ASPIRATION AND VOICING OF CHINESE AND ENGLISH PLOSIVES

David Deterding and Francis Nolan

National Institute of Education, Singapore and University of Cambridge
david.deterding@nie.edu.sg / fjn1@cam.ac.uk

ABSTRACT

The six plosives of Standard Chinese are compared with those of RP British English, to see if there is a difference in their aspiration and/or voicing. Recordings of 7 speakers from China reading words beginning with each of the 6 plosives are compared to similar recordings of 7 speakers of RP British English, and it is found that there is little difference in the aspiration of the plosives in the two languages, though there is a difference in the voicing during the closure when the plosive occurs between two vowels.

Keywords: Plosives, aspiration, VOT, voicing, Chinese, English.

1. INTRODUCTION

The six plosives of Standard Chinese are usually classified as voiceless unaspirated /p, t, k/ and voiceless aspirated /p^h, t^h, k^h/ [3, 5]. However, there is little evidence that these six plosives are much different from the English plosives which tend to be classified as voiced /b, d, g/ and voiceless /p, t, k/ [4, 6]. In reality, we know that, in most circumstances, none of the English plosives are voiced and it is aspiration that largely cues the difference between these two sets of consonants. So why is the classification in the two languages different?

In the current study, seven male Chinese speakers from various provinces were recorded reading a list of monosyllabic words, and the duration of aspiration of their plosives is compared with similar recordings of seven male speakers of RP British English.

2. SPEAKERS

The Chinese speakers all come from the People's Republic of China and, at the time of the recording, their average age was 19.3 years and they had been in Singapore for about twelve weeks. They were enrolled on an intensive English language programme in preparation to study for a university degree. They come from a range of provinces in Central, East, and North China. Their places of origin and ages are listed in Table 1.

Table 1: Chinese subjects.

Speaker	*Age*	*Home Province*
C1	20	Henan
C2	19	Hunan
C3	20	Shandong
C4	19	Anhui
C5	21	Shandong
C6	18	Liaoning
C7	18	Jiangsu

All the Chinese subjects claimed to speak Mandarin as their home language. It is possible that some of them also speak a regional variety of Chinese in some situations, though none of them stated this in the biodata questionnaire they filled in.

The British speakers are all lecturers in Singapore, and their average age was 50.6 years at the time of the recording. All have an RP or near-RP accent. Brief details of the British speakers are shown in Table 2.

Table 2: British subjects.

Speaker	*Age*	*Place of Origin*
B1	47	London
B2	49	London
B3	46	London
B4	60	Bristol
B5	57	Southampton
B6	52	London
B7	43	Cardiff

Although there is a clear difference in terms of background (students vs lecturers) and age (average 19.3 years vs 50.6 years) between the two groups of speakers, they constitute reasonably representative speakers of the two languages.

3. DATA

All the subjects read a list of ten monosyllabic words embedded in a carrier phrase. The Chinese words were just presented in characters. The carrier phrase in both Chinese and English has three syllables before the word under investigation and four syllables following it, and in both languages, the preceding word ends with a vowel and the following word starts with an affricate. The Chinese words and carrier phrase are listed in Table 3 (where the pronunciation is that suggested in [5]), and the English words and carrier phrase are shown in Table 4.

Table 3: The Chinese words and carrier phrase.

Carrier Phrase	请 你 把 ___ 再 说 一 遍 tɕʰiŋ ni pa ___ tsai ʃuo i pien please (obj) ___ again say once			
Words	1	夫	fu	('husband')
	2	半	pan	('half')
	3	离	li	('separate')
	4	判	pʰan	('judge')
	5	蛋	tan	('egg')
	6	书	ʃu	('book')
	7	炭	tʰan	('coal')
	8	干	kan	('do')
	9	看	kʰan	('read')
	10	马	ma	('horse')

Table 4: The English words and carrier phrase.

Carrier Phrase	Now please say ____ just one more time.	
Words	1	food
	2	bun
	3	leak
	4	pan
	5	done
	6	shoe
	7	tan
	8	gun
	9	can
	10	mark

We are only concerned with items 2, 4, 5, 7, 8, and 9 which in both sets of data consist of a plosive, an open vowel and a nasal. The remaining items, 1, 3, 6 and 10, are all distracters. The phonetic shape of the distracters is also similar in both languages except that some of the English items have a final consonant.

Despite the attempts to match the data from the two languages, there are inevitably some differences. In the Chinese, the word before the test word (*pa*) is a grammatical object marker, and it carries no tone (it has an empty tone). All the test words have a high-level or falling tone, which means that *pa* is spoken on a low pitch, and this is sometimes accompanied by creak. Furthermore, the fact that it is a function word while in English *say* is a content word may have some influence on the results.

Both sets of data were recorded directly onto a computer in the Phonetics Laboratory, with a high-quality microphone placed a few inches from the lips of the subjects.

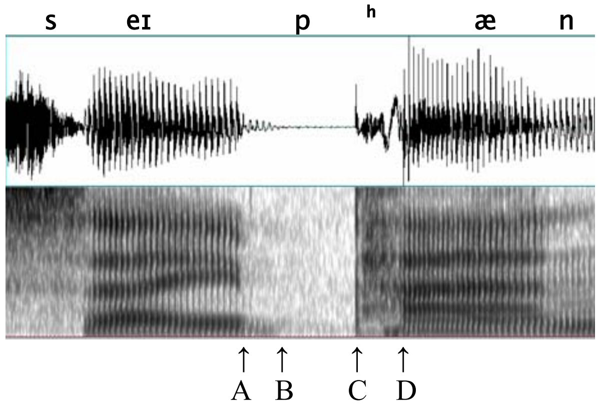
The Chinese subjects all read the data more quickly than the British subjects. The average for the Chinese is 20.6 sec while that for the British is 33.0 sec. This is partly related to the nature of the data (the English carrier phrase has final obstruents: /z/ in *please* and /st/ in *just*), but it is also probably related to the background of the speakers. As the British speakers are all lecturers, they are accustomed to speaking carefully.

4. MEASUREMENTS

All measurements were made independently by the two researchers by means of spectrograms generated using Praat [1]. For each token of each speaker, four separate time measurements were made, as illustrated in Fig 1.

- beginning of closure for the plosive (A)
- end of voicing during the closure (B)
- end of closure for the plosive (C)
- end of release burst (aspiration) (D)

Figure 1: Location of the four measurement points (for *pan* of B1.)



With some of the Chinese tokens, the end of the vowel in *pa* is breathy, which might be described as pre-aspiration of the plosive. In these cases, locations A and B are recorded as the same, so there is 0% of voicing during the plosive.

Most difficulty was encountered in determining the location of the end of voicing (B), and for some tokens it was something of a guess.

The end of the release burst (D) was measured by examination of the spectrogram at about 2000 Hz (basing it on the energy at the onset of the second formant). In many tokens, voicing starts earlier than D, so a measure of voice onset time (VOT) would give a smaller number. One might also select a slightly later point for D, as a high-frequency breathy start to the vowel sometimes continues longer than the blast of energy associated with the second formant. However, the location of D is relatively well defined and it gives a reasonable estimate for the duration of aspiration.

Two sets of results will be discussed: duration of aspiration at the release of the plosive, D – C; and percentage of the closure that is voiced:

$$\%voice = 100 \times \frac{B - A}{C - A}$$

There was a high level of agreement between the two measurers for the duration of aspiration, with on average a difference of 2.4 ms in one direction or the other between the two results. In only one case was there a substantial disagreement: a difference of 43 ms for the /g/ of B5. Further investigation confirmed that it is indeed difficult to determine the end of the closure (C) for this token, so it was excluded from the results for aspiration, together with the /k/ for B5. The results presented below are based on the average of the measurements for the two researchers for the remaining 82 tokens.

There was more discrepancy between the two measurers over %voice, mainly because of problems in determining the end of voicing (B). However, the overall correlation was 0.728, so there was reasonably good agreement in general. The results are based on the average %voice for the two measurers.

5. RESULTS

A summary of the results for aspiration is shown in Table 5. As expected, the figures for British English are a little larger than those reported elsewhere for VOT [2]. However, the pattern is similar to that reported in [2] with the greatest duration for velar plosives and least for bilabials, and this extends to the Chinese data as well.

Although the aspirated plosives of Chinese might seem to have slightly longer aspiration than the British ones (89.7 ms compared with 82.3 ms), the difference is not in fact significant (t=1.8, df=2, two-tailed, paired-sample, ns). And similarly there is no significant difference between the results for any of the three aspirated plosives.

For the unaspirated plosives, the overall value for the Chinese aspiration (16.6 ms) is slightly less than for the British (18.8 ms), but once again this difference is not significant (t=2.0, df=2, two-tailed, paired-sample, ns). With these values, the only difference found to be marginally significant is that for the alveolar plosives where the Chinese average of 13.6 ms is marginally less than the 17.7 ms for the British (t=2.369, df=12, two-tailed, independent samples, p<0.05), but even here the difference is small.

Table 5: Average duration of aspiration (in ms) for the Chinese and British plosives.

	Chinese		*British*		*Signif.*
aspirated	/p^h/	85.2	/p/	71.4	ns
	/t^h/	85.6	/t/	80.4	ns
	/k^h/	98.4	/k/	97.3	ns
	average	89.7		82.3	ns
unaspirated	/p/	11.0	/b/	11.0	ns
	/t/	13.6	/d/	17.7	*
	/k/	25.2	/g/	29.3	ns
	average	16.6		18.8	ns

The percentage of voicing (%voice) for the different plosives is shown in Table 6. It can be seen that, for the unaspirated plosives, there is a highly significant difference between the British and Chinese data both overall (t=5.87, df=82, two-tailed, independent samples, p<0.01) and also for each of the three plosives (for the bilabials, t=4.24, df=12, p<0.01; for the alveolars, t=4.11, df=12, p<0.01; for the velars, t=3.44, df=12, p<0.01). However, for the aspirated plosives, there is no difference between the two sets of data, either overall or for any of the three sounds.

Table 6: Percentage of voicing (%voice) during the closure for the Chinese and British plosives.

	Chinese		*British*		*Signif.*
aspirated	/p^h/	14.6	/p/	14.5	ns
	/t^h/	15.5	/t/	22.4	ns
	/k^h/	8.6	/k/	20.2	ns
	average	12.9		19.0	ns
unaspirated	/p/	12.2	/b/	47.3	**
	/t/	19.7	/d/	59.0	**
	/k/	19.9	/g/	67.0	**
	average	17.2		57.8	**

The differences for the unaspirated plosives can perhaps best be illustrated by comparing two example spectrograms. A Chinese bilabial token is shown in Fig. 2 and a British one in Fig. 3. Both measurers found %voice of about 22% for the token in Fig. 2 and 65% for that in Fig. 3.

Figure 2: %voice for *pan* of C1

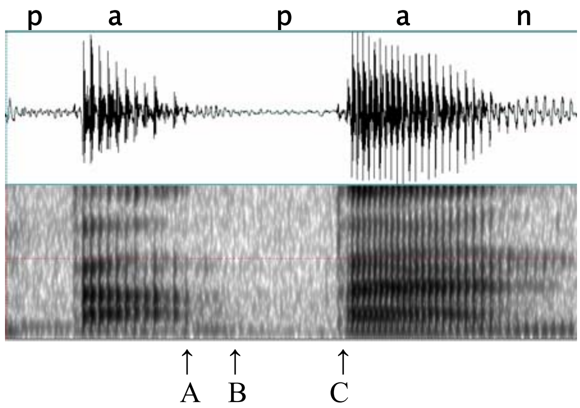
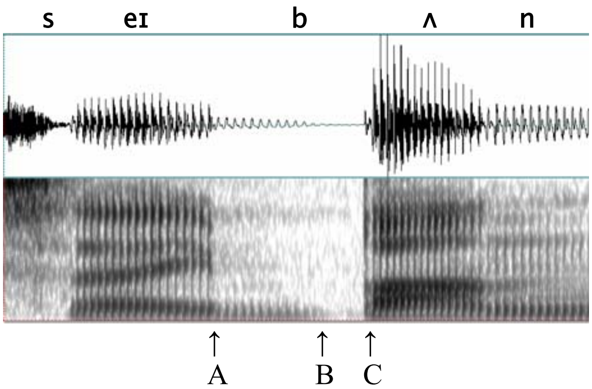


Figure 3: %voice for *bun* of B1



Of course, there may be other important differences between the Chinese and British data shown in Figs 2 and 3, such as the intensity of the burst; but they do illustrate the differences in voicing during the closure quite clearly.

6. DISCUSSION

Little difference in the duration of the release burst has been found between Chinese and English aspirated plosives. The main difference between the two sets of data lies in the voicing of unaspirated plosives in intervocalic position. However, we should remember that English /b, d, g/ would not generally be voiced following a voiceless sound or after a pause, and in those environments there would be little if any difference between Chinese and English plosives. In fact, it has recently been suggested that a contrast between aspirated voiceless and unaspirated voiceless plosives (T^h vs T) is the unmarked system of contrasts for plosives [7], and this seems to be the basic system that occurs in both Chinese and English. So why are the plosives in these two languages often classified differently?

One factor that may influence the classification lies in the patterning of other consonants. English has four voiced fricatives /v, ð, z, ʒ/ which contrast with the corresponding voiceless fricatives /f, θ, s, ʃ/, but Chinese has no voiced fricatives, only voiceless ones. So voicing has a wider role in the description of English consonants, but not Chinese ones, and that may partly explain why many analysts also use voicing for the classification of plosives in English but not Chinese. In contrast, in Chinese the unaspirated affricates /ts, tʃ, tɕ/ contrast with aspirated ones /ts^h, tʃ^h, tɕ^h/, and this is why aspiration is adopted for the classification of Chinese plosives.

In conclusion, although the evidence of duration of aspiration gives no support to a difference in phonological analysis between English and Chinese, the difference in intervocalic behaviour, with English unaspirated stops being more prone to voicing incursion, suggests the traditional analysis may be attempting to capture a real phonetic difference, as well as reflecting the broader patterning of obstruents.

7. REFERENCES

[1] Boersma, P., Weenink, D. Praat: doing phonetics by computer. http://www.fon.hum.uva.nl/praat/ visited 11-May-05
[2] Docherty, G. J. 1992. *The Timing of Voicing in British English Obstruents*. Berlin: Foris.
[3] Duanmu, S. 2000. *The Phonology of Standard Chinese*. Oxford: Oxford University Press.
[4] Ladefoged, P. 1999. American English. In IPA (ed.) *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 41–44.
[5] Lee, W-S., Zee, E. 2003. Standard Chinese (Beijing), *Journal of the International Phonetic Association*, 33, 109–112.
[6] Roach, P. 2004. British English: Received Pronunciation. *Journal of the International Phonetic Association*, 34, 239–245.
[7] Vaux, B., Samuels, B. 2005. Laryngeal markedness and aspiration. *Phonology*, 22, 395-436.

8. ACKNOWLEDGEMENTS

This work was partly supported by National Institute of Education Academic Research project NIE: RI 1/03 LEL: *Theoretical Speech Research and its Practical Implications*.