Acoustic Evidence for the Development of Aspiration Feature in Putonghua Stops

Ji-Yeon Han*

ABSTRACT

This study was investigated developmental temporal features in Putonghua-speaking children. The total of 212children between the ages 2;6 and 6;5 participated in Shanghai. Speech materials were constructed according to aspiration feature in stop sounds of Putonghua. Six words were selected in this study. A voice onset time was measured. Non-parametric procedures were employed for all the analyses. The VOT value across bilabial, alveolar, and velar stops was significantly differed between aspirated and unaspirated stops for each age group. Effect of age is significant for unaspirated stops. It is clear that each of Putonghua stops showed decreasing mean and standard deviation. The overshoot phenomenon of VOT was apparent from the age of 2;6-2;11 to 4;6-4;11. There was high variability in the production of lag time for aspirated stops.

Keywords: development, Putonghua, voice onset time, aspiration

1. Introduction

Acoustic data in children speech has two major trends before it has the values of adult-like range. The reduction of magnitude and variability is clearly noted as a function of age. Overshoot phenomena of acoustic parameters is appeared for covering to their final values.

In child's speech, VOT showed short-lag than that of adults. Macken and Barton (1979) found English-learning children first use voiceless unaspirated, and short-lag voiceless stops. Distributional characteristics of VOT was investigated in /p, t/ from 7 five-year-old children and 14 adults (Koenig, 2001). VOT distributions tended to have rightward skew values were more extreme for the children. It was, therefore, suggested that VOT development need to address the distributional characteristics of the data. Development of VOT appears to be fairly systematic. As Kent (1976) noted, changes in the VOT distributions is from unimodal to bimodal. Child's first word of VOT distribution sgenerally are unimodal having short-lag range. By the age of three years, adult liked the bimodal form begin. But it is mostly occur in the voiced range of the continuum. Greatly overlapped bimodal form compared to that of adults has

* Dept. of Speech and Hearing Sciences East China Normal University

distinct form by the age of eight.

The study of Macken and Barton (1980) formulate a three-step developmental sequence of VOT. During stage III mean values of VOT showed over than for adults. Barton and Macken (1980) reported overshoot phase in acquiring voiceless aspirates at the age of four. High variability persisted in the voiceless aspirated category until age 12 or later (Kent, 1976; Ohde, 1985). The variability of VOT decreases as getting complete bimodal distribution, also. In the study of the development of the aspiration in Cantonese-speaking children (Clumeck, Barton, Macken, & Huntington, 1981), the aspiration contrast was marked at age 2;4 by lengthening of the VOT that was not reached the range of adult value. At age 4;0 the place of articulation for labial, alveolar, and velar stops was acquired by the VOT with overly long lag compared to adults.

There are studies that reveal the difference between speech disorders and normal children in terms of acoustic analysis. The language-delayed children's control of the acoustic-phonemic details of the voicing contrasts was less mature than that of the normal-speaking children through the spectorgram analysis (Bond & Wilson, 1980). Although auditory-perceptual analysis remains the basis of speech disorder assessment, acoustic analysis of speech has potential clinical value because it provides quantities information that contributes to speech subsystem description and identifies objective and measurable of speech intelligibility and quality. By examine error patters of phonemic contrast which is related with acoustic and physiological information, speech pathologist is able to define exactly the factors give rise to speech deficits in terms of speech physiology (Nelson, Leeper, Blomgre, & Cameron, 2001).

Although perceptual analysis by transcription has been used the basic method in the study of speech development and disorder, acoustic analysis of speech has perspective underlying representation as well as clinical value because it provides quantitative information that contributes to identify objective and measurable of speech quality. This study will apply the method of acoustic measure, which contributes to provide objective information in speech development. It will provide acoustic data in the aspect of development of Putonghua speech sound.

2. Methods

2.1 Subjects

The total of 212 children between the ages 2;6 and 6;5 participated in Shanghai. Age interval was six month, and the age group was consisted of boys and girls. The subject was required as following criteria. The distribution of subjects by gender and age is shown Table 1.

(1) The absence of hearing and speaking problems.

(2) The absence of allergies, or upper respiratory infections at the time of testing.

(3) Screened for normal nasal air emission.

(4) Primary language used in home is Putonghua or Shanghaihua.

Table 1. Subjects by gender and age

Age group	Male	Female	Total
2;6 - 2;11	13	12	25
3;0 - 3;5	15	15	30
3;6 - 3;11	12	15	27
4;0 - 4;5	16	12	28
4;6 - 4;11	9	14	21
5;0 - 5;5	15	15	30
5;6 - 5;11	14	15	29
6;0 - 6;5	10	10	20
Total	104	108	212

2.2. Materials

1) Speech materials

Speech materials were constructed according to aspiration features in stop sounds of Putonghua. Six words were selected in this study, and target words were presented in Table 2.

Table 2. Target words for aspiration feature in stop sounds

Pair	aspiration feature	Target word
Bilabial stop	Aspiration	抛 /pʰaʊ1/
	Unaspiration	包 /paʊ1/
Alveolar stop	Aspiration	套 /tʰaʊ4/
	Unaspiration	稻 /taʊ4/
Velar stop	Aspiration	靠 /kʰu1/
	Unaspiration	菇 /ku1/

2) Recording Procedures

Direction for administrating test was settled with using subsequent questions, prompts, and imitations. Speech samples were recorded via a microphone to digital using KORG-DS1200 (KORG, Inc.) in an isolated a sound-attenuated room. The KORG-DS1200 unit was placed on anther table neared by examiner. The microphone was at a distance of around 5 cm 10 cm from the subject. A teacher who had rapport on children administrated the test. Each word was recorded as produced the 50 words, three times at each word. Subject repeated the word at approximately 1-2-s intervals. Whole test for one child was spent about 15 minutes.

2.3 Procedures

Acoustic analysis was conducted by Dr. Speech (Tiger electronics, Inc.). The sampling frequency was set to 44 kHz for data recording, which is considerably high. The chosen 44 kHz allows simple down sampling to 11 kHz. The window length used 24 ms, and overlapped 12 ms. Hamming window type was selected. Because some children, mostly among of the younger ones, speak very weakly and softly, dB was boosted to clearly see noise part with 6 dB interval.

In order to ensure the reliability of acoustic measurement, the data of 2;6-3;11 years was measured again by examiner. Person's correlation analyses showed a high degree of inter-rater reliability (r = 0.9, p < 0.001).

2.4 Data Analysis

VOT (msec) is defined as the time equivalent of the space on the waveform, from the onset of the release burst to onset of voicing of a following vowel (the presence of periodicity cycle). The point of the onset of vowel was decided near by up going zero crossing. VOT is crucial in determining the distinction between aspirated and unaspirated sound in Mandarine Chinese (Liu et al., 2000; Wu & Lin, 1989).

2.5 Statistical Analysis

Normal distribution was examined by test of skewness for each variable across all categories. The results showed in Table 5 of summary of numeric data. Non-parametric procedures were employed for all the analyses. Sex-difference was determined using the Mann-Whitney U test. As results there was no exist sex-related difference at alpa level .0001. Therefore, the factor of sex was not considered. To examine significant difference among of acoustic values at each age, For each sound, the effect of age was determined by the Kruskal-Wallis ANOVA at alpa level .05. On significant results, the Mann-Whitney U statistic was run at the Bonferroni correctness < .05/8 = .0063.

Acoustic variables were showed by the graph of boxplots and error bars categorized by aging. Descriptive statistics was entered mean, range (minimum : maximum), and skewness.

3. Results and Discussion

The VOT value across bilabial, alveolar, and velar stops was significantly differed between aspirated and unaspirated stops for each age group. Each minimal pair of bilabial, alveolar, and velar stops was contrasted in the aspects of aspiration vs. unaspiration. All variables showed significant difference in Table 3.

Table 3. The Wilcox signed ranks test of VOT

		$/p^h/$ - /p/	$/t^h/$ - /t/	$/k^h/$ - /k/	Aspirated unaspirated stop
2;6-2;11	Z	-4.104	-4.319	-4.029	-4.319
	Sig.	0.000*	0.000*	0.000*	0.000*
3;0-3;5	Z	-4.783	-4.700	-4.762	-4.782
	Sig.	0.000*	0.000*	0.000*	0.000*
3;6-3;11	Z	-4.541	-4.541	-4.541	-4.541
	Sig.	0.000*	0.000*	0.000*	0.000*
4;0-4;5	Z	-4.623	-4.623	-4.623	-4.623
	Sig.	0.000*	0.000*	0.000*	0.000*
4;6-4;11	Z	-4.198	-4.198	-4.198	-4.198
	Sig.	0.000*	0.000*	0.000*	0.000*
5;0-5;5	Z	-4.782	-4.782	-4.783	-4.782
	Sig.	0.000*	0.000*	0.000*	0.000*
5;6-5;11	Z	-4.704	-4.703	-4.703	-4.703
	Sig.	0.000*	0.000*	0.000*	0.000*
6;0-6;5	Z	-3.883	-3.921	-3.920	-3.920
	Sig.	0.000*	0.000*	0.000*	0.000*

* $p < .05$

To examine age-related differences on VOT for aspirated and unaspirated initial stops, the Kruskal-Wallis non-parameter ANOVA was used on each of 6 stops and each group of aspirated and unaspirated stops. Table 4 shows that effect of age is significant for unaspirated stops (χ^2 = 16.147, *df* = 7, p = 0.024). On the VOT values of unaspirated stops, Multiple comparisons were conducted by Mann-Whitney U-tests at the Bonferroni correction .05/8 = .00625. Results indicated there was statistically significant differences the groups of age 2;6-2;11 and two age groups 5;6-5;11 (U = 171.500, p = 0.001) and 6;0-6;5 (U = 114.500, p = 0.002).

Table 4. The Kruskal-Wallis ANOVA of age on VOT

	$/p^h/$	/p/	$/t^h/$	/t/	$/k^h/$	/k/	Aspirated stops	Unaspirated stops
χ^2	11.438	8.928	2.568	5.488	6.929	10.405	5.213	16.147
df	7	7	7	7	7	7	7	7
Sig.	0.121	0.258	0.922	0.601	0.436	0.167	0.634	0.024*

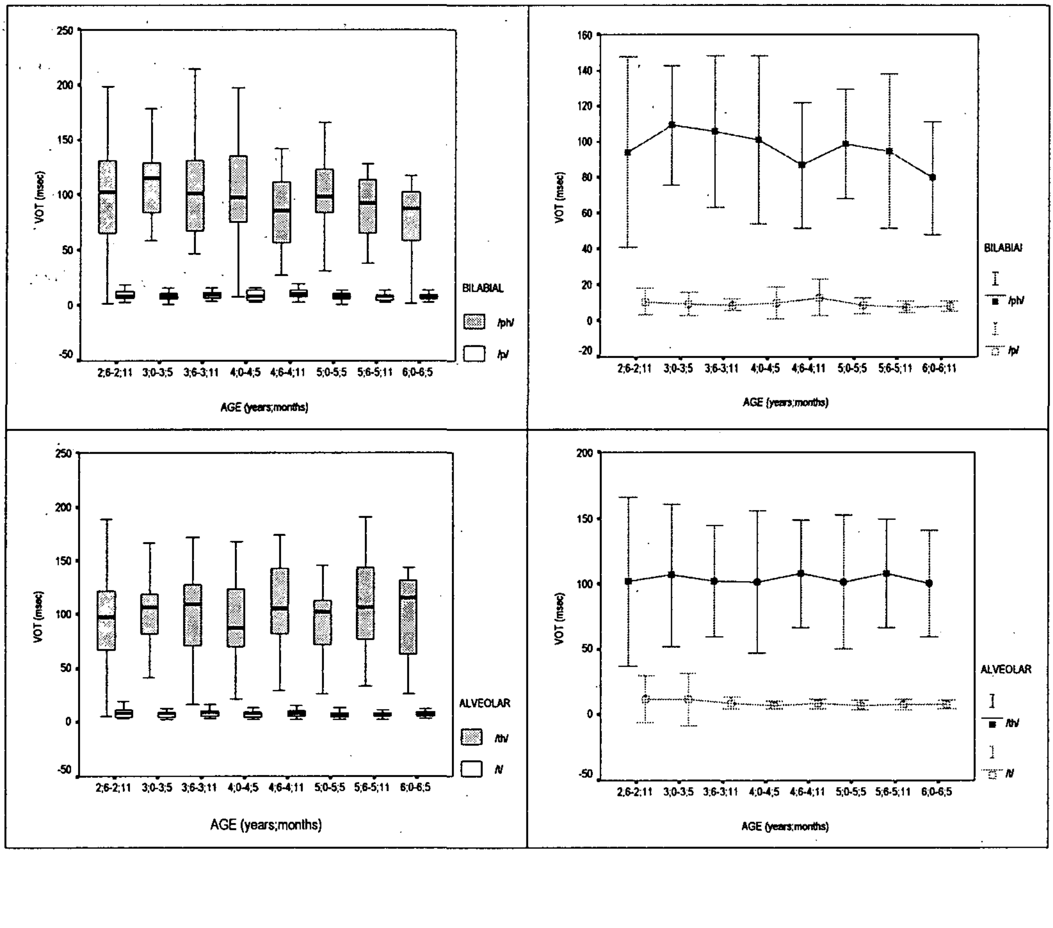
* $p < .05$

The following Figure 1 shows the comparison of the VOT value in the aspects of mean, range, and standard deviation. There was overlap with the range between aspirated and unaspirated stops at age 2;6-2;11, and slightly overlapped at age 4;0-4;5. The mean of VOT of aspirated stops (247 msec) was higher than that of unaspirated stops (28 msec). In Figure 1, it is clear that between-subject mean and standard deviation decrease as function of age. The mean and variability decreasing VOT was agreed with previous studies (Kent, 1976; Ohde, 1985). It was noted that high standard deviations at age 4;0-4;5 for aspirated stops and at age

4;6-4;11 for unaspirated stops. From the aspect of averaged VOT as older, The highest VOT value showed at age 3;0-3;5 for aspirated stops (265 msec), and unaspirated stops (36 msec) at age 2;6-2;11. These two findings could be explained with the reason of overshoot phase. It is well corresponded with previous study on developmental aspects of VOT in children speech (Barton & Marken, 1980).

Three minimal pairs exhibited overlap trends of VOT range in Figure 1. The bilabial stop minimal pair was showed the overlapped ranged at age 2;6-2;11, 4;0-4;5, and 6;0-6;5. The /pH/ (97 msec) is higher than /p/ (10 msec) forthe mean of VOT. At age 3;0-3;5 (110 msec) the aspirated bilabial stop /pH/ showed the highest the mean of VOT, at age 4;6-4;11 (13 msec) for unaspirated bilabial stop /p/ (Table 3).

In the case of alveolar stops contrasted by feature of aspiration, there was overlap at age 2;6-2;11 and 3;6-3;11. The mean of VOT for aspirated alveolar stop /tH/ (104 msec) is higher for unaspirated alveolar stop /t/ (9 msec). The highest VOT value was existed at the age of 4;6-4;11 (108 msec) for aspirated alveolar stop /tH/, and at the age of 2;6-2;11 (12 msec) for unaspirated alveolar stop /t/.



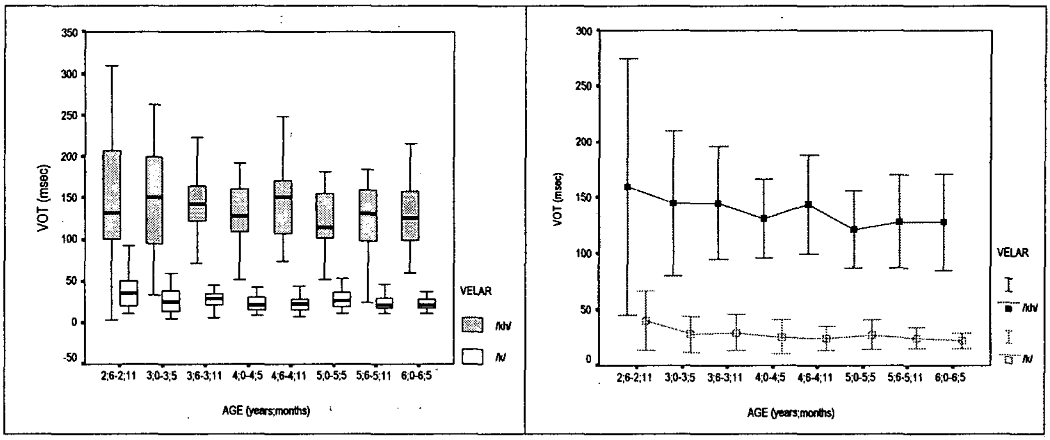


Figure 1. Duration and standard deviation of VOT for bilabial, alveolar, and velar stop. (a) and (b) for bilabial stop, (c) and (d) for alveolar stop, (d) and (e) for velar stop.

The velar stop showed highly overlap of range to compare with other bilabial and alveolar stops. It was overlapped at age 2;6-2;11, 3;0-3;5, 5;0-5;5 and 5;6-5;11 in Figure 1. The mean of VOT value showed the higher for aspirated velar stop /kH/ (138 msec) than for unaspirated velar was stop /k/ (28 msec). Velar exhibited the highest VOT value at age 2;6-2;11 for aspirated velar stop /kH/ (160 msec) and unaspirated velar stop /k/ (40 msec).

Table 5. Summary of numeric data of VOT for aspiration feature in stops

		/pʰ/	/p/	/tʰ/	/t/	/kʰ/	/k/	Aspirated stops	Unaspirated stops
2;6-2;11	Mean	94	11	102	12	160	40	249	36
	R.	2 : 199	3 : 35	6 : 317	4 : 95	4 : 593	12 : 116	13 : 504	17 : 114
	Skew	-0.283	2.043	1.483	4.547	2.281	1.291	0.388	2.687
3;0-3;5	Mean	110	9	107	12	146	28	265	31
	R.	13 : 178	1 : 30	8 : 319	3 : 114	34 : 262	5 : 60	101 : 491	11 : 131
	Skew	-0.657	1.632	1.894	5.098	-0.109	0.391	0.833	3.947
3;6-3;11	Mean	106	9	102	9	145	30	257	28
	R.	47 : 215	4 : 16	17 : 172	4 : 20	36 : 289	6 : 77	84 : 375	13 : 53
	Skew	0.637	-0.011	-0.312	0.994	0.33	1.407	-0.666	0.826
4;0-4;5	Mean	101	10	101	8	131	26	246	26
	R.	8 : 198	3 : 50	22 : 250	3 : 14	52 : 192	9 : 70	55 : 469	13 : 66
	Skew	-0.01	3.453	1.347	0.479	-0.246	1.68	0.567	1.973
4;6-4;11	Mean	87	13	108	9	144	24	242	30
	R.	27 : 142	3 : 43	29 : 174	3 : 16	74 : 248	8 : 44	123 : 376	15 : 66
	Skew	0.113	2.37	-0.23	0.708	0.469	0.429	0.199	1.794
5;0-5;5	Mean	99	9	101	7	122	28	241	25

	R.	31 : 166	1 : 23	26 : 307	3 : 21	52 : 182	11 : 71	92 : 525	10 : 45
	Skew	-0.012	1.172	2.135	1.9	0.058	1.391	1.802	0.217
5;6-5;11	Mean	95	8	108	8	129	25	246	24
	R.	38 : 238	3 : 17	33 : 190	3 : 26	24 : 184	11 : 47	125 : 477	16 : 47
	Skew	1.581	1.025	0.215	3.323	-0.562	0.924	0.917	1.79
6;0-6;5	Mean	80	8	100	8	129	23	223	24
	R.	2 : 118	3 : 14	26 : 144	4 : 17	60 : 215	12 : 38	90 : 317	15 : 34
	Skew	-0.972	0.581	-0.819	1.123	0.426	0.536	-0.612	0.295
Total	Mean	97	10	104	9	138	28	247	28
	R.	2 : 238	1 : 50	6 : 319	3 : 114	4 : 593	5 : 116	13 : 525	10 : 131
	Skew	0.186	3.118	1.148	8.329	2.32	1.828	0.451	3.895

4. Conclusion

It is clear that each of minimal pairs of AUS contrast showed decreasing mean and standard deviation. The overshoot phenomenon of VOT was apparent from the age of 2;6-2;11 to 4;6-4;11. There was high variability in the production of lag time for aspirated stops. Clumeck et al. (1981) reported Cantonese-speaking children did not reach for producing aspirated stops until by age 4. In this study, the VOT value of unaspirated stops had a significant age-dependent effect in compare with aspirated stops. This finding may be interpreted as improving of articulatory-timing coordination during that period until 5;0-5;6 for unaspirated stops.

References

Barton, J. A. & Macken, M. A. 1980. “An instrumental analysis of the voicing contrast in word-initial stops in the speech of four-year-old English-speaking children.” *Language and Speech*, 23, 159-169.

Bond, Z. S. & Wilson, H. F. 1980. “Acquisition of the voicing contrast language-delayed and normal-speaking children.” Journal of Speech, *Language and Hearing Research*, 23, 152-161.

Clumeck, H., Barton, D., Macken, M. A. & Huntington, D. A. 1981. “The aspiration contrast in Cantonese word-initial stops: data from children and adults.” *Journal of Chinese Linguistics*, 9, 210-225.

Koening, L. L. 2001. “Distributional characteristics of VOT in children's voiceless aspirated stops and interpretation of developmental trends.” *Journal of Speech, Language, and Hearing Research*, 44, 1058-1068.

Liu, H. M., Tseng, C. H. & Tsao, F. M. 2000. “Perceptual and acoustic analysis of speech intelligibility on Mandarin-speaking young adults with cerebral palsy.” *Clinical Lingusitics & Phonetics*, 14, 447-464.

Macken, M. A. & Barton, D. 1979. “A longitudinal study of the acquisition of voicing contrast in American-English word-initial stops, as measured by voice onset time.” *Journal of Child Language*, 7, 433-458.

Macken, M. A. & Barton, D. 1980. “The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonant.” *Journal of Child Language*, 7, 41-74.

Ohde, R. N. 1985. “Fundamental frequency correlates of stop consonanat vocing and vowel quality in the speech of preadolescent children.” *Journal of Acoustical Society of America*, 78, 1554-1561.

Wu, J. & Lin, M. 1987. *Shiyanyuyinxuegaiyao.* Beijing: Gaodengjiaoyu Chubanshe [in Chinese].

접수일자: 2005. 07. 30
게재결정: 2005. 09. 01

▲ Ji-yeon Han
Dept. of Speech and Hearing Sciences East China Normal University
9-208 Lucky APT Onchen 2-dong Dongrae-gu Busan, 607-753, KOREA
E-mail: lotus73@dreamwiz.com