Title	Expression and perception of gender in prepubertal children's voice : an acoustic study
Author(s)	Lau, Tsz-yan, Dorothy; 劉芷欣
Citation	Lau, T. D. [劉芷欣]. (2013). Expression and perception of gender in prepubertal children's voice : an acoustic study. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR.
Issued Date	2013
URL	http://hdl.handle.net/10722/238528
Rights	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.

Expression and Perception of Gender in Prepubertal Children's Voice - An Acoustic Study

Lau Tsz Yan, Dorothy

The University of Hong Kong

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, 2013.

Abstract

The present study described the fundamental frequency (f0) and the first two formant frequencies (F1 and F2) obtained from voice samples produced by 25 male and 26 female Cantonese-speaking prepubertal children under natural (neutral) condition and upon request to mimic the opposite gender voice (imitation condition), to investigate sexual dimorphism in prepubertal children's voice and to assess their implicit knowledge on voice gender. Average accuracy of voice gender identification by adult listeners was 81.7% and 41.1% for prepubertal children's voice gender produced under neutral and imitation conditions respectively. No significant difference in f0 was found between genders under neutral condition, suggesting similar vocal mechanism for prepubertal boys and girls. Average F1 and F2 associated with boys were lower than that with girls. It was suggested that both difference in vocal tract length, and the sex-specific articulatory behaviors contributed to the differences in formant frequencies, thus enhancing sexual dimorphism for gender voice identification. Under imitation condition, boys exhibited significantly higher f0 than girls. F1 associated with boys was also higher than that with girls. It could be concluded that prepubertal children had the implicit knowledge on the sexually dimorphic acoustic correlates (f0 and F1) and were capable in altering the vibration rate of vocal folds and the effective vocal tract length upon request to conform to vocal characteristics of the opposite gender.

Keywords: prepubertal children, voice, gender, acoustic, fundamental frequency, formant frequencies

Expression and Perception of Gender in Prepubertal Children's Voice - An Acoustic Study

Human sexual dimorphism manifested in voices contributes to gender identification. Adult males and females have well-defined anatomical differences including the vocal apparatus, and as a result perceptual and acoustical differences in the voices produced. It has been documented that average fundamental frequency (f0) associated with adult males and females are between 100 -146 Hz and 188-211 Hz respectively (e.g., Gelfer & Mikos, 2005). Comparing to adult female voices, adult male voices are also characterized by less variability in f0 and intonation contour, lower formant frequencies, less formant dispersion and generally higher vocal intensity. Such sexual dimorphism in vocal characteristics can be attributed to the anatomical and developmental differentiation that occur during puberty, during which the human phonatory system is modified and thus changing generation of sound by the vocal folds and the articulatory system (the subsequent filtering of sounds in vocal tract) (Fitch & Giedd, 1999). With respect to the sound source, the increased level of testosterone during male puberty leads to lengthening and thickening of male's vocal folds, which results in growth of laryngeal cartilages including thyroid and cricoid cartilages, and these changes consequently yield a lowering of the mean and range of f0 (Kahane, 1982). With respect to filter, male larynx outgrows that of female by 40% after puberty (Titze, 1994). The longer vocal tract and the secondary descent of the larynx in males give rise to the lower formant frequencies and less formant dispersion (Fitch & Giedd, 1999).

Apart from the anatomical differences, adult male and female voices also demonstrate significant acoustical discrepancies that can be attributed to sociolinguistic factors that affect articulatory behaviors (Lee et al., 1999). In general, adult males tend to speak with a lower voice pitch, conforming to the stereotypic characteristics of male voice (Sachs, Lieberman, & Erickson, 1973). Moreover, they may also make use of other speech characteristics such as

manipulating intonation contour and extent of mouth opening to enhance the perceptual representation of voice gender (Ohde & Sharf, 1992).

Prepuberty corresponds to Stage 1 of growth (Tanner, 1986). The prepubertal period ends when there is a growth spurt marked by rapid somatic growth, development of primary sexual organs and secondary sexual characteristics which are sexually dimorphic (Rogol et al., 2002). Anatomically, there are only negligible differences in the vocal apparatus between the two genders in prepubertal stage (Kaplan, 1971). Yet, sexual dimorphism still exists among children voices. Weinberg and Bennett (1971) presented samples of spontaneous speech produced by children of 5-6 years old to adult listeners, and found that adult listeners were able to correctly identify the gender of 78% and 71% of the boys and girls speakers respectively, although a large overlap in the f0 ranges of the two genders was revealed. Sachs, Lieberman, and Erickson (1973) also reported similar findings. In their study of voice samples produced by 26 preadolescent children of 4-14 years of age, it was found that adult listeners were able to correctly identify the gender of 81% of speech samples of the children speakers.

These studies showed that gender of children's voices could be identified with high accuracy, implying the presence of cues for gender distinction in the voices. To address the factors contributing to such apparent gender distinction, gender-specific f0 and vowel formants of preadolescent children have been examined. Busby and Plant (1995) measured the f0 and formant frequencies of Australian vowels produced by 40 prepubertal children with ages from 5-11 years old. Results revealed no difference in f0 between boys and girls, but girls exhibited higher formant frequencies when compared with boys. Eguchi and Hirsh (1969) also reported similar findings on f0, F1 and F2 of preadolescent children.

Similarly, Bennett (1981) investigated the vowels produced by 42 children of 7-8 years old, and found that vowel formants associated with boys were consistently lower than

those with girls. Bennett suggested that such formant difference could be attributed to differences in the vocal tract size, as well as gender-specific articulatory behaviors which further enhanced the distinction between boys' and girls' voices. For example, boys tended to speak with more lip rounding, and possibly with a lowered larynx than girls. These adjustments helped alter the length and configuration of the vocal tracts, yielding lower formants and enhancing voice gender identification.

Since behavioral factors have been suggested to contribute to sexual dimorphism in prepubertal children vocal production, such behaviors were likely to be acquired and might vary across cultures (Hasek et al., 1980), as it was suggested that concept of gender polarization differed across cultures (Williams & Best, 1990). As a matter of fact, it has been reported that gender differences in terms of personality traits were stronger in American and European cultures in which great progress had been made to reduce gender inequality than in African and Asian cultures (McCrae et al., 2005). Moreover, differences exist in the dimensions of vocal apparatus among difference racial groups. In the study carried out by Xue et al. (2006), Chinese males were found to exhibit significantly larger oral and vocal tract volumes when compared to African American and White American males. However, existing acoustical studies of voice gender identification in prepubertal children only focused on native English speakers, but not tonal languages. In previous studies, prepubertal f0 of native English speakers was reported to be non-sexually dimorphic (Weinberg & Bennett, 1971; Sachs et al., 1973). However, for tonal languages such as Cantonese, the relative f0 levels and f0 changes are the acoustic correlates for tone perception (Khouw et al., 2007). Possible f0 alternations for expression of gender in prepubertal children may affect their tone production. For example, a Cantonese-speaking boy may consciously or unconsciously lower his f0 and formants to facilitate listeners' perception of male's voice, but this may at the same time hinder production of high tones. This suggested that variation of f0 for voice

gender identification might be restricted by the tonal restriction associated with a tonal language. As such, there appeared a research gap on acoustic correlates of prepubertal children's voice and their implicit knowledge of gender distinction. The present study therefore investigated the acoustic correlates of Cantonese-speaking prepubertal children's voice, and examined their implicit knowledge on sexual dimorphism in voice.

To answer the research gap, the present study involved two experiments. Experiment 1 aimed to investigate the acoustical properties of the voices produced by prepubertal Cantonese-speaking children (neutral condition). Experiment 2 aimed to investigate the acoustical properties of the voices produced by prepubertal Cantonese-speaking children upon request to mimic the opposite gender voice (imitation condition), in order to assess the implicit knowledge on gender voice through the children altering the apparent perceived voice gender. In addition, f0 and formant frequencies of the vowel production by prepubertal Cantonese-speaking boys and girls in the two experiments were compared and the accuracy of voice gender identification under both conditions was investigated. These also helped examine the effectiveness of acoustic cues for voice gender identification.

Method

Participants

Fifty-one prepubertal children (25 boys and 26 girls) participated in the study. Their ages were from 6 to 11 years, with a mean age of 8.5 years. The participating children were randomly recruited from a local primary school. All children met the selection criteria which included: (1) they were native Cantonese speakers, (2) they had no known speech and/or voice disorders, and (3) they had no history of speech and/or hearing problems. All participants were willing to participate in the study.

Speech materials

The speech materials consisted of a short Cantonese passage and five Cantonese

monosyllabic words. All monosyllabic words were with a CV structure, each consisted of a Cantonese vowel (/a/, /ɛ/, /i/, /ɔ/, /u/). In maintaining the naturalness of production, each citation word was embedded into a carrier phrase /li1 kɔ3 tʊk6 ____ tsi6/ ("呢個讀___字"), meaning "*This word is* ____". The short passage was used to obtain average f0 values, while the citation words were used for the extraction of vowel formants.

Instrumentation and procedure

Experiment 1: Neutral condition. In this experiment, the children were instructed to read aloud the following speech materials: (1) a short passage, and (2) the five citation words embedded in the carrier phrases. They were required to produce the speech materials at their normal speaking voice with normal loudness, pitch and rate of speech.

Experiment 2: Imitation condition. In Experiment 2, the children were instructed to read aloud the same speech materials: (1) a short passage, and (2) the five citation words embedded in the carrier phrases by mimicking the voice of the opposite gender. That is, boys were asked to pretend to speak like a girl by feminizing their voices, while girls were asked to pretend to speak like a boy by masculinizing their voices.

All speech samples were recorded in a sound-treated classroom of low ambient noise level (about 30-40 dB) via a high-quality microphone (SM59A, Shure, USA) and a professional grade pre-amplification system (MobilePre, M-Audio, USA). To control for loudness, a constant microphone-to-mouth distance of about 5 cm was maintained. All recorded speech samples were digitized at a sampling rate of 20 kHz and quantization rate of 16 bits/sample using *Praat* (Boersma, 2001). The entire recording session lasted for about 15-20 minutes for each participant. To familiarize the children with the recording materials and recording environment, they were provided with sufficient practice period and instructions were provided before the experiments were carried out.

Measurement of fundamental frequency and formant frequencies

Average f0 were measured by using the pulse detection routine of *Praat*. In addition, the first two formant frequencies were measured using Linear Predictive Coding (LPC) of *Praat*. Only the medial 80% of the vowel portion of the CV syllables were used for formant analysis in order to avoid vowel initiation and termination effect.

Identification of voice gender

Voice gender identification experiment was carried out to determine whether expression and perception of voice gender were consistent. A group of 12 adult native Cantonese speakers who had no known speech and/or hearing problems were recruited as listeners. During the identification test, a total of 102 voice samples recorded by 51 prepubertal boys and girls under both neutral and imitation conditions were presented to the listeners. The voice samples were randomized in order to avoid order effect. During the listening experiment, the listeners were provided with answer sheets and were instructed to determine the gender of the voice of the speakers. To provide listeners with sufficient time to complete the gender identification task, an inter-stimulus pause of approximately three seconds was provided. A brief practice period was provided to the listeners in order to familiarize themselves with the experiment procedure and the experiment environment.

Data and statistical analyses

Neutral condition and imitation condition. Independent-samples t-tests were carried out to assess the gender difference on f0. In addition, to determine the possible effects of gender and vowel on F1 and F2, two 2 (gender) x 5 (vowel) two-way repeated-measures analysis of variance (ANOVA) were carried out for each production condition. The tests were followed by post-hoc multiple comparisons if necessary.

Comparison between neutral and imitation conditions. Independent-samples t-tests were carried out to assess the change in f0 under different conditions, one for boys and one for girls. To determine the possible effect of condition and vowel on F1 and F2, two 2

(condition) x 5 (vowel) two-way ANOVAs were carried out, one of each gender. The tests were followed by post-hoc multiple comparisons if necessary.

Results

Experiment 1: Acoustical properties of prepubertal children's voice under neutral condition

Perception of voice gender. Figure 1 shows the percent accuracy of voice gender identification under neutral condition. Results showed that listeners were able to correctly identify 80.5% and 82.8% of the voice gender of prepubertal boys and girls respectively from the recordings of passage reading under neutral condition, yielding an overall accuracy of 81.7%.

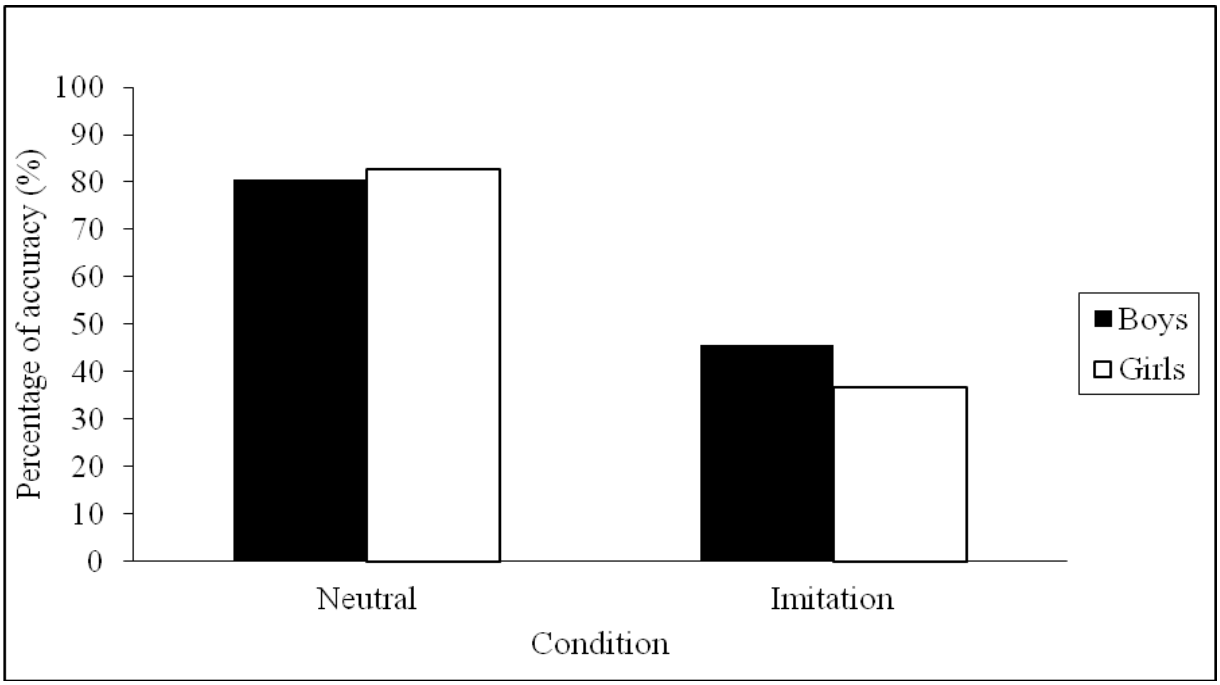

Figure 1. Percentage of accuracy in perception of intended voice gender under neutral and imitation conditions.

Fundamental frequency. Mean and standard deviation values of f0 associated with prepubertal boys and girls under neutral condition are shown in Table 1 and Figures 2. Prepubertal boys and girls exhibited average f0 values of 235.9 Hz and 274.9 Hz respectively. Independent-samples t-test was carried out to compare the difference in f0 between

prepubertal boys and girls under neutral condition, and no significant difference in f0 between the two genders was found ($t(25) = -1.284$, $p = .211$).

Table 1

Mean and standard deviation values of f0, F1 and F2 associated with prepubertal boys and girls under neutral and imitation conditions

	Boys		Girls	
	Neutral	Imitation	Neutral	Imitation
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
f0	235.9 (23.5)	240.1 (29.4)	274.9 (53.4)	224.4 (31.8)
F1	595.6 (38.2)	634.2 (48.5)	620.0 (79.8)	586.1 (72.8)
F2	1654.3 (167.7)	1737.1 (211.9)	1660.9 (183.5)	1754.2 (159.8)

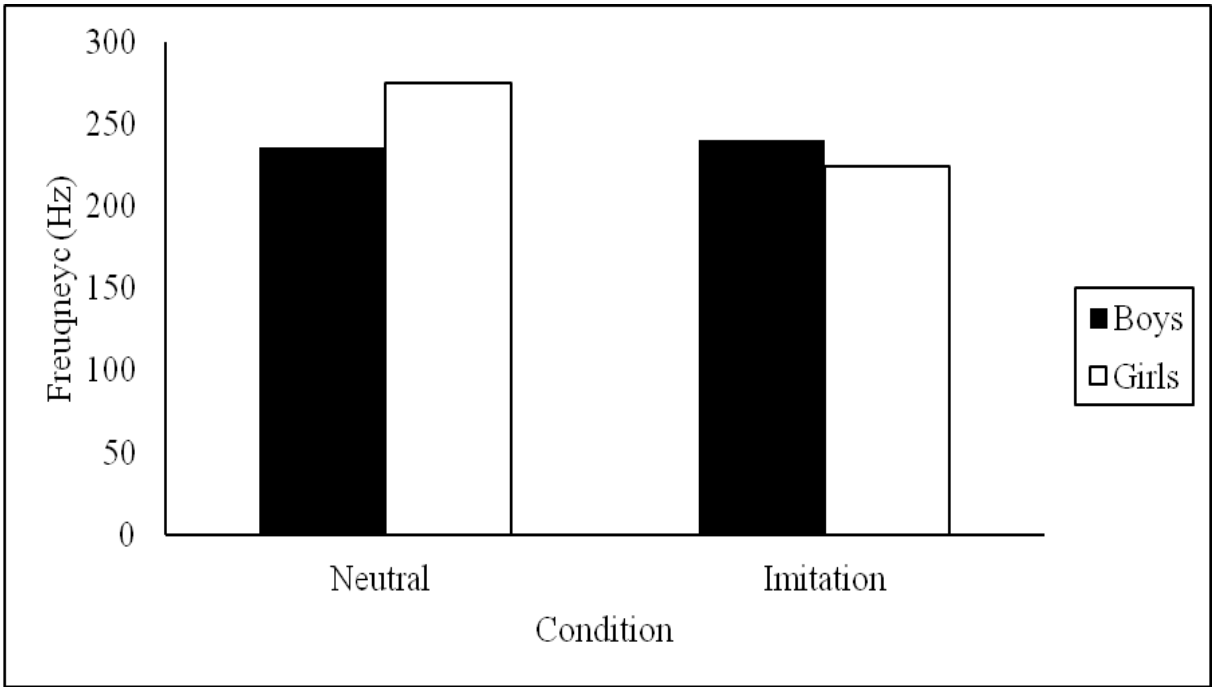

Figure 2. Average f0 associated wtih prepubertal boys and girls under neutral and imitation conditions.

Formant frequencies. Average F1 and F2 values associated with prepubertal children are shown in Figures 3 and 4. Mean and standard deviation values of F1 and F2 associated with different vowels produced by prepubertal boys and girls under neutral

condition are shown in Table 2 and Figure 5.

Regarding F1 values, prepubertal boys and girls exhibited average of 595.6 Hz and 620.0 Hz respectively. Since the assumption of sphericity was violated ($X^2(9) = 44.969$, $p < .001$), results of two-way repeated-measures ANOVA with Greenhouse-Geisser correction was used, which revealed no significant interaction between gender and vowel ($F(2.913, 142.721) = 1.608$, $p = .191$). There was significant main effect of vowel ($F(2.913, 142.721) = 403.522$, $p < .001$), but not gender on F1. Subsequent post-hoc multiple comparisons showed that the F1 value associated with /i/ was significantly lower than that with the other vowels ($p < .001$).

Regarding F2 values, prepubertal boys and girls exhibited average F2 values of 1654.3 Hz and 1660.9 Hz respectively. Since the assumption of sphericity was violated ($X^2(9) = 147.900$, $p < .001$), results of two-way repeated-measures ANOVA with Greenhouse-Geisser correction was used and revealed no significant interaction between gender and vowel for F2 ($F(2.065, 101.178) = 2.454$, $p = .089$). Significant main effect was found only for vowel ($F(2.065, 101.178) = 121.948$, $p < .001$), but not gender on F2. From the results of post-hoc multiple comparisons, the value of F2 associated with /u/ was significantly lower than that with the other vowels ($p < .001$).

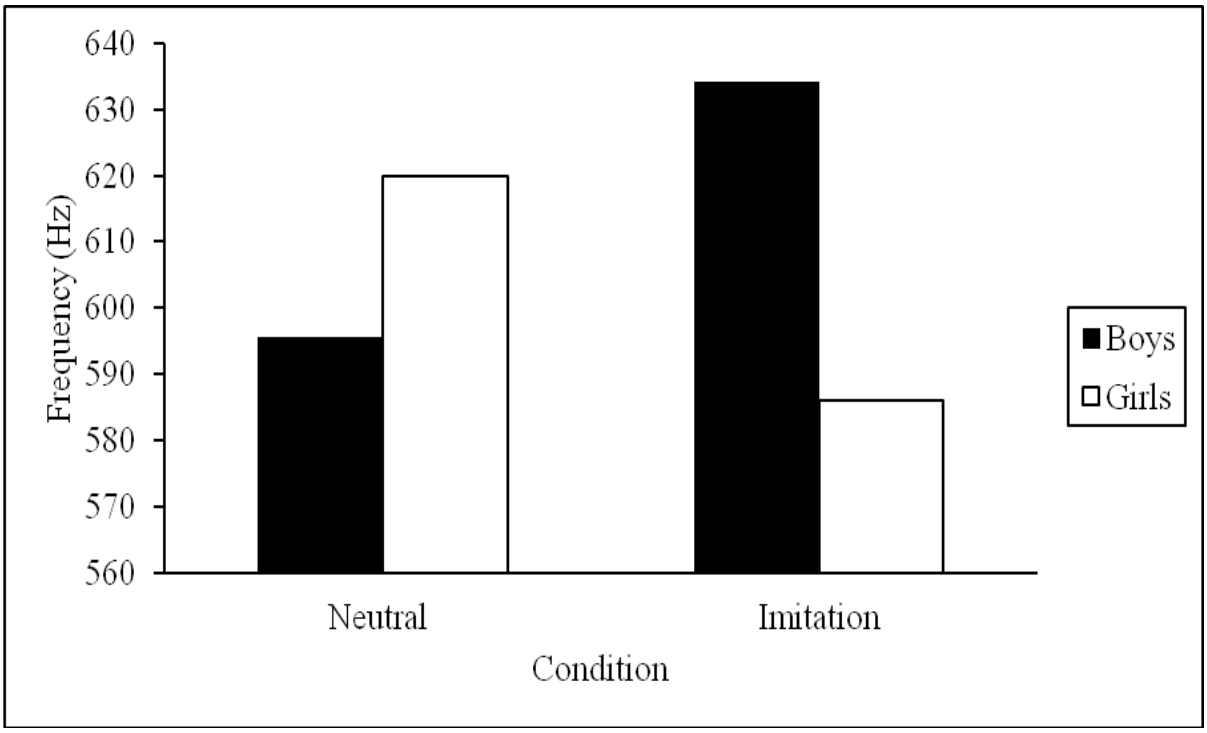

Figure 3. Average F1 values associated with prepubertal boys and girls under neutral and imitation conditions.

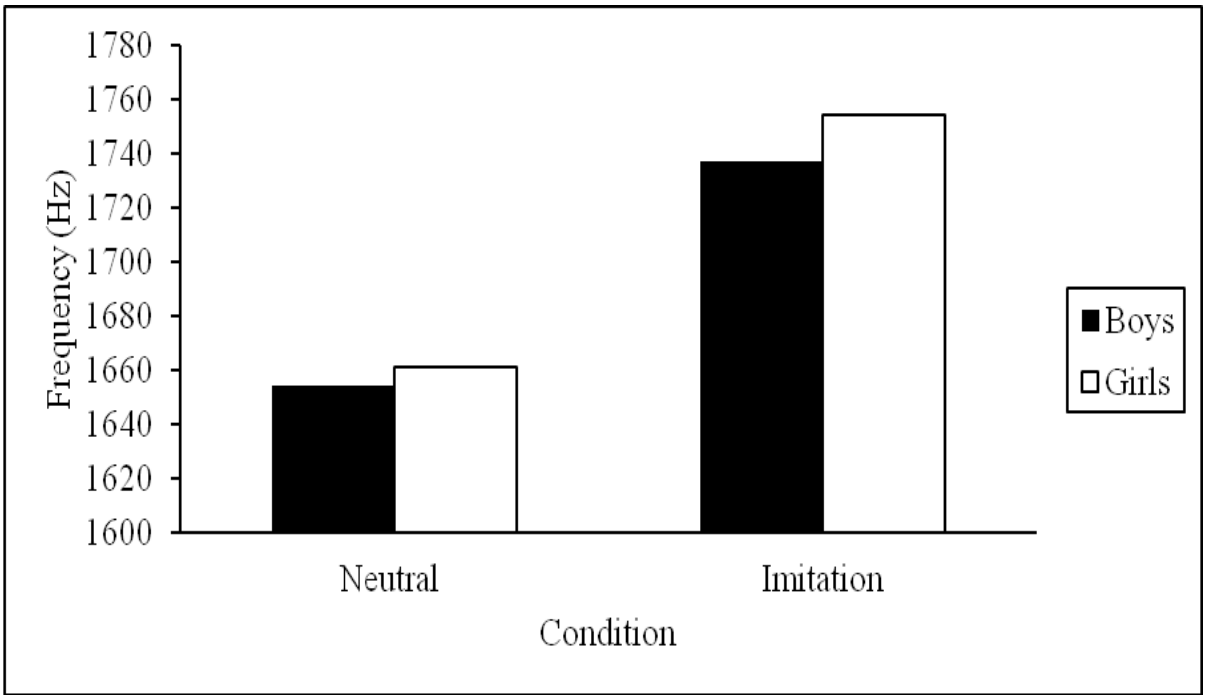

Figure 4. Average F2 values associated with prepubertal boys and girls under neutral and imitation conditions.

Table 2

Mean and standard deviation values of F1 and F2 associated with different vowels produced by prepubertal boys and girls under neutral condition

	Boys		Girls	
	F1	F2	F1	F2
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
/a/	978.7 (84.9)	1749.7 (202.1)	950.5 (130.5)	1694.6 (267.1)
/ɛ/	611.8 (84.0)	2234.1 (473.0)	658.4 (99.0)	2118.5 (531.9)
/i/	306.4 (45.2)	2202.4 (752.4)	348.8 (54.7)	2513.4 (632.8)
/ɔ/	660.1 (78.6)	1139.4 (156.3)	693.2 (102.8)	1269.0 (134.5)
/u/	450.8 (118.5)	976.8 (153.4)	480.4 (52.0)	984.6 (161.0)

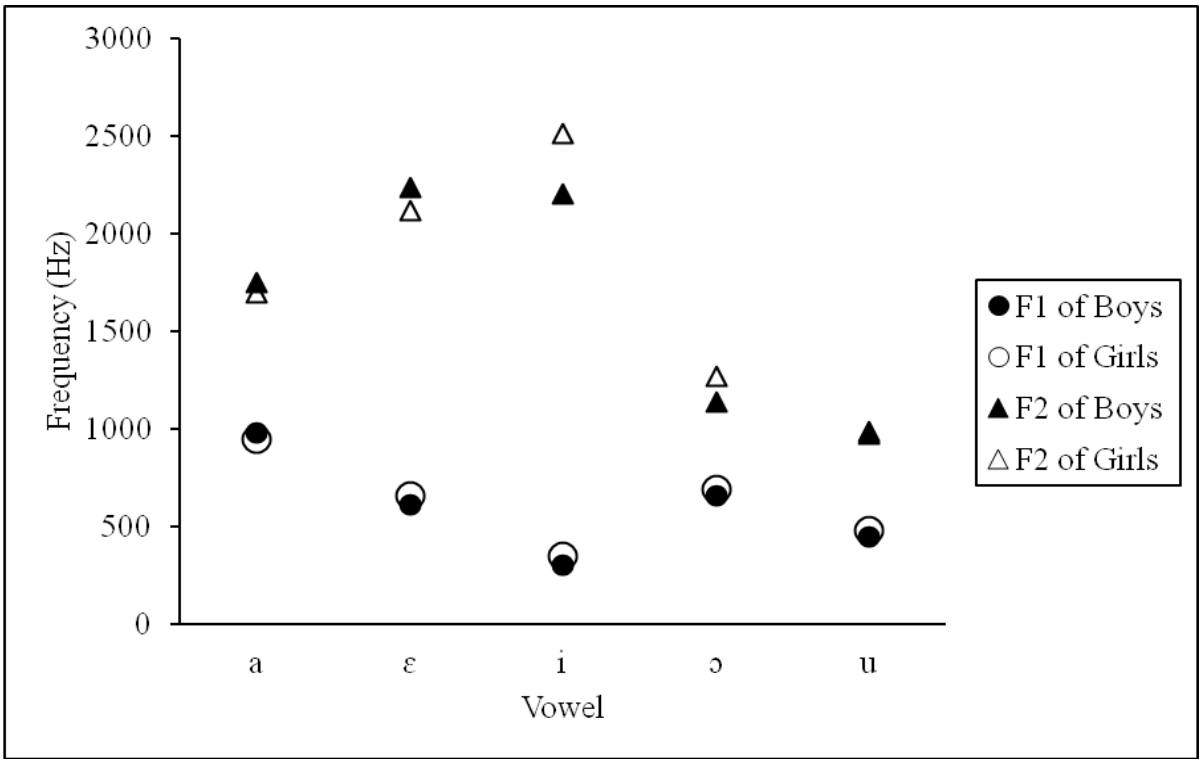

Figure 5. F1 and F2 of different vowels produced by prepubertal boys and girls under neutral condition.

Experiment 2: Acoustical properties of prepubertal children's voice under imitation condition

Perception of voice gender. Percent accuracy in voice gender identification under

imitation condition is shown in Figure 1. Results indicated that listeners were partially able to correctly identify the imitated voice gender from the voice samples produced by boys and girls. The percent correct gender identification rates for boys and girls were 45.5% and 36.6% respectively, with an overall accuracy of 41.1%. According to the data, boys appeared to be more successful in imitating a girl voice (45.5%) than girls pretending a boy's voice (36.6%).

Fundamental frequency. Mean and standard deviation values of f0 associated with prepubertal boys and girls under imitation condition are shown in Table 1 and presented in Figures 2. Prepubertal boys and girls exhibited average f0 values of 240.1 Hz and 224.4 Hz respectively. Results of an independent-samples t-test showed that f0 produced by boys was significantly higher than that of girls ($t(25) = 2.449$, $p = .022$).

Formant frequencies. Average F1 and F2 values associated with prepubertal children are shown in Figures 3 and 4. Mean values of F1 and F2 associated with different vowels produced by prepubertal boys and girls under imitation condition are shown in Table 3.

Regarding F1 under imitation condition, prepubertal boys and girls exhibited average F1 values of 634.2 Hz and 586.1 Hz respectively. Since the assumption of sphericity was violated ($X^2(9) = 73.133$, $p < .001$), results of two-way repeated-measures ANOVA with Greenhouse-Geisser correction was used and showed no interaction effect between gender and vowel ($F(2.498, 122.414) = .330$, $p = .676$). However, there was main effect for vowel for vowel ($F(2.498, 122.414) = 184.646$, $p < .001$), but not gender on F1. Subsequent post-hoc multiple comparisons revealed that the F1 value associated with /i/ was significantly lower than that with the other vowels ($p < .001$).

Regarding F2 under imitation condition, prepubertal boys and girls exhibited average F2 values of 1737.1 Hz and 1754.2 Hz respectively. Since the assumption of sphericity was

violated ($X^2(9) = 79.645$, $p < .001$), results of two-way repeated-measures ANOVA with Greenhouse-Geisser correction was used and revealed significant interaction between gender and vowel ($F(2.571, 125.955) = 14.576$, $p < .001$). One-way repeated-measures ANOVAs were then carried out to investigate the simple main effect of vowel on the value of F2, one for each gender. Since the assumption of sphericity was violated for both boys ($X^2(9) = 50.523$, $p < .001$) and girls ($X^2(9) = 34.618$, $p < .001$), Greenhouse-Geisser correction was used. Results revealed significant difference between vowels on the value of F2 for both boys and girls. Post-hoc multiple comparisons showed that the value of F2 associated with /ɔ/ ($p < .001$) and /u/ ($p < .001$) produced by boys were significantly lower than that with the other vowels; while the value of F2 associated with /i/ ($p < .001$) produced by girls was significantly higher than that of the other vowels.

Table 3.

Mean and standard deviation values of F1 and F2 associated with different vowels produced by prepubertal boys and girls under imitation condition

	Boys		Girls	
	F1	F2	F1	F2
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
/a/	915.0 (146.2)	1739.1 (225.2)	872.0 (207.2)	1580.1 (254.5)
/ɛ/	645.7 (97.0)	1942.3 (460.9)	624.7 (122.4)	2193.7 (376.2)
/i/	379.6 (107.0)	2124.2 (586.0)	353.9 (67.6)	2783.1 (449.0)
/ɔ/	678.5 (135.1)	1299.8 (229.8)	611.4 (115.9)	1154.6 (179.4)
/u/	498.2 (120.0)	1188.4 (209.0)	439.9 (61.7)	1085.0 (304.6)

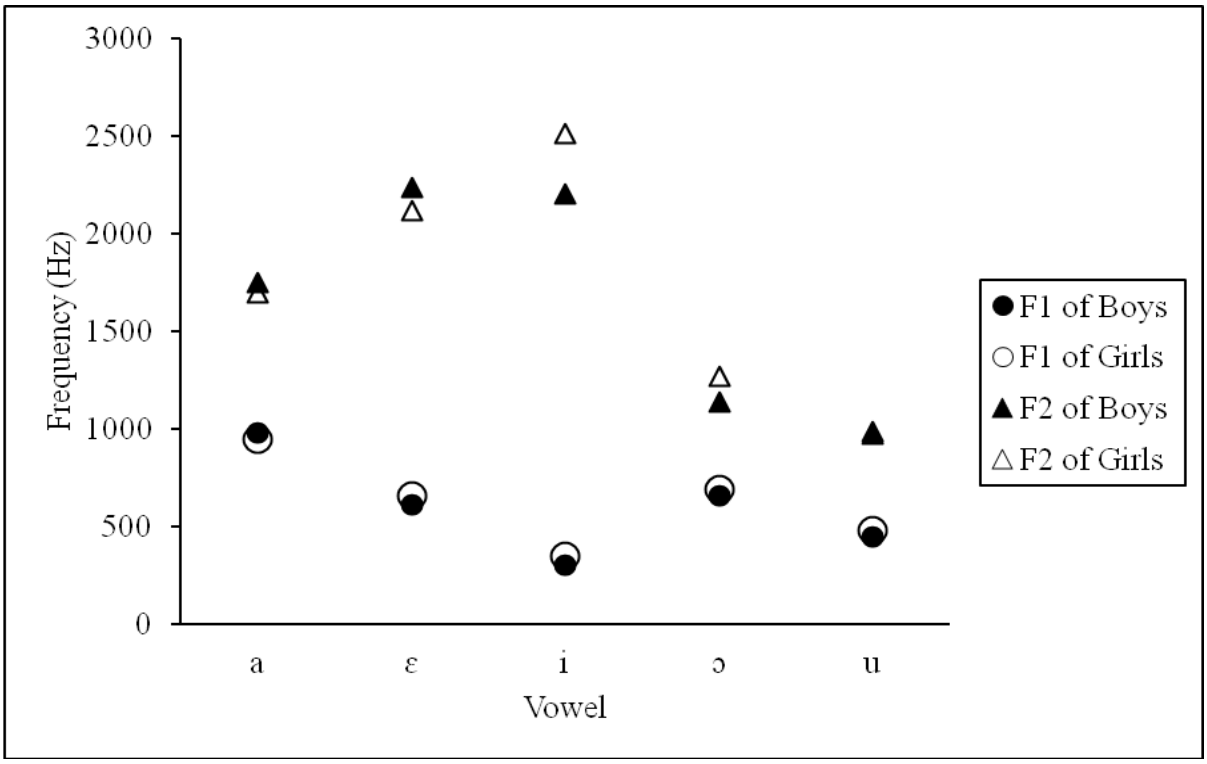

Figure 6. F1 and F2 of different vowels produced by prepubertal boys and girls under imitation condition.

Comparison between neutral and imitation conditions.

Fundamental frequency. Results of independent-samples t-test showed that f0 produced by boys under imitation condition was significantly higher than that under neutral condition ($t(25) = -3.808$, $p < .001$) and f0 produced by girls under imitation condition was significantly lower than that under neutral condition ($t(25) = 3.347$, $p = .003$).

Formant frequencies ANOVA test. Girls were found to exhibit significantly lower F1 under imitation condition than neutral condition ($F(1, 50) = 5.521$, $p = .023$); but no main effect of condition on F1 exhibited by boys ($F(1, 48) = 1.597$, $p = .212$). For F2, no main effect of condition for both boys ($F(1, 48) = .001$, $p = .976$) and girls ($F(1, 50) = .574$, $p = .452$).

Discussion

Experiment 1: Acoustical properties of prepubertal children's voice under neutral condition

Perception of voice gender. The present data indicated that identification of gender based on children's voices were above chance level (81.7%) for both prepubertal boys and girls. This was consistent with the findings by Sachs, Lieberman, and Erickson (1973) in which a high percentage of accuracy (81%) could be attained in the identification of gender in prepubertal children's voice by adult listeners. This indicates that, despite the little anatomical and morphological differences between prepubertal boys and girls, they exhibited differences in their voices that allowed listeners to correctly identify their gender.

Fundamental frequency. The finding of lack of f0 difference in prepubertal boys and girls was in line with the previous studies reported by Perry et al. (2001) and Busby and Plant (1995). They consistently reported little or no difference in f0 among prepubertal boys and girls (under 12 years of age). Recall that f0 refers to rate of vocal fold vibration. The similar f0 between prepubertal boys and girls suggested that their vocal mechanism were similar both functionally and anatomically. Combined with the findings obtained from the voice gender identification experiment, it could be concluded that the cue for listeners to correctly identify voice gender did not lie in f0. In other words, f0 appears to be a less reliable cue for voice gender identification in prepubertal children, and other cues must be present that helped listeners in perceiving voice gender correctly.

Formant frequencies. To further investigate, the first and second formants (F1 and F2) were obtained from prepubertal boys and girls. Results indicated a discernible difference in F1 and F2 between the two genders (see Figures 3 and 4). In general, girls were associated with higher F1 and F2 than boys. This is in line with what Busby and Plant (1995), and Bennett (1981) reported, in which they found that prepubertal girls were associated with higher F1 and F2 than boys. Accordingly, as formants are inversely related to body size and vocal tract length (Fitch & Giedd, 1999; Bennett, 1981), prepubertal girls were found to have a shorter vocal tract than boys. According to the cephalometric study reported by Walke and

Kowalski (1972), prepubertal boys and girls exhibited similar mandibular length. It follows that the anatomical major difference between prepubertal boys and girls that gave rise to vocal tract length difference lied in the length of pharynx. This was confirmed by King's (1952) who reported that the length of pharynx was approximately 2-8% longer for boys starting from early infancy when compared with counterparts.

That said, it should be noted that only vocal tract length could change during phonation. This could be facilitated by maneuvers such as lip protrusion or deliberate lowering of larynx. The term "effective vocal tract length" should be used to indicate the actual length of vocal tract for resonance during phonation, instead of anatomical vocal tract length. King (1952) and Walke and Kowalski (1972) reported findings about anatomical vocal tract length. In fact, the reported anatomical difference alone between prepubertal boys and girls would not be accountable for such a marked difference particularly in the value of F1 found in the present study. The sex-specific articulatory behaviors might also contribute to the differences in formant frequencies. Bennett and Weinberg (1979) suggested that prepubertal boys might adopt more pronounced lip rounding, smaller jaw opening and lower larynx position to attain lower formant frequencies as resulted from this study. As such, both the difference in vocal tract length among prepubertal boys and girls, as well as the sex-specific articulatory behaviors might have contributed to the differences in formant frequencies, thus enhancing sexual dimorphism for gender voice identification.

Recall that vowels can be classified into groups according to various dimensions which include tongue height and tongue advancement of the vowel for accurate production (Di Benedetto, 1989). Moreover, F1 corresponds to tongue height and F2 corresponds to tongue advancement (Ball & Rahilly, 1999; Borden, Harris, & Raphael, 2003; Pickett, 1999). The current results conform to the expected vowel formant pattern based on the Cantonese vowel chart that represents tongue height and advancement. As shown in Figure 5, /i/ being a

closed vowel with the highest tongue height had the lowest value of F1, while /a/ being an open vowel with the lowest tongue height exhibited the highest value of F1. With regard to F2, /i/ being a front vowel was associated with the highest F2, while /u/ being a back vowel with the lowest F2.

Experiment 2: Acoustical properties of prepubertal children's voice under imitation condition

Perception of the imitated voice gender. The voice gender identification experiment revealed that listeners were able to correctly identify the imitated voice gender with an overall percentage accuracy of 41.1%, resembling only about half of that under neutral condition. This indicates that prepubertal boys and girls were successful in imitating voices of opposite sex for only about 40% of the time. The low percentage reflects that the use of parameters and other strategies by the prepubertal children to conform to vocal characteristics of opposite gender were not very effective. This might be related to limitation associated with their immature and similar vocal apparatus. The prepubertal boys and girls had similar vocal apparatus and they were bound by such limitation when attempting to imitate voices of opposite sex.

Fundamental frequency. Despite the limitation mentioned above, the prepubertal children were able to alter f0 in an attempt to imitate voices of opposite gender. The present f0 data showed that prepubertal boys were able to produce a significantly higher f0 under imitation condition. On the contrary, girls were able to mimic boy's voices with a significantly lower f0. This finding suggests that prepubertal children have already developed an implicit knowledge about f0 which is one of the sexually dimorphic acoustic properties in the expression of the opposite gender voice. Even before puberty, they knew what was required of f0 in order to "sound" like a person of opposite sex. This was done by manipulating the tension of the vocal folds, and thus the vibration rate of the vocal folds.

According to anatomy and physiology of speech production, lateral cricoarytenoid muscle, cricothyroid muscle and vocalis muscle contribute to the regulation of longitudinal tension of vocal folds (Zemlin, 1998; Hirano et al., 1969). Particularly with activity of vertical fibers (pars recta) of cricothyroid muscle, the elevated vocal fold tension leads to an increase f0. Prepubertal boys were likely to involve in such maneuvers in order to increase their f0 when imitating girl's voices. On the contrary, when prepubertal girls attempted to lower their f0, they might do so by relaxing laryngeal tension. Physiologically, elevating vocal fold tension seems easier than lowering vocal fold tension. Habitual or optimal pitch, which is understood as the average pitch level during speech production, is located about at the bottom 25% of the total pitch range (Fairbanks, 1959; Stone, 1983). There seemed to be "more room" for the boys to increase their f0 than the girls to lower their f0. In fact, the present results also indicated this pattern: the extent of f0 increase in prepubertal boys was more than the extent of f0 decrease in girls.

Despite the alteration in f0, the listeners were only able to correctly identify the intended voice gender 41.1% of the time. This further indicates that f0 was not a reliable acoustic cue for voice gender identification in prepubertal children.

Formant frequencies. A comparison of Figures 5 and 6 shows that prepubertal boys and girls were able to change vowel formants in an attempt to mimic the voices of opposite sex, at least for F1. On average, prepubertal boys increased F1 from 507 Hz to 551 Hz, yet girls decreased F1 from 545 Hz to 507 Hz. Only negligible change was seen for F2. This showed that prepubertal children also had the implicit knowledge of F1 difference between genders and attempted to manipulate F1 to achieve voices of opposite sex.

Formant frequencies in general are inversely correlated to the length of vocal tract (Mermelstein, 1967). A girl's voice is generally associated with higher formant frequencies. Specifically, F1 is responsive to a change of tongue height (or oral cavity volume) during

vowel production (Ball & Rahilly, 1999; Borden, Harris, & Raphael, 2003; Pickett, 1999). The higher is the tongue height, the lower is the F1, and vice versa. To mimic a girl's voice, prepubertal boys attempted to increase F1 probably by lowering the tongue and increasing oral cavity volume, as perceptually the elevated F1 will lead to listeners' perception of a girl's voice. The opposite is true for prepubertal girls, in which they attempted to reduce F1 to facilitate listeners' perception of a boy's voice. The significant difference associated with F1 produced by girls may imply that prepubertal girls were more capable in modifying their effective vocal tract length, thus the value of F1 to conform to the acoustic characteristics of the opposite gender voice, when compared with boys. However, the listening experiment found that prepubertal boys were more successful in imitating opposite gender voice than girls, as reflected by the higher percent accuracy in perception of the imitated voice gender by the listeners. This suggests that F1 was only one of the acoustic cues available for voice gender identification. Other acoustic cues were present to help listeners identify voice gender.

On the other hand, F2 was not significantly different between neutral and imitation conditions for both boys and girls. This finding implies that boys and girls did not attempt to manipulate F2 to imitate voices of opposite sex. Recall F2 reflects tongue advancement; the more front is the tongue during vowel production, the higher is the F2 value (e.g., Ball & Rahilly, 1999). The lack of F2 change across condition may imply that both prepubertal boys and girls failed to use F2 to imitate their voices. Their inability to alter F2 for voice imitation might be related to the short vocal tract associated with prepubertal boys and girls. According to MRI study reported by Vorperian, Kent, Lindstrom, Kalina, Gentry, and Yandell (2005), children of age about 6 years possessed an average vocal tract length of 11.4 cm, only 75% of that of the adult. It is plausible that, with such a short vocal tract, the prepubertal boys and girls were not able to advance their tongue to alter F2, in an attempt to mimic voices of the opposite gender.

Further direction for investigation

The present study only focused on the acoustical properties (f0, F1 and F2) of the prepubertal children's voice under neutral and imitation conditions. In studies investigating the cues for gender identification in adult's voice, it was reported that there were differences in vowel duration (Hillenbrand et al., 1995), intonation contour (McConnell-Ginet, 1978), speech rate (Byrd, 1992), mean airflow and aerodynamic power (Titze, 1989) between adult males and females. Future investigations focusing on the above parameters should be carried out on prepubertal children to examine other possible factors that contribute to the sexual dimorphism in voice at prepubertal stage, thus leading to the high percentage of accuracy in voice gender identification.

Conclusion

The present study described f0, F1 and F2 of the Cantonese-speaking prepubertal children under neutral condition and imitation condition. Average accuracy of voice gender identification by adult listeners was 81.7% and 41.1% for prepubertal children's voice gender produced under neutral and imitation conditions respectively. No significant difference in f0 was found between genders under neutral condition, while average F1 and F2 associated with boys were lower than that with girls with discernible difference, suggesting formant frequencies being one of the cues for voice gender identification. Under imitation condition, boys exhibited significantly higher f0 than girls. F1 associated with boys was also higher than that with girls. It could be concluded that prepubertal children had the implicit knowledge on the sexually dimorphic acoustic correlates (f0 and F1) and were capable in altering the vibration rate of vocal folds and the effective vocal tract length upon request to conform to vocal characteristics of the opposite gender.

Acknowledgements

I would like to express my heartfelt gratitude to my supervisor, Dr. Lawrence Ng. Dr. Ng

has been very keen in offering inspiring advice and encouragement throughout. I am truly grateful and thankful for the supervision and invaluable feedbacks from Dr. Ng.

I would also like to express my appreciation to Baptist (Sha Tin Wai) Lui Ming Choi Primary School for the arrangement for subject recruitment and data collection.

Last but not least, I would like thank all subjects and listeners for participating in the recording procedures and the voice gender identification experiment.

References

Ball, M. J. & Rahilly, J. (1999). *Phonetics: the science of speech.* London: Arnold, the Hodder Headline Group.

Bennett, S. (1981). Vowel formant frequency characteristics of preadolescent males and females. *Journal of the Acoustical Society of America, 69,* 321-238.

Bennett, S., & Weinberg, B. (1979). Acoustic Correlates of Perceived Sexual Identity in Preadolescent Children's Voices. *Journal of Acoustical Society of America, 66,* 989-1000.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5,* 341-345.

Borden, G. J., Harris, K. S., & Raphael, L. J. (2003). *Speech Science Primer - Physiology, Acoustics, and Perception of Speech* (4th ed.). Baltimore, MD: Lippincott Williams and Wilkins.

Busby, P. A., & Plant, G. L. (1995). Formant frequency values of vowels produced by preadolescent boys and girls. *Journal of Acoustical Society America, 97,* 2603-2606.

Byrd, D. (1992). Preliminary results on speaker-dependent variation in the TIMIT database. *Journal of Acoustical Society of America, 92,* 593-596.

Chan, A. Y. W. & Li, D. C. S. (2000). English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture, and Curriculum, 13,* 67-85.

Di Benedetto, M. (1989). Vowel representation: Some observations on temporal and spectral properties of the first formant frequency, *Journal of the Acoustical Society of America, 86,* 55-66.

Eguchi, S. & Hirsh, I. J. (1969). Development of speech sounds in children. *Acta Oto-Laryngologica, 257,* 1-51.

Fairbanks, G. (1959). *Voice and Articulation Drillbook* (2nd ed.). New York, NY: Harper and Row.

Fant, G. (1973). *Speech sounds and features*. Cambridge: MIT Press.

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

Gelfer, M. P., & Mikos, V. A. (2005). The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. *Journal of Voice, 19,* 544-554.

Hasek, C., Singh, S., & Murry, T. (1980). Acoustic attributes of preadolescent voices. *Journal of the Acoustical Society of America, 68,* 1262-1265.

Hillenbrand, J., Getty, L. A., Clark, M. J. & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97,* 3099-3111.

Hirano, M., Ohala, J., & Vennard, W. (1969). The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. *Journal of Speech and Hearing Research, 12,* 616-628.

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

Kaplan, H. M. (1971). *Anatomy and physiology of speech*. New York, NY: McGraw-Hill.

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

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of Acoustical Society of America, 105*, 1455-1468.

McConnell-Ginet, S. (1978). Intonation in a man's world. *Signs: Journal of Women in Culture and Society, 3*, 541-559.

McCrae, R. R., Terracciano, A., & 78 members from the Personality Profiles of Culture Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88*, 547-561.

Mermelstein, P. (1967). Determination of vocal-tract shape from measured formant frequencies. *Journal of Acoustical Society of America, 41*, 1283-1294.

Ohde, R. N., & Sharf, D. J. (1992). *Phonetic analysis of normal and abnormal speech*. New York: Merrill.

Perry, T. L., Ohde, R. N., & Ashmead, D. H. (2001). The acoustic bases for gender identification from children's voices. *Journal of Acoustical Society America, 109*, 2988-2998.

Pickett, J. M. (1999). *The Acoustics of Speech Communication - Fundamentals, Speech Perception Theory, and Technology.* Needham Heights, MA: Allyn and Bacon.

Rogol, A. D., Roemmich, J. N., & Clark, P. A. (2002). Growth at Puberty. *Journal of Adolescent Health, 31*, 192-200.

Sachs, J., Lieberman, P., & Erickson, D. (1973). Anatomical and cultural determinants of male and female speech. In R. Shuy & R. W. Fasold (Eds.), *Language attitudes: Current trends and prospects* (pp. 74-84). Washington, DC: Georgetown University Press.

Stone, R. E. (1983). Issues in clinical assessment of laryngeal function: contraindications for

subscribing to maximum phonation time and optimum fundamental frequency. In D. Bless and J. Abbs (eds.), *Vocal Fold Physiology*. San Diego, CA: College-Hill Press.

Tanner. J. M. (1986). Normal growth and techniques of growth assessment. *Journal of Clinical Endocrinology & Metabolism, 15*, 411-451.

Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *Journal of Acoustical Society of America, 85*, 1699-1707.

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

Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., & Yandell, B. S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *Journal of the Acoustical Society of America, 117*(1), 338-350.

Walker, G. F., & Kowalski, C. J. (1972). On the growth of the mandible. *American Journal of Physical Anthropology, 36*, 111-117.

Weinberg, B., & Bennett, S. (1971). Speaker sex recognition of 5- and 6-year-old children's voices. *Journal of the Acoustical Society of America, 50*, 1210-1213.

Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A multination study.* Newbury Park, CA: Sage.

Xue, S. A., Hao, G. J. P., & Mayo, R. (2006). Volumetric measurements of vocal tracts for male speakers from different races. *Clinical Linguistics and Phonetics, 20*, 691-702.

Zemlin, W. R. (1998). *Speech and Hearing Science - Anatomy and Physiology* (4th ed.). Needham Heights, MA: Allyn and Bacon.