

Effect of Japanese Articulation of Stops on Pronunciation of Chinese Aspirated Sounds by Japanese Students

Akemi Hoshino and Akio Yasuda*

Toyama National College of Maritime Technology
1-2 Neriya, Ebie, Shinminato City, Toyama 933-0293
hoshino@toyama-cmt.ac.jp

*Tokyo University of Marine Science and Technology, Tokyo 135-8533

ABSTRACT

It is relatively easy for Japanese students studying Chinese to learn the pronunciation of the bilabial, alveolar, and velar groups of Chinese aspirated sounds compared to other sounds. We examined the voicing features for voice onset time and power during the utterance of Japanese stop consonants in bilabial, alveolar, and velar sounds, which are articulated in the same way as in Chinese. The sounds were uttered by 10 native Japanese who had not studied Chinese before. We then compared these features to the features of Chinese aspirated sounds uttered by 35 Japanese students, who had studied Chinese for one year, and eight native Chinese. The results showed that the features of the Japanese stops uttered by Japanese were similar to those for Chinese aspirated sounds. The features of the students' utterances were also similar to those of the Chinese speakers. This similarity appeared to be one of the main factors in students achieving good grades for their pronunciation of these Chinese sounds compared to other sounds.

1. INTRODUCTION

The pronunciation of Chinese is more complicated than that of any other language. Most of the pronunciation is quite different from Japanese pronunciation. Many Japanese students find it particularly difficult to pronounce Chinese aspirated sounds. It is difficult for one teacher to evaluate sounds uttered simultaneously by several students and there is also no way for students to evaluate their own pronunciation when practicing at home. After analyzing the pronunciation of Chinese aspirated sounds and extracting their voicing features, we proposed that the power during VOT (voice onset time) rather than the VOT itself was useful in evaluating pronunciation [1][2]. These research results will form the basis for developing an effective computer-aided instruction tool for Chinese pronunciation, which could be presented in a multimedia format.

In training Japanese students to pronounce Chinese aspiration sounds, we found they were able to pronounce the bilabial (pa[p'a], pi[p'i], po[p'o], and pu[p'u]), alveolar (ta[t'a] and ti[t'i]), and velar syllables (ka[k'a] and ku[k'u]), more easily than other Chinese aspirations. These syllables are uttered at almost the same articulation points as Japanese unaspirated stops for the bilabial ("pa", "pi", "po", and "pu"), alveolar ("ta" and "ti"), and velar syllables ("ka" and "ku"). In a previous study, we

proposed using measurements of the voicing features for VOT and power during VOT to evaluate Chinese pronunciation. In this study, we measured and compared these features for the pronunciation of Chinese aspirations uttered by nine native Chinese and 35 Japanese students, who had studied Chinese three hours per week for one year, and those of Japanese stops by 10 Japanese who had not studied Chinese before. The utterances of the Japanese students were evaluated by eight native Chinese in a hearing test. The results showed that features of the utterance of Japanese stops were similar to those for Chinese aspirated sounds uttered by Chinese, and those of Chinese sounds uttered by Japanese students were also similar to those uttered by Chinese. The students' utterances of Chinese aspirated sounds with the same articulating points as Japanese stops received good grades compared to those for the other sounds.

2. COMPARISON OF STOP CONSONANTS OF CHINESE AND JAPANESE

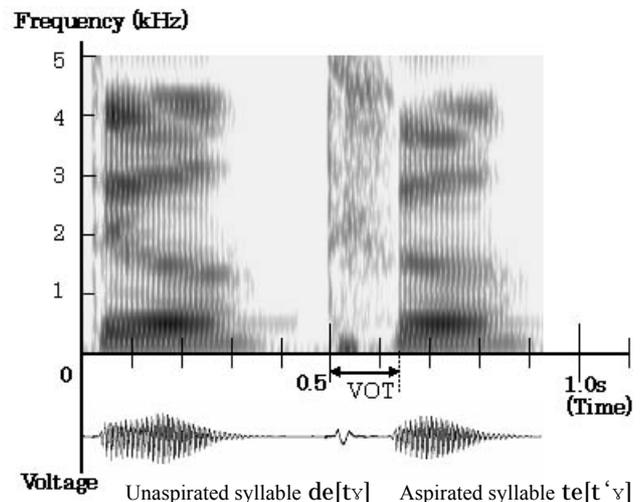


Figure 1: Spectrograms of unaspirated syllable de[tv] (left) and aspirated syllable te[t'v] (right) pronounced by Chinese speakers. Lower shows air-vibrations of the uttered sounds.

Stop releases of the breath are classified as aspirated or unaspirated. Aspiration is a breathy noise generated as air passes through the partially closed vocal folds and into the pharynx [3]. The points of articulation in stop consonants of Chinese are

bilabial, alveolar, and velar and consist of pairs of aspirated and unaspirated syllables, e.g., bilabial b[p^ˈ] (unaspirated) /p[p^ˈ](aspirated); alveolar d[t^ˈ] (unaspirated) /t[t^ˈ] (aspirated); and velar g[k^ˈ] (unaspirated) /k[k^ˈ] (aspirated).

Figure 1 shows the air vibrations of the uttered sounds (lower part) and spectrograms of the unaspirated syllable de[te] (left) and the aspirated syllable te[t^ˈʰ] (right). In the right spectrogram, the aspiration appears in a brief interval between the stop burst and the onset of vocal-fold vibrations followed by a vowel. This time interval is the VOT [3]. The VOT in this figure is 135 ms. The onset of the vocal-fold vibration is so close to the burst that no aspiration interval appears in the left spectrogram.

Stop consonants in Japanese are also called bilabial, alveolar, and velar depending on the articulating points just as in Chinese. Although the stop consonants in Japanese are similar to those in Chinese, the articulation methods are somewhat different. The power and length of the breath at the time of a burst depend on the individual [4]. The vocal fold does not vibrate and the articulation method differs slightly for different people. For example, "pan" may be pronounced as unaspirated [paN] or aspirated [pa^ˈN] in Japanese. Unaspirated and aspirated sounds are interchangeable in Japanese and do not differentiate meaning.

In Chinese, however, aspirated and unaspirated sounds serve to discriminate between meanings [5].

3. EVALUATION PROCEDURE

We examined the VOT and power during VOT for Chinese and Japanese single-vowel syllables with similar points of articulation. They included pairs of Chinese aspirated bilabial syllables (pa[p^ˈa], pi[p^ˈi], po[p^ˈo], and pu[p^ˈu]), and Japanese unaspirated ones ("pa", "pi", "po", and "pu"); Chinese aspirated alveolar syllables (ta[t^ˈa], ti[t^ˈi], te[t^ˈv], and tu[t^ˈu]) and Japanese ones ("ta" and "ti"); and Chinese aspirated velar syllables (ka[k^ˈa], ke[k^ˈv], and ku[k^ˈu]) and Japanese ones ("ka" and "ku"). The syllables were pronounced by nine native Chinese speakers, 35 Japanese students who had studied Chinese for three hours per week for one year, and 10 Japanese who had not studied Chinese. We calculated the VOT and relative average power during VOT from spectrograms using the procedure reported previously [1] [2]. Eight native Chinese speakers acted as examiners in a hearing test of the pronunciation of the 35 students. The grades were as follows: 3 = pronunciation that sounded aspirated; 2 = unclear sounds; and 1 = unaspirated sounds.

Table 1. Average VOT and relative average power (Pav) of Chinese bilabial aspirated sounds made by 35 students and 9 Chinese, and Japanese bilabial sounds made by 10 Japanese.

Chinese aspiration	pa[p ^ˈ a]			pi[p ^ˈ i]			po[p ^ˈ o]			pu[p ^ˈ u]		
	VOT	Pav	Grade									
Student	31.58	32.24	2.70	37.19	42.57	2.60	43.88	15.47	2.60	59.58	20.70	2.50
Chinese	71.12	38.70		78.09	16.28		65.04	24.78		59.91	47.56	
Japanese sound	pa			pi			po			pu		
	VOT	Pav	Grade									
Japanese	26.68	49.84		34.28	12.48		30.11	15.52		28.75	27.98	

Table 2. Average VOT (ms) and relative average power of Chinese aspirated alveolar sounds made by 35 students and 9 Chinese, and Japanese alveolar sounds made by 10 Japanese.

Chinese aspiration	ta[t ^ˈ a]			ti[t ^ˈ i]			te[t ^ˈ v]			tu[t ^ˈ u]		
	VOT	Pav	Grade									
Student	45.63	3.60	2.60	54.66	1.22	2.70	48.78	1.24	2.20	51.62	2.18	2.30
Chinese	76.54	5.12		87.25	0.87		61.23	4.86		77.58	4.70	
Japanese sound	ta			ti			non			non		
	VOT	Pav	Grade									
Japanese	23.33	0.25		21.10	0.05							

Table 3. Average VOT (ms) and relative average power (Pav) of Chinese aspirated velar sounds made by 35 students and 8 Chinese, and Japanese velar sounds made by 10 Japanese.

Chinese aspiration	ka[k ^ˈ a]			ke[k ^ˈ v]			ku[k ^ˈ u]		
	VOT	Pav	Grade	VOT	Pav	Grade	VOT	Pav	Grade
Student	71.37	0.23	2.80	77.6	0.08	2.40	85.53	0.07	2.60
Chinese	81.90	0.27		99.51	0.14		90.00	0.82	
Japanese sound	ka			non			ku		
	VOT	Pav	Grade	VOT	Pav	Grade	VOT	Pav	Grade
Japanese	46.25	0.01					54.66	0.007	

4. RESULTS OF EVALUATION

Tables 1, 2, and 3 show the average VOT and relative average power during VOT for the pronunciation of Chinese aspirations by 35 Japanese students and eight native Chinese and for the pronunciation of Japanese sounds with similar articulating points as those of Chinese aspirations by 10 Japanese. The average grades of the students for bilabial, alveolar, and velar syllables, respectively, are also indicated.

It is important in training Japanese students to pronounce Chinese to examine differences in the voicing features of VOT and relative average power during VOT between the students and Chinese speakers. It is also important to be aware of differences in the features of Chinese aspiration sounds and Japanese sounds with similar articulation points to Chinese. We have already reported that the VOT and relative average power during VOT are effective in evaluating the pronunciation of bilabial Chinese aspirations [1][2]. However, we did not examine Japanese bilabial sounds with the same articulating points.

Table 1 shows the results for the bilabial sounds. Although the average value of the students' VOT for pa[p'a] was somewhat longer than the Japanese "pa", it was much shorter than that of the Chinese speakers. The average power of the Japanese "pa" was higher than for the Chinese pa[p'a]. This means that Japanese breathe out more strongly when they utter the Japanese "pa". Thus, it is considered an easier sound for students to articulate. The average power was just 17% lower than for the Chinese speakers. This is the main reason that the students' utterances received a high average grade of 2.7. The relative average power of the Japanese "pi" was only 27% lower than that of the Chinese pi[p'i]. The students used more power than the Chinese speakers. The same tendency was seen for "pa" with this utterance receiving a good average grade of 2.6. The average VOT and power for the Japanese "po" are small compared to those for the Chinese po[p'o]. They were 67% and 62%, respectively, of those of Chinese speakers. However, this utterance produced several measurements with a sufficiently long VOT and with appropriate power in a specific range of the VOT [1][2]. It therefore received a good average grade of 2.6. Although the students' power was considerably less than that of the Chinese speakers for pu[p'u], the average value of the VOT was the same as for a Chinese speaker. Since this utterance produced a relatively large number of measurements with a sufficiently long VOT, it received a good average grade of 2.5.

Table 2 shows the results for the alveolar sounds. The articulating point of a Chinese alveolar is slightly different from that of Japanese. The Chinese articulating point is the alveolus. The tip of the tongue touches the upper alveolus and then the sound is pronounced [6]. In contrast, the Japanese articulating point involves both the alveolus and teeth. Quite a large portion of the front of the tongue is applied to the alveolus, the tip touches the teeth, and then the sound is pronounced. Thus, it is difficult to produce a fricative after a burst that is different from that of other Chinese aspirated sounds [7]. Although the average values of VOT and power for the Japanese "ta" are smaller than for the Chinese aspirated syllable ta[t'a] uttered by Chinese speakers, the students found it easy to articulate because the articulating point of the Chinese ta[t'a] is almost the same as for

the Japanese "ta". The students' power for the Chinese ta[t'a] was 70% of that of the Chinese speakers, which was not a large difference, and this utterance received a good average grade of 2.6. The Japanese sound "ti" comes originally from English and is familiar to Japanese students. The articulating point is almost the same as for the Chinese aspiration ti[t'i]. The average power of the students' utterance of ti[t'i] was high and was 1.4 times that of the Chinese speakers. Since this utterance by the students produced several samples with high power, it received a high average grade of 2.7.

There is no sound in Japanese with the same articulating point as the Chinese aspirated syllables te[t'e] and tu[t'u]. Thus, the students' power was as low as 26 and 46% of that of the Chinese speakers, and the utterances received low grades of 2.2 and 2.3, respectively.

Table 3 shows the results for the velar sounds. Since the articulating point for the Chinese ka[k'a] is the same as for the Japanese "ka", the average values of VOT and power of the students' utterances of the Chinese ka[k'a] did not differ much from that of the Chinese speakers. Figure 2 shows the distribution of data for the velar-aspirated syllable ka[k'a] in Chinese and the velar sound "ka" in Japanese with the VOT along the abscissa and the relative average power along the ordinate. As shown in the figure, the Japanese data are located in approximately the same domain as the Chinese speakers' data except for the leftmost and bottom data. This means that the phonetic features of the Chinese ka[k'a] and Japanese "ka" are similar. The students' utterances received a high average grade of 2.8. The Japanese "ku" has the same articulating point as the Chinese aspirated sound ku[k'u]. Although the average power of the students' utterance of the Chinese ku[k'u] was less than 10%, the average value of the VOT was as long as 86 ms, which differed from that of the Chinese speakers by only 4 ms. At 55 ms, the average VOT for the Japanese "ku" was longer than that for the other Japanese sounds. It is considered that the longer average VOT led to a good average grade of 2.6 for this utterance.

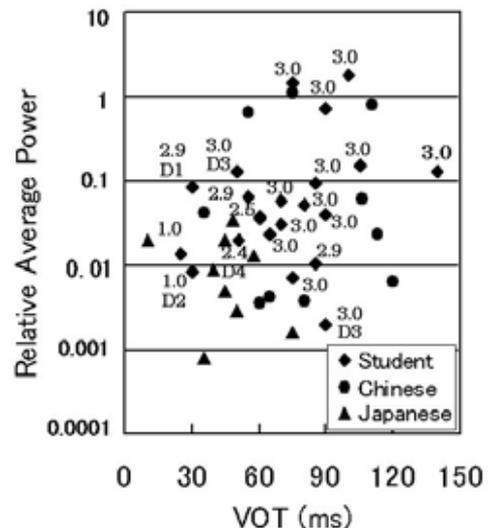


Figure 2: Distribution of data for Chinese velar aspirated syllable ka[k'a] and Japanese velar sound "ka" on the surface of VOT on abscissa and relative average power on ordinate. Numbers above students' data indicates grade.

No sound in Japanese has the same articulating point as the Chinese aspirated syllable $ke[k^{\text{h}}v]$. The average values of VOT and power of the utterances of the students were less than those of the Chinese speakers, which is probably why this utterance received an average grade of as low as 2.4.

5. EVALUATION OF ASPIRATED SOUNDS AND DISTRIBUTION OF DATA

Figures 3, 4, and 5 show the number of samples (ordinate) in each grade (abscissa) for the Chinese aspirated bilabial, alveolar, and velar sounds, respectively, made by the students. Since the four bilabial syllables in Figure 3 have the same articulating point as in Japanese, more than 70% of all the samples received an average grade of better than "2.6", which corresponds to good pronunciation. The syllables $ta[t^{\text{h}}a]$ and $ti[t^{\text{h}}i]$ in Figure 4 have almost the same articulating point as in Japanese, and 70 and 84% of these samples, respectively, received better than the average grade of "2.6". The syllables $te[t^{\text{h}}v]$ and $tu[t^{\text{h}}u]$, however, have a slightly different articulating point from that of the Japanese, and only 40 and 50% of these samples, respectively, received a grade of better than 2.6.

These results show that the students' utterances of the Chinese aspirated bilabial, alveolar, and velar syllables that have the same articulating points as in Japanese, though unaspirated, received high grades.

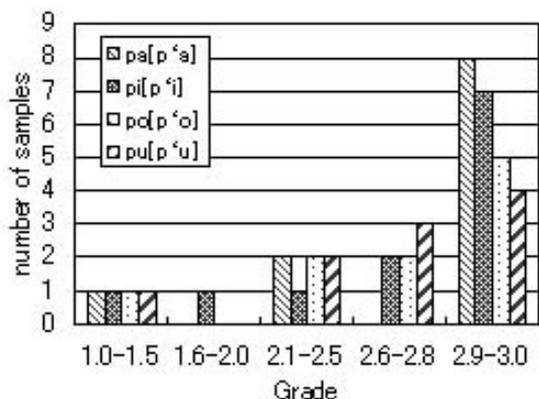


Figure 3 Distribution of data in each grade for pronunciation of Chinese aspirated bilabial sounds by students.

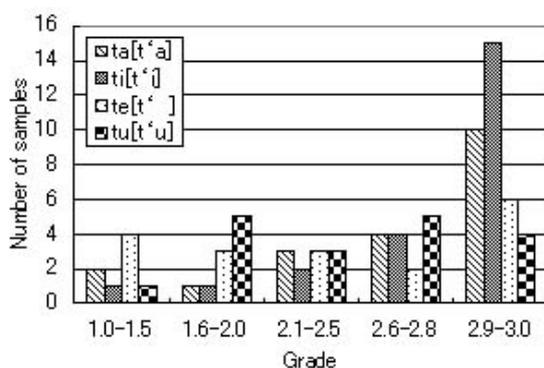


Figure 4 Distribution of data in each grade for pronunciation of Chinese aspirated alveolar sounds by students.

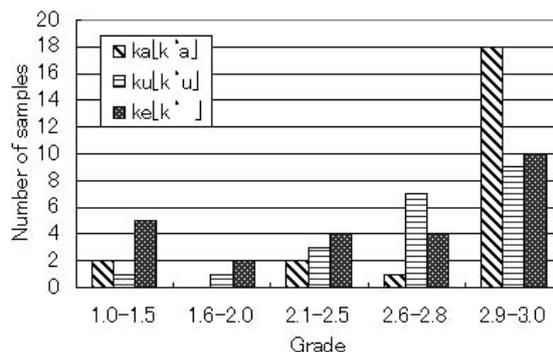


Figure 5 Distribution of data in each grade for pronunciation of Chinese aspirated velar sounds by students.

The aspirated syllables $ka[k^{\text{h}}a]$ and $ku[k^{\text{h}}u]$ in Figure 5 also have almost the same articulating point as in Japanese, and 83 and 80% of these samples, respectively, received grades of better than 2.6. The syllable $te[t^{\text{h}}v]$, however, has a different articulating point in Japanese, and only 61% of these samples received grades higher than 2.6, while 30% scored below 2.0.

6. CONCLUSION

In this paper, we examined the voicing features of VOT and relative average power during VOT for Chinese aspirated bilabial ($pa[p^{\text{h}}a]$, $pi[p^{\text{h}}i]$, $po[p^{\text{h}}o]$, and $pu[p^{\text{h}}u]$), alveolar ($ta[t^{\text{h}}a]$ and $ti[t^{\text{h}}i]$), and velar sounds ($ka[k^{\text{h}}a]$ and $ku[k^{\text{h}}u]$) with the same articulating points as the Japanese bilabial ("pa", "pi", "po", "pu"), alveolar ("ta" and "ti"), and velar sounds ("ka" and "ku"). The sounds were pronounced by 35 Japanese students of Chinese, eight Chinese speakers, and 10 Japanese. The features of the students' utterances of Chinese sounds with the same articulating point as similar Japanese sounds were close to those of the native Chinese speakers. In an evaluation, these utterances received good average grades of more than 2.6.

Our results showed that the relative average power during VOT was an effective measure to use in evaluating how well students pronounced Chinese aspirated sounds. We now plan to develop pronunciation training materials for students based on the articulation of Japanese stops and using the evaluation measures described above.

REFERENCES

- [1] A. Hoshino and A. Yasuda, 'Evaluation of Chinese aspiration sounds uttered by Japanese students using VOT and power (in Japanese),' *The Journal of the Acoustical Society of Japan*, Vol. 58, No. 11, pp. 689-695, Nov. 2002
- [2] A. Hoshino and A. Yasuda, 'Evaluation of Chinese aspiration sounds uttered by Japanese students based on VOT and power' *Proceeding of ICASSP 2003*, pp.472-475, Hong Kong, 2003
- [3] Ray D. Kent and Charles Read, 'The Acoustic Analysis of Speech,' Singular Publishing Group, Inc., San Diego and London, pp. 105-108, 1992
- [4] M. Enomoto, 'Japanese English- Phonetics in Spoken Language', Tamagawa University Publishing Co., Tokyo, pp. 44-48, 2000
- [5] Tamotsu Koizumi, 'An Introduction to Phonetics', Daigaku Shorin Publishing Co., Tokyo, pp. 73-74, 2003
- [6] S. Kawakami, 'Outline of the Japanese Pronunciation', Ouhuu Publishing Co. Tokyo, 2002
- [7] M. Enomoto, 'Japanese English- Phonetics in Spoken Language', Tamagawa University Publishing Co., Tokyo, pp. 95-97, 2000