Acoustic Study in Mandarin-Speaking Children: Developmental Changes in Vowel Production

Chih-Chi Chen, MD; Keh-chung Lin[1], ScD, OTR; Ching-Yi Wu[2], ScD;
Chia-Ling Chen, MD, PhD; Chia-Hui Chen, MEd; Hsieh-Ching Chen[3], PhD;
Wei-Hsien Hong[4], PhD; Alice May-Kuen Wong, MD

Background: Acoustic analysis had been well incorporated into clinical evaluation and management of children with speech disorders for many years. The aim of this study is to investigate developmental changes in vowel production in Mandarin-speaking children using acoustic study analysis.

Methods: A total of 22 children from 5-12 years old were analyzed in this study. Each child read a list of speech materials consisting of 6 dissyllabic words in Mandarin phonemes and the speech samples were recorded. The digitized acoustic recordings were submitted for acoustic analysis. The acoustic parameters in this study include the first and second formant frequencies (F1 and F2) of /a/, /i/ and /u/ and the vowel space. We used the Wilcoxon rank sum test and Spearman's rho correlation test for statistical analysis.

Results: The F1 values of the vowel /i/ were significantly lower in boys than those in girls ($p = 0.013$) by Wilcoxon ranksum test. The F1 value of the vowel /i/ was negatively correlated with children's age (rho = –0.601, $p = 0.003$) and their body height (rho = 0.478 $p = 0.045$). The F1 values of the other two vowels (/u/ and /a/), the F2 values of all three vowels and the vowel space had no association with age and gender.

Conclusions: F1 acoustic parameters have developmental and gender changes in vowel production in Mandarin-speaking children. The data in this study provide references for acoustic assessment of Mandarin-speaking children.
(Chang Gung Med J 2008;31:503-9)

Key words: acoustics, Mandarin, children, vowel, development, formant

Acoustic analysis has been used to evaluate the intelligibility of articulation for many decades.(1) Acoustics is the study of sound.(2) It compensates for the subjective deficit of perceptual measures by quantifying speech sound to several acoustic parameters. Acoustic parameters in vowels include vowel duration, fundamental frequency, formant frequencies, vowel space, and spectral-envelope variability.(2,3)

Resonant frequencies of the vocal tract are called formants. Different shapes and dimensions of the vocal tract cause different formants.(2,4) According

From the Department of Physical Medicine and Rehabilitation, Chang Gung Memorial Hospital, Taipei, Chang Gung University College of Medicine, Taoyuan, Taiwan; [1]Department of Occupational Therapy, National Taiwan University, Taipei, Taiwan; [2]Department of Occupational Therapy, Chang Gung University, Taoyuan, Taiwan; [3]Department of Industrial Engineering and Management, Chaoyang University of Technology, Taichung, Taiwan; [4]School of Sports Medicine, China Medical University, Taichung, Taiwan.
Received: Aug. 22, 2007; Accepted: Jan. 7, 2008
Correspondence to: Dr. Chia-Ling Chen, Department of Physical Medicine and Rehabilitation, Chang Gung Memorial Hospital. No. 5, Fusing St., Gueishan Township, Taoyuan County 333, Taiwan (R.O.C.) Tel.: 886-3-3281200 ext. 3846; Fax: 886-3-3281320; E-mail: clingchen@gmail.com

to the tube-resonator model, resonant frequencies are inversely related to tube length. The vocal tract is like a tube that opens at one end and closes at the other. It starts from the glottis and ends at the lip. The vocal tract can be further separated into three large resonating cavities by the separation of the elevated tongue and soft palate. These are the anterior oral cavity, posterior oropharyngeal cavity and nasal cavity.[2] In vowel production, the soft palate is lowered and the nasal cavity is not opened. The major determinants of formants of vowels are the anterior oral cavity and posterior oropharyngeal cavity.[2]

One classical description of the articulation-acoustic relationships of vowels was made by observing first formant frequency (F1) -second formant frequency (F2) diagrams and the highest position of the tongue in the back-front and low-high dimensions from X-ray findings: F1 decreased as the tongue elevated and F2 or F2-F1 increased as the tongue advanced.[4,5] The posterior oropharyngeal cavity lengthened as the tongue elevated, while the oral cavity shortened as the tongue advanced.[4,5] Further cavity-acoustic relationships were interpreted as F1, correlated negatively with the length of the back of the oropharyngeal cavity and F2 or F2-F1, correlated negatively with the length of the front of the oral cavity.[2,6,7] Although these relationships oversimplified the acoustic expressions of the compound resonator system in the vocal tract, they still widely used now.[2,5,8]

Vowel space is a measurement of the size formed by the extreme vowels in the F1/F2 space. The extreme vowels in Mandarin are /i/, /u/ and /a/.[9] Reduction in vowel space is considered to exist in many dysarthric speakers such as those who have amyotrophic lateral sclerosis,[10] closed head trauma[11] and cerebral palsy.[9]

With the increased availability of low-cost software that can exploit analysis of speech, acoustic analysis had been well incorporated into clinical evaluation and management of children with speech disorders. Acoustic studies of speech are variable in children because of their continuously developing articulator structure.[12] Some studies have dealt with the development-related anatomical and physiological changes in speech-related structures.[12-15] Many researchers have also concentrated on acoustic and functional developmental changes in English- speaking children.[1,16-18] However, there are no studies which have investigated the developmental changes in acoustic features in Mandarin-speaking children in Taiwan. The aim of this study is to investigate developmental changes in vowel production in Mandarin-speaking children in Taiwan using acoustic study analysis. The data provided in this study may help clinicians understand the mechanism in speaking vowels in children with normal development.

METHODS

We recruited children 5 to 12 years old with normal development in this study. The inclusion criteria were that the child was a native Mandarin-speaker and had adequate intelligence in performing the speech-recording tasks based on the experimenter's observations. Children were excluded from the study if they had documented problems with voice, language, hearing, development, the neurological system or the respiratory system. A total of 22 children, 11 boys and 11 girls, were recruited in this study. All children received acoustic analysis.

The speech materials consist of 6 high-frequency dissyllabic words in Mandarin (three corner vowels /i/, /a/, and /u/ × 2 variants). The dissyllabic words were arranged as consonant-vowel-consonant-vowel (/CVCV/) and the target vowel was placed in the first syllable for ease of acoustic analysis. Each word had its phonetic transcription at its side.

The subject was seated in a quiet room. Each of the 6 words was printed on a card. A microphone was placed on a table in front of the subject's lip. The distance between the subject's lips and the microphone was approximately 15 cm. The recording system consisted of an IBM ThinkPad 570E laptop computer and one general extra microphone. The model of the sound card of the IBM ThinkPad 570E was CS 4280. The subjects were asked to read the words on the card aloud at a normal speaking rate. If any word was unfamiliar, the experimenter would explain the word or ask the subject to try to read it with the assistance of its phonetic transcription. No modelling of the sound production was provided by the experimenter.

The digitized acoustic recordings were submitted to acoustic analysis. Speech analysis software, Pratt signal processing software,[19] was used to analyze these speech samples. F1 and F2 were identified by finding the steady part of the formant bars on wide band spectrograms and by using the auto-tracking program. The vowel space area was calculated as the absolute value of {[F1i*(F2a–F2u)+F1a*(F2u–F2i)+F1u*(F2i–F2a)]/2}.[9] Speech samples of 6 participants were randomly selected for inter-rater and intra-rater reliability measurements.

Differences between boys and girls were determined by the Wilcoxon ranksum test. The Pearson correlation test was used to examine the inter-rater and intra-rater reliability measurements from the speech samples. Spearman's rho correlation test was used to determine the relationships of age, body weight and body height with acoustic parameters. A value of $p < 0.05$ was considered statistically significant.

RESULTS

There were no significant differences in age, body height and body weight between the boys and the girls in this study (Table 1). The reliability test showed the speech samples from the 6 participants had high inter-rater (r = 0.874, $p < 0.001$) and intra-rater reliability (r = 0.999, $p < 0.001$).

The F1 values of the three vowels were different among these children, with the highest values for the vowel /a/, then the vowel /u/, and the lowest for the vowel /i/ (Table 2). There were similar F2 values for the three vowels in boys and girls (Table 2). The F1 value of the vowel /i/ showed a significant difference between boys and girls in the Wilcoxon rank-sum test. The F1 value of the vowel /i/ was significantly higher in girls than in boys ($p = 0.013$) (Table 2). There was no significant difference of the F1 values of the vowels /u/ and /a/ between boys and girls (Table 2). The F2 values of all three vowels (/i/, /u/, and /a/) did not show significant differences between boys and girls (Table 2). The F2-F1 values of the three vowels also showed no significant differences between boys and girls (Table 2).

Spearman's correlation analysis demonstrated that the F1 value of the vowel /i/ was negatively correlated with age (rho = –0.601, $p = 0.003$) and body height (rho = –0.478, $p = 0.045$). However, it was not correlated with body weight (Table 3). The F1 values of the vowels /u/ and /a/, and the F2 and F2-F1 values of the vowels /i/, /u/, and /a/ showed no signifi-

Table 1. Subject Data

	Boys (n = 11)	Girls (n = 11)	*p* value
Age (years)	9.09 ± 2.12	7.82 ± 2.18	0.181
Body height (cm)	127.67 ± 14.14	124.44 ± 15.28	0.649
Body weighty (kg)	28.44 ± 7.00	28.22 ± 8.51	0.953

Values are expressed as mean ± standard deviation.

Table 2. Acoustic Parameters of Three Vowels in Boys and Girls

	Vowel	Boys (n = 11)		Girls (n = 11)		
		Median	Range	Median	Range	
F1 (Hz)	/i/	360	320-506	458	358-548	0.013*
	/u/	497	386-563	531	460-642	0.133
	/a/	845	670-995	815	666-1058	0.699
F2 (Hz)	/i/	1335	1128-2014	1357	1048-2342	0.847
	/u/	1326	1099-1538	1337	1053-1544	0.748
	/a/	1338	1251-1472	1278	1227-1411	0.171
F2-F1 (Hz)	/i/	954	793-1654	895	644-1841	0.652
	/u/	816	581-986	780	670-990	0.699
	/a/	503	270-627	481	257-731	0.797
Vowel space (Hz*Hz)		34127	714-208576	13312	305-337996	0.270

*: $p < 0.05$

Table 3. Correlation Coefficients of Acoustic Parameters of Three Vowels and Subject Data

			F1			F2			F2-F1			Vowel
			/a/	/i/	/u/	/a/	/i/	/u/	/a/	/i/	/u/	space
Age												
	rho		–0.14	–0.60†	–0.22	–0.08	–0.27	–0.20	0.05	–0.16	–0.13	–0.18
	95% CI	L	–0.30	–0.82	–0.59	–0.48	–0.62	–0.57	–0.42	–0.52	–0.52	–0.56
		U	0.53	–0.24	0.22	0.36	0.17	0.24	0.43	0.31	0.31	0.27
Height												
	rho		–0.21	–0.48*	–0.17	–0.08	–0.2	–0.15	0.09	–0.06	–0.07	–0.08
	95% CI	L	–0.58	–0.75	–0.55	–0.49	–0.65	–0.53	–0.35	–0.49	–0.49	–0.49
		U	0.24	0.07	0.27	0.35	0.12	0.29	0.49	0.35	0.35	0.35
Weight												
	rho		–0.06	–0.25	0.05	–0.32	–0.33	0.19	–0.23	0.30	0.20	–0.25
	95% CI	L	–0.47	–0.60	–0.38	–0.65	–0.66	–0.24	–0.59	–0.64	–0.24	–0.61
		U	0.37	0.20	0.46	0.12	0.10	0.57	0.22	0.14	0.57	0.19

Abbreviations: CI: confidence interval; L: lower limit; U: upper limit; *: $p < 0.05$; †: $p < 0.01$.

cant correlations with age, body height or body weight (Table 3).

The vowel space areas varied a lot among these normal children (Table 2). There were no significant differences between boys and girls. Vowel spaces were not correlated with age, gender, body height or body weight (Table 3).

DISCUSSION

In this study, the F1 value of the vowel /a/ was the highest among these three vowels, followed by the vowel /u/ and the vowel /i/. The reason, perhaps, is that the F1 value has a negative relationship with the posterior oropharyngeal space in pronouncing vowels.[2,6,7] The posterior oropharyngeal space is longest in pronouncing vowel /i/, next longest when pronouncing the vowel /u/, and the shortest when pronouncing the vowel /a/. When uttering the vowel /i/, the anterior oral cavity becomes very short and the posterior oropharyngeal space is much elongated as the tongue advances and the tongue root elevates.[2,4,5] The vowel /a/ is articulated with the tongue depressed and protracted, the oral cavity lengthened and the posterior oropharyngeal cavity shortened.[2,4,5] The posterior oropharyngeal cavity is also somewhat elongated when pronouncing the vowel /u/ by raising the tongue root.[2,4,5]

We found the F1 value of the vowel /i/ in boys was significantly lower than in girls. These finding suggests that the posterior oropharyngeal space in boys is longer than in girls when pronouncing the vowel /i/. This finding is compatible with Greenberg's study which used magnetic resonance imaging (MRI) techniques in 200 children and found the oropharyngeal distance was longer in boys than in girls.[20]

In this study, the F1 value of the vowel /i/ was negatively correlated with the children's age and body height. However, the F1 values of the other vowels (/u/ and /a/) had no significant correlation with age. These results are compatible with Eguchi and Hirsh's study in which a linear tendency in an F2 versus F1 plot appeared only in the vowel /i/ as children developed, while the linear changes were quite poor in other vowels.[21] Using MRI, Vorperian et al. suggested that the increase in vocal tract length throughout development is predominantly due to growth of the posterior pharyngeal structures.[21] The pace of development for learning to produce and to coordinate articulatory gestures is not uniform.[22] F1 values are not only influenced by physical elongation during child development but also by different articulatory gestures in pronouncing vowels. Children may use different articulatory gestures in compensating for the gradually elongating posterior pharyngeal cavity in pronouncing different vowels.

We found the F2 and F2-F1 values of all three vowels (/i/, /a/ and /u/) had no association with age and gender. F2 and F2-F1 values are thought to be associated with the anterior oral space.[2,6,7] The anterior oral space is relatively small compared to the posterior oropharyngeal area and it elongates slightly during children development.[15] Therefore, the reason that F2 values do not change with age may result from limited anterior oral space elongation during child development.

In our study, the vowel space showed no significant change with development in these children. In addition, there was a great divergence in vowel space size among these children. These findings suggest that the vowel space area does not absolutely reflect the relative clarity in normally developing children with relatively good intelligibility. Previous studies found reduced vowel space was an index corresponding with impaired intelligibility.[9,23] However, an absolute value of the vowel space area was not uniquely associated with a given intelligibility estimate.[23] All children in this study were older than 5 years with relatively good intelligibility, although younger children often have less intelligibility than older children. Therefore, the vowel space did not change with age in the subjects in this study.

This study is a primitive study of acoustic analysis in Mandarin-speaking children. Investigations of more case numbers of children with normal development are needed in further studies. Future studies should focus on acoustic features in Mandarin-speakers with speech disorders of different etiologies such as cerebral palsy and specific language impairment.

Conclusions

The F1 value of the vowel /i/ in boys was significantly lower than in girls. It was correlated negatively with age in developing children. Other acoustic parameters used in this study were not associated with age or gender. The F1 acoustic parameter has developmental and gender changes in vowel production in Mandarin-speaking children. The data provided in this study can give clinicians references for acoustic assessment of Mandarin-speaking children.

REFERENCES

1. Kent RD. Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. J Speech Hear Res 1976;19:421-47.
2. Raphael LJ, Borden GJ, Harris KS. Speech Science Primer: Physiology, Acoustics, and Perception of Speech. 5th ed. Baltimore: Williams & Wilkins Co., 2007:105-30.
3. Lee S, Potamianos A, Narayanan S. Acoustics of children's speech: developmental changes of temporal and spectral parameters. J Acoust Soc Am 1999;105:1455-68.
4. Fant G. Acoustic Theory of Speech Production. The Hague: Mouton., 1970:107-35.
5. Kent RD, Kent JF, Vorperian HK, Duffy JR. Acoustic studies of dysarthric speech: methods, progress and potential. J Commun Disord 1999;32:141-86.
6. Fant G. Non-uniform vowel normalization. STL-QPSR 1975;16:1-19.
7. Fry DB. Acoustic phonetics: A course of basic reading. Hear Res 1976;38:1001-13.
8. Liu HM, Tseng CH, Tsao FM. Perceptual and acoustic analysis of speech intelligibility in Mandarin-speaking young adults with cerebral palsy. Clin Linguist Phon 2000;14:447-64.
9. Liu HM, Tsao FM, Kuhl PK. The effect of reduced vowel working space on speech intelligibility in Mandarian-speaking young adults with cerebral palsy. J Acoust Soc Am 2005;117:3879-89.
10. Kent RD, Sufit R, Rosenbeck J, Kent J, Weismer G, Martin R, Brooks B. Speech deterioration in amyotrophic lateral sclerosis: A case study. J Speech Hear Res 1991;34:1269-75.
11. Ziegler W, von Cramon D. Spastic dysarthria after acquired brain injury: An acoustic study. Br J Disord Commun 1986;21:173-87.
12. Sharkey SG, Folkins JW. Variability of lip and jaw movements in children and adults: implications for the development of speech motor control. J Speech Hear Res 1985;28:8-15.
13. Goldstein UG. An articulatory model for the vocal tracts of growing children. Ph.D. dissertation, MIT, Cambridge, MA, 1980.
Robbins J, Klee T. Clinical assessment of oropharyngeal motor development in young children. J Speech Hear Res 1987;52:271-7.
14. Vorperian HK, Kent RD, Lindstrom MJ, Kalina CM, Gentry LR, Yandell BS. Development of vocal tract length during early childhood: A magnetic resonance imaging study. J Acoust Soc Am 2005;117:338-50.
15. Assmann PF, Katz WF. Time-varying spectral change in the vowels of children and adults. J Acoust Soc Am 2000;08:1856-66.
16. Hazan V, Barrett S. The development of phonemic categorization in children aged 6-12. Phonetica 2000;28:377-96.

17. Menard L, Schwartz JL, Boe LJ, Aubin J. Articulatory-acoustic relationships during vocal tract growth for French vowels: Analysis of real data and simulations with an articulatory model. J Phon 2007;35:1-19.
18. Boersma P, Weenink D. Pratt 3.9.15 [Computer software]. Amsterdam: Institute of Phonetic Sciences, 2001.
19. Greenberg RS, Zahurak M, Belden C, Tunkel DE. Assessment of oropharyngeal distance in children using magnetic resonance imaging. Anesth Analg 1998;87: 1048-51.
20. Eguchi S, Hirsh IJ. Development of speech sound in children. Acta Otolarynx 1969;257:5-43.
21. Nittrouer S. The emergence of mature gestural patterns is not uniform: evidence from an acoustic study. J Speech Hear Res 1993;36;959-72.
22. Turner G, Tjaden K, Weismer G. The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. J Speech Hear Res 1995;38:1001-13.

臺灣兒童國語聲學分析之研究：母音發展的變化

陳芝琪 林克忠[1] 吳菁宜[2] 陳嘉玲 陳嘉惠 陳協慶[3] 洪維憲[4] 黃美涓

背 景： 聲學分析應用於臨床來評估與治療兒童言語障礙已爲之多年，本研究目的爲利用聲學分析探討臺灣說國語的兒童母音發展的變化。

方 法： 本次研究總共將 22 名年齡介於 5 至 12 歲的兒童納入此項研究。受試者要唸 6 組常見的雙音節中文字並錄製下來。透過數位化後的聲學語音材料去分析出母音的共振波頻率，並算出母音動態位圖面積。我們利用 Wilcoxon rank-sum 檢定與相關分析來做統計。

結 果： 以 Wilcoxon rank-sum 檢定母音/i/的第一振幅頻率男生較女生爲低，並有統計上顯著的差異 ($p = 0.013$)。母音/i/的第一振幅頻率亦與年紀 (rho = –0.601, $p = 0.003$) 與身高 (rho = 0.478, $p = 0.045$) 呈負相關。其它的聲學分析測量值與母音動態位圖面積在男女別或年紀別上均無統計上的意義。

結 論： 母音/i/的第一振幅頻率是用來識別男生與女生之間以及兒童成長不同年齡之間母音差異性的良好測量值。

(長庚醫誌 2008;31:503-9)

關鍵詞： 聲學分析，國語，兒童，母音，發展，共振鋒

長庚紀念醫院 台北院區 復健科；長庚大學 醫學院；[1]臺灣大學 職能治療系；[2]長庚大學 職能治療系；[3]朝陽技術大學 應用工業工程系；[4]中國醫藥大學 運動醫學系
受文日期：民國96年8月22日；接受刊載：民國97年1月7日
通訊作者：陳嘉玲醫師，長庚紀念醫院 復健科。桃園縣333龜山鄉復興街5號。Tel.: (03)3281200轉3846; Fax: (03) 3281320; E-mail: clingchen@gmail.com