SPECTRAL ANALYSIS OF SIBILANT FRICATIVES AND THE LING SOUND TEST FOR SPEAKERS OF CHINESE DIALECTS

Aijun LI[a] & Zhiqiang LI[b]

[a]Chinese Academy of Social Sciences, China; [b]University of San Francisco, USA
liaj@cass.org.cn; zqli@usfca.edu

ABSTRACT

The Ling six-sound test is widely used clinically to verify effectiveness of hearing aids or cochlear implant fitting. When the Ling test is administered to speakers of major Chinese dialects, decisions have to be made about the specific sounds to use based on dialect-specific phonological contrasts and frequency distribution patterns. Spectral properties of the sibilant fricatives and affricates were analyzed in Beijing, Wuhan and Xiamen. The results showed that in order to adequately encompass the whole frequency range to accommodate the fact that the alveolar /s/ in all three dialects is produced with a much higher center of gravity value than its English counterpart, a palatal /ɕ/ or /tɕʰ/ is introduced to represent the 5-7kHz range, along with the post-alveolar /ʂ/ in Beijing and the aspirated alveolar /tsʰ/ in Wuhan and Xiamen, leading to a revised seven-sound test. Justifications for using tones in lower pitch register are also provided.

Keywords: Ling sound test, sibilant fricatives, frequency distributions, Chinese dialects

1. INTRODUCTION

According to a national survey conducted in 2006 [13], there are 27.8 million people with hearing loss in China, and it is estimated that 23,000 infants were born with hearing impairment each year. As hearing screening techniques have become more readily available than ever before, early detection and intervention will be key to rehabilitation.

The Ling sound test [12] is a hearing test widely used in audiometry to assess children's ability to hear across the frequency range essential for spoken language development for people with hearing impairment, and to verify effectiveness of hearing aids or cochlear implant fitting in children. Compared to other assessment methods, the Ling test can be easily administered as either a detection or identification task with high validity [5, 10, 21, 23]. Originally developed for the American population, the test selects familiar speech sounds that would broadly represent the speech spectrum from 250 to 8000 Hz, with concentration of acoustic energy manifested in low, mid and high frequency ranges. The speech sounds used in the test include three vowels /i/, /u/, /a/, two sibilant fricatives /s/, /ʃ/, and a nasal /m/, which was added later due to its low frequency content. The test was subsequently referred to as the Ling six-sound test [12]. Table 1 shows frequency distributions of the six sounds [12, 14]. Based on the values of center of gravity, /m/ is usually used to assess hearing of the low-frequency sounds while /s/ and /ʃ/ are used to test hearing of the high-frequency sounds. The three corner vowels are used to test perception of the vowel space defined by the first two formants F1 and F2.

Table 1: Frequency distributions of the six sounds in Hz.

Sounds	F1	F2	Center of gravity
/m/			250-500
/u/	350	900	
/a/	700	1300	
/i/	300	2500	
/ʃ/			2000-4000
/s/			3500-7000

When the Ling sound test is used clinically outside North America, differences between the American English and the language of the population need to be accounted for. Agung et al. [1] recommended replacing /u/ with /ɔ/ to better represent the vowel space in Australian English because the vowel /u/ is articulated with the tongue in a more fronted position and often considered a high central vowel.

For Chinese speakers, previous studies have focused on production and perception of lexical tones in Mandarin-speaking children with cochlear implants [7, 18, 23, 24]. Li et al. [8] turned their attention to the Mandarin adaptation of the Ling six sounds. Mandarin does not have /ʃ/, so they used the post-alveolar /ʂ/ (also termed retroflex fricative) instead in their study. Their analysis based on data from 60 speakers showed that /s/ in Mandarin is produced with a much higher center of gravity than its counterpart in American English as reported in [12]. They proposed an expanded seven-sound test (/m, u, a, i, s, ʂ, ɕ/) to include /ɕ/, whose center of gravity is located in the 5-7kHz range.

Following the methodology in [8], the current study considers the fact that a significant portion of Chinese population, including those with hearing aids and cochlear implant users, are speakers of Chinese dialects or even minority languages, which are phonologically distinct from the official variety of Mandarin. Many of these dialects have different phonological contrasts and are mutually unintelligible. For example, the three-way contrast of /s/, /ʂ/, /ɕ/ in Mandarin is completely missing in Cantonese as the latter has only /s/ among the three fricatives. Other dialects lack /ʂ/ and its affricated forms. Phonological contrasts in fricatives have implications for establishing the dialect-specific sound inventory before conducting the Ling sound test for speakers of Chinese dialects.

In addition, all Chinese dialects use tones to distinguish lexical meanings and many have very complex tone systems. However, lexical tones are currently not considered in the Ling sound test. As Li et al. [8] suggested, lexical tones at the higher and lower register of the pitch range should be incorporated in the test of vowels.

The goal of this study is to analyze spectral properties of sibilant fricatives and affricates in two major Chinese dialects, Wuhan and Xiamen, to determine frequency distributions of these sounds and their eligibility to target the frequency bands essential for the Ling sound test. Comparisons will be made with the Mandarin data reported in [8].

2. PHONOLOGICAL SYSTEMS OF FRICATIVES AND AFFRICATES

Wuhan dialect is a major variety of Southwestern Mandarin, spoken in the most populous city in Central China. Xiamen, formerly known as Amoy, is home to the Southern Min dialect, which is also spoken in Taiwan and Southeast Asia. Table 2 summaries the phonological systems of fricatives, affricates and nasals in Wuhan [9], Xiamen [11, 21] and Mandarin [15-17]. As can be seen from Table 2, The phonological register of Mandarin Chinese has three voiceless fricatives: alveolar /s/, post-alveolar (retroflex) /ʂ/ and palatal /ɕ/ that also occur in affricated forms where they can be both aspirated (/ts^h, tʂ^h, tɕ^h/) and unaspirated (/ts, tʂ, tɕ/). Wuhan and Xiamen, however, lack the post-alveolar series. The palatal series in Xiamen is often treated as allophonic variants of the alveolar series, which becomes palatalized before the high vowel /i/ or the pre-nuclear glide /j/. In traditional descriptions like [11, 21], the articulation of /s/ is said to be more fronted and less palatalized than its affricated forms before the high vowel /i/. All three nasals can occur in the syllable onset, except the velar nasal /ŋ/ in Mandarin.

Like most Mandarin dialects, Wuhan and Mandarin both have four lexical tones, but they differ in tonal values. Xiamen has a seven-tone system: five long tones and two checked tones that only occur on syllables ending in /p, t, k/ or a glottal stop. To simplify the discussion, we label tones numerically and provide transcriptions on a five-point scale in the table below, where 1 represents the lowest pitch and 5 the highest. Checked tones are underlined.

Table 2: Fricatives, affricates and nasals in Wuhan, Xiamen and Mandarin.

	Wuhan			Xiamen			Mandarin		
fricatives	s		ɕ	s		(ɕ)	s	ʂ	ɕ
affricates	ts		tɕ	ts		(tɕ)	ts	tʂ	tɕ
	ts^h		tɕ^h	ts^h		(tɕ^h)	ts^h	tʂ^h	tɕ^h
nasals	m	n	ŋ	m	n	ŋ	m	n	

Table 3: Tones in Wuhan, Xiamen and Mandarin.

	T1	T2	T3	T4	T5	T6	T7
Mandarin	55	35	214	51			
Wuhan	55	213	42	35			
Xiamen	55	35	53	11	33	<u>11</u>	<u>55</u>

3. METHODS

3.1. Participants

40 speakers of Wuhan dialect and 44 speakers of Xiamen dialect were recruited for the experiment. There is an equal split of female and male speakers (total = 88) in both groups, with the average age of 22.5 (SD=1.86). They reported no hearing impairment and can communicate in Mandarin.

3.2. Materials

Recall that the three consonants in the Ling sound test are /m/, /s/ and /ʃ/. Wuhan and Xiamen have /m/ and /s/, but do not have post-alveolar fricatives, so we included palatal fricatives and affricates in the current study. We recorded both consonants and single syllables in which the vowels are chosen to be in the same place of articulation as the consonants when the combinations are phonologically permissible, such as /sɿ/, /ts^{h}ɿ/ in Wuhan and /ɕi/, /tɕ^{h}i/ in Xiamen. Palatal fricatives and affricates can be elicited in the context of the high vowel /i/ in Xiamen. Single syllables are often used in sound tests by audiologists. The list of recording materials is given in Table 4.

Table 4: Single consonants and monosyllables used in the experiment.

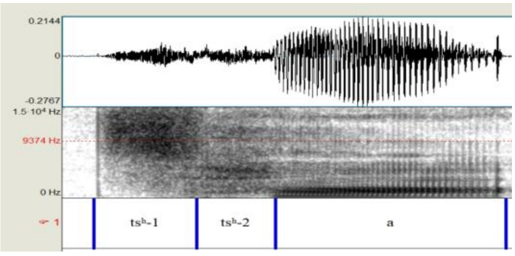
	Wuhan	Xiamen	Mandarin
fricatives	s, sɿ, ɕ, ɕi	s, sa, ɕ, ɕi	s, sɿ, ʂ, ʂʅ, ɕ, ɕi
affricates	ts^h, ts^{h}ɿ, tɕ^h, tɕ^{h}i	ts^h, ts^{h}a, tɕ^h, tɕ^{h}i	ts^h, ts^{h}ɿ, tʂ^h, tʂ^{h}ʅ, tɕ^h, tɕ^{h}i
nasals	m, mo	m, mi	m, mo

3.3. Recording and Data Extraction

The elicited production of the single consonants and syllables was digitally recorded at a sampling rate of 44.1 kHz with a 16-bit resolution in a sound-proof recording booth with ambient noise around 30dB SPL. Participants – half are males and half are females – were instructed to produce all materials with a high-level tone, which is adequate for the current study. In total, 2000 tokens were obtained (10 sounds × 5 times × 40 speakers) from Wuhan speakers, and 2200 tokens were obtained (10 sounds × 5 times × 44 speakers) from Xiamen speakers.

All sounds were manually segmented in Praat [3]. An example of segmentation and labelling from Xiamen is given in Fig. 1. Measurements were made of the frication noise in aspirated affricates, corresponding to the segment labelled "tsʰ-1" below. Spectral properties are characterized by the first four moments of the consonants, i.e. center of gravity, standard deviation, skewness and kurtosis, extracted using Praat and analyzed by gender of the speakers.

Figure 1: Segmentation and labelling of /tsʰa/.



4. RESULTS

4.1. Spectral properties of consonants in Wuhan

Average values of the four parameters for each test sound are presented in Table 5, with female average on top of male average in each cell. Results showed significant gender difference with respect to center of gravity (CG) values: female speakers have higher CG than male speakers in all 10 sounds tested ($p<0.01$). Fig. 2 plots the average CG values and one SD (standard deviation) for female and male speakers separately.

Due to influence of the following vowels, fricatives and affricates said in syllables generally show slightly lower center of gravity than those said in isolation. Significant differences were spotted between /s/ and /sɿ/ in both female and male speakers, and additionally, between /ɕ/ and /ɕi/ in female speakers, and between /tsʰ/ and /tsʰɿ/ in male speakers. In terms of frequency distribution, /s/ has its most acoustic energy concentrated in the high frequency range, followed by /tsʰ/, /tɕʰ/, /ɕ/ and /m/.

Table 5: Consonant moments in Wuhan, n=40, CG=center of gravity, female average on top of male average in each cell.

	CG (Hz)	SD (Hz)	skewness	kurtosis
m	253	110	54.86	8375.1
	195	146	24.45	1894.99
mo	244	88	75.49	13060.53
	191	108	31.93	4674.52
ɕ	6558	2310	0.81	2.72
	5408	2294	1.15	3.22
ɕi	6246	2352	0.8	2.51
	5280	2295	1.23	3.14
tɕʰ	6672	2496	0.68	2.08
	5202	2390	1.11	2.79
tɕʰi	6356	2435	0.76	2.34
	5286	2358	1.14	2.89
tsʰ	9774	2990	-0.67	2.5
	8408	3298	-0.25	0.75
tsʰɿ	9282	3006	-0.42	1.62
	7635	3237	-0.02	0.38
s	10656	2579	-0.84	3.72
	9190	3080	-0.34	1.25
sɿ	10207	2541	-0.72	3.31
	8634	3110	-0.2	0.77

Figure 2: CG for consonants in Wuhan, n=40, **: $p<0.01$, *: $p<0.05$.

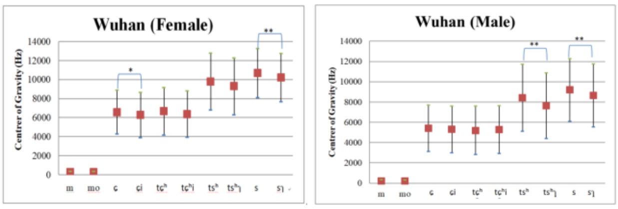


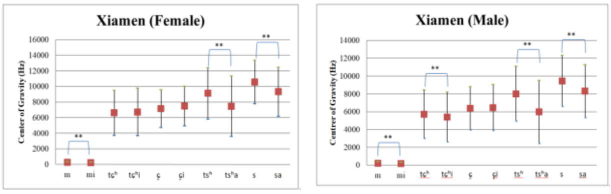
Table 6: Consonant moments in Xiamen, n=44, CG=center of gravity, female average on top of male average in each cell.

	CG (Hz)	SD (Hz)	skewness	kurtosis
m	241	117	60.66	10806.90
	192	165	19.46	1134.78
mi	225	89	93.95	29318.48
	177	114	26.65	3206.86
tɕʰ	6596	2882	0.34	1.63
	5703	2727	0.72	1.68
tɕʰi	6704	3025	0.08	1.3
	5365	2794	0.69	1.44
ɕ	7137	2435	0.43	2.85
	6362	2433	1.00	3.06
ɕi	7473	2539	0.16	2.56
	6451	2588	0.61	2.29
tsʰ	9076	3322	-0.86	2.28
	7989	3085	-0.34	1.91
tsʰa	7446	3875	-0.55	1.08
	5956	3519	0.13	1.52
s	10533	2815	-1.04	4.31
	9436	2877	-0.71	2.78
sa	9292	3139	-1.14	3.61
	8282	2965	-0.45	1.96

4.2. Spectral properties of consonants in Xiamen

Like Wuhan, Xiamen results in Table 6 showed similar significant gender differences with respect to center of gravity values with the only exception of /tsʰa/ ($p=0.57>0.05$). Other 9 sounds all showed higher CG values in female speakers than male speakers. Fig. 3 plots the average CG values for female and male speakers separately.

Figure 3: CG for consonants in Xiamen, n=44, **: p<0.01.



Similar patterns of differences in center of gravity between consonants said in syllables and in isolation were also observed in Xiamen. Between /m/ and /mi/, /tsʰ/ and /tsʰa/, /s/ and /sa/ all showed significant differences in female and male speakers at the p<0.01 level. Male speakers also showed significant differences between /tɕʰ/ and /tɕʰi/ at the p<0.01 level. In terms of frequency distribution, /s/ is followed by /tsʰ/, /ɕ/, /tɕʰ/ and /m/, from high frequency range to low.

Significant differences were also identified when we examined how phonologically similar sounds are phonetically realized across dialects. Three-way ANOVA tests revealed great variability in the values of center of gravity, especially in female speakers, as illustrated in Table 7. While /s/ has the most consistent phonetic realization across dialects, palatal sounds /ɕ/ and /tɕʰ/ exhibit greater variability than other sounds. As a result, hearing screening and assessment should be based on dialect-specific analyses of the frequency distributions of the consonants.

Table 7: Tukey HSD pairwise comparisons of phonologically similar sounds across dialects in Wuhan (WH), Xiamen (XM) and Mandarin Chinese (MC), F: Female, M: Male.

WH-XM-MC	WH–XM		WH–MC		XM-MC	
	F	M	F	M	F	M
m-m-m	*	--	--	**	**	**
ɕ-ɕ̡-ɕ	*	**	**	**	**	--
tɕʰ-tɕʰ-tɕʰ	--	**	*	**	**	**
tsʰ-tsʰ-tsʰ	**	**	**	--	--	--
s-s-s	--	--	**	--	**	--
mo-mi-mo	**	**	--	**	**	**
ɕi-ɕ̡i-ɕi	**	**	**	**	--	--
tɕʰi-tɕʰi-tɕʰi	--	--	**	**	**	**
tsʰɿ-tsʰa-tsʰɿ	**	**	--	--	**	**
sɿ-sa-sɿ	**	--	*	--	*	--

5. DISCUSSIONS AND RECOMMENDATIONS

In our analysis, the frequency band of /m/ in Wuhan is located in the lowest frequency range of 180-250 Hz. /ɕ/ and /tɕʰ/ are in the similar frequency range of 3.5-7 kHz. /tsʰ/ is much higher in the 6-11 kHz range, and /s/ is in the 7-12 kHz range. In Xiamen, the spectral properties of /ɕ̡/ are similar to that of /ɕ/ in Mandarin, as both show spectral prominence in the range of 4.5-8 kHz, higher than /tɕʰ/ in the range of 3.5-8 kHz. /tsʰ/ and /s/ are in the frequency ranges of 5-10 kHz and 7-12 kHz, respectively. It is worth noting that the alveolar fricative /s/ in Chinese dialects is articulated with much higher center of gravity frequency than its counterpart in American English. In other words, with /m/ in the similar low frequency range in English and Chinese, Chinese speakers have to deal with a more expanded frequency range than English speakers.

Based on the above analysis, we recommend testing /m/, /ɕ/, /tsʰ/ and /s/ in Wuhan and /m/, /tɕʰ/, /tsʰ/ and /s/ in Xiamen so as to adequately encompass the whole frequency range essential for speech development. A similar recommendation was made for Mandarin [8]. Combining them with three corner vowels /a/, /i/ and /u/ gives rise to a seven-sound test. Note that /sa/ should be used instead of /si/ in Xiamen when conducting syllable-based tests because of the palatalization process that involves /s/ before the high vowel /i/.

A list of test sounds and their mean center of gravity frequency values ± one SD are given in Table 8 as a reference for conducting hearing tests.

Table 8: Recommended test sounds and their mean center of gravity frequency values ± one SD in Wuhan, Xiamen and Mandarin.

Wuhan			Xiamen			Mandarin Chinese		
	Female	Male		Female	Male		Female	Male
m	253±110	195±146	m	241±117	192±165	m	261±109	209±162
ɕ	6558±2310	5408±2294	tɕʰ	6596±2882	5703±2727	ɕ	7797±1777	6778±2119
tsʰ	9774±2990	8408±3298	tsʰ	9076±3322	7989±3085	ʂ	5168±2187	4862±2233
s	10656±2579	9190±3080	s	10533±2815	9436±2877	s	9843±2190	1116±598
mo	244±88	191±108	mi	225±89	177±114	mo	252±88	203±121
ɕi	6246±2352	5280±2295	tɕʰi	6704±3025	5365±2794	ɕi	7629±1772	6495±2036
tsʰɿ	9282±3006	7635±3237	tsʰa	7446±3875	5956±3519	ʂʅ	4780±2184	4283±2185
sɿ	10207±2541	8634±3110	sa	9292±3139	8282±2965	sɿ	9677±2067	8558±2287

Tonal changes happen in the low frequency range. In tone languages like Chinese, lexical meanings are distinguished by tones. While we used only high tone here, we recommend incorporating test of tonal variations by including lexical tones that occur in the lower register of pitch range in hearing tests in future studies. For example, Tone 2 in Wuhan is a low dipping tone (213), and Tone 4 in Xiamen is a low level (11). The ability to perceive and produce lexical tones will be critical for cochlear implant fitting in children.

6. ACKNOWLEDGEMENTS

This research is supported by the Key NSSFC (No. 15ZDB103) and National Program on Key Basic Research Project (973 Program) under Grant 2013 CB329301, as well as Beijing Municipal Administration of Hospitals Clinical Medicine Development of Special Funding (No. XMLX201848).

7. REFERENCES

[1] Agung, K. B., Purdy, S. C., and Kitamura, C. 2005. The Ling sound test revisited. *The Australian and New Zealand Journal of Audiology* 27(1), 33-34.

[2] Bao, H. 1984. Putonghua danyuanyin fenlei de shengli jieshi [Physiological accounts of monophthongs in Standard Chinese], *Studies of the Chinese Language* (2).

[3] Boersma, P. and Weenink, D. 2017. Praat: Doing phonetics by computer, http://www.fon.hum.uva.nl/praat/.

[4] CASS (Chinese Academy of Social Sciences) and Australian Academy of the Humanities. 1987. *Zhongguo Yuyan Dituji [Chinese Language Atlas]*. Hong Kong, China: Hong Kong Longman (Far East) Co., Ltd.

[5] Chen X. 2011. Ying'er tingjue nengli pinggu [Infant hearing ability assessment]. *Chinese ENT News and Reviews* 26(2), 106-108.

[6] Dong, T. 1958. Sige Minnan fangyan [Four Minnan dialects]. *Bulletin of the Institute of History and Philology Academia Sinica* 30(2), 729-1042.

[7] Han, D., Liu B., Zhou, N., Chen, X., Kong, Y., Liu, H., Zheng, Y., & Xu, L. 2009. Lexical tone perception with HiResolution and HiResolution 120 sound-processing strategies in pediatric Mandarin-speaking cochlear implant users. *Ear and Hearing,* 30, 169-177.

[8] Li, A., Zhang, H., Sun, W. 2017. The Frequency Range of "The Ling Six Sounds" in Standard Chinese. *Proc. Interspeech 2017,* 1864-1868. DOI: 10.21437/Interspeech.2017-329.

[9] Li, R. 1995. *Dictionary of Wuhan Dialect*. Nanjing: Jiangsu Education Publishing House.

[10] Liu, J., Sun, X. 2012. Evaluations on Rehabilitation Requirement and Effect for Cochlear Implanted Children. *Disability Studies* (2), 39-44.

[11] Lou, C. 1930. *Xiamen yinxi [Phonology of Xiamen Dialect]*. Book 4 in Separate Book Series I of the Institute of History and Philology Academia Sinica.

[12] Ling, D. 1976. *Speech and the Hearing-impaired Child: Theory and Practice*. Washington, DC: Alexander Graham Bell Association for the Deaf.

[13] Sun, X., Yu, L., Qu, C. et al., 2008. Zhongguo tingli canji goucheng tedian ji kangfu duice [An epidemiological study on the hearing-impaired population identified in China and proposed intervention strategies]. *Chinese Scientific Journal of Hearing and Speech Rehabilitation* 27(2), 21-24.

[14] Tye-Murray N. 1998. *Foundations of Aural Rehabilitation: Children and Their Family Members*. San Diego, CA: Singular.

[15] Wu, Z. 1964. Putonghua yuanyin he fuyin de pinlü fenxi ji gongzhenfeng de cesuan [The spectral analysis of vowels and consonants and the measurement of formants in Standard Chinese], *Acta Acustica* 1.

[16] Wu Z. (ed.). 1986. *Hanyu Putonghua danyinjie yutuce [The Spectrogram of Monosyllables in Standard Chinese]*. Beijing: Chinese Social Science Press.

[17] Wu, Z. and M. Lin. 1989. *Shiyan yuyinxue gaiyao [An Introduction to Experimental Phonetics]*. Beijing: Higher Education Press.

[18] Xu, L. & Zhou, N. 2011. Tonal languages and cochlear implants. In Zeng, F.-G., Popper, A. N., & Fay, R. R. (Eds). *Auditory Prostheses: New Horizons* (pp. 341-364). Springer Science+Business Media, LLC, New York.

[19] Yang, J. 2006. P*utonghua seyin yu cayinpu de shengxue texing yanjiu [An Acoustic Study on the Spectral Properties of Stops and Fricatives in Standard Chinese]*. Thesis. Beijing: Chinese Academy of Social Sciences, Institute of Linguistics.

[20] Yuan, G. and Zhang, H. 2016. Chinese questionnaires of artificial effect assessment. *Int J Otolaryngol Head Neck Surg* 40(4), 215-219.

[21] Zhou, C. 1996. *Xiamen fangyan cidian [Dialect Dictionary of Xiamen]*. Nanjing: Jiangsu Education Publishing House.

[22] Zhou, X., Hua, Q. and Cao, Y. 2008. Ertong yanyu ceting cailiao [A review of children's speech audiometry materials]. *Journal of Audiology and Speech Pathology* 16(2), 160-163.

[23] Zhou, N., & Xu, L. 2008. Development and evaluation of methods for assessing tone production skills in Mandarin-speaking children with cochlear implants. *Journal of the Acoustical Society of America*, 123, 1653-1664.

[24] Zhou, N., Huang, J., Chen, X., & Xu, L. 2013. Relationship between tone perception and production in prelingually-deafened children with cochlear implants. *Otology & Neurotology*, 34(3), 499-506.