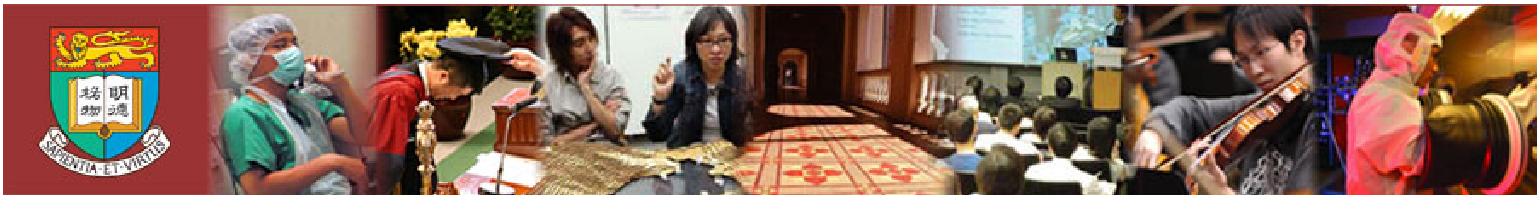


Title	The effects of vocal intensity on vocal attack time (VAT)
Author(s)	Lee, Tsz-yan, Taro; 李芷茵
Citation	
Issued Date	2010
URL	http://hdl.handle.net/10722/173717
Rights	Creative Commons: Attribution 3.0 Hong Kong License

The effects of vocal intensity on vocal attack time (VAT)

Lee Tsz Yan, Taro

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2010

ABSTRACT

Vocal attack time (VAT), is a parameter which could provide information about vocal folds attack gesture during pre-phonatory phase. The aim of the present study was to investigate the effects of vocal intensity on VAT. Simultaneously, sound pressure level and electroglottographic signals were collected from 50 participants (25 females and 25 males). The participants were asked to read aloud 12 disyllabic word stimuli using three different intensity levels, comfortable intensity first, then soft intensity and finally loud intensity. The data collected were then subjected to calculation of VAT values using a Matlab-based analysis program. Results indicated that both female and male participants demonstrated the shortest VAT values at loud intensity but the longest VAT values at soft intensity. Both gender groups show similar performance across the three intensity levels. This study provides some empirical data to support the use of VAT as a clinical tool for voice assessment.

Key words: vocal attack time (VAT), vocal fold attack gestures, pre-phonatory phase, vocal intensity level

INTRODUCTION

Phonation is divided into two phases which are the pre-phonatory phase and the attack phase (Orlikoff, Deliyski, Baken & Watson, 2007). During the pre-phonatory phase, various settings occur including adjusting appropriate tension of the vocal folds, setting the activity of laryngeal muscles, adjusting the gestures of phonatory structures and aerodynamic force for phonation. The attack phase involves the onset of vocal fold vibrations and production of sound. Attack gesture describes how the vocal folds vibrate, for example, whether it is producing a breathy voice or a hard glottal attack voice. The pre-setting in pre-phonatory phase determines the vocal fold attack gesture in the attack phase (Orlikoff et al., 2007).

According to Titze (1989), the vocal folds vibrate with small-amplitude oscillations before they make contact in the mid-line. These oscillations contribute to existence of small sound pressure levels before vocal fold adduction. Then the vocal folds continue its adduction to the mid-line and start the periodic adduction-abduction cycle and leads to production of sound pressure signals.

Sound pressure signal is related to vocal fold oscillations and vibrations. Its value correlates positively with vocal fold vibrations amplitude. Electroglottography measures the relative contact area between the vocal folds (Colton, Casper & Leonard, 2006). In electroglottography, low electrical current passes through two electrodes attached to thyroid lamina at the neck. When the vocal folds make contact, electroglottographic signals will be detected. Its amplitude also positively correlates with the relative contact area between the two vocal folds. Vocal attack time (VAT) is a measure associated with functioning of vocal fold at pre-phonatory phase. It is computed by the difference in time between the onset of sound pressure signal and electroglottographic signal (Orlikoff et al., 2007).

The VAT values vary in different phonatory situations. Orlikoff et al. (2007) compared the VAT values at three different types of phonations, namely breathy voice, comfortable voice and hard glottal attack voice. Five vocally healthy subjects (3 females and 2 males) were recruited. The subjects were instructed to produce the three types of voice quality mentioned above. The results indicated that the mean VAT value was the longest in breathy voice production and the shortest in hard glottal attack voice production, but the mean VAT value in comfortable voice production was in between the two. The mean VAT value was negative in hard glottal attack voice production, which implied that the vocal folds made contact with each other before sound pressure signals were generated. Different attack gestures give rise to different VAT values. Hence, the study concluded that the VAT value could provide information about one's vocal fold attack gesture. Its benefits included non-invasive in nature, common in clinical setting, quick measurement and close relation with acoustic energy generation (Orlikoff et al., 2007).

Li (2009) studied the effects of Cantonese tones on VAT, with an aim to set up a database for the Cantonese population. Sixty-two participants (29 males and 33 females) took part in the study. The participants were asked to read aloud 12 disyllabic words. The stimuli included words of the six Cantonese tones and two homophones were selected for each tone. Using their comfortable voice, the participants were asked to read aloud all the disyllabic words. Results indicated that the mean VAT values for contour tones were significantly longer than those for level tones. However, effects of tone were only found to be significant in female group. Li (2009) concluded that the tone features in Cantonese showed significant effects on vocal fold mechanism during pre-phonatory phase in females, but not in males. It was proposed that this difference in performance could be due to differences in vocal fold anatomical structures between the two gender groups (Li, 2009). Production of different tones is related to variation in fundamental frequency of vocal fold vibration (Ciocca & Khouw, 2007; Li, 1997; Li, 2006). This may suggest that males and females employ different fundamental frequency changing mechanisms for different tones production.

During phonation, muscle contractions of chest and abdominal wall lead to production of subglottic pressure. This subglottic pressure continues to increase until the adducted vocal folds open and start vibrations. This is then followed by a series of opening and closing of vocal folds (Colton et al., 2006). Subglottal pressure is related to aerodynamic force required for initiating phonation and it is one major factor to vary the intensity level. When the aerodynamic force for phonation is increased, the intensity will be higher. Conversely, if the aerodynamic force for phonation is decreased, the intensity will be lower (Holmberg, Hillman & Perkell, 1988). Plant and Younger (2000) states that the vocal intensity increases with higher subglottic pressure and that the relationship between vocal intensity and subglottal pressure was found to be almost linear in many cases.

At higher vocal intensity levels, the vocal folds adduct with greater force. By Newton's second law F=ma (F represents force, m and a represent mass and acceleration respectively), as the vocal folds adduct with greater force at higher intensity, the F in equation increases (Titze, 1994). While mass of vocal folds remain unchanged, acceleration will increase and contributes to higher velocity. The higher velocity should lead to shorter time for vocal fold closure. Therefore, it is expected that the VAT value will become shorter at higher intensity. Contrary, the vocal folds adduct with weaker force at lower vocal intensity. As the force in F=ma equation decreases, the acceleration also decreases (Titze, 1994). The smaller acceleration leads to slower velocity and vocal folds adduction takes longer time. Hence, longer VAT value is expected.

Therefore, in addition to magnitude of aerodynamic force required to initiate phonation, the vocal folds closure velocity has been proposed as another major factor that is related to intensity control. It is suggested the vocal folds adduct with a faster velocity at louder intensity voice production. On the other hand, the vocal folds adduct with slower velocity during softer intensity voice production (Titze, 1994).

In a previous study conducted by Li (2009) investigating the effects of six Cantonese tones on VAT, gender-related differences existed, that is, the two genders showed different performance towards the effects of tones and thus the change in fundamental frequency. In this study, the effects of different intensity levels on VAT are investigated. Also, the performance of the two genders in intensity control is observed, so as to find out if the two genders respond differently in intensity varying mechanism. This could provide more insights about vocal folds physiology and phonation mechanisms in different conditions for the two genders. For example, if the two genders show different performance in achieving loud intensity production.

Settings of aerodynamic force and vocal folds tension are completed in pre-phonatory phase before phonation begins (Orlikoff et al., 2007). Pre-setting of aerodynamic force is different in production of different vocal intensity, for example, greater aerodynamic force is set for louder intensity but less aerodynamic force is set for softer intensity (Titze, 1994). This may imply different vocal fold attack gesture is used at different intensity levels. Thus, the VAT values at different intensity levels may be different.

This study aims to investigate the effects of different vocal intensity level on VAT. Also, it could provide theoretical support for pre-setting of aerodynamic force in pre-phonatory phase. The findings could allow better understanding of vocal fold physiology and the mechanisms in varying intensity levels.

METHOD

Participants

Fifty native Cantonese speakers (25 males and 25 females) were recruited. They were within the age range of 20 to 25 years old (mean age=21.92 years old, SD=1.47). The participants were judged to be vocally healthy by an experienced speech therapist and a final-year speech therapy student of the Division of Speech and Hearing Sciences, The University of Hong Kong. All participants did not report any hearing impairments and previous/ current voice problems. All of them filled in a health questionnaire to ensure that they were physically healthy on the day of recording. Participants were excluded from the study if the result of the health questionnaire reviewed that they have case history of voice, hearing, and speech and language problems.

Instrumentation

M-Audio Fast Track recording interface device (Fast Track USB) with a Radio Shack 33-3012 headset microphone was used to gather sound pressure (SP) signals. KayPENTAX model 6103 was used to gather electroglottographic (EGG) signals. An AKG K601 open headphone was used to play babbling noise with four people reading news together to participants. Audacity (version. 1.2.6) was used to record the SP and EGG signals at the same time. 5 cm distance was maintained between participants' mouth and the microphone throughout the data collection process.

Materials

The stimuli were constructed by twelve disyllabic words with target syllable /Na/. The stimuli word list included words with target syllable /Na/ across the six Cantonese tones. Homophones, which share same pronunciation but different meaning, were included in the word list. Two homophones were selected for each tone. At each intensity level, the participants were asked to read aloud the stimuli for five times. Randomized order of stimuli was presented at each trial.

Procedures

The data collection process was implemented in a sound proof booth at the Division of Speech and Hearing Sciences, The University of Hong Kong. The participants' neck areas were cleaned with alcohol pads before the recording process so as to enhance the contact effectiveness between the electrode and neck. The investigator first searched the thyroid lamina position. The researched then placed two surface electrodes for measuring EGG signals on either side of the thyroid lamina. Each participant was asked to phonate the vowel /a/ with his/her comfortable voice to ensure the electrode-to-neck contact position contact effectiveness was maximized. An oscilloscope was used to display the EGG waveform. The researcher adjusted the vertical position of the electrodes until the amplitude of EGG waveform was the largest as observed from the oscilloscope. This procedure was taken to ensure that the electrode-to-neck contact was the highest and the electrodes position was the most appropriate (Ma & Love, 2010). A neckband was used to attach the electrodes to the back of each participant's neck. To ensure the electrode-to-neck contact was maximized, the neckband was adjusted to be adequately tight enough.

The participants were briefed about the recording procedure and the recording stimuli before the actual recording process began. This allowed the participants to have a general idea about what to do in the recording process. The participants were instructed to read all the disyllabic words stimuli once using their comfortable pitch and loudness for the researcher to get their intensity baselines at comfortable intensity level. This procedure also ensured that the participants would not read the disyllabic words stimuli incorrectly during the recording process.

Babbling noise with four people reading news together was played to the participants through an open headphone when they were reading aloud the disyllabic words with their comfortable voice and loud voice. The intensity level of babbling noise was higher when participants read aloud with their loud voice. The sound pressure levels of babbling noise used at comfortable intensity and loud intensity were 75dB and 90dB respectively. The intensity levels of babbling noise were standardized across all the participants. No babbling noise was played to the participants when they were reading aloud the words with their soft voice. The purpose of playing babbling noise to the participants was to facilitate the participants to vary their intensity level as instructed. The researcher did not instruct them to attain a specific intensity level in order to make their productions as natural as possible (Holmberg et al., 1988). Practice of reading the stimuli at the three intensity levels was carried out beforehand to ensure that the participants' productions were all appropriate. The recording procedure was recorded in the following order: comfortable intensity, soft intensity and finally loud intensity.

After reading the five trials at each speech intensity level, the participants were asked to have a subjective perceptual rating on their speech loudness of reading at that particular speech intensity level using a visual analog scale anchored with number 1-10. The self-rating was done separately for each speech intensity level and the participants were not allowed to look back to their previous rating so as to avoid bias caused by previous rating results.

Data analysis

VAT analysis software created by MATLAB was used to compute the VAT value of each participant automatically at different intensity level based on the sound pressure (SP) signals and electroglottographic (EGG) signals. There were altogether 108 SP signals and 108 EGG signals for analysis of each participant (12 disyllabic words X 3 trials X 3 intensity levels X 3 trials). Averaging the sum of the VAT values obtained from the three trials was done to calculate the mean VAT value for each homophone. At each tone, the mean VAT values for the two homophones were averaged to give the mean VAT values for that particular tone.

Therefore the mean VAT values of the six tones at the three intensity levels would be compared to see if significant differences existed between the VAT values at different intensity levels.

RESULTS

At each intensity level, the participants produced the twelve disyllabic word stimuli for five times. The mean VAT values of the six tones of the middle three trials (i.e., the 2^{nd}, 3^{rd} and 4^{th}) were chosen for statistical analysis. Data were rejected from the statistical analysis if the any one of the following criteria was met:

1) outliers spotted by plotting boxplot graphs: boxplot graph was plotted using each participant's VAT values of each homophone of the middle three trials for each trial. One boxplot graph was plotted individually for each gender. All outliers shown on the boxplot graphs were excluded from further statistical analysis;

2) abnormal display of resulting graph during calculations by the VAT analysis program: during the calculations by the VAT analysis program, the parameter SP-to-EGG delay represented the VAT value and the resulting graph would normally show a bell-shaped normal distribution graph. Data were excluded from further statistical analysis if multiple peaks were spotted from the resulting graphs.

Descriptive statistical analysis

Table 1 and 2 list the mean VAT values and standard deviations of the six Cantonese tones and the 12 homophones at the three intensity levels of the female group and the male group respectively.

Table 1: Mean VAT values in milliseconds (and standard deviations) of the six Cantonese tones and the 12 homophones at different intensity levels of the female group

Intensity level	Tone	Homophones	Mean (SD)	Averaged mean (SD)
Soft	1	/Na$_1$ ts^{h}a$_1$/ 丫叉	2.45(4.66)	3.67(4.46)
		/Na$_1$ p^{h}in$_3$/ 鴉片	4.88(4.83)	
	2	/Na$_2$ pa$_1$/ 啞巴	6.94(4.70)	7.04(3.85)
		/Na$_2$ liN$_4$/ 啞鈴	7.14(4.70)	
	3	/Na$_3$ kwAn$_1$/ 亞軍	4.78(3.84)	4.92(3.80)
		/Na$_3$ tsAu$_1$/ 亞洲	5.07(4.60)	
	4	/Na$_4$ ts^{h}i$_2$/ 牙齒	7.22(5.56)	6.70(3.94)
		/Na$_4$ ts^{h}Oi$_3$/ 芽菜	6.18(4.23)	
	5	/Na$_5$ tin$_2$/ 雅典	6.02(4.07)	5.88(3.51)
		/Na$_5$ p^{h}in$_2$/ 瓦片	5.74(3.70)	
	6	/Na$_6$ ts^{h}a$_6$/ 掗搾	3.82(4.89)	4.98(4.05)
		/NaN$_6$ pAi$_6$/ 硬幣	6.14(5.17)	
Pooled data (soft)				**5.53**
Comfortable	1	/Na$_1$ ts^{h}a$_1$/ 丫叉	-2.07(6.14)	-1.67(4.86)
		/Na$_1$ p^{h}in$_3$/ 鴉片	-1.26(4.55)	
	2	/Na$_2$ pa$_1$/ 啞巴	1.63(6.26)	2.26(5.18)
		/Na$_2$ liN$_4$/ 啞鈴	2.88(5.26)	
	3	/Na$_3$ kwAn$_1$/ 亞軍	-0.71(5.61)	-1.19(5.69)
		/Na$_3$ tsAu$_1$/ 亞洲	-1.67(6.00)	
	4	/Na$_4$ ts^{h}i$_2$/ 牙齒	3.39(5.25)	3.16(4.32)
		/Na$_4$ ts^{h}Oi$_3$/ 芽菜	2.92(4.11)	
	5	/Na$_5$ tin$_2$/ 雅典	3.45(4.44)	3.32(5.31)
		/Na$_5$ p^{h}in$_2$/ 瓦片	3.18(6.59)	
	6	/Na$_6$ ts^{h}a$_6$/ 掗搾	0.26(5.35)	0.85(4.80)
		/NaN$_6$ pAi$_6$. 硬幣	1.44(5.01)	
Pooled data (comfortable)				**1.12**
Loud	1	/Na$_1$ ts^{h}a$_1$/ 丫叉	-6.60(6.95)	-6.61(5.99)
		/Na$_1$ p^{h}in$_3$/ 鴉片	-6.63(6.55)	
	2	/Na$_2$ pa$_1$/ 啞巴	-2.71(4.48)	-2.85(4.42)
		/Na$_2$ liN$_4$/ 啞鈴	-2.98(5.22)	
	3	/Na$_3$ kwAn$_1$/ 亞軍	-6.32(4.52)	-6.32(4.31)
		/Na$_3$ tsAu$_1$/ 亞洲	-6.32(4.98)	
	4	/Na$_4$ ts^{h}i$_2$/ 牙齒	-3.55(4.84)	-4.06(4.93)
		/Na$_4$ ts^{h}Oi$_3$/ 芽菜	-4.56(5.82)	
	5	/Na$_5$ tin$_2$/ 雅典	-2.77(5.70)	-3.39(5.45)
		/Na$_5$ p^{h}in$_2$/ 瓦片	-4.01(5.98)	
	6	/Na$_6$ ts^{h}a$_6$/ 掗搾	-5.47(6.03)	-5.23(4.69)
		/NaN$_6$ pAi$_6$. 硬幣	-4.99(5.77)	
Pooled data (loud)				**-4.74**

Table 2: Mean VAT values in milliseconds and standard deviations of the six Cantonese tones and the 12 homophones at different intensity levels of the male group

Intensity level	Tone	Homophones	Mean (SD)	Averaged mean (SD)
Soft	1	/Na$_1$ ts^{h}a$_1$/ 丫叉	5.86(3.69)	5.43(3.56)
		/Na$_1$ p^{h}in$_3$/ 鴉片	5.00(3.85)	
	2	/Na$_2$ pa$_1$/ 啞巴	5.58(5.17)	6.05(4.16)
		/Na$_2$ liN$_4$/ 啞鈴	6.53(4.09)	
	3	/Na$_3$ kwAn$_1$/ 亞軍	6.36(5.04)	5.92(4.68)
		/Na$_3$ tsAu$_1$/ 亞洲	5.48(4.92)	
	4	/Na$_4$ ts^{h}i$_2$/ 牙齒	7.23(8.21)	6.21(5.31)
		/Na$_4$ ts^{h}Oi$_3$/ 芽菜	5.20(4.67)	
	5	/Na$_5$ tin$_2$/ 雅典	6.50(5.52)	6.73(5.66)
		/Na$_5$ p^{h}in$_2$/ 瓦片	6.97(7.00)	
	6	/Na$_6$ ts^{h}a$_6$/ 掗搾	4.72(5.77)	5.40(5.37)
		/NaN$_6$ pAi$_6$. 硬幣	6.09(5.74)	
Pooled data (soft)				**5.96**
Comfortable	1	/Na$_1$ ts^{h}a$_1$/ 丫叉	-0.90(5.93)	0.14(5.08)
		/Na$_1$ p^{h}in$_3$/ 鴉片	1.17(5.12)	
	2	/Na$_2$ pa$_1$/ 啞巴	0.84(4.98)	1.51(4.24)
		/Na$_2$ liN$_4$/ 啞鈴	2.18(4.32)	
	3	/Na$_3$ kwAn$_1$/ 亞軍	-0.58(6.54)	-0.60(6.31)
		/Na$_3$ tsAu$_1$/ 亞洲	-0.61(6.35)	
	4	/Na$_4$ ts^{h}i$_2$/ 牙齒	1.46(4.23)	1.27(3.72)
		/Na$_4$ ts^{h}Oi$_3$/ 芽菜	1.09(3.67)	
	5	/Na$_5$ tin$_2$/ 雅典	-0.04(4.55)	0.63(3.86)
		/Na$_5$ p^{h}in$_2$/ 瓦片	1.31(3.91)	
	6	/Na$_6$ ts^{h}a$_6$/ 掗搾	-0.43(6.04)	-0.08(5.25)
		/NaN$_6$ pAi$_6$. 硬幣	0.28(5.14)	
Pooled data (comfortable)				**0.48**
Loud	1	/Na$_1$ ts^{h}a$_1$/ 丫叉	-6.07(5.58)	-5.74(4.91)
		/Na$_1$ p^{h}in$_3$/ 鴉片	-5.41(5.05)	
	2	/Na$_2$ pa$_1$/ 啞巴	-3.59(5.89)	-3.56(5.50)
		/Na$_2$ liN$_4$/ 啞鈴	-3.52(5.94)	
	3	/Na$_3$ kwAn$_1$/ 亞軍	-6.76(6.11)	-7.11(6.00)
		/Na$_3$ tsAu$_1$/ 亞洲	-7.45(6.78)	
	4	/Na$_4$ ts^{h}i$_2$/ 牙齒	-5.05(6.50)	-4.75(6.16)
		/Na$_4$ ts^{h}Oi$_3$/ 芽菜	-4.44(6.57)	
	5	/Na$_5$ tin$_2$/ 雅典	-5.44(5.89)	-4.25(5.28)
		/Na$_5$ p^{h}in$_2$/ 瓦片	-3.06(5.37)	
	6	/Na$_6$ ts^{h}a$_6$/ 掗搾	-6.63(6.31)	-5.93(5.79)
		/NaN$_6$ pAi$_6$. 硬幣	-5.23(5.99)	
Pooled data (loud)				**-5.22**

Table 3 lists the mean and standard deviations of participants' perceptual rating of their speech loudness across the three intensity levels. Table 4 lists the mean differences between the mean scores across the three intensity levels. Significant differences were found between all the perceptual rating score pairs in the three intensity levels.

Table 3: Mean scores and standard deviations of the perceptual rating score across the three intensity levels

Intensity level	Mean scores (SD)
Soft	2.18 (0.78)
Comfortable	4.75 (0.80)
Loud	7.92 (0.75)

Table 4: Mean differences between the mean scores across the three intensity levels

Intensity level (X)	Intensity level (Y)	Mean differences (X-Y)	Significance level
Comfortable	Soft	2.57	0.0001*
	Loud	-3.17	0.0001*

* = p<0.01 level.

Effects of vocal intensity, tones and gender on VAT

The effects of vocal intensity were determined by comparing the mean VAT across the three intensity levels. A three-way mixed ANOVA with 3(intensity levels) x 6(Cantonese tones) x 2(gender) was used to investigate the effects of vocal intensity on VAT values. The dependent variable of the study is the VAT value. The within-subject variables included vocal intensity (soft, comfortable and loud intensity) and tone (six Cantonese tones). The between-subject variable included gender (male and female). For all the analyses in this part, the *alpha* level was set at 0.05.

Intensity main effect. The sphericity for intensity was assumed as shown from the result of the Mauchly's Test of Sphericity ($p > 0.05$). The main effect of intensity on mean VAT was found to be significant [$F(2, 8628)=175.57$, $p=0.0001$]. Post-hoc, pairwise comparison with Bonferroni adjustment was carried out to study the mean VAT of all the participants at different intensity levels. Significant differences existed in the mean VAT between soft and comfortable intensity levels ($p=0.0001$), loud and comfortable intensity levels ($p=0.0001$), and soft and loud intensity levels ($p=0.0001$). For all six tones, the mean VAT was found to be the shortest at loud intensity level and the longest at soft intensity level, while the mean VAT at comfortable intensity level fell in between the two. Table 5 summarized the significance level between the intensity level pairs.

Table 5: Mean differences of VAT and the significance level of intensity pairs of all participants

Intensity level (X)	Intensity level (Y)	Mean differences (X-Y)	Significance level
Comfortable	Soft	-4.95	0.0001*
	Loud	5.77	0.0001*

* = p<0.01

Tone main effect. The sphericity for tone was not assumed as shown from the result of the Mauchly's Test of Sphericity ($p= 0.0001$). Therefore, the degrees of freedom of tone were corrected by Greenhouse-Geisser estimates of sphericity. The main effect of tone on mean VAT value was found to be significant [$F(3, 319)=19.94$, $p=0.0001$]. Post Hoc, pairwise comparison with Bonferroni adjustment was carried out to investigate the differences in mean VAT values between the three level tones and the three contour tones. Significant differences were found between all the level tone and contour tone pairs. Table 6 summarized the significance levels between the tone pairs.

Table 6: Mean differences in milliseconds and the significance level of tone pairs of the six Cantonese tones

Level tones (X)	Contour tones (Y)	Mean difference (X-Y)	Significance level
1	2	-2.57	0.0001*
	4	-2.22	0.0001*
	5	-2.29	0.0001*
3	2	-2.50	0.0001*
	4	-2.15	0.0001*
	5	-2.22	0.0001*
6	2	-1.77	0.0001*
	4	-1.43	0.002*
	5	-1.49	0.001*

* = p<0.01 level

Gender main effect. There was no significant main effect of gender on mean VAT values [F(1,11)=0.05, p>0.05].

Interactions. The tone-by-gender interaction was significant [F(3, 53)=3.30, p=0.02], i.e., the two gender groups showed different performance on mean VAT value across the six Cantonese tones. Figure 1 shows the mean VAT values of the male and female groups across the six Cantonese tones. None of the other interactions reached a significance of 0.05 level (tone-by-intensity interaction, [F(7, 20)=1.95, p>0.05]; intensity-by-gender interaction, [F(2, 24)=0.50, p>0.05]; tone-by-intensity-by-gender interaction, [F=(7, 13)1.28, p>0.05].

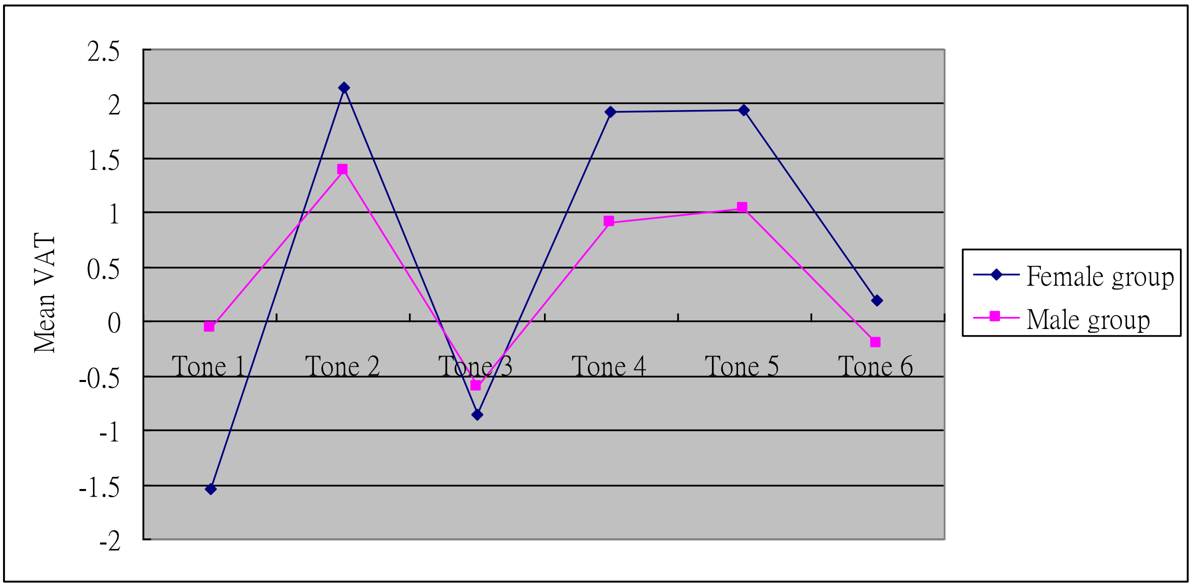


Figure 1: Mean VAT values of male and female groups across six Cantonese tones.

To summarize, tone and intensity level was shown to have significant effect on the mean VAT values of all participants. Significant interaction effect was shown between tone and gender. Moreover, significant differences were found between all the perceptual rating score pairs across the three intensity levels.

DISCUSSION

The present study aimed to investigate the effects of vocal intensity on vocal attack time (VAT). This investigation will allow better understanding of the vocal fold functioning and how vocal intensity will affect the value of VAT. It will also provide theoretical support for pre-setting of aerodynamic force in pre-phonatory phase. It was hypothesized that the VAT value will vary according to different intensity level, in which the VAT value at loud intensity will be the shortest but the VAT value at soft intensity will be the longest.

Effects of vocal intensity on VAT

Among the three intensity levels, the mean VAT value at loud intensity level is significantly shorter than those obtained at comfortable and soft intensity levels. This applied to both the female groups and the male groups. The results suggested that vocal intensity exerts a significant effect on VAT value, and were consistent with the hypothesis. Intensity control is associated with the magnitude of subglottal pressure and subglottal pressure can be considered as aerodynamic force needed for vocal fold vibration. Therefore, to achieve higher intensity, greater aerodynamic force is needed to initiate vocal fold vibration (Colton et al., 2006; Titze, 1989). Besides greater aerodynamic force, Titze (1994) proposed velocity of vocal folds closure is a factor associated with intensity control. He suggested that at louder intensity production, the vocal folds will close with a faster closing speed. Conversely at softer intensity production, the vocal folds will close with a slower closing speed.

VAT is a parameter measuring the time difference between the onset of sound pressure (SP) signal and electroglottographic (EGG) signal (Orlikoff et al., 2007). Intensity control is associated with magnitude of aerodynamic force and vocal fold closure speed. These two factors will in turn affect the value of VAT. At louder intensity level, the vocal folds adduct with greater aerodynamic force and higher velocity (Titze, 1994). In this case, the time differences between the onset of SP signals and EGG signals will be reduced, that is, the VAT are shortened. From our results, it showed that the mean VAT values at loud intensity are negative in both gender groups. This may suggest that, at loud intensity level, the vocal folds approximated before voice onset starts in pre-phonatory phase (Orlikoff et al., 2007). However, to achieve a softer intensity level, the aerodynamic force for vocal folds adduction and the vocal folds closing speed becomes weaker and slower respectively (Titze, 1994). As a result, the VAT is lengthened. Results in this study revealed that the mean VAT value is the longest at soft intensity. This supported with the hypothesis that the VAT values vary with different intensity levels due to different pre-settings of aerodynamic force during pre-phonatory phase.

Effects of tones on VAT

Cantonese is a tonal language which is consisted of six tones. Tone 1, tone 3 and tone 6 are level tones while tone 2, tone 4 and tone 5 are contour tones (Matthews & Yip, 1994; Yip, 2002). Each tone carries different fundamental frequency (Ciocca & Khouw, 2007; Li, 1997; Li, 2006). From the results, tone-by-gender interaction was found to be significant, i.e., two gender groups have different performance across the six tones on VAT. Results in present study were consistent with previous findings in Li's study (2009). This gender-related difference on mean VAT values across the six Cantonese tones may imply that the anatomical structure differences in vocal folds affect the frequency changing mechanism in the two gender groups.

It is known that males generally have thicker, longer and tenser vocal folds than females (Sulter, Schutte & Miller, 1996). Change in tones is related to changes in fundamental frequency. Due to this structural difference, the stiffer and more resistant nature of male vocal folds makes phonatory gestures adjustments required in changing fundamental frequency during pre-phonatory phase more difficult to be accomplished (Chan & Titze, 1999). Therefore, the degree of adjustments in vocal fold configurations in males may be less than that in females. Because of the reduced degree of adjustments in vocal fold configurations, no significant difference on the mean VAT values is found across the six tones in male groups.

For female groups, significant differences in mean VAT values existed between all the level and contour tone pairs except between tone 2 and tone 6. Anatomically, females have thinner, lower-mass and less tense vocal folds than males. Because of the greater elasticity and higher viscosity nature of vocal folds in females, the degree of adjustments in vocal folds configurations in changing fundamental frequency during pre-phonatory phase is relatively easier to be accomplished in females (Chan & Titze, 1999). The degree of adjustments in vocal folds configurations may be more prominent in females and therefore females demonstrated significant differences between the mean VAT values of level tones and contour tones.

VAT in females and males

From the results, it showed that gender alone has no significant effect on the mean VAT values but the tone-by-gender interaction effect is found to be significant. Moreover, intensity-by-gender interaction effect is found to be not significant. As mentioned in previous paragraphs, differences exist in vocal folds anatomy and physiology between the two gender groups, in which males generally have thicker, longer and higher-mass vocal folds than females (Chan & Titze, 1999). However based on findings in current study, it showed that this gender-related difference affects only the fundamental frequency changing mechanism but not the intensity changing mechanism. That is, males and females have different performances in changing the fundamental frequency. Due to stiffer and more resistant vocal folds, it seems that it is more difficult for males to adjust their degree of displacement of their vocal folds in response to frequency changing. Yet this is not the case for females. Females are comparatively easier to adjust their degree of displacement of their vocal folds due to greater elasticity and lower viscosity. This explains why males and females respond differently to the mean VAT values across the six Cantonese tones (Li, 2009).

Based on the current findings, both males and females show similar performance in mean VAT values across the three intensity levels. This may imply that differences in vocal folds anatomy and physiology between the two gender groups do not affect their way in varying intensity.

Theoretical implications

The present study aims to provide theoretical support for pre-setting of aerodynamic force in pre-phonatory phase. Orlikoff et al. (2007) stated that the pre-phonatory phase is associated with adjusting vocal fold tension and aerodynamic force for vocal fold adduction. It is known that intensity control is related to changes in subglottal pressure, in which subglottal pressure can be considered as the aerodynamic force required for vocal fold adduction (Titze, 1994). From the results, it showed that the mean VAT values vary according to different intensity levels. The mean VAT is found to be the shortest at loud intensity but the longest at soft intensity. This suggests that to vary the intensity level, pre-setting of aerodynamic force has to be completed during the pre-phonatory phase. Results in this study could provide additional support about existence of pre-setting of aerodynamic force in pre-phonatory phase.

Clinical implications

VAT values vary in different phonatory situations and tone and intensity are found to have significant effects on mean VAT values. From the results, it is proposed that when using VAT as an objective diagnostic tool clinically to examine the vocal fold attack gesture, factors that may affect the VAT values, such as vocal intensity should be controlled.

Limitations of this study and further research directions

Cautions should be taken before applying the VAT measure clinically. A larger sample size of Cantonese speakers will make the results to be more representative. Moreover, aging is known to be a factor to affect the anatomy and physiology of vocal folds (Ma & Love, 2010). Therefore effects of age on mean VAT values will be of great research interest and worth investigating. Investigations with different age groups, such as young adults versus elderly, can be conducted in order to make sure the differences in anatomy and physiology is large enough for investigating. Equal number of participants should be guaranteed in all age groups and gender groups. Besides, individual differences exist in mean VAT values. Therefore to overcome with this, it is possible to conduct a study with larger sample size in future so as to construct a more representative database for VAT in Cantonese population. Furthermore, the intensity should be one factor to control in constructing the database since it is found to be a factor which will affect the mean VAT values. This could make sure that a more reliable diagnosis could be made based on the values of the VAT.

CONCLUSION

The present results suggest that intensity is a factor that significantly affects the mean VAT values in both male and female speakers. This suggests that both gender groups use similar intensity changing mechanism to achieve a desired intensity level. Also, Cantonese tone features, involving three level tones and three contour tones, are shown to have significantly effect only in female speakers, but this effect is absent in male speakers. Significant differences are found between level tones and contour tones in female speakers yet no significant difference is found between all the tone pairs in male speakers. The present study provides some empirical data to support the use of vocal attack time (VAT) as a clinical tool for voice assessment, and to understand different phonation mechanism.

ACKNOWLEDGEMENT

I would like to thank my dissertation supervisor, Dr. Estella Ma, for her clear guidance and helpful comments throughout the whole period of doing the dissertation project. She has made this dissertation project to be a valuable learning process for me. Great thanks are given to my friends, fellow classmates and all the participants for their participations in this study. Their contributions determined the successful completion of this study. Special thanks are given to Miss Kity Wong for her assistance in data analysis process. Last but not least, I would like to express my immense gratitude to my family for their warmest love which makes me feel I am always blessed.

REFERENCE

Chan, R. W., & Titze, I. R. (1999). Viscoelastic shear properties of human vocal fold mucosa: measurement methodlogy and empirical results. *Journal of the Acoustical Society of America*, *104*(4), 2008-2021.

Ciocca & Khouw (2007). Perceptual correlates of Cantonese tones. *Journal of Phonetics, 35*(1), 104-117.

Colton, R. H., Casper, J. K., & Leonard, R. (2006). (3rd Ed) *Understanding voice problems: A physiological perspective for diagnosis and treatment*. Baltimore: William & Wilkins.

Erickson, D. (1993). Laryngeal muscle activity in connection with Thai tones. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics*, 27, 135-149.

Stathopoulos, E.T. (1986). Relationship between intraoral air pressure and vocal intensity in children and adults. *Journal of Speech and Hearing Researc*h, 29(2), 71-74

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *Journal of the Acoustical Society of America*, *84*(2), 511-529.

Li, L.C. (1997). *Voice quality of Cantonese tones: an acoustic study*. Unpublished BSc Honours Thesis. Division of Speech and Hearing Sciences. The University of Hong Kong

Li, P.M. (2009). *The effects of Cantonese tones on vocal attack time (VAT). Unpublished BSc Honours Thesis*. Division of Speech and Hearing Sciences. The University of Hong Kong

Li, Y. (2006). *Tone ratio combined with F_0 register in Cantonese as speaker-dependent characteristics.* Retrieved April 1, 2010, from http://aune.lpl.univ-aix.fr/~sprosig/sp2006/contents/papers/PS2-13_0206.pdf

Ma, E. P.-M. & Love, A. L. (2010). Electroglottographic evaluation of age and gender effects during sustained phonation and connected speech. *Journal of Voice*, 24(2), 146-152

Matthews, S. & Yip, V. (1994). *Cantonese: a comprehensive grammar.* Lodon: Routledge

Orlikoff, R. F., Deliyski, D. D., Baken, R. J. & Watson, B. C. (2007). Validation of a glottographic measure of vocal attack. *Journal of Voice*, *23*(2), 164-168.

Plant, R. L. & Younger, R. M. (2000). The interrelationship of subglottic air pressure, fundamental frequency, and vocal intensity during speech. *Journal of Voice*, *14*(2), 170-177.

Rose, P. (1996). Between- and within- speaker variation in the fundamental frequency of Cantonese citation tones. In Davis, P. J. & Fletcher, N. H. (Eds.). *Vocal fold physiology: controlling complexity and chaos.* (pp. 307-324). San Diego: Singular Pub.

Sulter, A.M., Schutte, H.K. & Miller, D.G. (1996). Standardized laryngeal videostroboscopic rating: differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency, and age. *Journal of Voice*, 10(2), 175-189

Titze, I. R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *Journal of the Acoustical Society of America*, *85*(2), 901-906.

Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, N. J.: Prentice Hall.

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

Appendix 1 – Consent form and personal information

香港大學

言語及聽覺科學部

四年級畢業論文

[廣東話聲量與發聲原理的關係]

参與同意書

本人是香港大學言語及聽覺科學部四年级學生李芷茵,現正在馬珮雯博士的督導下進行一項關於廣東話聲量與發聲原理之關係的學術研究。旨在更深入了解發聲原理,以幫助發展日後聲線學用途。

此項研究中,参與者須根據研究員的指示讀出一些詞語。本項研究, 是利用電腦儀器分析不同的廣東話聲量與發聲原理的關係。在過程中, 研究員會將兩塊鐵片貼於你的頸部兩旁的位置,以用作錄音用途,整個過程將不會有任何不適的感覺。你將會讀出一些詞語,研究員將會錄音以用作研究用途,整個過程將會維持約二十分鐘。

是次研究並不為閣下提供個人利益,但所搜集數據將對研究聲線問題提供寶貴的資料。是次研究並不涉及任何風險及不會對参與者的身心理造成任何危險的後果。閣下可隨時提出終止,有關決定將不會引致任何不良後果。参與純屬自願性質,個人資料將絕對保密。如你對是項研究有任何問題,請現在提出。

如日後你對是項研究有任何查詢,請與研究員李芷茵小姐聯絡(電話: 98688661/電郵: tarolee@gmail.com)。如你對是項研究有任何查詢,亦可與馬珮雯博士聯絡(電話: 28590594/ 電郵: estella.ma@hku.hk)。

如你想知道更多有關研究参與者的權益,請聯絡香港大學非臨床研究操守委員會(22415267)。

如你明白以上內容,並願意参與是項研究,請在下方簽署。

参與者姓名: ____________________ 参與者簽署:____________________

研究員姓名: ____________________ 研究員簽署:____________________

参與研究日期: ______________________

Appendix 2- health questionnaire

香港大學

言語及聽覺科學

四年級畢業論文-[聲線學術研究]

健康問卷調查

請回答以下有關你的健康狀況的問題:

1. 請問你現在/ 曾經有沒有患過聲線上的問題?

☐ 有　　如有，請詳細列明情況:____________________________

☐ 沒有

2. 你現在/ 曾經有否患過言語及/或 咬字不清的問題?

☐ 有　　如有，請詳細列明情況:____________________________

☐ 沒有

3. 請問你有沒有任何聽覺問題?

☐ 有　　☐ 沒有

4. 請問你在過去兩星期曾否有過以下病症? (可選擇多於一項)

☐ 喉嚨發炎　☐ 傷風　☐流行性感冒

☐ 竇炎　☐ 鼻炎　☐過敏性鼻炎(鼻敏感)

☐ 氣喘(病),哮喘　☐甲狀腺機能亢進

☐甲狀腺機能減退　☐ 扁桃腺炎

☐ 不適用，我沒有任何以上病症

5. 請問一個月前到現在，你有否服食任何藥物?

☐ 有　　如有，請詳細列明情況:____________________________

☐ 沒有

6. 你從前或現在有沒有患過哮喘或其他相關的呼吸系統疾病?

☐ 有　　☐ 沒有

7. 請問你的喉嚨或附近位置曾否受過任何損傷?

☐ 有　　☐ 沒有

8. 請問你有沒有胃酸倒流的情況?

☐ 有　　☐ 沒有

9. 請問你今天/近幾天有沒有上呼吸道受感染的情況?
   ☐ 有 ☐ 沒有

10. 請問你有否吸煙?
   ☐ 有 ☐ 沒有
   如有，吸煙的次數及數量是:________________________________

11. 你是否在嘈雜的環境下工作?
   ☐是 ☐不是

12. 平日工作或消閒時，你是否須要大聲說話?
   ☐是 ☐不是

13. 廣東話是否你的母語? ☐是 ☐不是

個人資料:

姓名: ____________________ 性別:________________________
聯絡電話:____________________
出生日期:____________________ 年齡:__________________________

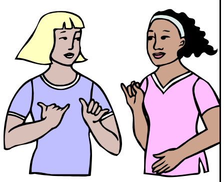
Appendix 3- word stimuli and presentation of stimuli

Word stimuli

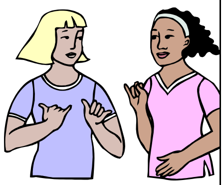
Tone					
1	2	3	4	5	6
丫(丫叉)	啞(啞巴)	亞(亞洲)	芽(芽菜)	瓦(瓦片)	硬 (硬幣)
鴉 (鴉片)	啞(啞鈴)	亞(亞軍)	牙(牙齒)	雅(雅典)	掗(掗拃)

Presentation of stimuli- Stimulus: /Na/

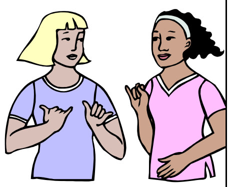
Trial 1

					
丫叉	亞洲	亞軍	瓦片	掗拃	芽菜
					
牙齒	啞鈴	啞巴	鴉片	硬幣	雅典

Trial 2

					
亞軍	雅典	亞洲	瓦片	牙齒	芽菜
					
掗拃	啞鈴	丫叉	鴉片	啞巴	硬幣

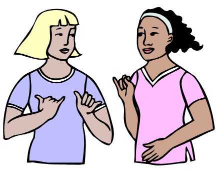
Trial 3

					
硬幣	亞洲	掗拃	丫叉	牙齒	鴉片
					
啞鈴	芽菜	雅典	亞軍	啞巴	瓦片

Trial 4

					
啞鈴	芽菜	亞洲	丫叉	硬幣	瓦片
					
牙齒	啞巴	雅典	亞軍	鴉片	揠拃

Trial 5

					
亞洲	丫叉	啞鈴	瓦片	硬幣	芽菜
					
啞巴	鴉片	揠拃	亞軍	雅典	牙齒