Proceedings of the Institute of Acoustics

VOCAL CHARACTERISTICS: A DEVELOPMENTAL STUDY

S. P. Whiteside & C. Hodgson

Department of Human Communication Sciences, University of Sheffield S10 2TA, United Kingdom

1. INTRODUCTION

Fundamental frequency (F0) differences between adult males and females are well documented [1, 2]. Research into children's F0 characteristics has shown that rapid changes in laryngeal structure occur during the 0-3 year age band and adolescence [3, 4, 5, 6]. Studies reporting the rate of F0 development between the ages of 3 and 10, suggest that changes in laryngeal structure continue to occur during this period [7, 8, 9]. Because of differences in research design, results across studies have not always been comparable. In addition, studies that have been carried out on this age group have reported inconsistent findings with regard to sex differences, and the rate of F0 change with age. Fairbanks et al. [10] for example, found no significant differences between the F0 values of 7 and 8 year old boys. They also reported the same findings for 7 and 8 year old girls [11]. In addition there were no significant differences between the males' and females' F0. McGlone and McGlone [12] later replicated the results reported by Fairbanks et al. [11], which suggests that for these ages, there may not be sex or age differences in laryngeal size. However, there is some evidence to suggest the contrary. Hasek et al. [9] for example, found a decrease in F0 between 5 to 6 years and 7 to 8 years, which was significant only for the male children in the study.

Formant frequency characteristics play an important role in the perceived sex of a preadolescent child [13]. A number of developmental studies have reported that vowel formant frequencies decrease with age and that in general, female children have higher formant frequencies than their male peers [8, 14, 15]. Although age and sex differences occur in vowel formant frequencies in general, by using Fant's *kn*-factor [16, 17], this has been found to be dependant upon both the vowel and the formant frequency in question [5, 14]. This therefore suggests that the vocal tract dimensions between preadolescent boys and girls differ in a non-uniform way and replicates data on adult speaker sex differences [16, 17]. Similarly, both age and sex related *kn*-factor values could reflect the rate of relative change between the vocal tract dimensions of preadolescent male and female children. However, while formant frequency patterns and *kn*-factor values could be interpreted as evidence for the physical differences and maturation in the vocal apparatus they cannot fully account for those differences that maybe a product of the linguistic and stylistic conventions [18].

Formant frequency bandwidths reflect both voice source spectrum characteristics and the degree of acoustic damping in the vocal tract. Women's voices are often described as being more breathy [19, 20], and this has been found to be largely attributable to the characteristics of the glottal source [19, 20, 21]. It has been shown for example that women display more symmetry in their glottal waveforms and have larger open quotients [21]. This therefore means a steeper spectral slope and weaker excitation of the vocal tract and less intense and more diffuse and wider formant bandwidth patterns both in the lower and higher frequency regions, for women speakers [19, 22]. On this basis it is therefore proposed that the pattern of formant frequency bandwidths in the vocal patterns of the children in this study may be indicative of both developmental changes and sex-related patterns in voice source characteristics. This preliminary study aims to examine the vocal characteristics of a group of male and female children between the ages of 6 and 10, separated by approximately two year intervals, and see whether any age-related or sex related differences emerge.

2. METHODOLOGY

2.1 Subjects

Three boys and three girls from each of two age bands: 6 years (mean age 6.0 years) and 8 years (mean age 8.1 years), and four boys and four girls from the 10 years age group (mean age 10.0 years), served as subjects in the study. This made a total of 20 children (10 boys and 10 girls) who participated in the study. The 20 children used in the study were matched for height and weight according to their age bands and were all from the same local primary school in the north east of England (Tyneside). All children had lived in the local area all their life and spoke English as their first language. All the children were healthy and had no speech, language or hearing difficulties.

2.2 Data collection

The recordings were made in a quite room using a portable AIWA HDS1 DAT recorder and a stereo microphone. The microphone was laterally offset at a distance of 30cm from the child's mouth.

Picture cards were used to elicit a possible total of nine target phrases from the children. The target phrases were *The red/blue/green bar/jar/car*. Also included in the elicitation were four distracters, for example *The red boat, The green balloon.* Prior to the actual collection of the data, children were asked to name the colours and the pictures separately. This was done to ensure that none of the children were colour blind and also to establish picture names and therefore avoid confusion during the recording session. If a child hesitated or made an error in naming a particular picture, the picture was simply presented again later in the naming sequence.

2.3 Acoustic Analysis

The elicited data were digitised with a sampling rate of 10 kHz using a Kay Computerised Speech Lab (CSL- Model 4300). All acoustic analyses were performed on the KAY CSL using routines that are outlined below.

2.3.1 F0 analysis. An autocorrelation method (20 ms frame length and 20 ms frame advance) was used to calculate F0 for the entire duration of the vowel /ɑː/. The start of the vowel was taken from the start of the voicing for /ɑː/ following /b k dʒ/, as determined by the onset of periodic excitation of F2/F3 [23, 24].

2.3.2 Formant frequency and formant bandwidth analysis. Formant frequency and formant bandwidth data for the entire duration of the vowel /ɑː/ were derived from automatic LPC analysis (10 ms frame length, 12th order filter, with pre-emphasis (0.9), with triangular window weighting, autocorrelation method, bandwidth cut off of *500Hz*). The same sections that were used for the F0 analysis were used to derive LPC-based formant frequencies and formant frequency bandwidths.

3. RESULTS & DISCUSSION

The mean and standard deviation values of F0 for the vowel /ɑ/ are given in Table 1 according to age and sex. Also given in Table 1 are the means and standard deviations of the standard deviation values for the vowel. Table 2 gives the results of a two-factor ANOVA for the fundamental frequency parameters. In the case of the F0 data for the phrase-final vowel /ɑː/, there were significant sex and age differences, but no age-by-sex interaction. A post-hoc Scheffé test revealed significant differences (at $p<0.05$) between the F0 data of 10-year olds and the 8-year olds (mean difference -19.81Hz) and between those of the 10-year olds and the 6-year olds (mean difference -23.40 Hz). No significant differences were found between the data of the 6-year olds and the 8-year olds (mean difference –3.59 Hz), therefore amplifying the marked decrease in F0 that is taking place between the ages of 8 and 10 for the boys (see Figure 1).

		6 years		8 years		10 years	
		Mean	s.d.	Mean	s.d.	Mean	s.d.
Males	F0	248.7	13.3	255.4	9.8	223.3	8.3
Females	F0	248.3	26.3	234.4	11.6	226.9	18.6
Males	F1	737.3	90.7	791.6	102.7	671.1	61.5
	B1	192.3	61.4	171.1	50.8	153.6	44.4
Females	F1	823.8	135.1	861.2	85.1	772.8	61.9
	B1	191.3	61.9	225.9	54.7	221.0	54.3
Males	F2	1248.9	97.3	1347.0	145.6	1278.6	109.7
	B2	176.6	36.0	185.5	42.1	203.18	71.0
Females	F2	1365.0	164.4	1455.9	91.7	1317.4	123.7
	B2	193.0	45.8	168.6	39.3	211.9	60.9
Males	F3	1891.5	208.1	2154.2	254.1	2059.3	142.0
	B3	247.6	81.8	269.4	67.7	305.9	307.3
Females	F3	2089.3	222.5	2165.0	153.0	2075.1	184.7
	B3	221.7	83.2	271.5	71.7	260.8	70.9
Males	F4	3213.8	194.7	3441.2	182.1	3095.4	521.1
	B4	368.1	130.3	280.8	103.6	297.8	491.7
Females	F4	3177.7	289.5	3279.6	219.0	3249.0	119.1
	B4	334.3	97.4	230.6	55.4	277.3	65.6

Table 1: Mean and standard deviation fundamental frequency, formant frequency (F1 to F4) and formant bandwidth (B1 to B4) values (Hz) for the vowel /ɑː/ by age and sex.

If we attempt to seek a physiological explanation for these findings, they may be indicative of the rate of laryngeal growth. For the female children this appears to be gradual between the ages of 6 and 10, whereas for the male children, the rate of laryngeal growth is gradual between the ages of 6 and 8 but more rapid after this point. Kent [5] summarises that *both* male and female children undergo gradual F0 decreases until approximately, the ages of 11 and 12, at which point males experience a substantial F0 decrease. This study suggests that this substantial decrease may begin earlier - around the age of 8 years. However, we are still left with the question of whether the drop at this stage is entirely due to physiological factors or whether linguistic, social and cultural conventions may also be contributory factors. Hasek, Singh and Murry [9] for example found a decrease in F0 in male children after the age of 7 years, While they suggest that this could be partly attributable to growth, they suggest that the male child may be adopting a lower than optimum pitch to sound 'more male'. There is some sociolinguistic evidence to suggest that different intonation patterns are adopted by boys and girls from Tyneside, with girls exhibiting more rises than boys [25]. Furthermore, it would appear that this pattern also occurs in the intonation patterns of adult speakers from Tyneside [26]. The patterns of F0 reported here, suggest that the accelerated F0 changes for the boys from the age of 8 may be a result of the acquisition of these intonation patterns. It is therefore proposed, that the development of F0 in children is not only determined by physical growth and changes in the vocal organs but are also a product of the sociolinguistic factors that may be operating in a specific speech community.

Parameter	Sex	Age	Sex by Age
/ɑ:/ mean F0	F(1, 173)=6.26, p<.05	F(2, 173)=41.61, p<.001	F(2, 173)=10.34, p<.001
/ɑ:/ F1	F(1, 173)=40.12, p<.001	F(2, 173)=21.26, p<.001	F(2, 173)=0.50, ns
/ɑ:/ B1	F(1, 173)=24.28, p<.001	F(2, 173)=0.66, ns	F(2, 173)=6.47, p<.005
/ɑ:/ F2	F(1, 173)=22.15, p<.001	F(2, 173)=12.31, p<.001	F(2, 173)=1.92, ns
/ɑ:/ B2	F(1, 173)=0.25, ns	F(2, 173)=5.63, p<.005	F(2, 173)=10.98, p<.001
/ɑ:/ F3	F(1, 173)=6.48, p<.05	F(2, 173)=10.16, p<.001	F(2, 173)=4.17, p<.05
/ɑ:/ B3	F(1, 173)=0.98, ns	F(2, 173)=1.58, ns	F(2, 173)=0.37, ns
/ɑ:/ F4	F(1, 173)=0.11, ns	F(2, 173)=6.92, p<.005	F(2, 173)=4.52, p<.05
/ɑ:/ B4	F(1, 173)=0.95, ns	F(2, 173)=2.28, ns	F(2, 173)=0.06, ns

Table 2. P values of a two-factor ANOVAs (sex and age) for fundamental frequency (F0: Hz), formant frequencies (F1 to F4: Hz) and formant frequency bandwidths (B1 to B4: Hz). (ns: not significant at p<0.05)

The mean and standard deviation values for the first four formant frequencies (F1 to F4) and their bandwidths (B1 to B4) of the vowel /ɑ:/ are given in Table 1. The results of a two-factor (sex and age) ANOVA are given in Table 2. These results show that for both the male and female children, the first four formant frequencies rose slightly between the ages of 6 years and 8 years before falling again between the ages of 8 years and 10 years. Data for the first four formant frequencies showed significant overall sex differences (see Table 2) with females having higher formant frequencies. Using Fant's scale factors [16] these overall sex differences equate to values of 12.03, 6.44, 3.39 and 0.08 for k_1, k_2, k_3 and k_4 respectively (see Table 1 for formant frequency data). Given that k_1 and k_2 reflect both front and back vocal cavity dimensions and k_3 reflects front cavity dimensions it would appear that there are more pronounced differences between the back cavity dimensions of the male and female children, with the lowest value being observed for k_3. These results replicate previous findings for low back vowels [8, 14].

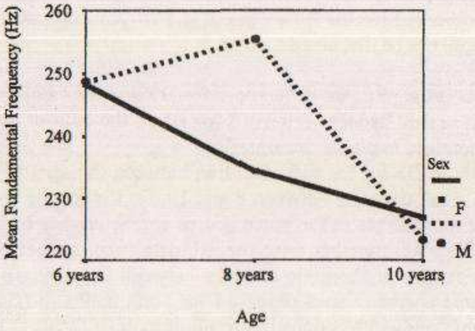

Figure 1: Mean fundamental frequency for the boys and girls by age.

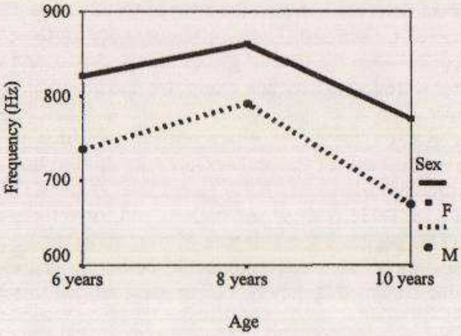

Figure 2: Mean frequency values - first formant (F1) for the boys and girls by age.

It has been found and therefore proposed, that in the case of low vowels (like /ɑː/), F1 may be more sensitive to age and sex related differences in the size of the pharyngeal and oral cavities [8]. This suggestion is supported by the data in this study where significant sex related differences were only found for F1 in *all group* comparisons using a post-hoc Scheffé test (see Figure 2). In addition, the higher $k1$ value reported here could be interpreted as further evidence for this suggestion.

VOCAL CHARACTERISTICS

The formant frequency bandwidth data showed significant sex differences for B1, significant age differences for B2, and both age and age by sex differences were found for B1 and B2 (see Table 2). No significant age, sex or age-by-sex interactions were found for B3 or B4. This however, could have been due to errors in LPC formant bandwidth estimation which cannot be ruled out especially in light of the large standard deviation values observed for the B4 values of the 10-year old males (see Table 1). The results for B3 and B4 will therefore not be discussed further.

Post-hoc Scheffé tests showed no sex significant differences for B1 in the data of group three (age six). There were however, significant sex differences observed for B1 in the data of both groups two (age eight) and one (age ten). This therefore explains the significant age-by-sex interaction for B1, which is reflected by a systematic decrease in B1 for the male children between the ages of 6, 8 and 10, and an increase between 6 and 8 with a small decrease between 8 and 10, for the female children. The B1 data could be interpreted as evidence for changes in the voice source characteristics of the female and male children. The F0 data discussed above, together with the B1 data here, suggest that there are more marked changes in the vocal apparatus of the male children between the ages of six and ten as they begin to acquire the glottal source characteristics observed for adult males. It has been reported that adult females have wider formant bandwidths compared to adult males [22]. It is therefore suggested, that the lower values of B1 observed here for the 8 and 10 year old male children, are a reflection of the changes in glottal source properties that are accompanying changes in fundamental frequency and formant frequencies, which are a result of physical maturation.

Data for B2 showed that there was a systematic increase in the values of the male children, with the female children showing a significant decrease between the ages of 6 and 8, before increasing again to age ten. Post-hoc Scheffé tests showed no sex significant differences for all three groups. There were however, significant age differences between the data of groups one and two, and between groups one and three. No significant differences were found between group two and group three.

Discriminant analysis tests were run on the vocal characteristic measures to see whether these parameters could classify the data on the basis of age and sex. Results showed that the parameters were successful in giving correct classification scores of 78.7% (see Table 3) and 83.5% (see Table 4) for age and sex respectively. The data in Table 6 show the highest and lowest classification scores for groups one and two respectively. This suggests some degree of overlap in the vocal characteristics of group two with both groups one and three. This therefore, could be interpreted as being indicative of the degree of change in vocal characteristics that may be taking place around this age, in the voices of the children reported here.

		Predicted Group Membership		
Actual Group	No. of cases	Age 10	Age 8	Age 6
Age 10	71	60 (84.5%)	4 (5.6%)	7 (9.9%)
Age 8	49	8 (16.3%)	36 (73.5%)	5 (10.2%)
Age 6	44	2 (4.5%)	9 (20.5%)	33 (75.0%)

Table 3. Results of discriminant analysis showing correct classification according to age. (78.7% of cases correctly classified)

		Predicted Group Membership	
Actual Group	No. of cases	M	F
M	85	70 (82.4%)	15 (17.6%)
F	79	12 (15.2%)	67 (84.8%)

Table 4. Results of discriminant analysis showing correct classification according to sex. (83.5% of grouped cases correctly classified)

In the case of sex classification (see Table 4), the female data were correctly classified only slightly higher at 84.8% compared to 82.4% for the male data. This high level of correct classification suggests that there are relatively stable sex differences across all three age groups.

4. CONCLUSIONS

The data reported here for a group of 6- to 10-year old children suggest that the development and physical maturation in the vocal apparatus of pre-adolescent children result in changes in fundamental frequency, formant frequencies and their corresponding bandwidths. It also appears that there are sex specific changes that take place during this period with changes in laryngeal dimensions co-occurring with changes in pharyngeal dimensions. These sex specific differences could be purely attributed to physical changes. However, these differences may also be a product of social and linguistic conventions.

5. ACKNOWLEDGEMENTS

We wish to thank the Headmaster, staff and pupils of Priory Primary School, Tynemouth, Tyneside, for their co-operation and help with this study.

6. REFERENCES

1. Peterson, G., & Barney, H. Control methods used in the study of vowels, *Journal of the Acoustical Society of America*, 24 629-637 (1952).
2. Titze, I. Physiologic and acoustic differences between male and female voices, *Journal of the Acoustical Society of America*, 85 1699-1707 (1989).
3. Hollien, H., & Malcik, E. Adolescent voice change in southern Negro males, *Speech Monographs*, 29 53-58 (1962).
4. Hollien, H., Malcik, E., & Hollien, B. Adolescent voice change in southern white males, *Speech Monographs*, 32 87-90 (1965).
5. Kent, R. D. Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies, *Journal of Speech and Hearing Research*, 19 421-446 (1976).
6. Robb, M. P., & Saxman, J. H. Developmental trends in vocal fundamental frequency of young children, *Journal of Speech and Hearing Research*, 28 421-427 (1985).

7. Bennett, S. A 3-year longitudinal study of school-aged children's fundamental frequencies, *Journal of Speech and Hearing Research*, 26 137-142 (1983).
8. Busby, P. A., & Plant, G. L. Formant frequency values of vowels produced by preadolescent boys and girls, *Journal of the Acoustical Society of America*, 97 2603-2606 (1995).
9. Hasek, C. S., Singh, S., & Murry, T. Acoustic attributes of preadolescent voices, *Journal of the Acoustical Society of America*, 68 1262-1265 (1980).
10. Fairbanks, G., Wiley, J., & Lassman, F. An acoustical study of vocal pitch in 7 and 8 year old boys, *Child Development*, 20 63-69 (1949).
11. Fairbanks, G., Herbert, E., & Hammond, J. An acoustical study of vocal pitch in 7 and 8 year old girls, *Child Development*, 20 71-78 (1949).
12. McGlone, R. E., & McGlone, J. Speaking fundamental frequency of 8 year old girls, *Folia Phoniatrica*, 24 313-317 (1972).
13. Bennett, S., Weinberg, B. Acoustic correlates of perceived sexual identity in preadolescent children's voices, *Journal of the Acoustical Society of America*, 66, 989-1000 (1979).
14. Bennett, S. Vowel formant frequency characteristics of preadolescent males and females, *Journal of the Acoustical Society of America*, 69 231-238 (1981).
15. Eguchi, S., & Hirsh, I. Development of speech sounds in children, *Acta Otolaryngologica*, 257 5-48. Suppl. (1969).
16. Fant, G. A note on vocal tract size factors and non-uniform F-pattern scalings, *QPSR-STL, KTH-Stockholm*, 4 22-30 (1966)
17. Fant, G. Non-uniform vowel normalization, *QPSR-STL, KTH-Stockholm*, 1-19 (1975).
18. Mattingly, I. Speaker variation and vocal tract size, *Journal of the Acoustical Society of America*, 39 S1219A (1966).
19. Karlsson, I. Glottal waveform parameters for different speaker types, *STL-QPSR, KTH-Stockhlom* 2/3 61-67 (1988).
20. Klatt, D. H. & Klatt, L.C. Analysis, synthesis and perception of voice quality variations among female and male talkers, *Journal of the Acoustical Society of America*, 87 820-857 (1990).
21. Holmberg, E. B., Hillman, R. E., & Perkell, J. S. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal and loud voice, *Journal of the Acoustical Society of America*, 84 511-529 (1988).
22. Fant, G. Temporal fine structure of formant damping and excitation, *Speech Communication Papers, Acoustical Society of America*, 161-165 (1979).
23. Fischer-Jørgensen, E. & Hutters, B. Aspirated stop consonants before low vowels, a problem of delimitation, - its causes and consequences, *ARIPUC* (University of Copenhagen Working Papers) 15 77-102 (1981).
24. Klatt, D. H. Voice onset time, frication and aspiration in word initial consonant clusters, *Journal of Speech and Hearing Research*, 18 686-706 (1975).
25. Local, J. Modelling intonational variability in children's speech. In S. Romaine (Ed.) *Sociolinguistic Variation in Speech Communities*. London: Edward Arnold. 1982.
26. Pellowe, J. & Jones, V. On intonational variability in Tyneside speech. In P. Trudgill (Ed.) *Sociolinguistic patterns in British English*. London: Edward Arnold, 1978.