<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Lexical tone production in Cantonese alaryngeal speech</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Other Contributor(s)</strong></td>
<td>University of Hong Kong</td>
</tr>
<tr>
<td><strong>Author(s)</strong></td>
<td>Chow, Tak-yu, David; 周德裕</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>1998</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/56389">http://hdl.handle.net/10722/56389</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.; The author retains all proprietary rights, such as patent rights and the right to use in future works.</td>
</tr>
</tbody>
</table>
Lexical Tone Production in Cantonese Alaryngeal Speech

Chow Tak Yu, David

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, April 30, 1998
Abstract

This study investigated the acoustic properties of six Cantonese contrastive tones produced by three groups of alaryngeal speakers: four pneumatic, seven tracheoesophageal (three using the Blom-Singer type and four using the Provox type), and five normal laryngeal speakers. The six contrastive tones of the Cantonese word /si/ were used. The duration of vowel, the average fundamental frequency, the range of fundamental frequency and the slopes of initial and final portions of the vowel produced by different groups were measured. Results showed that the average fundamental frequency of alaryngeal groups was significantly higher than that of the normal laryngeal group. The range of fundamental frequency produced by pneumatic group was smaller than that of the normal laryngeal and tracheoesophageal groups. The patterns of the initial and final slopes of the vowel produced by the alaryngeal groups were roughly similar to that of normal laryngeal group.
Introduction

Background

Laryngectomy is a surgery that involves the removal of all or part of the larynx because of structural abnormalities. After the surgery, the alaryngeal patients no longer have the vocal folds, which means they cannot use their original anatomical and physiological mechanism to produce voices. Therefore, they have to make use of some compensatory methods to speak. There are four main types of alaryngeal speech, which are the esophageal, tracheoesophageal, pneumatic and electrolaryngeal speech. The mechanism of esophageal speech is using an inhalation method (voluntarily open the pharyngoesophageal (PE) segment and draw air from the mouth and pharynx) or injection method (force the air in the mouth and pharynx into the PE segment) to store about 80 cubic centimeters of air (van den Berg & Moolenaar-Bijl, 1959) in the upper part of esophagus and releases the air back to the oral cavity. During the process, the vibration of the PE segment produces a voice for the esophageal speech. The tracheoesophageal method is similar to the esophageal one, but a prosthesis is inserted into the fistula between the trachea and esophagus so that pulmonary air becomes the air source of voice. Pneumatic artificial larynx is a funnel-like device that consists of a reed or rubber membrane. The funnel diverts the pulmonary air from the stoma back to the oral cavity, and vibrates the reed or rubber membrane to produce the voice simultaneously. The electrolarynx is an electronic device that produces mechanical vibrations and transmits them to the vocal tract for speech production. (Casper & Colton, 1993)

Literature Review

In the past 20 years, many studies had been done on the different alaryngeal speech methods. Bridges (1990) showed no significant differences in the intelligibility of different
alaryngeal speakers, but Bennett & Weinberg (1973) found that the pneumatic speech received significantly higher ratings in the acceptability than other alaryngeal speech methods. Other studies (Gandour & Weinberg, 1983; Gandour & Wejinberg, 1984; Gandour, Weinberg & Petty, 1985) showed that esophageal and tracheoesophageal speakers could produce intonation and contrastive stress with acoustic properties similar to that of normal laryngeal speakers. Acoustic studies (Robbins, 1984; Robbins, Fisher, Blom & Singer, 1984; Robbins, Christensen & Kempster, 1986; Qi & Weinberg, 1995) also showed the acoustic properties of different alaryngeal speech methods. However, those studies were all English-based. As a result, the findings obtained may not be totally applicable in Hong Kong, which Cantonese is the main spoken dialect.

Cantonese is a tone language. Differences in tones can bring a change in the meaning. According to Fok Chan (1974), there are 6 contrastive lexical tones, which are the ‘Upper even’ (Tone 1), ‘Upper ascending’ (Tone 2), ‘Upper departing’ (Tone 3), ‘Lower even’ (Tone 4), ‘Lower ascending’ (Tone 5) and ‘Lower departing’ (Tone 6). Different tones are mainly characterized by different fundamental frequency contours (Fok Chan, 1974). As the tonal differences cause a change in the meaning, the measuring parameters of speech intelligibility in the English-based studies may not be adequate for Cantonese alaryngeal speakers.

In 1990’s, some perceptual studies on the speech intelligibility of Cantonese alaryngeal speakers have already been done. Yiu, van Hasselt, Williams & Woo (1994) showed that esophageal speakers received the highest rating in speech intelligibility, but pneumatic and tracheoesophageal speech could convey tonal contrasts most effectively. Ching, Williams & van
Hasselt (1994) also found that pneumatic speech could convey lexical tones most successfully, and the second best group was the esophageal speech. However, Ng, Kwok & Chow (1997) showed that all kinds of alaryngeal speech demonstrated no significant difference in speech intelligibility, which is contradictory to the findings of Yiu et al. (1994). The differences in the conclusions of these studies may be due to the differences in methodology, stimuli, number of subjects and subjects selection criteria. For example, in the studies of Ching et al. (1994) and Ng et al. (1997), only superior alaryngeal speakers were recruited, but there was no particular selection criteria in the study of Yiu et al. (1994). The study of Yiu et al. (1994) involved fifty-three alaryngeal speakers, while Ching et al (1994) had only nine subjects. Nevertheless, a conclusion can still be drawn from the studies, that lexical tones were produced most effectively by pneumatic speakers, and least successfully by the electrolaryngeal group.

A study of lexical tone production of alaryngeal speakers of another tone language—Thai was done by Gandour, Weinberg, Petty & Dardarananda (1988). Both perceptual and acoustic analyses were used in their study. It was shown that an esophageal speaker could produce tone differences that could be identified by normal subjects with 70% accuracy, but the tones produced by an electrolaryngeal speaker were only identified with about 20% accuracy. The acoustic analysis showed that the esophageal speaker produced the tones with a fundamental frequency range of 20 semitones, which was larger than that of the normal speakers (about eight semitones). Besides, the fundamental frequency contours produced by esophageal speakers for the same tone were also found inconsistent. For the electrolaryngeal speech, systematic change of fundamental frequency in tone production was not found, which supported to the results of the perceptual test in the same study. The study showed the importance of acoustic analysis in the
study of alaryngeal speech methods in tone language. It tried to reveal the acoustic features of different alaryngeal speech methods and provide possible explanations to their performance on the tone production ability. However, the study investigated only the esophageal and electrolaryngeal speech methods. Besides, there were only two subjects in the esophageal group, and one subject in the electrolaryngeal group, which the number of subjects were obviously too few to draw a conclusion.

**Aim of Study**

This study is going to investigate the acoustic characteristics of the tones produced by Cantonese alaryngeal speakers. In the study, tracheoesophageal (using Blom-Singer and Provox types) and pneumatic speakers were chosen because past studies (Yiu, van Hasselt, Williams & Woo, 1994; Ching, Williams & van Hasselt, 1994) showed that they could produce lexical tones perceptually with relatively high percentage accuracy. The study of the acoustic characteristic of the speech methods may also provide objective evidence to their tone production abilities. The esophageal group was not included because of the difficulties of the subject recruitment. Electrolaryngeal group was not chosen because past studies (Gandour & Weinberg, 1984; Gandour, Weinberg, Petty & Dardarananda, 1988) showed that electrolaryngeal speakers could not produce a change in pitch and tone acoustically and perceptually.

Four dimensions of the acoustic characteristics of the alaryngeal speech were measured: the duration of vowel, the average value of fundamental frequency, the range of fundamental frequency and the change of fundamental frequency. It was hypothesized that:
1. The duration of the vowels produced by tracheoesophageal speakers was longer than that of normal laryngeal speakers (Robbins, Christensen & Kempster, 1986).

2. The average fundamental frequency produced by tracheoesophageal speakers was similar to that of normal laryngeal speakers (Robins, Fisher, Blom & Singer, 1984).

3. The alaryngeal (pneumatic and tracheoesophageal) speakers had a restricted range of fundamental frequency because of the anatomical and physiological limitation.

4. The tracheoesophageal and pneumatic speakers could produce changes in the fundamental frequency contour, but inconsistently (Gandour, Weinberg, Petty & Dardarananda, 1988).

Method

Subjects

Four groups of speakers were selected in the study: the pneumatic, tracheoesophageal (1)—using Blom-Singer prosthesis, tracheoesophageal (2)—using Provox prosthesis, and the normal laryngeal groups. There were totally four Pneumatic speakers, three Blom-Singer speakers, six Provox speakers and five normal laryngeal speakers involved in the study (Appendix A).

All the speakers were male Chinese, and aged between 43 to 73 (mean = 62, SD = 9.03). They were all volunteers, native speakers of Cantonese and able to read Chinese. For the normal laryngeal group, subjects had no speech and language problem. For the alaryngeal speech groups, subjects had no speech and language problem prior to the laryngectomy, were tested at least nine months post-surgery, and had at least three months experience in using the speech method. Both types of tracheoesophageal speakers were using the digital occlusion method.
(using the finger to occlude the stoma to divert the pulmonary air to the PE segment during speech). The pneumatic speakers were using the Taiwan Tube. All the alaryngeal speakers were recruited from the patients of the Speech Therapy Department of the Queen Mary Hospital and the Pamela Youde Nethersole Eastern Hospital.

**Procedure**

**Stimuli**

The following set of 6 Cantonese words that only differ by tone was used.

/si/ 詩 /si/ 史 /si/ 試 /si/ 時 /si/ 市 /si/ 是

‘Poem’ ‘History’ ‘Test’ ‘Time’ ‘Market’ ‘Yes’

The target words were embedded in a carrier phrase: ‘呢個係X字’ (This is a ‘X’). The stimuli were printed on white cards and the size of each character was 3 cm X 3 cm. The word /si/ was selected because the fricative /s/ before the vowel /i/ could enhance the identification of the vowel for performing the acoustic analysis.

**Recording Procedure**

Subjects were informed about the aim of the study, and signed on a standardized consent form. They were questioned according to a standardized question list so as to make sure each subject suited the criteria of the study. The recordings took place in a quiet room. The noise level of the room was measured with a Quest Electronics 215 Sound Level Meter, and was below 40dBA for each recording. Recordings were made on SONY Digital Audio MiniDiscs with a SONY MZ-R50 MD Digital Recorder (sampling rate: 44.1kHz). An AKG C525S microphone was used to collect the acoustic signals, and was placed at a distance of 15 cm from the mouth of
the speakers. Subjects were required to read out a list of stimuli. There were five trials for each target word, and the stimuli were presented in random order. Therefore, each subject was required to read aloud 30 written stimuli. Before the recording, the experimenter would speak out on the six target stimuli. Then, the subjects were required to produce six unrecorded trials so that the experimenter could make adjustment on the intensity level of the recording.

Data Analysis

Acoustic Analysis

The recordings were sampled at 44.1kHz and stored in sound files by using the software SoundDesigner II. The Soundscope software was then used for the acoustic analysis of the speech samples. The vowels of the target words were identified by inspecting the wide-band spectrogram and the waveform plot display, and edited for the analysis of the fundamental frequency. The portion of the vowel /i/ was selected from the third cycles to the last third cycles of the vowel. Then, the fundamental frequency contours were plotted by using the Peak-Picking or the Autocorrelation method. The Peak-Picking was the main method used. For the recording which a continuous fundamental frequency contour could not be plotted, the Autocorrelation method would then be used. The estimated fundamental frequency was manually checked by reviewing the amplitude waveform plot display to prevent miscalculation. The time values and the fundamental frequencies of the initial, middle and the final points of the vowel were then taken for the following measurements.

Four acoustic characteristics were investigated: the duration of vowel, the average value of fundamental frequency, the range of fundamental frequency and the change of fundamental
frequency. The duration of vowel was obtained by calculating the time difference between the initial point and the final point of the vowel. The average value of fundamental frequency was obtained by using the program in the Soundscope. The range of fundamental frequency was obtained by calculating the difference between the maximum and the minimum value of the average fundamental frequencies of each speaker. The change of fundamental frequency was obtained by calculating the average slope (Hz/s) of the first half of and the second half of the fundamental frequency contour.

**Data Reduction**

The speech samples of two Provox speakers were excluded from the acoustic analysis because the speech consists of a large amount of noise, which made the identification of the position of the vowel and the analysis of the fundamental frequency impossible. Figure 1 showing the amplitude waveform plot display of an excluded Provox speech, which the peaks of the waveform were inconspicuous and the waveform of the cycles were inconsistent.

*Figure 1. Amplitude waveform plot display of an excluded Provox speech*
Reliability

The reliability of selection of the data points in vowel identification was measured. Ten percent of the speech samples were reanalyzed after three weeks by the same experimenter. The points selected in the vowel identification were compared to the previous data. Total agreement was found on seven percent of the reanalyzed stimuli, and fifty-five percent of the samples were within the range of ± five mini-seconds. All points were within the range of ± twenty mini-seconds.

Statistical Analysis

Data obtained in the acoustic measurement was analyzed by using a one-way (types of speech) Kruskal-Wallis analysis of variance (ANOVA) non-parametric test.
Results & Interpretation

Duration

The average duration of vowels of the normal laryngeal group was 0.320 sec. \((SD = 0.067)\), pneumatic group was 0.365 sec. \((SD = 0.088)\), Blom-Singer was 0.387 sec. \((SD = 0.135)\), and Provox was 0.353 sec. \((SD = 0.080)\). All the alaryngeal groups produced slightly longer duration of vowels than that of normal laryngeal group, but the difference is not statistically significant. (Kruskal-Wallis ANOVA (Overall effect): \(x^2 = 0.533, df = 3, p < 0.912\)). Figure 2 shows the average duration of vowel /i/ produced by the four groups.

*Figure 2. Average Duration of Vowels Produced by the Four Groups of Speakers*
Average Fundamental Frequency

The average fundamental frequency of normal laryngeal group was about 114 Hz ($SD = 12.880$), pneumatic group was about 152 Hz ($SD = 16.865$), Blom-Singer group was about 175 Hz ($SD = 25.534$) and Provox group was about 184 Hz ($SD = 15.429$). The average fundamental frequencies of the alaryngeal groups (pneumatic, Blom-Singer and Provox) were significantly higher than that of normal laryngeal group. (Kruskal-Wallis (Overall effect): $x^2 = 9.333$, $df = 3$, $p < 0.025$; (Normal Vs Pneumatic): $x^2 = 9.000$, $df = 1$, $p < 0.002$; (Normal Vs Blom-Singer): $x^2 = 4.800$, $df = 1$, $p < 0.029$; (Normal Vs Provox): $x^2 = 9.000$, $df = 1$, $p < 0.003$) However, there was no significant difference of average fundamental frequency between the three alaryngeal groups. Figure 3 shows the average fundamental frequency produced by the four groups. Consistent pattern across the tones could be found, which the Blom-Singer group produced highest average fundamental frequency across the tones, and followed by the Provox, pneumatic and normal laryngeal group.

Figure 3. The Average Fundamental Frequency Produced by the Four Groups of Speakers
Range of Fundamental Frequency

The mean range of fundamental frequency of normal laryngeal group was about 71 Hz ($SD = 16.709$), pneumatic group was about 30 Hz ($SD = 12.803$), Blom-Singer group was about 80 Hz ($SD = 10.853$) and Provox group was about 65 Hz ($SD = 21.796$). It was found that the range of fundamental frequency produced by the pneumatic group was the smallest, and the range of fundamental frequency of normal laryngeal group is similar to that of Provox and Blom-Singer group. However, no significant difference between the groups was found (Kruskal-Wallis ANOVA: $x^2 = 7.200$, $df = 3$, $p < 0.066$). The $p$ value (0.07) was close to the significant value (0.05), which reflected a high possibility that the range of fundamental frequency of pneumatic group was different from the other three groups. The small number of subjects (only three) in the pneumatic group may affect the accuracy of obtaining the $p$ value.

Fundamental Frequency Contour

The individual variation within the group was large and it was difficult to use the statistical analysis to investigate the effect of types of speech on the slopes of tones. Kruskal-Wallis one-way ANOVA showed no significant difference of the slopes between different groups. However, it was found that the standard deviation of the slopes of disordered group was much larger than that of normal laryngeal speakers (Pneumatic: $SD = 84.93$, Blom-Singer: $SD = 96.28$, Provox: $SD = 124.66$, Normal laryngeal: $SD = 36.66$). The average slopes of the first and second half sections of the fundamental frequency contours was measured and compared. Figure 4 and 5 representing the first and second slope of the fundamental frequency produced by the four groups.
Figure 4. The First Slope of the Fundamental Frequency Produced by the Four Groups of Speakers

![First Slope Graph]

Figure 5. The Second Slope of the Fundamental Frequency Produced by the Four Groups of Speakers

![Second Slope Graph]
The directions of the initial and final slopes for the tones produced by normal laryngeal and tracheoesophageal group are roughly similar in shape for the various tones. When considering the first slope, normal laryngeal group generally produced a more positive value than the tracheoesophageal groups across the tones. For the second slope, the values of the normal laryngeal group were similar to that of the tracheoesophageal one. Particular discrepancy was found on the first slope of tone one produced by Provox group and the first slope of tone six produced by the Blom-Singer group. Besides, it was found that the pattern of fundamental frequency contour of the pneumatic group was different from the other three groups. The pneumatic speakers produced all the tones with a roughly increasing trend in the first slope and then followed by a decreasing trend in the second slope. Figure 6, 7, 8 and 9 shows the typical fundamental frequency contours of tone one produced by normal laryngeal, pneumatic, Blom-Singer and Provox groups respectively. The fundamental frequency contours of other tones produced by the four groups were shown in Appendix B, C, D and E. A relatively large jitter ratio was found in the tracheoesophageal (Blom-Singer & Provox) speech.

Figure 6. The Fundamental Frequency Contour of Tone One Produced by a Normal Laryngeal Speaker
Figure 7. The Fundamental Frequency Contour of Tone One Produced by a Pneumatic Speaker

Figure 8. The Fundamental Frequency Contour of Tone One Produced by a Blom-Singer Speaker
Discussion

Duration

The result shows that there was no significant difference between the vowel duration of different types of speech. It does not agree with the hypothesis of this study, which expected that the duration of vowel produced by tracheoesophageal speakers would be significantly longer than that produced by normal laryngeal speaker. It also contradicts with the findings of Robbins, Christensen & Kempster's (1986), which found the tracheoesophageal speakers produced significantly longer vowel than normal subjects. The discrepancy may be due to the small subject pool of tracheoesophageal speakers in this study, whereas they had recruited fifteen subjects for each group in their study. There may be also possible that the cross-linguistic and cross-cultural differences, which both Cantonese normal laryngeal and alaryngeal subjects need or prefer to produce an average longer vowel in order to carry out the lexical tone, and the average duration does not affect by the types of speech.
According to Fok Chan (1975), the duration of the vowel may be used for speaker identification, but not tonal identification. It was also found that the average duration of individual speakers largely differ within normal laryngeal group, in which one of the subject produced average vowel duration of 0.232s and another subject within the group with 0.410s. It was suggested that it was the speaking styles of individuals affect the duration of the vowel production. Although it could be possible that longer vowel can provide more time for the speaker to convey a change in fundamental frequency and produce a tone differences, the result of this study did not show any evidence that the duration of vowel produced by different types of speakers are significantly differ.

**Average Fundamental Frequency**

The findings of the average fundamental frequency of the alaryngeal speakers were quite interesting. Past studies (Robbins, 1984), (Gandour & Weinberg, 1984) and (Qi & Weinberg, 1995) has showed that the average fundamental frequency of male tracheoesophageal speech was about 100 Hz. However, the findings of this study showed that the average fundamental frequency of tracheoesophageal (both Blom-Singer and Provox) was about 179 Hz. One of the possible explanations is that the tracheoesophageal speakers produce a stronger air stream in order to produce a turbulence for the fricative /s/, and the excessive air stream cause the PE segment to vibrate more vigorously and produce a higher fundamental frequency.
Range of Fundamental Frequency

It was found that the range of fundamental frequency of pneumatic speakers is limited (about 30 Hz). The fundamental frequency of pneumatic artificial larynx is mainly determined by the vibration of the rubber membrane in the device, which can be controlled by adjusting the tension of the membrane before the using and the changing the breath pressure during the speech. It was suggested that the change of the breath pressure was limited by the respiratory support, and then restricted the range of vibration rate of the rubber membrane. For the tracheoesophageal speakers, the range of fundamental frequency was similar to that of normal laryngeal speakers. It shows that the PE segment of tracheoesophageal speakers provides them a large range of fundamental frequency.

Fundamental frequency contour

As past studies (Ching, Williams & Hasselt, 1994; Ng, Kwok & Chow, 1997; Yiu, van Hasselt, Williams & Woo, 1994) showed that tracheoesophageal and pneumatic speakers could produce Cantonese lexical tones perceptually, it is not surprising that the two groups of speakers could produce changes in fundamental frequency. The patterns of the initial and final slopes of the tones were similar between the tracheoesophageal (Blom-Singer & Provox), pneumatic and normal laryngeal groups, which provided an evidence to the tone production ability of the alaryngeal groups. The large variation found in tracheoesophageal groups and the pneumatic group may be due to the limitation in controlling their fundamental frequency. Particular high values of initial slope of tone one produced by Provox group and tone six produced by the Blom-Singer group were noted. Besides, the shape of the fundamental frequency contour of pneumatic speakers was consistently a hill like, which consists of a rising of fundamental frequency in the
first slope and followed by a falling of fundamental frequency in the second slope. This characteristic can be explained by the physical properties of the vibration of a rubber membrane, which the membrane takes a few ms to reach and leave its optimal frequency of vibration. Though the main trend of the fundamental frequency was described as above, minor differences were observed between different tones. It was suggested that the minor changes of fundamental frequency provided essential information for the perception of tones.

**Overall**

It was found that the pneumatic speech was characterized by a relatively high fundamental frequency (about 152 Hz) but with restricted range (30 Hz). The tracheoesophageal speech has high fundamental frequency (about 179 Hz) and also a large degree of inconsistency of fundamental frequency contour. The difference of vowel duration, average fundamental frequency, range of fundamental frequency and change of fundamental frequency between the Blom-Singer and Provox type was not significantly different. As different types of prosthesis require different air pressures to force open the slit (Weinberg & Moon, 1984), it is hypothesized that the different air pressures required may affect the strength of air stream through the PE segment and produce the differences in the average fundamental frequency, the range of fundamental frequency and other acoustic characteristics.

The results obtained in this study seemed to be contradictory to that of the perceptual studies on the tone production of alaryngeal speech. As the fundamental frequency range restricts the change of fundamental frequency, it was assumed that a larger range of fundamental frequency could provide a condition for larger changes of fundamental frequency, which may
enable the speaker to carry out the tone production more effectively. The result of this study shows the range of fundamental frequency of pneumatic speech was limited (about 30 Hz), and the tracheoesophageal speech has a range of fundamental frequency similar to that of normal (about 70 Hz). However, past studies showed that the pneumatic speech was the best kind of alaryngeal speech methods in producing Cantonese lexical tones perceptually (Ching, Williams & Hasselt, 1994; Yiu, van Hasselt, Williams & Woo, 1994). It was suggested that even though the fundamental frequency contours produced by pneumatic speakers for different tones are similar, the small changes in the fundamental frequency contour could still provide the essential information for the tone perception by others. Besides, the tracheoesophageal speech was characterized with a large jitter and shimmer ratio (Robin, Fisher, Blom & Singer, 1984; Trudeau & Qi, 1990). The inconsistency of fundamental frequency of tracheoesophageal speech may act as a distracting factor which interfered the perception of tone by the listeners, so the perceptual studies showed an inferior rank for the tone production ability of tracheoesophageal speakers.

Further Studies

As this study only involves the normal laryngeal, pneumatic and tracheoesophageal speakers, acoustic characteristic of tone production in esophageal speech has not been revealed yet. Besides, only male subjects were involved in the study. The difference in gender of alaryngeal speakers may have an effect on the acoustic characteristics of the various speech methods because of the differences in the length of vocal tract. In this study, only one set of Cantonese words /si/ was used. It will be interesting to know how the acoustic properties of the speech methods will be affected when different vowels and preceding consonants are used. Besides, as stated in the discussion, it was suggested that the inconsistency of fundamental
frequency of tracheoesophageal affected the ability of tone production. An acoustic
measurement of the jitter and shimmer ratio of the alaryngeal speech methods may be able to
provide objective evidence to this assumption. Finally, it would be interesting to carry out a
study using both perceptual and acoustic analysis with the same set of data, so as to investigate
the relationship between the ease of identification of the tone and the acoustic characteristics of
the alaryngeal speech methods.

Clinical Implications

The acoustic result showing that the fundamental frequency of both pneumatic and
tracheoesophageal speech was high, and confusion of the speaker’s gender may appear. This
may become an extra factor which needed to be consider when a new patient trying to make
decision on the types of alaryngeal speech. The range of the fundamental frequency of pneumatic
speech is particularly limited, extra teaching of the usage of the device may be needed in order to
maximize the use of the fundamental frequency range.

In a long-term aspect, the result reveals the acoustical limitation of the two alaryngeal
speech methods. Improvement on the design of the speech methods may be needed in order to
suit the particular characteristics of tone languages like Cantonese.
Acknowledgment

I would like to express my sincere thanks to all the staffs in the department, especially to Dr. Valter Ciocca for his kind supervision and advice on my study. Thanks are extended to Dr. Edwin Yiu for his helpful suggestion; Professor Paul Fletcher and Ms Tara Whitehill for their insightful comments during the initial stage of the study; and the technicians for their technical supports of the acoustic analysis.

I also want to thank all the subjects participated in this study, and the speech therapists of the Queen May Hospital and Pamela Youde Nethersole Eastern Hospital who kindly refer appropriate subjects to me.

Finally, special thanks are devoted to my classmates, especially to Vivian, who give me psychological support and helpful suggestions during this period of time.
References


### Appendix A

**Subjects**

<table>
<thead>
<tr>
<th>Number</th>
<th>Sex</th>
<th>Age</th>
<th>Type</th>
<th>Sub-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>M</td>
<td>71</td>
<td>Pneumatic</td>
<td>Taiwan</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>73</td>
<td>Pneumatic</td>
<td>Taiwan</td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>58</td>
<td>Pneumatic</td>
<td>Taiwan</td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>51</td>
<td>Pneumatic</td>
<td>Taiwan</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>70</td>
<td>Tracheoesophageal</td>
<td>Blom-Singer</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>64</td>
<td>Tracheoesophageal</td>
<td>Blom-Singer</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>49</td>
<td>Tracheoesophageal</td>
<td>Blom-Singer</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>61</td>
<td>Tracheoesophageal</td>
<td>Provox</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>52</td>
<td>Tracheoesophageal</td>
<td>Provox</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>69</td>
<td>Tracheoesophageal</td>
<td>Provox</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>71</td>
<td>Tracheoesophageal</td>
<td>Provox</td>
</tr>
<tr>
<td>1</td>
<td>M</td>
<td>67</td>
<td>Tracheoesophageal</td>
<td>Provox</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>65</td>
<td>Tracheoesophageal</td>
<td>Provox</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>43</td>
<td>Normal Laryngeal</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>63</td>
<td>Normal Laryngeal</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>63</td>
<td>Normal Laryngeal</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>56</td>
<td>Normal Laryngeal</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>68</td>
<td>Normal Laryngeal</td>
<td></td>
</tr>
</tbody>
</table>
Appendix B

Fundamental Frequency Contour of Tone One Produced by a Normal Laryngeal Speaker

Fundamental Frequency Contour of Tone Two Produced by a Normal Laryngeal Speaker

Fundamental Frequency Contour of Tone Three Produced by a Normal Laryngeal Speaker
Appendix B (continued)

*Fundamental Frequency Contour of Tone Four Produced by a Normal Laryngeal Speaker*

![Graph showing the fundamental frequency contour of Tone Four.](image)

*Fundamental Frequency Contour of Tone Five Produced by a Normal Laryngeal Speaker*

![Graph showing the fundamental frequency contour of Tone Five.](image)

*Fundamental Frequency Contour of Tone Six Produced by a Normal Laryngeal Speaker*

![Graph showing the fundamental frequency contour of Tone Six.](image)
Appendix C

Fundamental Frequency Contour of Tone One Produced by a Pneumatic Speaker

Fundamental Frequency Contour of Tone Two Produced by a Pneumatic Speaker

Fundamental Frequency Contour of Tone Three Produced by a Pneumatic Speaker
Appendix C (continued)

Fundamental Frequency Contour of Tone Four Produced by a Pneumatic Speaker

Fundamental Frequency Contour of Tone Five Produced by a Pneumatic Speaker

Fundamental Frequency Contour of Tone Six Produced by a Pneumatic Speaker
Appendix D

Fundamental Frequency Contour of Tone One Produced by a Blom-Singer Speaker

Fundamental Frequency Contour of Tone Two Produced by a Blom-Singer Speaker

Fundamental Frequency Contour of Tone Three Produced by a Blom-Singer Speaker
Appendix D (continued)

Fundamental Frequency Contour of Tone Four Produced by a Blom-Singer Speaker

Fundamental Frequency Contour of Tone Five Produced by a Blom-Singer Speaker

Fundamental Frequency Contour of Tone Six Produced by a Blom-Singer Speaker
Appendix E

Fundamental Frequency Contour of Tone One Produced by a Provox Speaker

Fundamental Frequency Contour of Tone Two Produced by a Provox Speaker

Fundamental Frequency Contour of Tone Three Produced by a Provox Speaker
Appendix E (continued)

Fundamental Frequency Contour of Tone Four Produced by a Provox Speaker

Fundamental Frequency Contour of Tone Five Produced by a Provox Speaker

Fundamental Frequency Contour of Tone Six Produced by a Provox Speaker