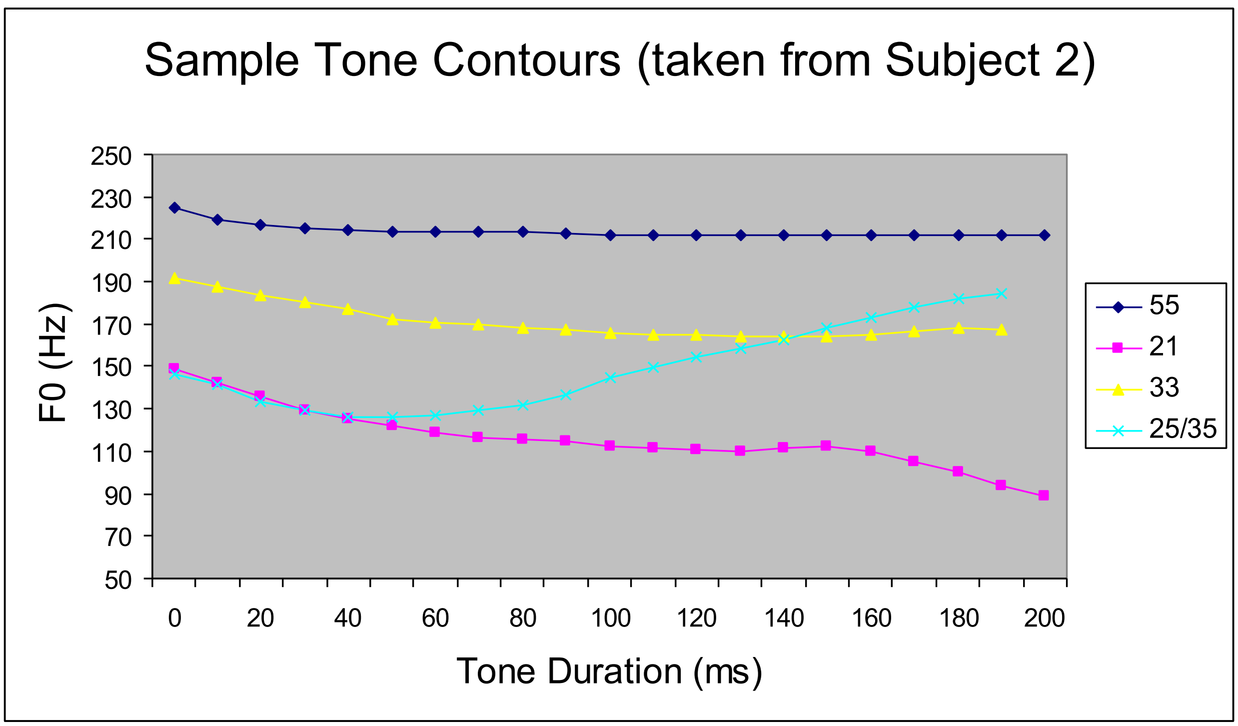

Figure 1: Sample Tone Contours

5. Methods/Procedures

5.1 Stimuli

Tone	pʰa	pa
55	趴地‘lie down on the floor’	巴士 ‘bus’
25/35	(豬)扒飯 ‘pork chop with rice’	把手 ‘handle’
33	怕怕 ‘scared’	霸位 ‘to hog seats’
21	爬山 ‘climb a mountain’	爸爸 ‘daddy’

Table 5: Experimental Stimuli

As can be seen in Table 5, all the stimuli have a rime of /-a/ and have bilabial onsets that contrast in aspiration across four tone categories. This controls for the effects of vowel and place of articulation.

5.2 Subjects

Six subjects (five male and one female), all native Cantonese speakers from Hong Kong who were in their early 20’s at the time of recording, were paid to participate in this study. All have lived in the United States for less than five years and were students at a Midwestern university. None of them reported any speech or hearing problems. Recordings of their speech were made using a solid-state recorder in a sound proof booth.

5.3 Procedures

Each subject was asked to read from a prepared list that included the 8 words in Table 5 as well as 16 additional words that were used for a separate study (Tse, 2005). Each of these words appeared 10 times in a randomized order and placed in a carrier phrase. A computer program automatically generated the list allowing each subject to read the phrase from a screen and to click the mouse when s/he was ready for the next word. Since written vernacular Cantonese is not a standardized language in the sense of it being accepted for use in official domains, it was important to have each subject read through the list before recording to make sure s/he knew how to pronounce each word correctly with the intended tone. For cases where ambiguity in meaning or pronunciation could potentially arise and for consistency, the target word was put in a contextual phrase XY so that it would be easier for subjects to correctly identify each word. For example, the character 鋪 (p^{h}ou) could either be pronounced with a 55 or with a 33 tone. The character that follows, 床 (t͡s^{h}ɔŋ) or 頭 (t^{h}ɐu), disambiguates the meaning and pronunciation of 鋪 (p^{h}ou). Thus, 鋪床 (p^{h}ou55 t͡s^{h}ɔŋ33) is pronounced with a 55 and means 'to make the bed' while 鋪頭 (p^{h}ou33 t^{h}ɐu25) is pronounced with a 33 and is a noun referring to 'a shop' or 'a store'. The contextual phrase always preceded the following carrier phrase with the first syllable in the contextual phrase (this was always the target word) inserted into the carrier sentence:

(1) XY, 我會讀 X 俾你聽
XY, ŋɔ23 wui23 tuk22 X pei25 nei23 t^{h}ɛŋ55. [5]
'XY, I will read X to you.'

The following is an example of a line of text that subjects were expected to read:

(2) 鋪床, 我會讀 鋪 俾你聽
p^{h}ou55 t͡sɔŋ33, ŋɔ23 wui23 tuk22 p^{h}ou55 bei25 nei23 t^{h}ɛŋ55
'XY, I will read X to you.' (X = p^{h}ou55)

---

[5] The phonetic pronunciation spelled out here is actually a conservative pronunciation of the phrase. In recent years, Hong Kong Cantonese has experienced a series of sound changes including the loss of the velar nasal onset, /ŋ/ > Ø or /ŋ/ > /ʔ/ for some speakers, the merger of the alveolar nasal and liquid /n/ > /l/ onsets, and alveolarization of velar codas, /ŋ/ > /n/ and /k/ > /t/ (Zee, 1999). All six subjects participated in these sound changes to varying degrees. Subjects who were the most advanced in these changes would have pronounced the carrier phrase as ʔɔ23 wui23 tut22 X pei25 lei23 t^{h}ɛn55.

The first element underlined is the target word placed in a contextual phrase XY. The target word, X, is then placed into the carrier phrase with the same tone as in the context phrase. A total of 200 tokens were recorded for each speaker adding up to a grand total of 1200 tokens. This paper reports on an analysis of a subset of these tokens. The subset to be examined contains 80 tokens per speaker for a total of 480 tokens total for all 6 subjects.

5.4 Measurements

Analysis first involved segmenting all 1200 tokens using the software PRAAT (Boersma & Weenink, 2005). An example of segmentation to determine the VOT boundaries is shown in Figure 2 below.

Sudden vertical bursts in the waveform were generally used to mark the boundary for the beginning of both aspirated and unaspirated stops. The end of VOT was determined by finding the point on the time-axis where periodicity clearly begins. In addition to VOT, vowel duration was also measured so that word duration could be determined from the sum of VOT and vowel duration.

Following segmentation of both VOT and vowel duration, a script was run on the segmented files that automatically calculated the duration of the VOT and vowel duration boundaries defined for each token. For the final stage of analysis, the ratio of VOT/Word duration was used rather than absolute VOT values so that the effects of speaking rate could be controlled. F0 values were also measured at regular intervals of 10% of the duration of each vowel by using a separate script on the same files[6].

The statistical analysis tested two hypotheses:

1) Tone category affects VOT.

2) The height of the F0 onset affects VOT.

[6] I would like to thank Mi-Ch'i Chen (陳米琪) and Yuwen Lai for assisting me with this task.

To address the first hypothesis, a one-way ANOVA test was run to determine if there is a significant effect of tone category on VOT values. A significant effect could either mean that VOT can act as a secondary cue for distinguishing between different tone categories or it could mean that this effect is due to F0 rather than tone category, which would be the second hypothesis. To address the second possibility, a Pearson correlation test was run between F0 onset height and VOT.

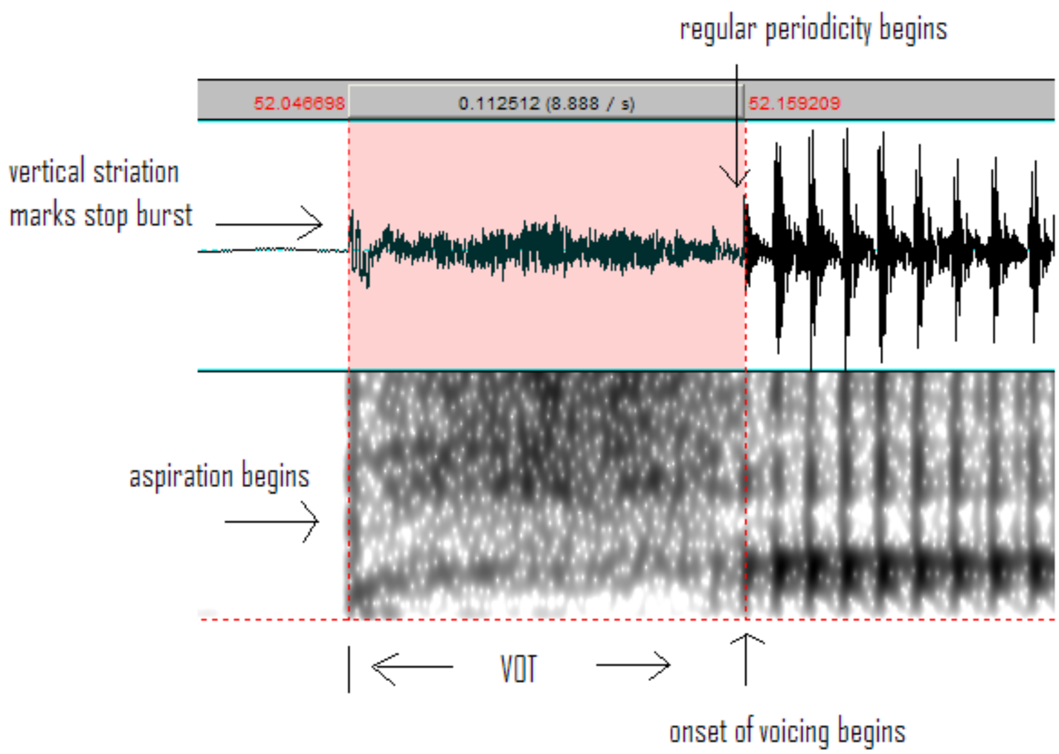

Figure 2: An Example of Segmentation on PRAAT

6. Results

Tables 6 and 7 include all the average values of the VOT/Word Duration ratio.

Tone	VOT/Word Duration Average (in seconds)	Std. Deviation	VOT (in seconds)	Std. Deviation	N
55	0.2487	0.06648	0.06270571	0.19569242	60
33	0.2619	0.06912	0.06577756	0.02152983	60
25	0.3002	0.06110	0.07716892	0.019320273	60
21	0.3240	0.07782	0.078581	0.023467483	60
All Tones	0.2837	0.07476	0.070105844	0.022026429	240

Table 6: VOT/Word Duration Averages for Aspirated Stops

Tone	Average (in seconds)	Std. Deviation	N
55	0.0551	0.02382	60
33	0.0555	0.02834	60
25	0.0602	0.03353	60
21	0.0610	0.03357	60
All Tones	0.0580	0.03002	240

Table 7: VOT/Word Duration Averages for Unaspirated Stops

Table 8 below shows the results from the ANOVA test for the effect of tone on the VOT/Word Duration ratio of aspirated stops while Table 9 shows the results for unaspirated stops. As shown in Table 8, tone has a significant effect on VOT/Word Duration for all subjects except for one. Table 9 shows that the effect is not significant for unaspirated stops with the exception of one subject. Results from a post-hoc test to determine which tone categories are significantly different from each other for aspirated stops are shown in Table 10. These results show that there is no significant difference between the 55 and the 33 tones and between the 25 and the 21 tones. All other possible pairs, however, do show significant differences in VOT. If we were to organize all the tone categories on a scale that corresponds to the different VOT values associated with each tone, we would get the following: 21, 25 > 33, 55.

To better illustrate the effects of different tone categories, boxplots showing the range of VOT/Word Duration values for each individual subject are shown below. These boxplots (Figures 3 through 9) show that although 5 out of 6 speakers show a significant effect of tone, there is quite a bit of variation in how different tonal categories are grouped for different speakers. The most consistent pattern seems to be that the 21 tone occurs with longer VOT.

Some speakers have VOT values associated with the 25 tone that are in a similar range as for the 21 tone while for others, this relationship does not exist.

Subject	p value	significance
1 (male)	< 0.000	*
2 (male)	0.001	*
3 (male)	< 0.000	*
4 (male)	0.003	*
5 (male)	0.02	*
6 (female)	0.469	n.s.
All Subjects	< 0.000	*

Table 8: ANOVA results (long-lag stops)

Subject	p value	significance
1 (male)	0.068	n.s.
2 (male)	0.470	n.s.
3 (male)	0.183	n.s.
4 (male)	0.462	n.s.
5 (male)	0.023	*
6 (female)	0.223	n.s.
All Subjects	0.833	n.s.

Table 9: ANOVA results (short-lag stops)

Tonal Pairs	**p =**	**Significance**
55 & 33	0.215	n.s.
25 & 21	0.059	n.s.
55 & 25	< 0.000	*
55 & 21	< 0.000	*
33 & 25	0.002	*
33 & 21	< 0.000	*

Table 10: Post-Hoc Analysis for Aspirated Stops [F = (3, 235) = 15.966, p < 0.000]

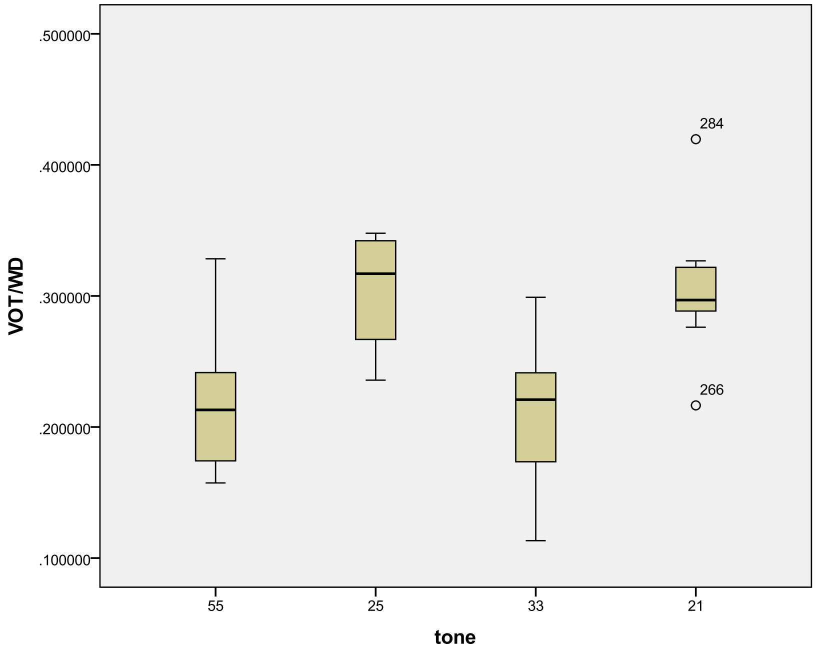

Figure 3: Subject 1 ANOVA (aspirated stops only)

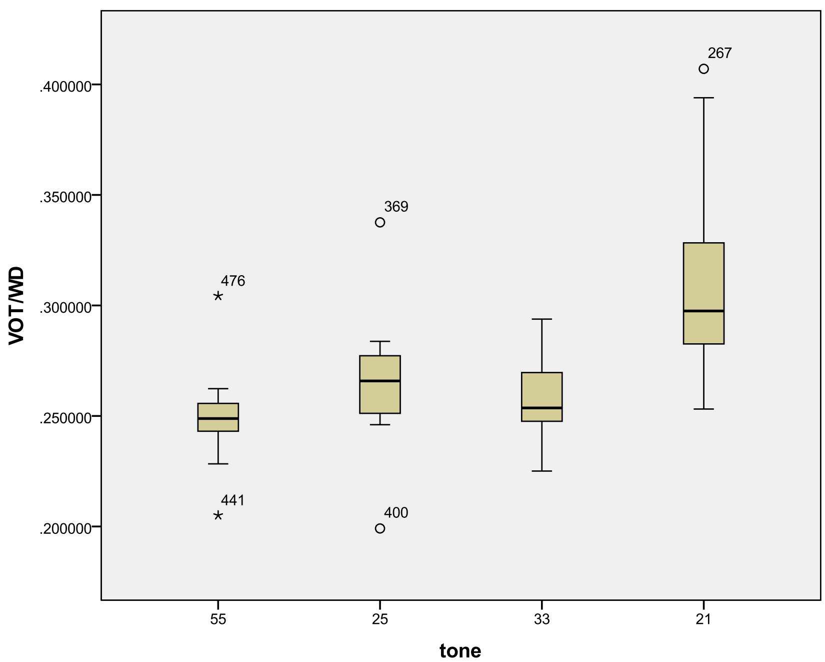

Figure 4: Subject 2 ANOVA (aspirated stops only)

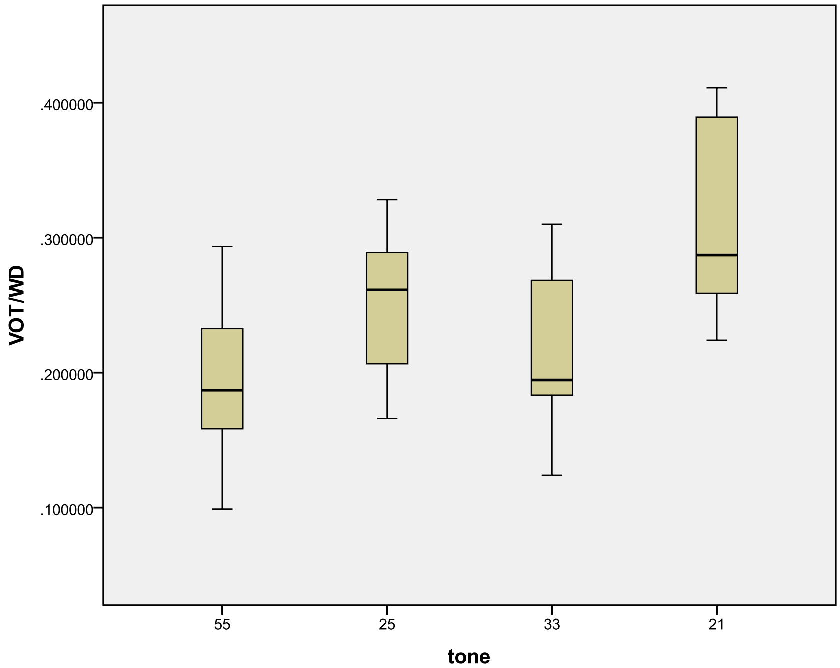

Figure 5: Subject 3 ANOVA (Aspirated stops only)

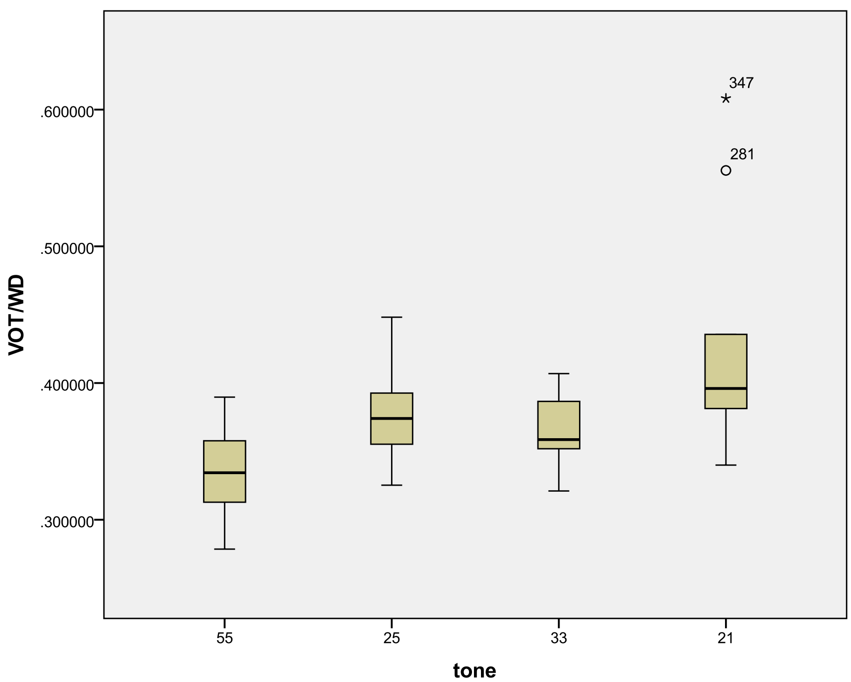

Figure 6: Subject 4 ANOVA (Aspirated stops only)

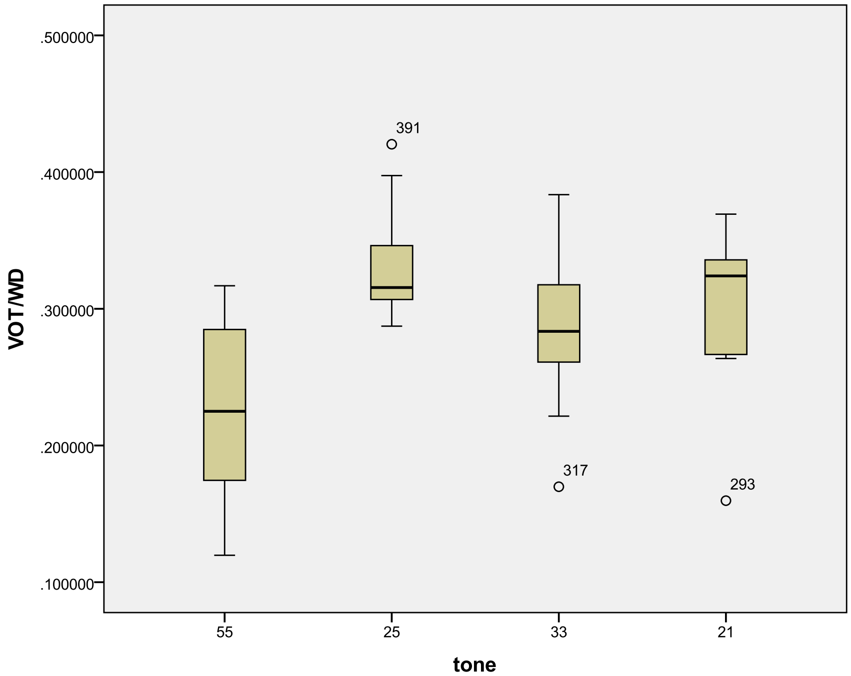

Figure 7: Subject 5 ANOVA (aspirated stops only)

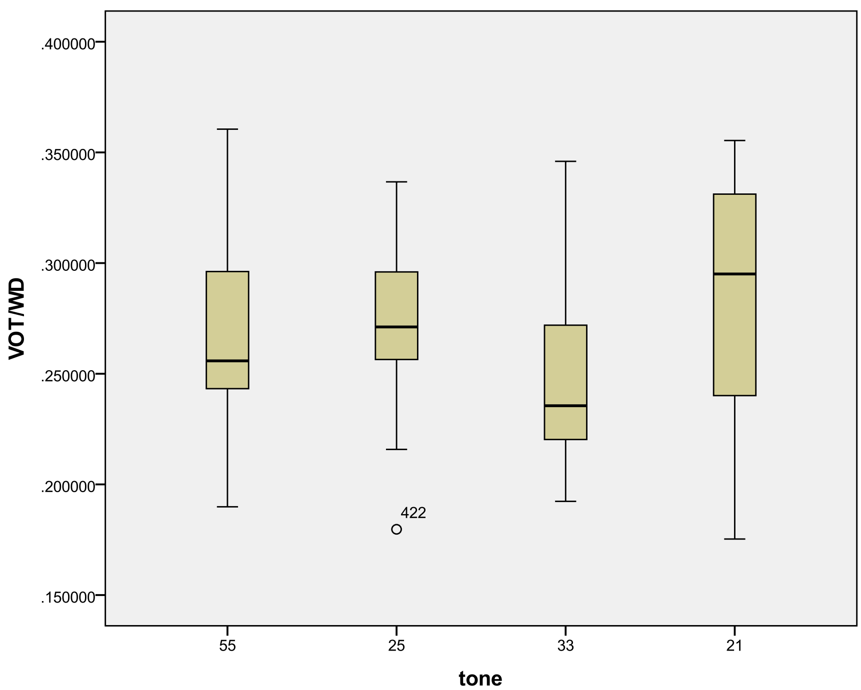

Figure 8: Subject 6 ANOVA (aspirated stops only)

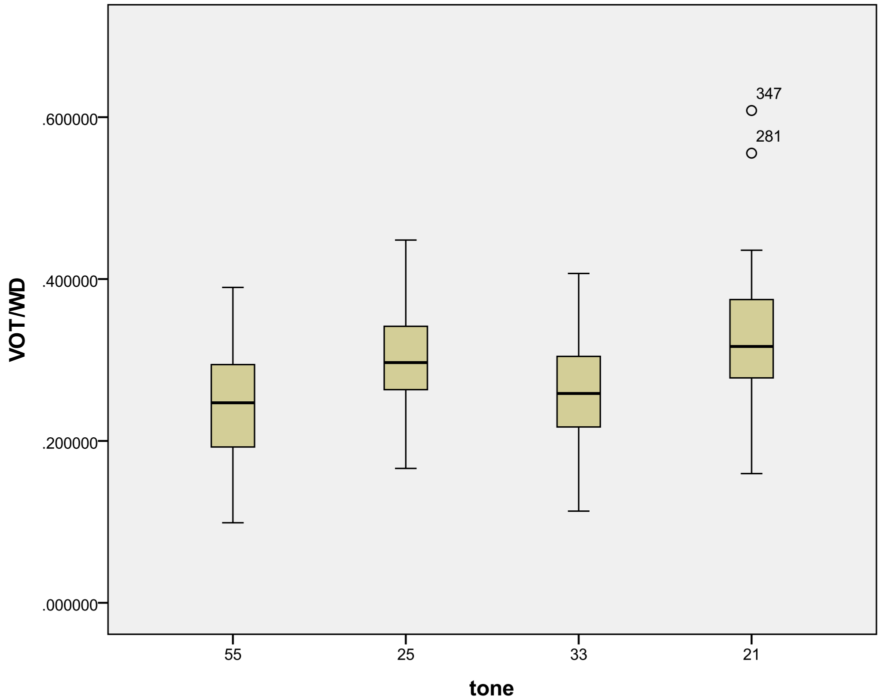

Figure 9: All Subjects ANOVA (aspirated stops only)

Tables 11 and 12 show the results of the Pearson Correlation test between F0 onset and VOT/Word Duration. For aspirated stops, there is a significant correlation if all of the subjects are pooled together. When looking at individual subjects, however, only 3 out of the 6 subjects showed a significant correlation. This is fewer than the number that showed a significant effect of tone based on the ANOVA test. Unaspirated stops showed no correlation with the exception of the individual results of one subject. A scatterplot for each individual subject that showed a significant correlation for aspirated stops is also included below.

Subject	Pearson Coefficient	p value	N	significance
1 (male)	-0.053	0.744	40	n.s.
2 (male)	-0.326	0.040	40	*
3 (male)	-0.350	0.027	40	*
4 (male)	0.007	0.963	40	n.s.
5 (male)	-0.313	0.049	40	*
6 (female)	0.129	0.426	40	n.s.
All Subjects	-0.210	0.01	240	*

Table 11: Correlation Test Results (Aspirated Stops) VOT ratio

Subject	Pearson Coefficient	p value	significance
1 (male)	-0.150	0.357	n.s.
2 (male)	-0.151	0.354	n.s.
3 (male)	-0.187	0.248	n.s.
4 (male)	0.329	0.036	*
5 (male)	-0.207	0.200	n.s.
6 (female)	-0.007	0.965	n.s.
All Subjects	0.018	0.778	n.s.

Table 12: Correlation Test Results (Unaspirated Stops)

VOT values

Subject	Pearson Coefficient	p value	N	significance
1 (male)	-0.114	0.484	40	n.s.
2 (male)	-0.477	0.002	40	*
3 (male)	-0.195	0.229	40	n.s.
4 (male)	-0.228	0.158	40	n.s.
5 (male)	-0.312	0.050	40	n.s.
6 (female)	0.288	0.071	40	n.s.
All Subjects	-0.075	0.249	240	n.s.

Table 13: Correlation Test Results (Aspirated Stops) with Raw VOT Values

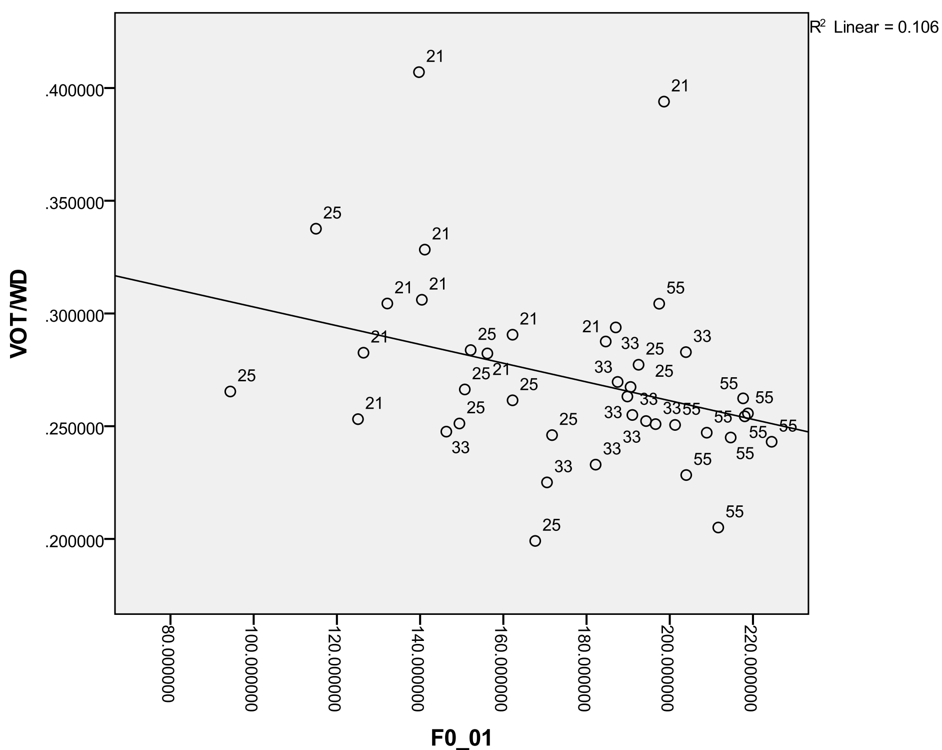

Figure 10: Subject 2 Correlation for Aspirated Stops

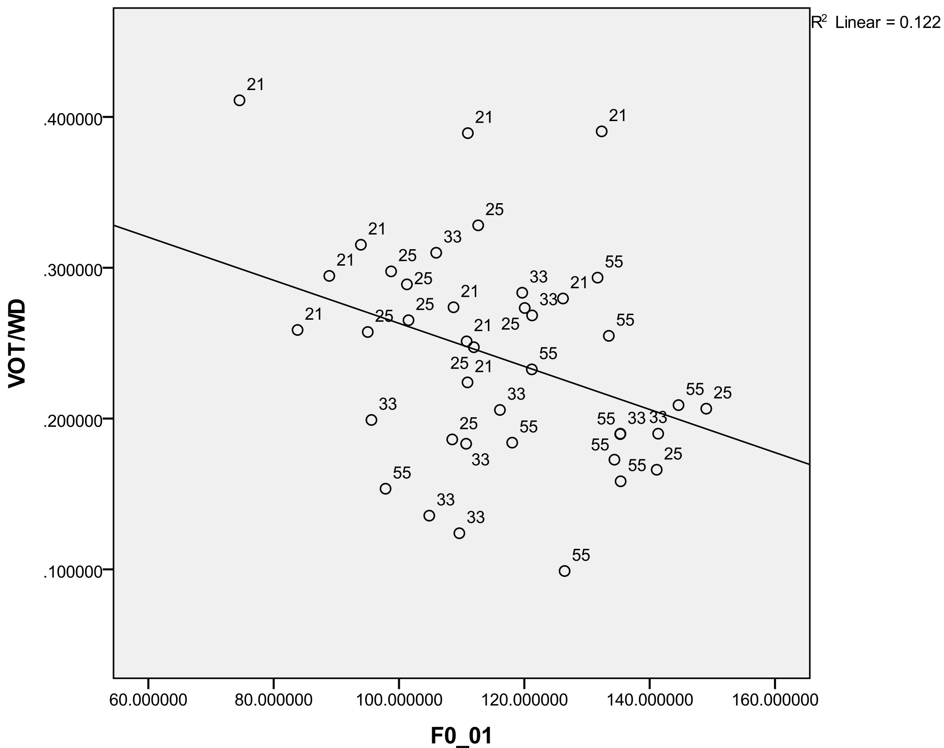

Figure 11: Subject 3 Correlation for aspirated stops

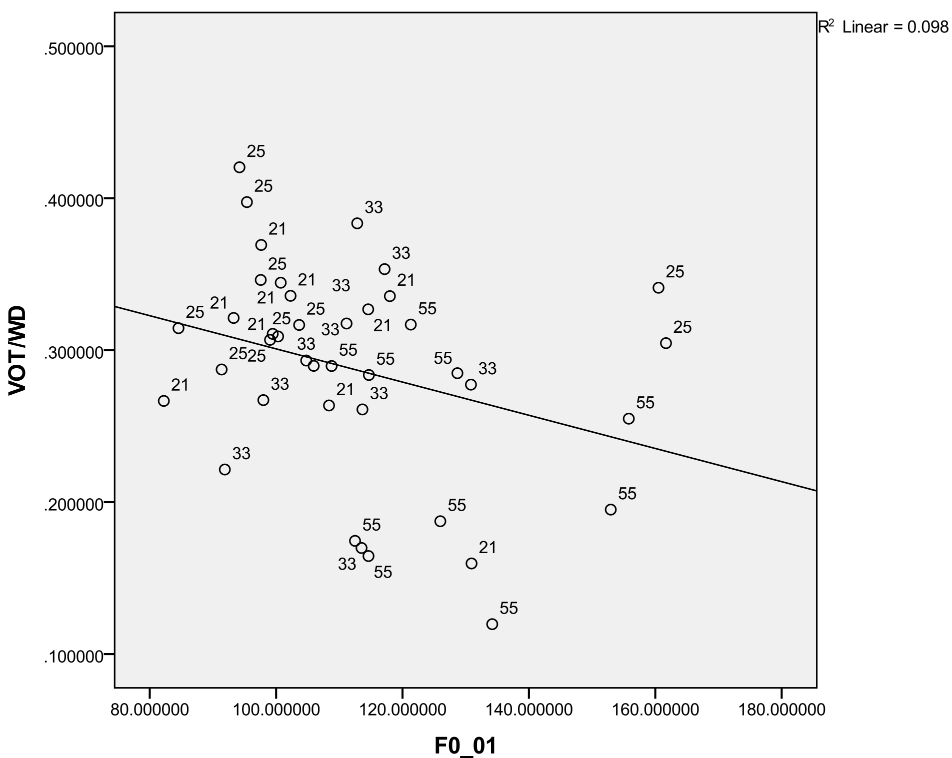

Figure 12: Subject 5 correlation all aspirated

7. Discussion

Based on the results of the experiment presented in this paper, we can address the three questions presented at the beginning of this paper.

(1) Does tone have an effect on VOT in Cantonese?

The answer to this question is "yes" for at least some speakers. The overall results from the ANOVA test on all six subjects combined showed that there is a significant effect of tone on the VOT/Word Duration values. When the ANOVA test was run on each individual subject, five out of the six subjects also showed a significant effect. The one subject that did not show significant differences was also the only female speaker in this study.

More female subjects would be needed to determine whether this individual speaker is simply an outlier or whether there is a general pattern of female speakers not making use of VOT to differentiate tonal categories. One possibility is that the wider pitch range of female speakers may make additional cues to distinguish between different tones less needed than for male speakers.

(2) If so, what kind of an effect does it have?

The results show that the way VOT is affected is quite variable. With all of the subjects aggregated together, there is a two-way split in tone categories based on VOT values. In one group are the 21 and 25 tones, which have longer VOT than in the other group, which includes the 55 and the 33 tones. Subject 1 shows the same kind of split. The other subjects, however, seem to show different kinds of splits. Subjects 2, 3, and 4, for example show a split with the 21 tone in one group and the other three tones in a separate group. This split also happens to coincide with the Yin and Yang tone register distinction. The 21 would belong in the Yang (low) register while the others would belong in the Yin (high) register, following traditional Chinese tone classification schemes. As mentioned in the literature review, the Yang register is associated with historically voiced stops and hence low VOT values. Interestingly, the results here show this tone category corresponding to the highest VOT values, which is the opposite of what would be expected.

Finally, Subject 5 showed yet a different pattern. For this subject, the 55 tone forms one group while the other three tones form another group. The 55 tone is associated with shorter VOT/Word Duration than the other tones. Though the 5 male subjects show different kinds of splits, one general pattern that seems to hold across all of them is that the 55 and the 21 tones are the most distinctly different from each other. This suggests that perhaps the relationship between VOT and tone may actually be an inverse correlation between VOT and F0 onset. This leads to the next research question of this paper.

(3) Is this effect purely an automatic articulatory consequence of F0 modulation or is this effect also mediated by lexical tone and hence a secondary cue that contributes towards maintaining phonological contrasts between different tonal categories?

To address this question, the results from both the correlation test between VOT and F0 and the ANOVA test of the effect of tone on VOT need to be looked at. A significant correlation between VOT and F0 would support an automatic articulatory effect account, while significance found through ANOVA tests would support a phonological cue account. The results show that when all of the speakers are aggregated together, there is in fact an inverse correlation between F0 height and VOT values. When looking at each subject individually, however, this inverse correlation holds for only 3 out of the 6 subjects. In contrast, all of the male subjects showed significant differences on the ANOVA test of tone categories. The fact that the ANOVA test yielded significant results for more subjects than did the correlation test seems to lend more support to a phonological cue account to explain the longer VOT with the 21 tone and the shorter VOT for the 55 tone. Although individual speakers showed different kinds of tone category splits, the fact that there is such variation and the fact that individual speakers do seem able to create organized patterns would suggest that VOT differences have the potential to be phonologized. If these VOT differences were simply about automatic articulatory processes, there would have been more consistency in the VOT and F0 correlation tests.

8. Conclusion

To conclude, this study has shown that tone has a stronger effect on VOT than F0 onset in Cantonese. Speakers, thus, have the ability to phonologize VOT differences as secondary cues to tonal distinctions. The fact that there is variability among the six speakers in this study suggests that the effects are not only language specific but also individually specific. Potential F0 onset effects on VOT should be viewed as a local perturbation effect that does not change the status of tone as the triggering mechanism. There may be other cues as well. For example, one next step in this research would be to look at why there is a correspondence between low tone and longer VOT in Cantonese. What else could be triggering the consonant lengthening other than tone itself? One possibility is phonation type. Impressionistically, the male speakers in this study seemed to be more likely to pronounce the 21 tone with creaky voice phonation. If this turns out to be true, it may explain why the one female speaker did not show any statistical significance. A follow-up study would look to see to what extent phonation plays a role in affecting VOT values.

Acknowledgements
This working paper is a revised version of my MA thesis. In preparing this revised version, I would like to thank Mi-Ch'i Chen (陳米琪) and Yuwen Lai for helping me extract F0 measurements and David Mortensen for his advising and support.

References

Boersma, P., & Weenink, D. (2005). Praat: doing phonetics by computer [Computer program] (Version 4.3.27). Retrieved from http://www.praat.org/

Boucher, V. J. (2002). Timing relations in speech and the identification of voice-onset times: A stable perceptual boundary for voicing categories across speaking rates. *Perception & Psychophysics*, *64*(1), 121–130.

Bradshaw, M. M. (1999). *A Crosslinguistic Study of Consonant-Tone Interaction* (Ph.D. Dissertation). The Ohio State University, Columbus, OH.

Chao, Y. R. (1930). A system of tone letters. *Le Maître Phonétique*, *45*, 24–27.

Chen, L., Peng, J.-F., & Chao, K.-Y. (2009). The effect of lexical tones on voice onset time. In *11th IEEE International Symposium on Multimedia, 2009. ISM'09.* (pp. 552–557). IEEE.

Chen, M. Y. (2000). *Tone Sandhi: Patterns across Chinese Dialects*. Cambridge: Cambridge University Press.

Cho, T. (2003). Lexical Stress, Phrasal Accent and Prosodic Boundary in the Realization of Domain-initial Stops in Dutch. In *Proceedings of 15th International Congress of Phonetic Sciences. Barcelona* (pp. 3–9).

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

Ewan, W. G. (1976). *Laryngeal behavior in speech* (Ph.D. Dissertation). University of California, Berkeley, Berkeley, CA.

Gandour, J. (1974). Consonant types and tone in Siamese. *Journal of Phonetics*, *2*, 337–350.

Gandour, J., & Maddieson, I. (1976). Measuring larynx movement in Standard Thai using the cricothyrometer. *Phonetica*, *33*(4), 241–267.

Haudricourt, A.-G. (1954). De l'origine des tons en vietnamien. *Journal Asiatique*, *24*, 69–82.

Haudricourt, A.-G. (1972). Two-way and three-way splitting of tonal systems in some Far Eastern languages (Translated by Christopher Court). *Tai Phonetics and Phonology*, 58–86.

Herrera Z., E. (2003). Tono, VOT y sonoridad en el mazateco de Santa Clara, Oaxaca. In E. Herrera Z. & P. M. Butragueño (Eds.), *La tonía: dimensiones fonéticas y fonológicas* (Vol. 74). Mexico City: El Colegio de México Centro de Estudios Lingüísticos y Literarios.

Hombert, J.-M. (1975). *Towards a theory of tonogenesis: an empirical, physiologically and perceptually based account of the development of tonal contrasts in languages* (Ph.D. Dissertation). University of California, Berkeley, Berkeley, CA.

Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In V. Fromkin (Ed.), *Tone: A Linguistic Survey* (pp. 77–112). London: Academic Press.

Hombert, J.-M., & Ladefoged, P. (1977). The Effect of Aspiration on the Fundamental Frequency of the Following Vowel. *UCLA Working Papers in Phonetics*, *36*, 33–40.

Hombert, J.-M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic Explanations for the Development of Tones. *Language*, *55*(1), 37–58.

Hyman, L. (1976). On some controversial questions in the study of consonant types and tone. *UCLA Working Papers in Phonetics: Studies on Perception and Production of Tone*, 90–98.

Hyman, L., & Schuh, R. (1974). Universals of Tone Rules: Evidence from West Africa. *Linguistic Inquiry*, *5*(1), 81–115.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, 286–319.

Kessinger, R. H., & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics*, *25*(2), 143–168.

King, L., & Schiefer, L. (1990). Tone Effect on Voice-Onset Time (VOT)?-Results from Shanghai. *Institut Fur Phonetik Und Sprachliche Kommunikation Der Universitat Munchen - Forschungsberichte*, *28*, 253–295.

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech, Language, and Hearing Research*, *18*(4), 686–706.

Lai, Y. (2004). *The effect of aspiration on fundamental frequency in Taiwanese syllables* (M.A. Thesis). University of Kansas, Lawrence, KS.

Lisker, L., & Abramson, A. (1964). A cross-linguistic study of voicing in initial stops: acoustical measurements. *Word*, *20*, 384–422.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*, *10*(1), 1–28.

Liu, H., Ng, M. L., Wan, M., Wang, S., & Zhang, Y. (2008). The effect of tonal changes on voice onset time in Mandarin esophageal speech. *Journal of Voice*, *22*(2), 210–218.

Maddieson, I. (1974). A note on tone and consonants. *The Tone Tome: Studies on Tone from the UCLA Tone Project (UCLA Working Papers in Phonetics)*, *27*, 18–27.

Maddieson, I. (1976). A further note on tone and consonants. *UCLA Working Papers in Phonetics*, *33*, 131–159.

Maddieson, I. (1977). Tone Effects on Consonants. *UCLA Working Papers in Phonetics*, *36*(July), 91–110.

Matthews, S., & Yip, V. (1994). *Cantonese: a Comprehensive Grammar*. London and New York: Routledge.

Pearce, M. (2005). Kera Tone and Voicing. *UCL Working Papers in Linguistics (UCLWPL)*, *17*, 61–82.

Pearce, M. (2009). Kera tone and voicing interaction. *Lingua*, *119*(6), 846–864.

Summerfield, Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology. Human Perception and Performance*, *7*(5), 1074–1095.

Tse, H. (2005). *The Phonetics of VOT and Tone Interaction in Cantonese* (M.A. Thesis). University of Chicago, Chicago, IL.

Van Dam, M. (2003). VOT of American English stops with prosodic correlates. *The Journal of the Acoustical Society of America*, *113*(4), 2328–2328. http://doi.org/10.1121/1.4780819

Xu, C. X., & Xu, Y. (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association*, *33*(2), 165–181.

Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication*, *46*(3), 220–251.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

Yue-Hashimoto, O. (1972). *Phonology of Cantonese.* Cambridge: Cambridge University Press.

Zee, E. (1977). The effect of F0 on the duration of [s]. *The Journal of the Acoustical Society of America*, *61*(S1), S30–S30. http://doi.org/10.1121/1.2015558

Zee, E. (1980). The Effect of Aspiration on the F0 of the Following Vowel in Cantonese. *UCLA Working Papers in Phonetics*, *49*, 90–97.

Zee, E. (1999). Change and Variation in the Syllable-Initial and Syllable-Final Consonants in Hong Kong Cantonese. *Journal of Chinese Linguistics*, *27*(1), 120–167.