Effect of Tone on Vocal Attack Time in Cantonese-speaking Children

Lui Pui Ching

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2014.

Effect of Tone on Vocal Attack Time in Cantonese-speaking Children

Lui Pui Ching

Abstract

Cantonese tones were shown to have significant effect on vocal attack time (VAT) in adult Cantonese speakers, with males produced greater VAT values than females (Ma et al., 2012). The present study aims to investigate the effect of tone on VAT in Cantonese-speaking children. Sound pressure (SP) and electroglottographic (EGG) recordings were collected from 55 native Cantonese-speaking children. 26 six-year-old and 29 nine-year-old children were asked to read aloud six monosyllabic or disyllabic words which contained all the Cantonese tones. One word was presented at the same time and children were asked to read the word immediately after the presentation. Results revealed significant differences between some contour tone (tone 2 and tone 4) and level tone (tone 1) pairs. Age and gender showed no significant effect on VAT values. Children demonstrated a clear different VAT profile compared with Cantonese adult speakers. The results support the idea that contour tones require more complicated pre-phonatory laryngeal settings. Different VAT patterns between children and adults suggest that they adopt different laryngeal adjustment strategies during phonation onset.

Introduction

Tone as a Linguistic Feature in Cantonese

Tone languages are spoken by over half of the world's population (Yip, 2002). Cantonese, which is the predominant language in Hong Kong, is one of the tone languages. In Cantonese, lexical meaning of a monosyllabic word is determined by tone, which is the perceptual correlate of pitch (Khouw & Ciocca, 2007). Therefore, any change in pitch could affect the meaning of a syllable. Cantonese contains six contrastive tones, in which tones 1, 3 and 6 have relatively steady pitch level within the vocalic segment and are referred to as level tone. In contrast, tones 2, 4, 5 are characterized by a change in pitch contour across the vocalic segment and are referred to as contour tone.

Overview of Vocal Attack Time

Baken and Orlikoff (1998) developed a noninvasive technique to evaluate vocal attack by comparing sound pressure (SP) and electroglottographic (EGG) signal during phonation onset. The underlying principle is that, due to the presence of transglottal airflow, the vocal folds begin to oscillate with small amplitude long before vocal fold contact is achieved. These small amplitude changes at the borders of the glottis lead to a weak SP signal, which grows and attains its maximum upon the completion of vocal fold contact. On the other hand, EGG signals are associated with the relative vocal fold contact area. No significant growth would be observed until vocal fold adduction is accomplished. Therefore, there is always a time lag between the onset of the SP and EGG signals. This time lag is termed as vocal attack time (VAT).

Tone and Vocal Attack Time

Tone is the acoustic correlate of fundamental frequency and the physiological correlate of the rate of vocal fold vibration. Different tones require different laryngeal settings, so vocal folds have to be adjusted to produce a specific tone. Ma and colleagues (2012) examined the relationship between VAT and Cantonese tones in 59 adult speakers. Their study revealed that contour tones (tone 2, 4, 5) yielded significantly greater VAT values than level tones (tone 1, 3, 6). The authors attributed this finding to two possibilities. The first one was that the production of contour tones required more complicated laryngeal adjustment than the production of level tones. Another possibility was the learned feature of the Cantonese tones. Moreover, their study revealed a gender difference, with males producing significantly greater VAT value than that of females. This finding was suggested to be due to the anatomical difference between the vocal folds of females and males. This study was the first study to examine the relationship between tone and VAT. However, it was limited to the adult population. Whether similar patterns can be found in Cantonese-speaking children has yet to be proved. Previous research has shown that 90% of children can achieve an accurate production of Cantonese tones by the age of two-and-a-half years old (Cheung, Ng & To, 2006). Although children produce tones that are perceptually the same as that of adults, the pre-phonatory vocal fold settings in producing various tones remains unknown. Investigation into the VAT pattern in children helps understanding the similarity and difference between children and adult in their underlying laryngeal physiology at the pre-phonatory phase.

Developmental changes of laryngeal structures occur as children grow up. These changes include increase in laryngeal size and vocal fold length. These developmental changes of laryngeal structures do not complete until 16 to 20 years of age (Hirano, Kurita, & Nakashima, 1983). As children approach puberty, their laryngeal size and vocal fold characteristics will be of more resemblance to that of adults. Therefore, the functioning of their vocal folds during the pre-phonatory phase is possible to differentiate throughout the school years. Because of the use of different mechanisms in producing tones, different VAT values within the children population is speculated. This leads to the hypothesis that pre-phonatory adjustment mechanism may change across age, leading to different profile of VAT values.

Gender effects on VAT were revealed in previous studies (Ma et al., 2012; Roark, Watson, Baken, Brown & Thomas, 2010). This gender effect may lie in the fact that the laryngeal sizes of males are greater and their vocal folds are longer and more elastic in nature. Since the sizes of larynx and laryngeal structures are similar between boys and girls prior to puberty (Fitch & Giedd, 1999), a gender effect on VAT may not be significant in young children.

Aim of Present Study

The aim of the present study was to investigate the effects of tones on VAT in young Cantonese-speaking children. Children of six-year-old and nine-year-old were recruited. It was hypothesized that, first, the VAT values of contour tones would be higher than that of level tones due to more complex laryngeal adjustment required by contour tones. Second, six-year-old children and nine year-old children would yield different VAT values for the same tone due to the physiological maturation in the laryngeal structures. The last hypothesis was that gender effect would not exist because boys and girls have similar laryngeal structures.

Results from this study would contribute to a better understanding of the physiological change of vocal folds in the process of vocal fold maturation. Comparison to adults' VAT data can also be done and the pattern of VAT in the two categories of tones can be verified.

Method

Participants

Fifty children, including 26 six-year-old (8 girls, 18 boys) and 29 nine-year-old (11 girls, 18 boys) native Cantonese-speaking children participated in the study. None of the children had a prior history of voice, speech, language or hearing problems. In order to ensure children's good health on the day of testing, their parents were asked to complete a health questionnaire for their children before the study.

Equipment

Data was collected using an *M-Audio Fast-Track interface* (Avid Technology, Irwindale, CA) at a sampling rate of 44.1 kHz per channel and a resolution of 32 bits. Sound pressure signals were recorded by a *Radio Shack 33-3012 headset* (Fort Worth, TX) with a fixed distance of 5cm from the labial commissure maintained throughout the data recording. To measure the vocal fold contact area, EGG was used and its data was collected by a digital *FourcinLaryngograph Processor* (Laryngograph Ltd, London, UK). The recording and storage of signals was managed by *Audacity* (version 1.2.6).

Stimuli

A pilot study was carried out to evaluate the linguistic appropriateness of stimuli for eliciting VAT in young children. Five children aged six and nine were recruited. The same set of stimulus word was adopted from a similar study of Ma et al. (2012) on Cantonese-speaking adults. The word list had 12 disyllabic words, which consisted of six pairs of homophones. Each pair represented one Cantonese tone, with the syllable /a/ at the word-initial position (Appendix

I). However, the pilot study showed that most of the words in the word list were inappropriate owing to the lower literacy level of young children.

To ensure a natural and accurate production of all stimulus words, the word list was then modified to suit the linguistic level of children. Only words that were found to be more familiar to children in the pilot study were kept. Extra monosyllabic words with initial consonant /f/, /s/, /h/, /l/ or /m/ were added to the list (Appendix II). The modified list contained 16 monosyllabic or disyllabic words, all of which contained the same vowel /a/. This was to ensure all stimulus words in the list were real words under the same vowel environment. The initial consonants chosen were associated with the least laryngeal preset adjustment by a natural production. The amended word list was tested by another eight children aged at six or nine years old. Among all stimuli, words with initial consonants /s/ and /l/ showed frequent rejection due to the failure to meet the Figure of Merit (FOM) ≥ 0.75 acceptability criteron of VAT values (Roark, Watson & Baken, 2012). Based on the familiarity of the words and the effect of initial consonant on VAT, a word for each tone was chosen from the 16 words to make up the finalized word list.

The finalized set of stimulus words consisted of four monosyllabic words and two disyllabic words (Table 1). Only one word was chosen for each individual tone. This aimed to reduce the memory load of the children and ensure an accurate. An overall greater number of trials were performed to obtain a reliable average VAT value of each tone.

Table 1

The six Cantonese tones and the associated tone characteristics

Tone	Pitch	Contour	Word	Lexical Meaning
1	High	Level	呀 /a1/	An interjection (used as an exclamation to pain, joy, surprise, etc.)
2	High	Rising	啞 /a2/	Dumb
3	Mid	Level	亞洲 /a3 tsɐu1/	Asia
4	Low	Falling	牙齒 /a4 ts^hi5/	Teeth
5	Low	Rising	馬 /ma5/	Horse
6	Low	Level	下 /ha6/	Down

Procedure

Recordings were made in quiet environments with background noise kept below 30dBA. Pictures and written forms of the stimuli were prepared on six 9cm x 8cm note cards. The experimenter shuffled the note cards and then turned over one card and faced it in front of the subjects. The subjects were instructed to read the words at a comfortable pitch, loudness and rate. They were specifically instructed to read each word immediately following the presentation of the words. This was to ensure a natural production of all the words without presetting of laryngeal structures. The task was repeated ten times with the cards presented in a randomized order. The children were introduced with all the words before the actual recording with the assistance of the pictures until they became familiar with all the stimuli.

Data Analysis

In the present study, VAT computation method followed the procedures described in the normative study reported by Roark et al. (2012). The mean VAT of each tone for each participant was calculated. Sixty VAT data were subjected to analysis for each participant (6 tones x 10 trials). Of the (6 tones x 10 trials x 55 subjects=) 3300 data points, 592 (17.9%) were rejected because of any one of the following reasons:

1. Failed to meet the Figure of Merit (FOM) ≥ 0.75 acceptability criterion (Roark, Watson & Baken, 2012)
2. Could not be analyzed by the VAT analysis program due to either insufficient SP or EGG amplitude or unconscious SP or EGG signal generated by the subjects before word onset
3. Outliers identified by plotting boxplot graph

To ensure the reliability of the result, subjects with tokens of any one of the tones being rejected for more than 50% were further excluded. Following the rejection, the remaining analysis was based on data from 15 six-year-old participants (4 females, 11 males) and 20 nine-year-old participants (8 females, 12 males). There were altogether 2100 data points, with 280 (13.3%) being rejected because of one of the abovementioned conditions. The remaining (2100-280=) 1820 tokens were entered into the statistical analysis.

Result

The mean VAT values were subjected to a three-way mixed analysis of variance involving one within-subject variable and two between-subject variables. The within-subject variable was tone (6 lexical tones) and the between-subject variables were gender (female versus male) and age (six years old versus nine years old). Table 2 showed the information concerning the body size of the two age groups of children. Independent samples t-test was carried out to compare the physique of the two age groups. The mean height, weight and head circumference of the two groups were summarized in Table 2. The differences were all significance (height: $t(33) = -8.85$,

$p < .05$; weight: $t(33) = -4.79$, $p < .05$; head circumference: $t(23.71) = -2.65$, $p < .05$). Table 3 summarized the means and SD of VAT as a function of tone, age and gender.

Table 2

Mean height, weight and head circumference (and standard deviations) of six-year-old and nine-year-old groups

Mean (standard deviation)	Six-year-old	Nine-year-old
Height (cm)	121.1 (3.99)	137.3 (6.17)
Weight (kg)	21.9 (2.22)	30.9 (6.98)
Head circumference (cm)	49.5 (0.50)	50.5 (0.35)

Table 3

Mean VAT values in milliseconds (and standard deviation) of the six Cantonese tones of the six-year-old and nine-year-old group

Tone	Boy	Girl	TOTAL
Six-year-old			
1	-7.28 (3.81)	-6.18 (2.84)	-6.99 (3.51)
2	-4.97 (3.11)	-4.86 (2.59)	-4.94 (2.89)
3	-4.65 (3.60)	-2.78 (2.64)	-4.15 (3.39)
4	-6.67 (3.09)	-1.98 (3.23)	-5.42 (3.69)
5	-9.59 (10.14)	0.74 (5.13)	-6.83 (10.08)
6	7.18 (5.67)	10.27 (7.11)	8.01 (5.98)
Nine-year-old			
1	-5.56 (2.71)	-5.96 (4.59)	-5.74 (3.58)
2	-3.77 (3.87)	-4.96 (4.07)	-4.31 (3.90)
3	-3.29 (3.64)	-3.53 (3.97)	-3.39 (3.69)
4	-4.45 (4.07)	-2.62 (3.42)	-3.62 (3.81)
5	-8.58 (11.50)	-4.23 (5.07)	-6.62 (9.24)
6	5.78 (2.10)	5.76 (3.57)	5.77 (2.77)

Main Effect of Tone

The Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(14) = 125.19$, $p < .001$. Therefore, the results of within-subject effects with Greenhouse-Geisser epsilon correction were reported. Results revealed a significant main effect of tone on mean VAT values, $F(2.07, 64.20) = 34.41$, $p < .001$. Post-hoc, pairwise comparison with Bonferroni adjustment was performed to compare the mean VAT between different tones. The pairwise comparison showed a significant difference in the mean VAT between some level and contour tone pairs. The mean VAT of a level tone (tone 1) was significantly smaller than that of two contour tones (tone 2 and tone 4) while the mean VAT of one level tone (tone 6) was significantly higher than that of all other tones. Significant difference was also noted within the level tone pairs (tone 1and tone 3; tone 1 and tone 6; tone 3 and tone 6). There were no significant differences between any two of the contour tones. Table 4, 5 and 6 summarized the significance level between different tone pairs while Figure 1 showed the relationship of mean VAT values of the six Cantonese tones between the two age groups.

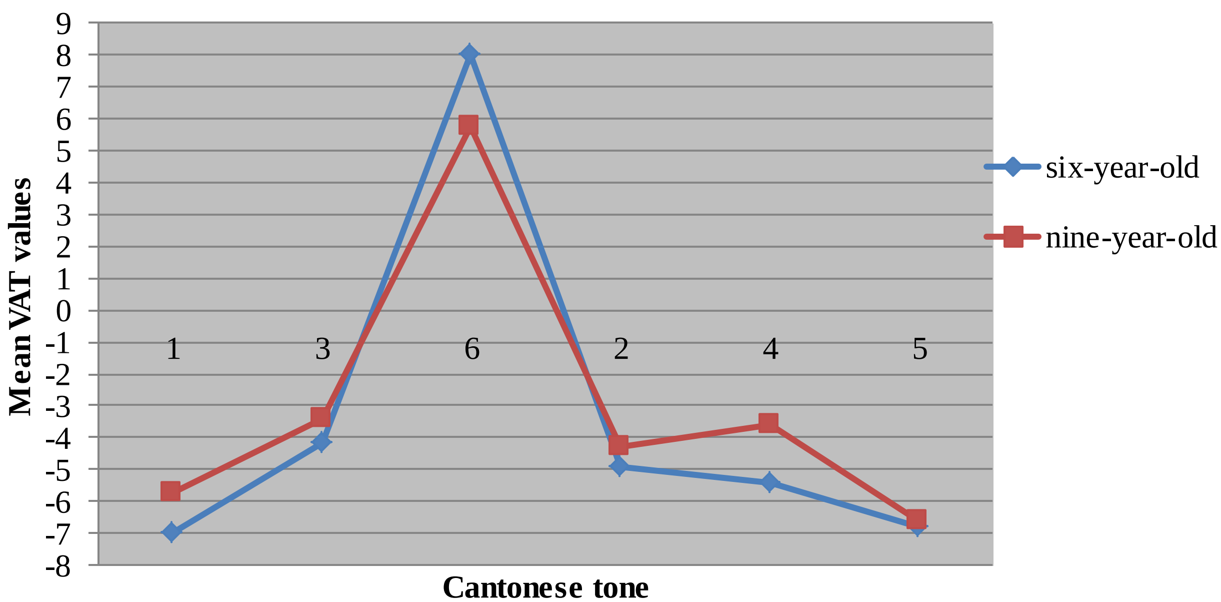

Figure 1. Mean VAT values of six-year-old and nine-year-old group across six Cantonese tones

Table 4

Mean Differences (in milliseconds) and the significance level between level and contour tone pairs

Level tone (X)	Contour tone (Y)	Mean differences (X-Y)	Significance level
1	2	-1.61	.017*
	4	-2.31	.001*
	5	-0.831	1.00
3	2	1.08	.451
	4	-0.37	1.00
	5	1.86	1.00
6	2	11.89	.0001*
	4	11.18	.0001*
	5	12.66	.0001*

Table 5

Mean Differences in milliseconds and the significance level between level and level tone pairs

Level tone (X)	Level tone (Y)	Mean differences (X-Y)	Significance level
1	3	-2.69	.001*
3	6	-10.807	.0001*
6	1	13.49	.0001*

Table 6

Mean Differences in milliseconds and the significance level between contour and contour tone pairs

Level tone (X)	Level tone (Y)	Mean differences (X-Y)	Significance level
2	4	-0.71	1.00
4	5	1.48	1.00
5	2	-0.77	1.00

** The mean difference is significant at the .05 level*

Main Effect of Gender

There was no significant main effect of gender on mean VAT values, $F(1, 31) = .314$, $p = .717$.

Main Effect of Age

There was no significant main effect of age on mean VAT values, $F(1,31) = 4.01$, $p = .054$.

Interactions

No significant interaction was revealed between any two factors (tone-by-age interaction: $F(2.07, 64.20) = .943$, $p = .397$; tone-by-gender interaction: $F(2.07, 64.20) = 2.81$, $p = .066$; tone-by-age-by gender interaction: $F(2.07, 64.20) = .251$, $p = .786$).

Discussion

The present study set out to investigate the effect of tone on VAT in young Cantonese-speaking children. The VAT in children of different ages between girls and boys were compared. It was hypothesized that contour tones would yield a higher VAT values than that of level tones. Secondly, VAT values would differ between the two age groups of children. Another hypothesis was that there would be no gender effects on the VAT values of young children. The present study would allow better understanding in the vocal physiology during the production of the two types of tone and the physiological change of vocal folds in the process of vocal fold maturation.

Effect of Tones on VAT

It was hypothesized that the VAT values of contour tones would be larger than those of level tones. This was suggested to be due to the more complex vocal fold adjustment during the phonation of contour tones, which led to slower vocal fold adduction and hence longer VAT (Ma et al., 2012). The present finding was partly consistent with this hypothesis. In this study, the mean VAT values of tone 2 and tone 4 were significantly greater than that of tone 1. However, this pattern did not exist between tone 3 and the other contour tones. A contradicting result was also noted, with the mean VAT of tone 6 being significantly greater than those of all other tones.

As explained before, the greater mean VAT of contour tone (tone 2 and tone 4) to that of level tone (tone 1) was possibly related to the rapid changes in F_0 within the syllable or word during the production of contour tones (Fujisaki, Ohno & Gu, 2004). Due to this changing F_0, more complicated laryngeal adjustment or ongoing vocal fold adjustment is required during the phonation. However, tone 5, which was also a contour tone, was not significantly higher than tone 1. A possible reason for this is the effect of the initial consonant /m/ of tone 5 stimuli. As a result of the initial consonant, speakers may need to adopt different strategies for achieving the consonantal contrasts. These strategies may involve a preset laryngeal system, with vocal fold adduction preceding the onset of sound pressure flow in the glottis. The early vocal fold contact leads to an overall shorter VAT for tone 5 stimuli.

Significant difference between level tone and contour tone was only revealed with tone 1 but not tone 3. This finding may be explained by the fundamental frequency of tone 1. Of the three level tones, the production of tone 1 was associated with the highest F_0 contour (Rose, 2000). Chhetri, Neubauer & Berry's study (2012) showed that, at higher F_0, the CT muscles would activate and elongate the vocal folds. Therefore, during the production of tone 1, the elongation of the vocal folds is the greatest and the glottal opening is the narrowest. As a result, the time for the vocal folds to approximate is reduced and the VAT is shortened. The significant difference between the VAT values of tone 1 and tone 3 can also be explained by this fact. That is, the higher F_0 contour of tone 1 resulted in narrower glottis opening and therefore a shorter vocal adduction time than that of tone 3.

Contradictory result with tone 6 being remarkably greater than all other contour tones was seen in the present study. The same pattern was even noted between tone 6 and other level tones. This result is likely to be a consequence of the initial consonant effect of the stimulus word for tone 6. The production of words with /h/ as initial consonant induce soft gentle onset of voice. The vocal fold contact time was possibly lengthened and the EGG signals were not generated until the onset of vowel. This accounts for the findings that tone 6 yielded a significantly greater VAT than all the other tones

Effect of Age on VAT

Two age groups of children (six-year-old and nine-year-old) were selected for the present study. It was hypothesized that the VAT values between the two age groups were different. This was based on the fact that laryngeal structure maturation takes place throughout the school years, leading to changing vocal fold physiology at different ages.

In this study, the two groups of children were found to differ significantly in their body size (height, weight and head circumference). However, statistical analysis showed only an approaching-significant result ($p = .054$) between the mean VAT of the two age groups. This may imply that there is no appreciable difference in the underlying laryngeal physiology between these two groups of children. The possible explanations are as follow.

First, children's laryngeal structures and body sizes may follow different growth rate. It was possible that the development of laryngeal structures during the years of six to nine was slower with respect to the growth of their body sizes. Growth is not linear in nature. Scammon (1930) described the three types of growth schedules of the head and neck region: neural, somatic and lymphatic. The body stature (height and weight), for instance, follows a somatic rate of growth, that is, a rapid growth following birth in early childhood to achieve about 25-40% of adults' size, followed by a regular and slow steady growth in the years prior to puberty. Due to the nonlinear nature of human growth, the growth rate of body size was possibly greater than that of laryngeal size during the years from six to nine. Sorenson (1989) found that there was no significant difference in the fundamental frequency of children aged from six to ten years old. This further supported that laryngeal growth was not evident during the years of six and nine. These little changes in laryngeal structures therefore cannot be reflected by a different VAT pattern.

In addition to the small laryngeal differences between the two groups, another possible reason for the insignificant result is the relatively small sample size of this study. The growth in laryngeal region may have an effect on laryngeal physiology during the production of tone as the difference between the two age groups was found to be close to significant ($p = .054$). However, a small sample size may not be sufficient to reveal a remarkable effect in VAT values.

Effect of Gender on VAT

Gender effect on VAT was previously found to exist in both Cantonese and English adult population (Ma et al., 2012; Roark et al., 2012). They suggested that thicker and more massive nature of males' vocal fold resulted in longer time to build up sufficient vocal fold closure and this leads to generally greater VAT values in males. However, since the sizes of larynx and laryngeal structures are similar between boys and girls prior to puberty (Fitch & Giedd, 1999), we hypothesized that a gender effect on VAT may not be significant in young children. The present result agrees with the hypothesis.

The finding suggests that there should be no obvious difference in the vocal fold structures between young boys and girls, and this is consistent with previous studies. Titze (1989) reviewed and summarized studies on the thickness and membranous length of developing larynx. He concluded that male vocal folds did not grow disproportionately in thickness in developmental years. Other studies (Ishii, Akita, Yamashita & Hirose , 2000; Wind, 1970) also suggested that the development of the vocal fold occurs late in comparison with other laryngeal structures.

Comparison to the VAT Data in the Adult Population

The effect of tones on VAT was previously examined in the Cantonese adult population (Ma et al., 2012). Although a direct statistical comparison cannot be made with the adults' VAT data because of the use of different stimulus set, a clear difference is seen between the two groups. While adults produced VAT values that are all positive, the VAT values obtained from children are mostly negative, with the exception of tone 6. Negative VAT values mean vocal fold contact is achieved prior to the oscillation of the vocal folds. These values were comparable to the VAT values from hard glottal attack of adults (Orlikoff et al., 2009). The negative VAT values in children may lie in the differences in their vocal fold structures with those of adults.

First of all, the vocal fold lengths and glottal area of children are smaller than those of adults (Hirano, Kurita, & Nakashima, 1983). Because of the smaller vocal fold lengths and glottal areas, the time required for completing vocal fold adduction is reduced. This vocal fold adduction time is so short that vocal adduction is accomplished before the generation of SP signals. Consequently, a negative VAT is obtained.

Secondly, the vocal fold compositions in children are different from those of adults. Vocal folds consist of layers of tissues, which keep differentiating from some connective tissue fibres since birth. This differentiation is not completed until after puberty (Ishii et al., 2000; Hirano, 1983). From 5 to 10 years of age, the superficial layer remains undifferentiated, with only partial development of the elastic fibres and collagen fibres density in the lower layers (Hartnick, Rehbar & Prasad, 2005; Hirano, Kurita & Nakashima, 1983). The elastic fibres enhanced the mobility and flexibility of vocal fold movement. With incomplete development of elastic fibres, children have less controlled in vocal fold closure. It was suggested that the lack of superficial layer in children lead to severe mechanical stress to the vocal folds because they have less phonatory control over their vocal folds (Ishii, Zhai, Akita & Hirose, 1996). This may explain our present observation on the VAT difference between children and adults. A severe mechanical stress to the vocal folds may occur in an early vocal fold contact. This contributed to the early onset of EGG signals and therefore a negative VAT pattern. In addition, while the elastin fibres control the mobility and flexibility of the vocal folds, the collagen fibres, which were made up of collagen, were able to bear strong tension (Ishii et al., 1996). The collagen provides strength and structural support to the vocal folds in order to withstand stress and resist deformation under force (Gray, 2000). Adults' vocal folds have fully developed collagen fibres and hence have greater stiffness than that of children. Under the aerodynamic force at the onset of phonation, children's vocal folds are less resistant to deformation. The vocal folds approximate each other more easily to achieve complete glottal closure.

Clinical Implication

The present study revealed a significant difference between the two categories of tones in Cantonese, but only in some pairs. This result partly agreed with previous study on Cantonese adult population (Ma et al., 2012). This provided additional support to the idea that contour tone required more complicated laryngeal settings at the pre-phonatory phase.

An age effect was found to be significant in the body size (head circumference, height and weight) but not the VAT values of the two age groups. This probably implied that laryngeal growth rate during the school years from six to nine was not as rapid as other physical growths.

On the other hand, an insignificant gender effect was suggestive that sexual difference of laryngeal structures was not obvious between boys and girls at nine-year-old or before.

Although a direct statistical comparison between adults and children could not be made because of the use of a different set of stimuli in this study, clear different VAT pattern was seen. This probably reflected different laryngeal or vocal fold behaviors being adopted by adults and children due to their anatomical differences.

Limitation and Future Direction

The set of stimulus word was modified from Ma et al.'s (2012) study to suit the literacy level of children. Two words with initial consonants /m/ and /h/ were added into the set. However, the former one sometimes led to preset laryngeal structures while the latter induced a soft phonatory onset. As a result, the effect of tone was sometimes masked.

It was patent that children and adults yielded very different VAT profiles. Since different sets of stimuli and methodology were adopted in this study, a direct statistical comparison could not be made. We suggest that the difference in VAT values between children and adults lie in their anatomical differences. Further research is recommended to verify this possibility and gain a better understanding in the differences in their laryngeal physiology.

Age effect was not seen in this study. The small sample size of this study may be one of the possible explanations for this insignificant result. Further research with greater sample size is therefore warranted. Moreover, cross-sectional study on different ages of children may provide some insights into the rate of laryngeal developmental growth and the process of sexual dimorphism in children.

Conclusion

The present study suggests that there exists a significant VAT difference between some level tone and contour tone pairs. This supported the idea that more complex pre-phonatory gestures are involved in the production of contour tone. There was no significant age difference in this study. Studies with larger sample size are warranted to give more representative results. The finding also confirmed that there is no obvious difference in the vocal fold structures between boys and girls at nine-year-old or before. Different VAT patterns between children and adults may suggest that they adopt different laryngeal adjustment strategies at phonatary onset.

Acknowledgment

I would like to express my heartfelt thanks to my dissertation supervisor, Dr. Estella Ma, for her guidance and comments throughout the whole period of dissertation. Many thanks are given to my friends Ching, Yuki and Jasmine for their help in the recruitment of subjects. I would also like to express my deepest gratitude to Mrs. Yuen, Ms. Eva and Ms. Hui for their keen assistance in the data collection process. Special thanks are given to Mr. Jason Chan for his help in the data analysis process and his warmest and unconditional support throughout the dissertation project. Finally, I would like to express my immense thanks to all the subjects for their participation in this study.

References

Baken, R.J., Orlikoff, R.F. (1998). Vocal fold adduction time estimated from glottographic signals. Presented at: 25th Mid-Winter Meeting of the Association for Research in Otolaryngology. St. Petersburg, FL.

Cheung, P. S. P., Ng, A., & To, C. K. S. (2006). *Hong Kong Cantonese Articulation Test*.

Chhetri, D. K., Neubauer, J. & Berry, D. A (2012). Neuromuscular control of fundamental frequency and glottal posture at phonation onset. *Journal of Acoustic Society of America, 131*, 1401-1412

Fitch,W., Giedd, J. (1999). Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *Journal of Acoustic Society of America*, *106*, 1511-1522

Fujisaki, H., Ohno, S. & Gu, W. (2004). Physiological and Physical Mechanisms for Fundamental Frequency Control in Some Tone Languages and a Command-Response Model for Generation of Their F0 Contours. Retrieved 7th April, 2014, from http://www.iscaspeech.org/archive/tal2004/tal4_061.pdf

Gray, S. D. (2000). Cellular physiology of the vocal folds. *Otolaryngologic Clinics of North America, 33*, 679-697

Hartnick, C. J., Rehbar, R., & Prasad, V. (2005). Development and maturation of the pediatric human vocal fold lamina propria. *The Laryngoscope*, *115,* 4-15

Hirano, M., Kurita, S., & Nakashima, T. (1983). Growth, development and aging of the human vocal folds. In D.M. Bless & J.H. Abbs (Eds.), *Vocal fold physiology* (pp. 22-43). San Diego, CA: College Hill Press

Ishii, K, Zhai, W.G., Akita, M., Hirose, H. (1996). Ultrastructure of the lamina propria of the human vocal fold. *Acta Otolaryngol, 116*, 778-782

Ishii, K., Akita, M., Yamashita, K., Hirose, H. (2000). Age-related development of the arrangement of connective tissue fibers in the lamina propria of the human vocal fold. *The Annals of Otology, Rhinology & Laryngology, 109,* 1055-1064

Khouw, E. & Ciocca, V. (2007). Perceptual correlates of Cantonese tones. *Journal of Phonetics, 35*, 104-117.

Ma, E. P.-M., Baken, R.J., Roark, R.M. & Li, P. -M. (2012). Effect of Tones on Vocal Attack Time in Cantonese Speakers. *Journal of Voice, 26,* 670.e1-6

Orlikoff, R.F., Deliyski, D.D., Baken, R.J., Watson, B.C. (2009). Validation of a glottographic measure of vocal attack. *Journal of Voice, 23*, 164-168

Roark, R.M., Watson, B.C., Baken, R.J., Brown, D.J., & Thomas, J.M. (2012).Measures of Vocal Attack Time for Healthy Young Adults. *Journal of Voice, 26,* 12-17

Roark, R. M., Watson, B. C., Baken, R. J. (2012). A Figure of Merit of Vocal Attack Time Measurement. *Journal of Voice*, 26, 8-11

Scammon, R. E. (1930). The measurement of the body in childhood, edited by J. A. Harris, C. M. Jackson, D. G. Patterson, and R. E. Scammon, *The Measurement of Man* (University of Minnesota Press, Minneapolis). pp. 173-215

Sorenson, D. N. (1989). A fundamental frequency investigation of children ages 6-10 years old. *Journal of Communication Disorders, 22*, 115-123

Watson, B.C., Baken, R.J., Roark, R. M., Reid, S., Ribeiro, M. & Tsai, W. (2013). Effect of Fundamental Frequency at Voice Onset on Vocal Attack Time. *Journal of Voice*, *27*, 273-277.

Wind, J. (1970). *On the Phylogeny and Ontogeny of the Human Larynx.A Morphological and Functional Study*. Groningen,the Netherlands: Wolters-Noordhoff Publishing.

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

Appendix I

Word List of Stimuli adopted from Ma et al. (2012)

Tone	Pitch	Contour	Word	Lexical meaning
1	High	Level	丫叉 /a1 ts^ha1/ 鴉片 /a1 p^hin3/	Slingshots Opium
2	High	Rising	啞巴 /a2 pa1/ 啞鈴 /a2 liŋ4/	A dumb person Dumbbell
3	Mid	Level	亞洲 /a3 tsɐu1/ 亞軍 /a3 kwɐn1/	Asia First runner-up
4	Low	Failing	芽菜 /a4 ts^hɔi3/ 牙齒 /a4 ts^hi2/	Sprout Teeth
5	Low	Rising	瓦片 /a5 p^hin2/ 雅典 /a5 tin2/	Tile Athens
6	Low	Level	硬幣 /an6 pɐi6/ 揸拃 /a6 tsa6/	Coins Occupying lots of space

Appendix II

Modified Word List of Stimulus

Tone	Pitch	Contour	Word	Lexical meaning
1	High	Level	丫叉 /a1 ts^ha1/ 花 /fa1/ 沙 /sa1/ 蝦 /ha1/	Slingshots Flower Sand Shrimp
2	High	Rising	啞 /a2 pa1/ 耍 /sa2/	Dumb Play
3	Mid	Level	亞洲 /a3 tsɐu1/ 亞軍 /a3 kwɐn1/ 傘 /sa3/	Asia First runner-up Umbrella
4	Low	Failing	芽菜 /a4 ts^hɔi3/ 牙齒 /a4 ts^hi2/	Sprout Teeth
5	Low	Rising	馬 /ma5/ 懶 /la5/	Horse Lazy
6	Low	Level	那 /la6/ 下 /ha6/ 爛 /la6/ 罵 /ma6/	That Down Broken scold