# Automatic Recognition of Chinese Aspirated Sounds Pronounced by Japanese Students

Akemi Hoshino*
National Institute of Technology, Toyama College, 1-2 Ebie, Neriya, Imizu City, Toyama, 933-0293 Japan.

* Corresponding author. Tel +81-766-86-5215; email: hoshino@nc-toyama.ac.jp
Manuscript submitted June 15, 2017; accepted September 10, 2017.
doi: 10.17706/ijcce.2017.6.4.221-228

**Abstract:** Chinese aspirates are usually difficult to pronounce for Japanese students. In particular, discriminating between the utterances of aspirated and unaspirated sounds is the most difficult to learn for them. For self-learning, an automatic judgment system was developed that enabled students to check their pronunciations using a computer. We extracted the features of correctly pronounced single-vowel bilabial aspirated sounds pa[p‘a], pi[p‘i], po[p‘o], and pu[p‘u] and unaspirated sounds of ba[pa], bi[pi], bo[po], and bu[pu] by observing the spectrum evolution of breathing power during both voice onset time (VOT), and the voiced period when uttered by 50 native Chinese speakers. We developed a high performance 35-channel computerized filter bank to analyze the evolution of the breathing power spectrum using MATLAB and automatically evaluated the utterances of 50 Japanese students. Using a high-resolution spectrogram, we closely examined the features in VOT closely and improve the criteria for a proper pronunciation. We applied our developed automatic recognition system with improved criteria to the utterances of the students, which passed the screening of native-Chinese speakers. Although our system rejected several samples passing the native speakers’ screening, the success rates were higher than 95% and 98% for aspirated and unaspirated sounds, respectively.

**Key words:** Automatic recognition, Chinese aspirated sounds, extracted the features, self-learning system.

## 1. Introduction

Aspirated syllables in Chinese are generally difficult to pronounce for Japanese students, because the Japanese language does not have such sounds. In particular, how to discriminate between utterances with aspirated and unaspirated sounds is difficult for Japanese students. We observed many of the students sounded like unaspirated sounds to the instructor, which implies that the students were unable to breathe out with sufficient power to articulate the sounds correctly.

To enable the students to check their pronunciations while self-learning, we must establish certain criteria for correctly pronounced syllables. Some reports [1] showed that breathing power during the voice onset time (VOT) is a useful measure for evaluating the correct pronunciation of Chinese aspirates. The results indicated that the quality is closely correlated with the breathing power used in uttering a sound. We thus concluded that breathing power is also an important factor in evaluating the quality of the pronunciation. In addition, they also developed an automatic evaluation system [2] for the utterances of Chinese aspirated sounds in accordance with the two parameters, VOT and the breathing power during the VOT. They also developed a self-learning system for the automatic discrimination of Chinese aspirated and retroflex affricates [3], [4].

To develop the system to discriminate between aspirated and unaspirated pronunciations automatically,

we extracted the features of correctly pronounced single-vowel bilabial aspirated sounds pa[p‘a], pi[p‘i], po[p‘o], and pu[p‘u], and unaspirated sounds ba[pa], bi[pi], bo[po], and bu[pu], by analyzing the spectrum of breathing power during the VOT of the sounds when uttered by 50 Chinese speakers. For this purpose, we developed a computerized, high-performance 35-channel frequency filter bank. To improve the discrimination performance of these bilabial sounds, we extracted the features of correctly pronounced aspirated and unaspirated sounds by analyzing the frequency spectrum of breathing power during both the VOT and the voiced period of the sounds, and established improved evaluation criteria according to the current analysis. In this paper, we inspect closely the relation between good pronunciation and deduced parameters to improve the performance, then include the additional criteria and discuss the results in terms of how the developed system successfully discriminates between bilabial aspirated sounds and unaspirated sounds when pronounced by Japanese students [5], [6].

For the Chinese applied in this study, we used Standard Chinese or Modern Standard Chinese (Putonghua), based on the Beijing dialect.

## 2. Difference between Bilabial Aspirated and Unaspirated Sounds

In this section, we define the distinctive features discriminating between single vowel bilabial aspirated sound of [p‘] and unaspirated sound of [p] by examining the spectrogram of the pairs of pa[p‘a] - ba[pa], pi[p‘i] - bi[pi], po[p‘o]] -bo[po], and pu[p‘u] - bu[pu] uttered by a native- Chinese speaker.

Fig. 1 shows the temporal evolution of spectrograms of the bilabial unaspirated sound ba[pa] (left) and the bilabial aspirated sound pa[p‘a] (right) uttered by a Chinese speaker. The lower part of the figure shows the waveform of the voltage evolution as picked up by a microphone. The ordinate, extended upward, shows the frequency component and the shade of the stripes, implying the approximate power level by the darkness at the respective times and frequencies.

The aspirate appears in the brief interval at the right of the spectrogram pa[p‘a], indicated by the light, thin vertical stripes during VOT, between the stop burst and the onset of vocal fold vibrations followed by a vowel sound. This time interval is called VOT; in this case it is long, 120 ms. The dark gray vertical stripes in the center were observed between 1050 and 1850 Hz, 2450 and 3050 Hz, and 3650 and 4450 Hz, during VOT. This is caused by the release of a burst of energy that is created as the impounded air escapes. The onset of the vocal fold vibration is so close to the burst in the left of the spectrogram of ba[pa] that the VOT is zero. The thick horizontal bands in the voiced period on the right part of the spectrogram represent the formants that help listeners discriminate between the four bilabial unaspirated sounds. The criteria for discrimination are discussed later.

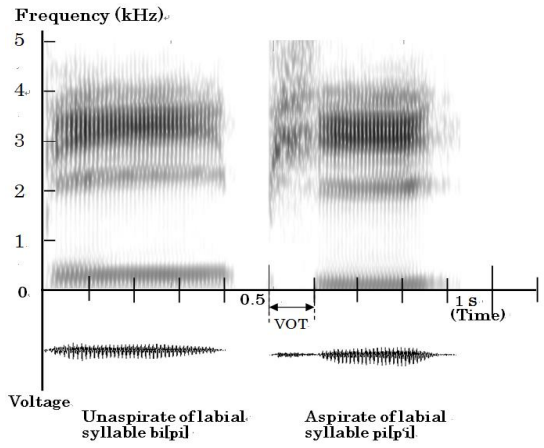


Fig. 1. Spectrograms of unaspirated syllable ba[pa] (left), and aspirated syllable pa[p‘a] (right) pronounced by a Chinese speaker.

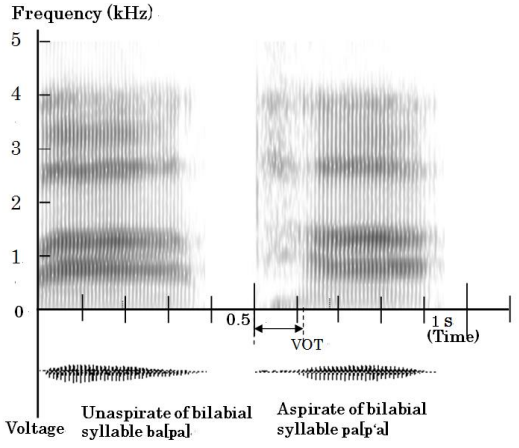


Fig. 2. Spectrograms of unaspirated syllable bi[pi] (left), and aspirated syllable pi[p‘i] (right) pronounced by a Chinese speaker.

Fig. 2 shows those for the bilabial unaspirated sound bi[pi] (left) and the bilabial aspirated sound pi[p'i] (right). VOT was as long as 100 ms. Stripes from 2200 to 5000 Hz are darker, implying the strongest breathing power there. For the frequencies lower than 2000 Hz in the VOT, almost no vertical stripes indicate weak breathing power. The left spectrogram shows the unaspirated sound bi[pi]; the VOT is zero.

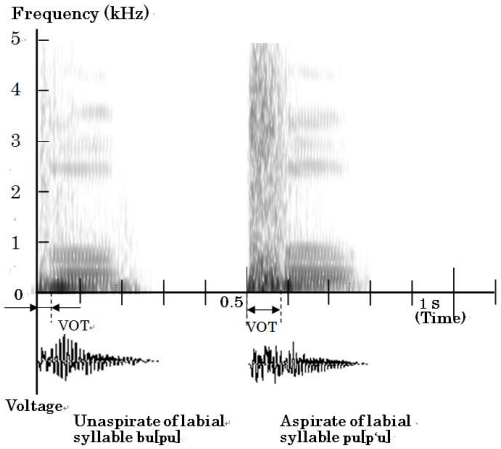


Fig. 3. Spectrograms of unaspirated syllable bo[po] (left), and aspirated syllable po[p'o] (right) pronounced by a Chinese speaker.

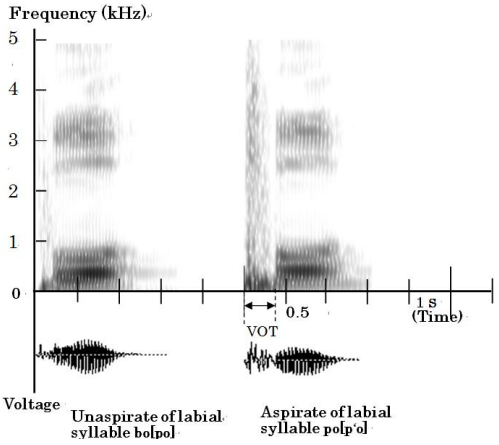


Fig. 4. Spectrograms of unaspirated syllable bu[pu] (left), and aspirated syllable pu[p'u] (right) pronounced by a Chinese speaker.

Fig. 3 shows those for the unaspirated sound bo[po] (left) and aspirated sound po[p'o] (right). The VOT of po[p'o] was as long as 80 ms. Throughout almost the entire VOT, darker gray vertical stripes were observed in the frequencies between 50 and 5000 Hz. The onset of the vocal fold vibration is so close to the burst in the left of the spectrogram that no aspiration interval of bo[po] appears, thus the VOT was 20ms.

Fig. 4 shows those for the bilabial unaspirated sound bu[pu] (left) and aspirated sound pu[p'u] (right). The VOT of pu[p'u] was as long as 95 ms. Throughout almost the entire VOT, dark vertical stripes were observed in the frequencies between 50 and 5000 Hz. It implies strongest breathing power there. The left spectrogram shows the unaspirated sound bu[pu]. Although it is unaspirated, we observe the VOT, which is the shortest 22 ms, where the vertical stripes are light grey, indicating weak breathing power.

## 3. Automatic Measurement of VOT and Breathing Power

We showed that the correct utterance of a bilabial aspirated sound is closely related to the frequency spectrum during the VOT period. We Previously developed an automatic measurement system for the VOT and the breathing power using a personal computer containing a 35-channel frequency filter bank, in which the center frequency ranges from 50 to 6850 Hz with a bandwidth of 200 Hz Powers at each of the 35 channels were checked at intervals of 5 ms with 11.025 kHz sampling frequency. We can extract the spectral of bilabial aspirated and unaspirated pronunciations in both the VOT and voiced periods.

We automatically detected the onset of a plosive release or burst. Pronounced signals were introduced into the filter bank and split to power at each center frequency every 5 ms. The start time of the VOT, $t_1$, was determined by comparing the powers for adjacent time frames when the number of temporally increasing channels is at maximum. The end of the VOT, $t_2$, was the starting point of the formant. Thus, $t_2$-$t_1$ is defined as the VOT. In Section 2, we described the features of the correct pronunciation of Chinese aspirated and unaspirated sounds by observing temporal variations of in the breathing power spectra during the VOT.

Average power during VOT is defined as follows. Power is deduced every 5ms and are referred to as $P_{i,j}$, meaning power at $j$×5ms of the channel $i$(1-35) where Pi is the integration of the power at each interval for

VOT of the channel *i*, as shown in Equation (1).

$$P_i = \sum_{j=1}^{J} P_{i,j}\,(t_j) \tag{1}$$

Thus the energy, $W_i$, of the channel i is defined as

$$W_{i,VOT} = P_i \times 5\text{ms} \tag{2}$$

The average power, $P_{i,av}$, of each frequency channel during VOT is defined as

$$P_{i,av}=W_{i,VOT}/VOT \tag{3}$$

Finally, average power $Pv_{i,av}$ at channel i in the voiced period, $T_{vs}$, can be defined similarly, as

$$Pv_{i,av}=W_{i,vs}/T_{vs} \tag{4}$$

## 4. Relationship between Breathing Power and Its Frequency Dependency during VOT, and the Pronunciations of Quality

We define the discrimination criteria of bilabial aspirated sounds by examining the VOT and the breathing power spectrum during the VOT for the pronunciations uttered by 50 Japanese students and 50 native- Chinese speakers. We used our automatic measuring system to define the parameters.

### 4.1. Scoring of Pronunciation Quality of Students

To investigate the correct pronunciation criteria of the bilabial aspirated sounds pa[p‘a], pi[p‘i], po[p‘o], and pu[p‘u], the sounds uttered by 20 Japanese students were ranked using a listening test made up of sounds reproduced, by nine native Chinese speakers. The scores were as follows: 3 = a correctly pronounced aspirated sound; 2 = an unclear sound; and 1 = a pronunciation in which an aspirated sound was judged to be unaspirated or vice versa. We defined an average score of more than 2.6 as good.

### 4.2. Relationship between Student Pronunciation Score and Evaluation Parameters

Fig. 5 shows the data distribution for aspirated sounds pa[p‘a] with the VOT on the abscissa and $P_{av}$ on the ordinate. The power of each utterance in this figure was automatically calculated at frequencies between 1050 Hz and 4450 Hz averaged during the VOT. The pronunciations of students with a good score and the Chinese speakers are gathered in the upper-right area of the figure.

The utterances of students with an average power of greater than 12, and a VOT of longer than 50 ms, received the highest score of 3.0. Data D1 and D2 at the lower-left had a short VOT of 6 and 10 ms, respectively, and a received score of 1.0. Datum D3 had a VOT of 36 ms and the highest score of 3.0, whereas data D5 and D4, which were far below D3, had almost the same VOT but received scores of 2.0 and 1.2, respectively. A pronunciation with a higher $P_{av}$ and VOT duration of 30 to 60 ms received a higher score in the figure. The cases of aspirated sounds pi[p‘i], po[p‘o] and pu[p‘u] had the almost the same tendencies with the sound of pa[p‘a] as shown in Fig. 5. When the duration of the VOT was within a certain range, pronunciations using a greater amount of breathing power received higher scores. These criteria are applied in the discrimination procedure as shown in the next section.

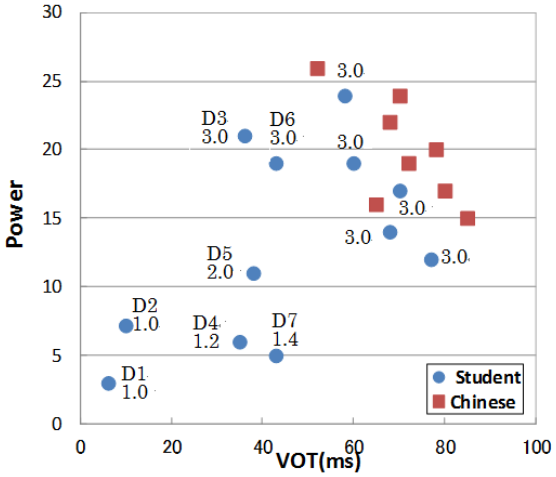


Fig. 5. Data distribution for aspirated syllable pa[p'a] with VOT on the abscissa and $P_{av}$ at on the ordinate.

## 5. Automatic Discrimination of Aspirated and Unaspirated Syllables

### 5.1. Scoring of Pronunciation Quality of Students

We extracted the features of correctly pronounced Chinese bilabial aspirated sounds and unaspirated sounds, by observing the spectrum evolution of the breathing power during the VOT and the voiced period of the sounds when uttered by 50 native-Chinese speakers.

Tables 1 and 2 show cases in which the VOT was longer than 60 ms, and between 20 and 60 ms, respectively, using the evaluation criteria for bilabial aspirated syllables. Table 3 shows the values of F1, F2, F3, and F4 during the voiced period after the VOT, which discriminate four different aspirated syllables. For a VOT of longer than 60 ms, as shown in Table 1, the power is greater than 4 at between 1050 and 1850 Hz, greater than 3 at between 2450 and 3050 Hz, and greater than 3 at between 3650 and 4450 Hz, as averaged during the VOT; in addition, when a high power appeared between 450 and 1050 Hz, 1050 and 1650 Hz, 2450 and 3050 Hz, and 3250 and 4450 Hz, as shown in Table 3, the utterances were judged to be the bilabial aspirated sound, pa[p'a]. When the VOT was between 20 and 60 ms, as shown in Table 2, the power was greater than 7 at between 1050 and 1850 Hz, greater than 6 at between 2450 and 3050 Hz, and greater than 6 at between 3650 and 4450 Hz, and when a high power appeared at between 450 and 1050 Hz, 1050 and 1650, 2450 and 3050 Hz, and 3250 and 4450 Hz, as shown in Table 3, the utterances were judged to be the bilabial aspirated sound pa[p'a].

When the VOT was longer than 60 ms, and the power was greater than 28 at between 1850 and 5650 Hz, as shown in Table 1; when the VOT was between 20 and 60 ms, and the power was greater than 35 at between 1850 and 5650 Hz, as shown in Table 2; and when a high power appeared at between 50 and 450 Hz, 1850 and 2450 Hz, 2650 and, 3650 Hz, and 3650 and 4250 Hz, as shown in Table 3, the utterances were judged to be the aspirated sound pi[p'i]. Similarly, the utterances are judged to be the aspirated sound po[p'o] and pu[p'u].

When a distinctive feature does not appear during the VOT period, we refer to Table 4, which lists the evaluation criteria for bilabial unaspirated sounds, namely, the values of F1, F2, F3, and F4 during the voiced period. When a high power appears at between 650 and 1050 Hz, 1050 and 1650 Hz, 2450 and 3050 Hz, and 3650 and 4450 Hz, the utterances are judged to be the unaspirated sound ba[pa]. Similarly, the utterances are judged to be the aspirated sound bi[pi],po[po] and pu[pu].

Table 1. Evaluation Criteria on Utterance of Bilabial Aspirated Syllables During VOT (VOT > 60ms)

| Syllable | Pattern | Channels(CH) | Frequency domain(Hz) | VOT range | Ave.Power in VOT |
|---|---|---|---|---|---|
| pa[p‘a] | ① | CH06～CH09 | 1050～1850 | 60ms or more | 4 or more |
| | ② | CH13～CH15 | 2450～3050 | 60ms or more | 3 or more |
| | ③ | CH19～CH21 | 3650～4450 | 60ms or more | 3 or more |
| pi[p‘i] | | CH10～CH26 | 1850～5650 | 60ms or more | 28 or more |
| po[p‘o] | | CH01～CH26 | 50～5650 | 60ms or more | 34 or more |
| pu[p‘u] | | CH01～CH26 | 50～5650 | 60ms or more | 42 or more |

Table 2. Evaluation Criteria on Utterance of Bilabial Aspirated Syllables During VOT (2oms<VOT <60ms)

| Syllable | Pattern | Channels(CH) | Frequency domain(Hz) | VOT range(ms) | Ave.Power in VOT |
|---|---|---|---|---|---|
| pa[p‘a] | ① | CH06～CH09 | 1050～1850 | 20<VOT<60 | 7 or more |
| | ② | CH13～CH15 | 2450～3050 | 20<VOT<60 | 6 or more |
| | ③ | CH19～CH21 | 3650～4450 | 20<VOT<60 | 6 or more |
| pi[p‘i] | | CH10～CH26 | 1850～5650 | 20<VOT<60 | 35 or more |
| po[p‘o] | | CH01～CH26 | 50～5650 | 20<VOT<60 | 40 or more |
| pu[p‘u] | | CH01～CH26 | 50～5650 | 20<VOT<70 | 48 or more |

Table 3. Evaluation Criteria for Formants of Bilabial Aspirated Syllables Voiced Period After VOT

| Syllable | F1(Hz)/(CH) | F2(Hz)/(CH) | F3(Hz)/(CH) | F4(Hz)/(CH) |
|---|---|---|---|---|
| pa[p‘a] | 450～1050/(CH4-CH5) | 1050～1650/(CH6-CH8) | 2450～3050/(CH13-CH15) | 3250～4450/(CH19-CH21) |
| pi[p‘i] | 50～450/(CH1-CH2) | 1850～2450/(CH10-CH12) | 2650～3650/(CH14-CH18) | 3650～4250/(CH19-CH20) |
| po[p‘o] | 50～250/(CH1) | 250～650/(CH2-CH3) | 650～1050/(CH4-CH5) | 2650～3050/(CH14-CH15) |
| pu[p‘u] | 50～450/(CH1-CH2) | 450～850/(CH3-CH4) | 2250～2650/(CH12-CH13) | 3050～3450/(CH16-CH17) |

Table 4. Evaluation Criteria for Formants of Bilabial Unaspirated Syllables Voiced Period After VOT

| Syllable | F1(Hz)/(CH) | F2(Hz)/(CH) | F3(Hz)/(CH) | F4(Hz)/(CH) |
|---|---|---|---|---|
| ba[ba] | 650～1050/(CH3-CH5) | 1050～1650/(CH6-CH8) | 2450～3050/(CH13-CH15) | 3650～4450/(CH19-CH21) |
| bi[bi] | 50～450/(CH1-CH2) | 1850～2450/(CH10-CH12) | 2650～3650/(CH14-CH18) | 3650～4250/(CH19-CH20) |
| bo[bo] | 50～250/(CH1) | 250～650/(CH2-CH3) | 650～1050/(CH4-CH5) | 2650～3050/(CH14-CH15) |
| bu[bu] | 50～450/(CH1-CH2) | 450～850/(CH3-CH4) | 2250～2650/(CH12-CH13) | 3050～3450/(CH16-CH17) |

## 5.2. Automatic Evaluation and Results

Fig. 6 illustrates the flow of our system for the automatic discrimination of the aspirated and unaspirated pairs of single-vowel bilabial sounds. In step 1, the uttered sounds are input into the computer. In step 2, the sounds are automatically analyzed using our 35-channel filter bank to create a database of temporal variations in the power spectra. In step 3, the VOT is deduced using the algorithm described in section 3. In step 4, when the VOT of an utterance is longer than 20 ms, step 5 is applied. Otherwise, the utterance is judged as unaspirated. In step 5, when the VOT of the utterance is longer than 60 ms, step 6a is applied. Otherwise, step 6b is applied. In steps 6a and 6b, the average power, $P_{i,av}$, is automatically calculated for each channel during the VOT, as described in section 3. In steps 7a and 7b, if any distinctive features are found during the VOT, the sample is tentatively judged to be aspirated, and differentiated from the others when referring to Tables1 and 3, and Tables 2 and 3, respectively. If an utterance fulfills both of the above conditions, it is successfully identified as one of the four aspirated bilabial sounds. Otherwise, it is judged as not bilabial aspirated. If a sample, judged to be unaspirated in step 7a, fulfills the conditions listed in Table 4, it is successfully identified as one of the four unaspirated bilabial sounds. Otherwise, it is not.

The 50 Japanese students were 15-16 years old who had studied Chinese for three hours per week for a period of six months. All of the 50 selected native Chinese speakers were raised in Beijing, China. They speak the exact Standard Chinese (Putonghua) dialect.

Some of the data that passed the screening test of the fifty native-Chinese speakers were rejected owing to the strict discrimination criteria. Table 5 lists the number of utterances of the 50 students that were judged as correct by the native-Chinese speakers for pa[p‘a], pi[p‘i], po[p‘o], and pu[p‘u], which were 44, 46, 45,and 45, respectively, and those by our judgment system, which were 42, 45, 43, and 43. The ratio of correct judgment between our system and the native-Chinese speakers for an aspirated sound were 95%, 98%, 95%, and 95%, respectively.

The utterances by the 50 Japanese students for the unaspirated bilabial sounds ba[pa], bi[pi], bo[po], and bu[pu] that were judged as correct by the native-Chinese speakers were 45, 47, 45, and 46 in number, respectively, whereas those determined as correct by our judgment system were 44, 46, 43, and 44 in number, respectively. The ratios of correct judgments between our system and the native-Chinese speakers for unaspirated sounds were 98%, 98%, 96%, and 95%, respectively. Again, some of the data that passed the screening test by the native-Chinese speakers were rejected by our system.

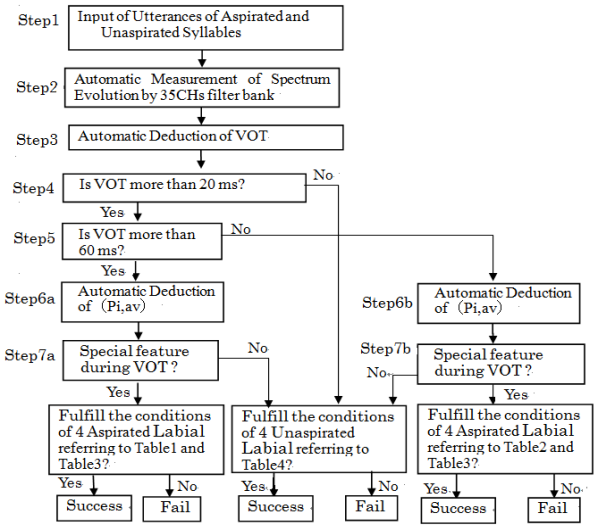


Fig. 6. Discrimination diagram of aspirated and usaspirated syllables.

Table 5. Utterance Number Which Passed Screening Test of 50 Native Chinese Speakers and Our Judgment System among Those by 50 Japanese Students

| | Unspirated labial syllables | | | | Aspirated labial syllables | | | |
|---|---|---|---|---|---|---|---|---|
| | ba[ba] | bi[bi] | bo[bo] | bu[bu] | pa[p'a] | pi[p'i] | po[p'o] | pu[p'u] |
| Number of correctly judged ones by native Chinese speakers (1) | 45 | 47 | 45 | 46 | 44 | 46 | 45 | 45 |
| Number of correctly judged ones by our judgment system (2) | 44 | 46 | 43 | 44 | 42 | 45 | 43 | 43 |
| Correct judgment ratio (2)/(1) | 98% | 98% | 96% | 95% | 95% | 98% | 95% | 95% |

## 6. Conclusion

We have been studying the pronunciation instruction system of Chinese aspirated sounds, which are generally difficult for Japanese students to perceive and reproduce. We closely examined the spectrograms of these sounds when uttered by native-Chinese speakers and by Japanese students, and determined the criteria for the correct pronunciations of various aspirated sounds. We previously developed an automatic system for measuring and calculating the VOT and the level of power during the VOT for the students' pronunciation.

In this paper, to develop a self-learning system for Chinese aspirated and unaspirated sounds, we aimed at the automatic distinction of the four pairs of aspirated and unaspirated bilabial sounds.

To improve the discrimination performance of these single-vowel bilabial sounds, we automatically calculated their frequency spectra during the VOT and voiced periods, and extracted the distinctive features of each sound. Based on the data obtained, we established the criteria for automatically discriminating aspirated and unaspirated sounds.

We then conducted an experiment on the automatic discrimination capability of our judgment system for the utterances of 50 Japanese students. The results of the test showed that the system rejected some of the pronunciations that were judged correct by the native-Chinese speakers. However, the rates of proper determination of our system for four pairs of aspirated and unaspirated sounds were 95% to 98%. We plan to apply our system to other samples of pronounced sounds to improve the level of accuracy.

## Acknowledgment

The authors appreciate the financial support provided by the Japan Society for the Promotion of Sciences.

## References

[1] Hoshino, A., & Yasuda, A. (2002). Evaluation of Chinese aspiration uttered sounds by Japanese students using VOT and power. *Acoust. Soc. Jpn., 58(1)1,* 689‐695.

[2] Hoshino, A., & Yasuda, A. (2011). Pronunciation training system for Japanese students learning Chinese aspiration. *Proceedings of the 2nd Intl. Conf. on Society and Information Technologies (ICSIT)* (pp. 288-293). Orlando, Florida, USA.

[3] Zhou, L., Segi, H., & Kido, K. (1998). The investigation of Chinese retroflex sounds with time-frequency analysis. *Acoust. Soc. Jpn., 54(8),* 561‐56.

[4] Hoshino, A., & Yasuda, A. (2014). *Automatic Discrimination of Pronunciations of Chinese Retroflex and Dental Affricates.* Springer LNAL 8202, 303-314.

[5] Kent, R. D., & Read, C. (1992). *The Acoustic Analysis of Speech*. San Diego and London: Singular Publishing Group, Inc.

[6] Zhu, C. (1997). *Studying Method of the Pronunciation of Chinese Speech for Foreign Students,* Yu Wu Publishing Co.

**Akemi Hoshino** was born in Shanghai, China on November 1, 1960. She received her doctorate in engineering degree from Tokyo University of Marine Science and Technology, Japan in 2005. She has been working with National Institute of Technology Toyama College since 1997. She is a professor in Liberal Studies.

One of her publications is “Evaluation of Chinese aspiration sounds uttered by Japanese students using VOT and power (in Japanese), *Acoust. Soc. Jpn.*, 58, No. 11, pp.689-695, (2002).” Her research interests include development of pronunciation training system of Chinese by CAI, and development of computer-aided automatic judgment system, and evaluation of Chinese aspiration sound uttered by Japanese Students based on VOT and power, and Automatic Discrimination of Pronunciations of Chinese Retroflex and Dental Affricates. Dr. Hoshino is a member of Acoustical Society of Japan and Institute of Electronics, Information, Communication Engineers in Japan.