Title	**Acoustic measure of fundamental frequency during three speech tasks in vocally healthy children**
Author(s)	**Lam, Lai-na; 林麗娜**
Citation	
Issued Date	**2011**
URL	**http://hdl.handle.net/10722/192885**
Rights	**The author retains all proprietary rights, (such as patent rights) and the right to use in future works.**

Acoustic measure of fundamental frequency during three speech tasks in vocally healthy children

Student number: 2007238797

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, 2011.

Abstract

The present study examined the fundamental frequency (F0) during three speech tasks in a group of vocally healthy children. The study also compared the reliability of different speech tasks for eliciting F0. Fifty-six vocally healthy children (31 boys and 25 girls) between the ages of 7.0 and 10.11 years participated in this study. Each child completed three speech tasks used to elicit a voice sample for subsequent analysis of fundamental frequency (F0). The tasks included: (a) sustained vowel /a/ prolongation, (b) repeating a sentence, and (c) reading aloud a passage. Two types of reliability, between-trial and between-day reliability, were compared across speech tasks. Results revealed significant difference in F0 between the three speech tasks (p = 0.01). Post hoc comparisons revealed that vowel task elicited significantly higher F0 values than the passage task. Passage reading task yielded the highest intra-class correlation coefficient values for both between-trial and between-day reliability. The results provide some empirical data for standardizing voice assessment protocol for school-age children.

Key words: fundamental frequency; speech task effect; between-trial reliability; between-day reliability; Cantonese-speaking children

Introduction

Acoustic analysis is a non-invasive and quantifiable method commonly used as part of voice assessment protocol (Hunter, 2009). Fundamental frequency (F0) is a commonly reported acoustic parameter in voice research (Guimaraes & Abberton, 2005). F0 refers to the rate of vocal folds vibration, and it is affected by length, mass and tension of vocal folds during speech production. The contraction or relaxation of vocal fold muscles causes changes of their tension resulting in a change in F0. Mean F0 reflects habitual pitch and helps to determine the appropriateness of frequency levels according to one's age and sex (Colton, Casper & Leonard, 2006).

The literature has documented that speaking F0 obtained from different speech tasks can vary significantly. Fitch (1990) investigated the consistency of speaking F0 and frequency perturbation across speech tasks in normal young adults. Results revealed that speaking F0 obtained from sustained vowel prolongation was significantly higher than those obtained from passage reading and connected speech tasks. Guimaraes and Abberton (2005) reported similar speech task effects. In their study, speaking F0 values were obtained from 109 adult speakers of Portuguese using different speech tasks. The results indicated that speaking F0 obtained from sustained vowel prolongation was significantly higher than that obtained from connected speech. Murry, Brown and Morris (1995) also examined the patterns of F0 for three types of voice samples in two age groups of normal male and female speakers and a vocal fold paralysis group. The results of the study revealed that there were F0 variations in the normal speakers. Young men produced reading and extemporaneous speech at a significantly higher F0 than their sustained phonation. Women of two age groups and older men produced reading and extemporaneous speech at a significantly lower F0 than their sustained phonation. For vocal fold paralysis group, they produced sustained phonation with a higher F0 than their reading or extemporaneous samples. Brown and Shrivastav (2007)

investigated the comfortable effort level in pre-school children's speech. Thirty boys and girls aged between 3 and 4 years were required to produce three speech tasks (i.e., repeat a vowel, a sentence and four English words). The results revealed a significant speech task effect on speaking F0. Speaking F0 for phonation of the vowel was significantly lower than that of the sentence and one of the words. A recent study examining the effects of task type on F0 in children reported that speech tasks could significantly influence speaking F0 values in young children aged between 5.0 and 7.11 years (Baker, Weinrich, Bevington, Schroth, & Schroeder, 2008). In that study, four speech tasks were examined namely sustaining vowel /a/ prolongation, sustaining the vowel embedded in a word at the end of a phrase, repeating a sentence, and counting from 1 to 10. Counting was found to elicit a significantly higher speaking F0 value than phrase and sentence tasks. However, their study did not include the speech task of connected speech, which may yield more representative results of habitual pitch than those obtained during sustained phonation.

Apart from speech tasks, speaking F0 can vary in different linguistic and racial groups. Awan and Mueller (1996) compared the speaking F0 of kindergartners of different racial groups (White, African American and Hispanic groups). Results indicated that there were significant differences between racial groups on the measure of mean speaking F0. The Hispanic kindergartners have higher mean speaking F0 than the African American children. The authors speculated that the variations in mean speaking F0 values might be related to anatomical, linguistic and societal differences among different racial groups. O'Neil, Jones and Nye (1997) compared the acoustic characteristics between Arabic, African American and Caucasian children. Their results revealed that the Arabic group produced a higher F0 than the other two groups of children in vowel task. The authors proposed cultural influence as the possible explanation for the differences of F0 among three racial groups. Fitch and Holbrook (1970) examined the modal vocal fundamental frequency of young adults. The authors

suggested that African American adults elicited a lower F0 and a greater speaking F0 range than that of Whites was due to larger body size. Graddol and Swann (1983) also supported the hypothesis of body size influence. In addition, Hirano, Kurita and Nakashima (1981) and Kahane (1982) found that the length of vocal folds in Caucasian Americans was longer than that of Japanese adults. Xue and Hao (2006) compared vocal tract dimensions of White American, African American, and Chinese male and female speakers. Vocally healthy adult speakers from three racial groups participated in the study. The study results revealed that race was an important factor affecting oral volume and total vocal tract volume. Mean oral volume of White American speakers were significantly smaller than Chinese speakers; while mean oral volume and mean total vocal tract volume of African American speakers were significantly smaller than Chinese speakers. It was concluded by the authors that there were differences in the patterns of vocal tract configurations among speakers from different racial groups. It was possible that the difference in vocal tract dimensions might result in different speaking F0 for speakers with different races. Apart from anatomical and racial differences, language was reported to have an influence on fundamental frequency. Chen (2005) studied the effect of tones on speaking F0 in Mandarin and Min dialects. Being tonal languages, Mandarin and Min were found to have greater maximum range of speaking F0 than non-tonal language of English. Cantonese is a tonal language; it is likely that the speaking F0 in Cantonese-speaking children is not comparable to that of English-speaking pediatric population.

Despite the fact that the acoustic characteristics of English-speaking pediatric population were well documented, English pediatric normative database demonstrates limited usefulness in the assessment of Cantonese children with voice disorders due to the abovementioned cultural, anatomical and linguistic differences on F0. There is a lack of normative database of Cantonese child population. In order to utilize acoustic analysis for

identifying dysphonic voices in Cantonese-speaking children, further studies are needed to develop normative database of Cantonese pediatric voice production.

Establishing the reliability of an assessment procedure is essential for an accurate assessment of voice conditions. The reproducibility of values in repeated measurements on the same subjects is referred as reliability (Hopkins, 2000). Repeated measures of analysis of variance (ANOVA) and t-tests are often used to compare the variations of F0 across experimental sessions. Intra-class correlation coefficient (ICC) indicates the correlation between two measurements within the same individuals (Bartlett & Frost, 2008) and it is a common measure used to evaluate the reliability (Vogel, 2011). ICC was used to evaluate the test-retest reliability in a study examining the reliability and sensitivity to change of acoustic measures of voice quality (Carding, Steen, Webb, Mackenzie, Deary, & Wilson, 2004). In addition, ICC was used to determine the inter-rater, between-days and trials reliability in a study investigating the reliability of maximum phonation time (Speyer, Bogaardt, Passos, Roodenburg, Zumach, Heijnen, Baijens, Fleskens, & Brunings, 2010). Several studies have investigated variations of F0 elicited from various speech tasks in vocally healthy individuals. However, the results were inconclusive. Brown, Murry and Hughes (1976) reported that sustained vowels had greater within-day F0 variation than short phrases. Garrett and Healey (1987) also examined the fluctuations of voices of twenty normal adult speakers across three times a day. That study revealed that male participants showed a significant change in speaking F0 from morning to noon to late afternoon, while no significant changes were observed in female participants across these three experimental sessions. Brown, Morris and Murry (1996) examined the variability of F0 of voice samples within and across experimental sessions. Three age groups (young, middle-aged and elderly) of adults were requested to repeat vowel /a/, read a standard passage, and speak extemporaneously three times a day over a 3-day period. For vowel, there was significant difference of F0 for middle-aged women

across days while no significant difference of F0 was shown across days for younger and older women, and men within the three age groups. For reading task, only younger women exhibited a significant difference in F0 across days. For the speaking task, women within the three age groups showed significant differences in F0 across days, but not across sessions. Study by Brown et al. (1996) revealed that there was variation of F0 across experimental sessions, though the results might not be extensive and conclusive enough. However, Fitch (1990) did not found any significant test-retest data of speaking F0 for vowels, reading and spontaneous speech in a between-day measurement. In addition, there was another investigation reported that F0 was not affected by the time-of-day in which the voice samples were collected (Nittrouser, McGowan, Milenkovic, & Beehler, 1996). The results revealed that similar speaking F0 values were obtained from two sessions of the same day, with one in the morning and the other one in the afternoon. Brown et al. (2007) reported similar result as Fitch (1990) and Nittrouser et al. (1996) that there were no significant variations of F0 across three test sessions in a group of pre-school children. In order to find out a more reliable speech task to aid in accurate diagnosis, establish baseline before treatment and document progress after treatment, the reliability of speech tasks should be an area to be investigated.

In summary, the present study aimed to establish normative database of F0 in Cantonese children. The study also aimed to investigate the effects of speech tasks on F0 in young children. In this study, Cantonese-speaking children aged from 7.0 to 10.11 years were recruited as participants. Three types of speech task were employed including sustained vowel /a/ prolongation, repeating a sentence and passage reading. Sustained vowel /a/ prolongation was chosen as it is a quick and easy way to obtain speech samples. Connected speech (sentence and passage reading) was chosen as it can yield more representative results of habitual pitch. In this study, between-trial reliability and between-day reliability for vowel /a/, sentence and passage reading were examined in Cantonese-speaking children. Between-

trial reliability measurement evaluated the reliability within the five repeated trials of a speech task. Each participant was requested to repeat five samples for each speech task, and the five repeated speech samples were used for the evaluation of between-trial reliability. For between-day reliability measurement, the reliability of speech tasks across days was evaluated. Some of participants were requested to perform the recording on two occasions. Results from the study would provide empirical data for establishing the clinical voice assessment protocols for young children.

Methods

Participants

A total of 56 school-age children (31 boys and 25 girls) participated in the study. Their mean age was 8.36 years (SD = 0.98; range = 7.0 to 10.11 years) (Table 1). This age range was targeted to ensure the children have the ability to follow instructions and have the ability to read a standardized passage. In addition, this age range helped to avoid inclusion of children at puberty stage that are prone to voice changes. All children were recruited from a local primary school in Hong Kong.

All children were native speakers of Cantonese. Children with previous history of voice problems and oral structural abnormalities were excluded. Children with speech and language problems were also excluded. All children were confirmed to be vocally healthy by the researcher and an experienced speech therapist. In addition, no signs of upper respiratory infections were shown in participants and participants reported themselves to be free from colds or other upper respiratory problems.

Table 1. Age and gender distribution of participants.

Age (years)	Boys	Girls	Total
7.0-7.11	6	7	13
8.0-8.11	11	6	17
9.0-9.11	8	11	19
10.0-10.11	6	1	7
Total	31	25	56

Procedures

All experimental procedures were carried out in a quiet room at the participants' school with background noise kept around 45 dB throughout the recording. A head-mounted professional-grade, condenser microphone (AKG Acoustics C420, Vienna, Austria) was worn by each participant and was connected to a digital recorder (H4next Handy Recorder, ZOOM Corporation, Japan) to record voice samples. The microphone was placed and maintained 5 cm from the child's mouth corner throughout the recording.

Each participant was asked to carry out three speech tasks including sustained vowel prolongation, reading aloud a sentence and a passage. In the vowel prolongation task, each child was asked to sustain the vowel /a/ for five seconds. In the sentence task, each child was asked to read aloud a Cantonese sentence / ba_1 ba_1 da_2 $gɔ_1$ $gɔ_1$/ (meaning 'Father hits elder brother.'). In the passage task, each child was asked to read aloud a Cantonese passage of 說話不簡單 (meaning 'Speaking is not a simple task.'). The passage had around 100 characters and was selected from a local textbook of primary-one level. The passage was printed in an A4 paper with font size 24. The researcher explained and demonstrated each of the speech tasks. Participants were given the opportunity to practice the speech tasks before actual

recordings started. Participants were instructed to complete all speech tasks using their comfortable pitch and loudness level. All participants completed five trials for each speech task. Each recording session lasted for approximately 20 minutes.

Among all participants, 21 of them were randomly selected to perform the recording on a second occasion four days after the first recording was obtained. This was to evaluate between-day reliability of the speech tasks. These participants were required to perform the same experimental procedures as the first occasion. As a measure of intra-rater and inter-rater reliability of acoustic analysis, 14 participants (25% of the total participants) were selected randomly for reanalysis.

Data Analysis

The recorded voice samples were analyzed using Praat (Version 5.2.10). For each sustained vowel /a/ prolongation, the middle one second of the sample was used for analysis. For each sentence, the sample from the onset of the first word (i.e., /ba_1/) to the offset of the last word (i.e., /$gɔ_1$/) was included for analysis. The entire passage was included for analysis. The middle three (i.e., the second, third, and fourth) trials of each speech task were used for data analysis of mean F0.

Two-way mixed analysis of variance (ANOVA) was carried out with speech task as within-subject factor and gender as between-subject factor. Two-way mixed ANOVA was used to evaluate the interaction effect between speech task and gender, and the main effect of gender and speech task. Post hoc tests for speech task were completed with Bonferroni adjustment.

One-way repeated measures ANOVA was used to compare the F0 difference among the five repeated trials of each speech task. A paired sample T-test was used to examine the F0 difference between day 1 and day 5 measurement of each speech task (i.e., test-retest).

There were two data sets for measuring reliability of speech task. One of the data sets was the between-trial measurement; and the other data set was the between-day measurement. The between-trial measurement examined the reliability within the five repeated trials of each speech task, while between-day measurement examined the reliability of speech task over successive sessions. All five trials for each speech task were used to obtain the reliability of between-trial and between-day measurement. Intra-class correlation coefficient (ICC) was used to evaluate the reliability of between-trial between-day measurement. For between-day measurement, the five samples of each speech task per participant during one day were averaged. ICC was then used to establish the relationship of the average measurement over the 4-day interval. ICC values can range from 0 (that is, no correlation) to 1 (that is, perfect correlation). The larger the overall ICC value, the more reliable the speech task. An ICC value of greater than 0.8 suggests high degree of correlation or high reliability (Bough, Heuer, Sataloff, Hills, & Cater, 1996).

Pearson *r* correlations were calculated to establish the intra-rater and inter-rater reliability. The significance level for the Pearson *r* correlation tests was set at 0.05.

Results

Mean Fundamental frequency (F0) across speech types

Table 2 lists the mean fundamental frequency values by gender and by speech tasks. The overall mean F0 value of vowel (mean = 250.98 Hz) was the highest, while the overall mean F0 value of passage (mean = 240.44 Hz) was the lowest.

For all speech tasks, the mean F0 values of girls were higher than that of boys, with the greatest mean F0 difference (around 3 Hz) found in sentence task, and smallest mean F0 difference (around 0.65 Hz) found in passage task. The means for F0 comparing vowel and

passage were very consistent, with a 10-Hz difference for voice samples of both boys and girls.

Speech task by gender effect. Two-way mixed analysis of variance (ANOVA) revealed that there was no significant interaction effect of speech task by gender [$F(1.6, 86.39) = 0.14$, $p > 0.05$]. This indicated that gender did not have an influence on F0 variations across speech tasks.

Speech task effect. Mauchly's test of sphericity for within-subject factor of speech task was shown to be significant ($p < 0.001$), the assumption of compound symmetry was violated. Hence, the results of within-subject effects with Greenhouse-Geisser epsilon correction were reported. Analysis of variance for the effect of speech task revealed that there was a significant difference among the three speech tasks [$F(1.6, 86.39) = 5.29$, $p = 0.01$]. Bonferroni post hoc comparisons indicated that the mean F0 of vowel was significantly higher than that obtained from the passage task. There was no significant difference found between the mean F0 of vowel and that of sentence. In addition, the mean F0 of sentence was not significantly different from that of passage.

Gender effect. The main effect of gender was not significant [$F(1, 54) = 0.07$, $p > 0.05$] indicating that the mean F0 of girls was not significantly different from the mean F0 of boys of the same age group.

Table 2. Summary statistics of 56 participants (31 boys and 25 girls) for vowel, sentence and passage.

	Boys	Girls	Overall
	Mean (SD)	Mean (SD)	Mean (SD)
Vowel			
Mean F0 (in Hz)	250.75 (27.55)	251.26 (23.94)	250.98 (25.41)
Sentence			
Mean F0 (in Hz)	243.34 (23.91)	246.85 (25.74)	244.91 (24.58)
Passage			
Mean F0 (in Hz)	240.15(23.56)	240.81(32.03)	240.44 (27.40)

F0 difference in repeated trials and test-retest measurements

One-way repeated measures ANOVA was used to compare the F0 difference among the five repeated trials of each speech task. A paired sample T-test was used to examine the F0 difference between day 1 and day 5 measurements of each speech task (i.e., test-retest).

F0 difference among the five repeated trials. Mauchly's test of sphericity for within-subject factor was shown to be significant ($p < 0.001$), the assumption of compound symmetry was violated. Hence, the results of within-subject effects with Greenhouse-Geisser epsilon correction were reported. There were no significant differences in the F0 value among the five repeated trials for vowel task [$F(1.74, 95.43) = 1.85$, $p > 0.05$] and passage task [$F(2.21, 121.42) = 1.49$, $p > 0.05$]. There was significant difference shown in the F0 value in five repeated trials of sentence task [$F(2.70, 145.76) = 15.49$, $p < 0.001$]. Bonferroni post hoc comparisons revealed that there were significant difference in F0 between the first and the

third trial, the first and the fourth trial, the first and the fifth trial, the second and the fifth trial, the third and the fifth trial, and, the fourth and fifth trial in sentence task.

F0 difference between day 1 and day 5 measurements (test-retest). There were no significant differences in the F0 value of test-retest data in sustained vowel phonation (t = -0.30, p > 0.05), sentence (t = -0.67, p > 0.05) and passage (t = 0.07, p> 0.05).

Table 3. Summary statistics for between day measurement of 21 participants (12 boys and 9 girls) for vowel, sentence and passage sampled with a 4-day interval.

Mean F0 (in Hz)	Boys		Girls		Overall	
	Test	Retest	Test	Retest	Test	Retest
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Vowel	247.27 (16.34)	248.12 (29.35)	245.30 (30.13)	247.62 (27.17)	246.42 (22.60)	247.91 (27.73)
Sentence	233.85 (18.00)	236.85 (26.64)	235.01 (29.32)	236.62 (17.26)	234.34 (22.86)	236.75 (22.57)
Passage	233.38 (24.28)	230.43 (24.01)	227.45 (32.75)	230.80 (34.36)	230.84 (27.61)	230.59 (28.10)

Speech task reliability

Intra-class correlation coefficient (ICC) was used to evaluate the reliability of between-trial and between-day data. Table 4 lists the Intraclass correlation coefficient for between-trial reliability measurement of 56 participants and between-day reliability measurement of 21 participants.

Between-trial reliability. There were high ICC values of between-trial measurement for sentence task and passage task, with the ICC value of passage task being the highest (ICC value = 0.96) of the single measures with 95% confidence interval. The result indicated that both sentence and passage tasks produced reliable objective measurements for each participant.

Between-day reliability. Both sentence and passage tasks yielded high ICC values, with passage showed the highest coefficient (ICC value =0.85) of the single measures with 95% confidence interval.

Table 4. Intraclass correlation coefficient for between-trial reliability measurement of 56 participants (31 boys and 25 girls) and between-day reliability measurement of 21 participants (12 boys and 9 girls).

Intraclass correlation coefficient	vowel	sentence	passage
Between-trial reliability	0.59	0.87	0.96
Between-day reliability	0.57	0.76	0.85

Reliability of acoustic analysis

Table 5 lists the Pearson *r* correlations of the intra-rater and inter-rater reliabilities for each speech task. The average *r* value of intra-rater reliability was 0.999; the average *r* value of inter-rater reliability was 0.924. The results showed that both the intra-rater and inter-rater reliabilities for acoustic analysis of the three speech tasks were very high.

Table 5. Intra-rater and inter-rater reliabilities.

	Pearson *r* correlations		
	Vowel	Sentence	Passage
Intra-rater	1.00	0.999	1.00
Inter-rater	0.998	0.869	0.904

Discussion

Normative acoustic measures on F0 characteristics were better established in English-speaking adults and children. However, there was a lack of information on the F0 characteristics of Cantonese-speaking children across speech tasks. The present study was set out to achieve three objectives. The first objective was to develop a set of preliminary normative acoustic measures for Cantonese-speaking children. The second objective was to examine the effects of speech tasks on the F0 produced by school-age children. The third objective of this study was to evaluate the reliability of different speech tasks. Two types of reliability, between-trial and between-day, were compared across speech tasks of sustained vowel prolongation, sentence and passage reading.

Speaking fundamental frequency characteristics

The acoustic characteristics of mean F0 in the voices of Cantonese-speaking children were defined. The mean F0 values for each of the speech tasks were as follows: vowel (Mean = 250.98 Hz); sentence (Mean = 244.91 Hz); passage (Mean = 240.44 Hz). The mean F0

value of vowel was greater than that of sentence and passage reading. Kent (1976) indicated that the mean F0 of English-speaking boys and girls in the age range of six to ten were between 200Hz and 325 Hz. The mean F0 values of the participants in this study were within this F0 range.

Table 6. Summary of the present study and previous investigation on speaking F0 in different speech tasks.

Investigators	Language /Race	Age range of participants (years)	Sustained vowel phonation	Sentence	Passage reading
			Mean (in Hz)	Mean (in Hz)	Mean (in Hz)
Present study	Cantonese	7-10	Boys: 250.75 Girls: 251.26	Boys: 243.34 Girls: 246.85	Boys: 240.15 Girls: 240.81
Sorensen (1989)	English	6-10	Boys: 258.06 Girls: 286.94	(not examined)	Boys:254.42 Girls:274.74
Whiteside & Hodgson (1999)	English	6-10	Boys: 223.25-248.68 Girls: 226.90-248.27	Boys: 231.87-258.71 Girls: 241.01-258.71	(not examined)

When compared with the study of Sorensen (1989), Cantonese-speaking boys in this study demonstrated similar F0 values of sustained vowel phonation with English-speaking boys; while Cantonese-speaking girls demonstrated a lower F0 value of sustained vowel phonation than English-speaking girls (see Table 6). For passage reading, both Cantonese-

speaking boys and girls demonstrated a lower F0 than that of English-speaking children. When comparing the present study with the study of Whiteside and Hodgson (1999), both Cantonese-speaking boys and girls showed a higher F0 value of vowel than that of English-speaking children. For sentence, Cantonese-speaking children had similar F0 values as that of English-speaking children. The difference in F0 values of different racial groups might suggest that the acoustic characteristics of one racial group could not be compared with another racial group due to anatomical, linguistic and societal differences among different racial groups. Normative database of Cantonese-speaking children was established and could be used to detect various pediatric vocal pathological abnormalities and to monitor the effects of voice therapy in Cantonese pediatric population.

The present study indicated that the mean F0 of boys was not significantly different from the mean F0 of girls in the age range of seven to ten years. The result suggested that the acoustic characteristics of school-age children were uniform across all prepubescent boys and girls. The results of this study were in agreement with previous studies that no significant difference was observed in mean F0 of prepubescent boys and girls (Sorenson, 1989; Linders, Massa, Boersma, & Dejonckere, 1995; Baker et al., 2008). Literature suggested that there is a gradual decrease of F0 from 3 years of life to puberty (Kent, 1976; Baken & Orlikoff, 2000). There is a marked voice change during puberty and laryngeal growth causes voice changes. During puberty, there is an increase in vocal fold length and pharynx size (Colton et al., 2006). Different laryngeal growth rate of boys and girls during puberty results in differences in vocal fold vibration and hence differences in mean F0 values. Prepubescent boys and girls have similar larynx size and vocal fold length, and hence have similar mean F0 values. It is possible that the school-age children in this study have not yet reached their puberty stage and significant difference in mean F0 of their voice would not be expected between boys and girls.

Results of this study revealed a significant speech task effect on the speaking F0 of a group of school-age children. Sustained vowel phonation produced a significantly higher mean F0 than that of passage and it elicited a relative higher mean F0 than that of sentence. The result of this study agreed with several studies examining speech task effect on adult population. Previous study by Fitch (1990) suggested that the measurement of F0 varies as a function of speech sample analyzed, with the F0 of vowel higher than that of sentence and passage tasks. The result of this study also corroborated with the findings of Murry et al. (1995). The authors concluded that young men produced reading and extemporaneous speech at a significantly higher F0 than their sustained phonation, while young and older women and older men produced reading and extemporaneous speech at a significantly lower F0 than their sustained phonation. Brown et al. (2007) and Baker et al. (2008) have also reported speech task effect. Brown et al. (2007) investigated speech task effect in pre-school children aged between 3 and 4 years; while Baker et al. (2008) examined task type effect on children aged from 5 to 7 years old. It was concluded that speech task effect could be found both in adult and pediatric population.

According to Sorensen (1989), the mean F0 difference of speech tasks are influenced by the anatomy and physiology of laryngeal structure. When producing sustained vowel phonation, a greater tension was observed in laryngeal muscles, which would increase the rate of vocal folds vibration and result in a higher F0 value. When reading a sentence and passage, the tension in laryngeal muscles is reduced and results in a lower fundamental frequency values. In addition, Murry et al. (1995) suggested that the high F0 value of vowel might be because the participants had the feeling of "performing" when having sustained vowel phonation. Therefore, sustained vowel phonation might not provide an accurate representation of the individual's habitual F0; while F0 of reading might be more representative of the individual's habitual F0. As there was speech task effect on mean F0 of

participants in this study, the idea proposed by Baken et al. (2008) that the same speech task should be used to monitor voice changes over time was also applicable in Cantonese school-age children.

F0 difference in repeated trials and test-retest measurements across speech tasks

When the F0 difference was examined within the five repeated trials for each speech task, there was no significant difference shown in the F0 value in vowel task and passage task; while there was significant difference in the F0 value in sentence task. When the F0 difference of test-retest data was evaluated, no significant differences of F0 were shown in the test-retest data in all the three speech tasks. Similar results have been reported by Fitch (1990). In the study, the speaking F0 of vowel, reading and spontaneous speech tasks were recorded on different days, and the study result indicated that no significant difference was found in the test-retest data.

Between-trial and between-day reliabilities across speech tasks

Among the three speech tasks, sustained vowel phonation has the lowest intra-class correlation coefficient (ICC) values for both between-trial and between-day reliability measures, with ICC value of 0.59 and 0.57 respectively. Passage reading has the highest ICC values with ICC value of 0.96 for between-trial and 0.85 for between-day reliability measure. The ICC value of between-day reliability measurement was smaller than that of between-trial reliability measurement; this indicated that the variability of speech task increases as time interval increases. ICC value of passage reading was greater than 0.80 in both between-trial and between-day reliability measures.

Overall reliability of the three speech tasks

When evaluating the overall reliability of the three speech tasks in this study, both the F0 difference among the five repeated trials and F0 difference between day 1 and day 5 measurements would be considered. In addition, the between-trial and between-day reliabilities across speech tasks would also be taken into account. The speech task that has the smallest F0 difference (i.e., no significant difference in F0 values in this study) within the five repeated trials and between the test-retest data, and has the highest ICC values in both between-trial and between-day reliability measurements would be regarded as the most reliable speech task in this study. For vowel task, there was no significant difference of F0 value shown in both the five repeated trials and the test-retest data, but, it yielded the lowest ICC values in both between-trial (ICC value = 0.59) and between-day (ICC value = 0.57) reliability measurements. For sentence task, there was significant difference of F0 value shown in the five repeated trials while no significant difference of F0 value was shown in test-retest data; the ICC values was 0.87 and 0.76 for between-trial reliability and between-day reliability respectively. For passage task, there was no significant difference of F0 value shown in both the five repeated trials and the test-retest data, and, it yielded the highest ICC values in both between-trial (ICC value = 0.96) and between-day (ICC value = 0.85) reliability measures. To sum up, it is concluded that passage yielded the highest overall reliability in repeated measurements, and it should be regarded as a reliable speech task to aid diagnosis, establish baseline before treatment and document progress after treatment in voice evaluation.

Clinical implications

Some preliminary data of the acoustic characteristics of mean F0 in the voices of Cantonese-speaking children were defined and can be used to detect various pediatric vocal pathological abnormalities and to monitor the effects of voice therapy. This study also

developed some guidelines for establishing clinical voice assessment protocol for school-age children. As speech task effect was identified in children population, it is suggested that the same speech task should be used to evaluate and monitor voice changes. In addition, passage reading has proven to be a highly reliable measure among the three speech tasks, passage reading is recommended as the standard measure in clinical evaluation and treatment progress of voice disorders.

Limitations and directions for further studies

As these acoustic data were collected from vocally healthy children, generalization to clinical population may not be valid enough. The vocal performance and the acoustic characteristics of children with voice problems may be different from that of vocally healthy children. Hence, further research on establishing the database of acoustic characteristics in children with voice problems is recommended. It is further suggested that acoustic data collected from children with voice problems can be compared to that of vocally healthy children. By collecting both the acoustic data from vocally healthy children and that from children with voice problems, a comprehensive database of acoustic characteristics of Cantonese pediatric population can therefore be established.

In this study, speech task effect was observed in vocally healthy children, it is also worth investigating the speech task effect in children with voice problems. The comparison of the similarity and/or difference of speech task effect between vocally healthy children and children with voice problems can hence be made.

Moreover, further research can also examine the reliability of speech task in discriminating children with voice problems from vocally healthy children. The information gained would further provide valuable empirical data for standardizing voice assessment protocol for school-age children.

Conclusions

The purpose of the present study was to establish the normative data of F0 in vocally healthy Cantonese-speaking children. The acoustic characteristics of mean F0 in the voices of Cantonese-speaking children were defined. The mean F0 values of sustained vowel phonation, sentence and passage reading were obtained, with mean F0 of vowel being the highest among the three speech tasks. In addition, no gender effect was found in the mean F0 of prepubescent boys and girls in this study. The second aim of this study was to investigate the effect of speech tasks on F0 in young children, and it was concluded that speech task effect could be found in both adult and pediatric population. Hence, it is suggested that the same speech task should be used to evaluate and monitor voice changes. Furthermore, the reliability of speech tasks has been investigated. There were no significant differences in F0 value within the five repeated trials and between the test-retest data for vowel and passage task. However, passage task yielded the highest ICC values for both between-trial and between-day reliability measurements; while vowel task yielded the lowest ICC value in both measurements. The results indicated that passage reading was the most reliable measure in voice evaluation and should be recommended as the standard assessment measure in clinical practice.

Reference

Awan, S. N., & Mueller, P. B. (1996). Speaking fundamental frequency characteristics of White, African American, and Hispanic kindergartners. *Journal of Speech and Hearing Research, 39*, 573–577.

Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice* (2nd ed.). San Diego, CA: Singular Publishing Group.

Baker, S., Weinrich, B., Bevington, M., Schroth, K., & Schroeder, E. (2008). The effect of task type on fundamental frequency in children. *International Journal of Pediatric Otorhinolaryngology, 72*, 885-889.

Bartlett, J. W., & Frost, C. (2008). Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound in Obstetrics and Gynecology, 31*(4), 466-475.

Bough, D. Jr., Heuer, Jr. R. J., Sataloff, R. T., Hills, J. R., & Cater, J. R. (1996). Intrasubject variability of objective voice measures. *Journal of Voice, 10*, 166-174.

Brown, W. S., Morris, R., & Murry, T. (1996). Comfortable effort level revisited. *Journal of Voice, 10*, 299-305.

Brown, W. S., Murry, T., & Hughes, D. (1976). Comfortable effort level: an experimental variable. *Journal of Acoustical Society of America, 60*, 696-699.

Brown, W. S., & Shrivastav, R. (2007). Comfortable effort level in young children's speech. *Folia Phoniatrica et Logopaedica, 59* (5), 227-233.

Carding, P. N., Steen, I. N., Webb, A., Mackenzie, K., Deary, I. J., & Wilson, J. A. (2004). The reliability and sensitivity to change of acoustic measures of voice quality. *Clinical Otolaryngology, 29*(5), 538-544.

Chen, S. H. (2005). The effects of tones on speaking frequency and intensity ranges in Mandarin and Min dialects. *Journal of Acoustical Society of America, 117*, 3225-3230.

Colton, R., Casper, J., & Leonard, R. (2006). *Understanding Voice Problems: A physiological perspective for diagnosis and treatment* (3rd ed.). New York, NY: Lippincott Williams and Wilkins.

Fitch, J. L. (1990). Consistency of fundamental frequency and perturbation in repeated phonations of sustained vowels, reading, and connected speech. *Journal of Speech and Hearing Disorders, 55*, 360-363.

Fitch, J. L., & Holbrook, A. (1970). Modal vocal fundamental frequency of young adults. *Archives of Otolaryngology - Head and Neck Surgery*, *92*, 379-382.

Garrett, K. L., & Healey, E. C. (1987). An acoustic analysis of fluctuations in the voices of normal adult speakers across three times a day. *Journal of Acoustical Society of America, 82*, 58-62.

Graddol, D., & Swann, J. (1983). Speaking fundamental frequency: some physical and social correlates. *Journal of Speech and Hearing Research, 26*, 351-366.

Guimarães, I., & Abberton, E. (2005). Fundamental frequency in speakers of Portuguese for different voice samples. *Journal of Voice*, *19*, 592-606.

Hirano, M., Kurita, S., & Nakashima, T. (1981). *Growth, development, and aging of human vocal folds*. Vocal Physiology Conference, Madison.

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, *30*, 1-15.

Hunter, E. J. (2009). A comparison of a child's fundamental frequencies in structured elicited vocalizations versus unstructured natural vocalizations: A case study. *International Journal of Pediatric Otorhinolaryngology, 73*(4), 561-71

Kahane, J. C. (1982). Growth of the human prepubertal and pubertal larynx. *Journal of Speech and Hearing Research*, *25*, 446-455.

Kent, R. (1976). Anatomical and neuromuscular maturation of the speech mechanism:

evidence from acoustic studies. *Journal of Speech and Hearing Research*, *19*, 421-447.

Linders, B., Massa, G. G., Boersma, B., & Dejonckere, P. H. (1995). Fundamental voice frequency and jitter in girls and boys measured with electroglottography: influence of age and height. *International Journal of Pediatric Otorhinolaryngology*, *33*, 61-65.

Murry, T., Brown, W. S., & Morris, R. J. (1995). Patterns of fundamental frequency for three types of voice samples. *Journal of Voice*, *9*(3), 282-289.

Nittrouer, S., McGowan, R. S., Milenkovic, P. H., & Beehler, D. (1990). Acoustic measurements of men's and women's voices: A study of context effects and covariation. *Journal of Speech and Hearing Research*, *33*, 761-775.

Norusis, M. J. (1992). *SPSS/PC+professional statistics, Version 5.0* (Chicago, IL: SPSS Inc).

O'Neil, E. N., Jones, G. W., & Nye, C. (1997). Acoustic characteristics of children who speak Arabic. *International Journal of pediatric Otorhinolaryngology*, *42*, 117-124.

Sorenson, D. N. (1989). A fundamental frequency investigation of children ages 6-10 years old. *Journal of Communication Disorders*, *22*(2), 115-123.

Speyer, R., Bogaardt, H. C. A., Passos, V. L., Roodenburg, N. P. H. D., Zumach, A., Heijnen, M. A. M., Baijens, L. W. J., Fleskens, S. J. H. M., & Brunings, J. W. (2010). Maximum Phonation Time: Variability and Reliability. *Journal of Voice*, *24*(3), 281-284.

Vogel, A. P., Fletcher, J., Snyder, P. J., Fredrickson, A., & Maruff, P. (2011). Reliability, stability and sensitivity to change and impairment in measures of timing and frequency, *Journal of Voice*, *25*(2), 137-149.

Whiteside, S. and Hodgson, C. (1999). Acoustic characteristics in 6-10 year old children's voices: some preliminary findings. *Logopedics Phoniatrics Vocology*, *24*, 6-13.

Xue, S. A., & Hao, J. G. (2006). Normative standards for vocal tract dimensions by race as measured by acoustic pharyngometry. *Journal of Voice*, *20*, 391-400.

Zechmeister, J. S., Zechmeister, E. S., & Shaughnessy, J. J. (2001). *Essentials of research*

methods in psychology. New York: McGraw-Hill.

Appendix 1-Standard passage for oral reading

〈說話不簡單〉

上課了，山羊老師要教大家說話。小牛和小馬聽了，都覺得好笑，心裏想：誰不會說話呢？

山羊老師請同學們出來練習說話。小牛第一個舉手要說笑話。他越說越快，越說聲音越小，還沒說到一半就笑個不停。

小馬出來給大家講故事。同學們都專心聆聽，可是小馬前言不搭後語，大家越聽越不明白。

小牛和小馬終於知道，說話真不簡單，也要好好學習。

《快樂學語文》一上第1冊　教育出版社

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Estella Ma, for giving me valuable advice and guidance throughout the study. I would also like to thank Ms Wong for her kind help in the recruitment of participants. Special thanks are given to all participants who helped in this study. Thanks are also given to my beloved family, friends and classmates for their endless support.