Perception of English syllable-final consonants by Chinese speakers and Japanese speakers

Yaming Zhang

Kobe University
longzeyaming@gmail.com

Ryoko Hayashi

Kobe University
rhayashi@kobe-u.ac.jp

ABSTRACT

This study investigates perception by Mandarin Chinese and Japanese native speakers of English consonants in syllable-final position, and the effect of vowel duration as a cue to voicing in syllable-final stops.

Two experiments were conducted in the study and the results revealed that Mandarin Chinese speakers performed better than Japanese speakers but the position of the consonant in the syllable affects the perception rate. Japanese speakers can perceive the voicing in syllable-final consonants better than Mandarin Chinese speakers; also they may make better use of vowel duration as a cue to perceive voicing in syllable-final stops.

Keywords: syllable-final consonants, vowel duration, perception, voicing.

1. INTRODUCTION

It has been long recognized that the syllable is the basic unit in the perception of spoken language. The difference of syllable structure between L1 and L2 might have potential influence on acquiring L2 sounds.

In Mandarin Chinese and Japanese, consonants seldom occur in syllable-final position. It is presumed, therefore, that Mandarin Chinese speakers and Japanese speakers will have difficulties in perceiving English syllable-final consonants.

Hillenbrand [3] and Lawrence [4] discussed extensively the importance of the duration of the preceding vowel in the perception of voicing in syllable-final consonants. Since long-short vowel contrasts do not exist in Mandarin Chinese, it is presumed that Chinese speakers are less sensitive to the change of vowel duration than are Japanese speakers, and cannot use vowel duration as a cue to perceive voicing in syllable-final consonants. In order to test the hypothesis, two experiments were conducted.

2. EXPERIMENT 1

This experiment will examine the perception of English syllable-final consonants by Mandarin Chinese speakers and Japanese speakers by using nonsense words.

2.1. Subjects

20 subjects participated in the experiment: ten Mandarin Chinese native speakers and ten Japanese native speakers. Each subject had normal hearing.

2.2. Materials

Along the lines of the study by Enomoto [1], stimuli used in this experiment were nonsense English words containing the following 13 American English consonants: /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /s/, /z/, /m/, /n/ and /l/. All words have the structure of $C_1V_1C_2C_3V_2$, with /e/ and /i/ in V_1 and V_2 position and /p/ in C_1 position. The 13 target sounds appear in C_2 and C_3 position alternately, from which 169 words were made. Among them, when the same consonant appears in C_2 and C_3 at the same time, this word is removed from the list. In total, 156 nonsense words were used.

The stimuli were recorded by an American English native speaker in a soundproof room. The speaker was asked to pronounce all the nonsense words naturally so that the consonants in C_2 position were reduced. All the words were recorded twice and the one with better quality was chosen as the stimulus. The recording was edited by Cool Edit (Version 2.1.3097.0) to randomize and to insert pauses between each word.

2.3. Procedures

The subjects were told to listen to 156 nonsense English words with a structure of /peC_2C_3i/ and asked to concentrate on the two consonants in C_2 and C_3 positions, and then to write down the sounds they had identified on the answer sheet.

2.4. Results

The average identification rate of syllable-initial consonants is 82.6%, while syllable-final consonants show a lower rate of 67.8%. This result reflects that the position of a consonant within a syllable does influence perception. Recasens [6] indicated that consonant reduction happens relatively more easily

for final syllables than initial syllables. For consonants in syllable-final position, the loss of acoustic cues could cause difficulties for the listeners' identifications.

Chinese subjects have a higher average score (63.5%) than Japanese subjects (52.3%) in the perception of all consonants in both syllable-initial position and syllable-final position.

Table 1 shows the perception rate of each consonant by both Chinese speakers and Japanese speakers. The results of each case were tested by a Chi-square test. Significant differences are marked with one asterisk ($p<0.05$), or two asterisks ($p<0.001$).

In comparison with Japanese speakers, the perception of Chinese speakers is more affected by syllable position. These results are a similar to the results reported by Flege (1989), which showed that Chinese speakers can perceive syllable-initial /t/-/d/ almost perfectly but perceive edited syllable-final /t/-/d/ poorly. Perhaps Chinese speakers use the same strategy to perceive differences between /t/-/d/ in both syllable-initial and syllable-final position, and are not sensitive to possibly different cues, such as preceding vowel length, in their identification of final stops.

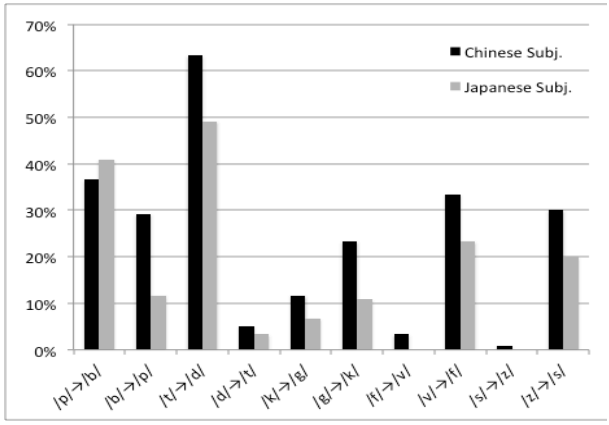
A significant difference in the perception of /v/ and /l/ sounds between Chinese and Japanese speakers, both in syllable-final and syllable-initial positions, is seen. Japanese speakers perform less well, possibly due to phonological restrictions: /v/ is considered an allophonic variant of /b/, and /l/ a variant of /r/.

Table 1: Perception rate of each consonant in syllable-initial and syllable-final position (%).

Syllable-final	Chinese subj.	Japanese subj.	Syllable-initial	Chinese subj.	Japanese subj.
/t/*	15.0	28.3	/t/	95.8	85.8
/d/*	62.5	75.0	/d/**	95.0	70.8
/p/*	50.0	32.5	/p/*	99.2	95.0
/b/	59.2	53.3	/b/**	86.7	65.8
/k/	87.5	90.0	/k/	83.3	83.3
/g/*	74.2	88.3	/g/*	75.8	89.2
/f/	95.8	90.8	/f/*	59.2	73.3
/v/**	55.0	24.2	/v/**	94.2	37.5
/s/*	91.7	97.5	/s/	93.3	94.2
/z/*	58.3	72.5	/z/*	86.7	76.7
/m/**	76.7	26.7	/m/	90.8	92.5
/n/	90.8	93.3	/n/*	90.8	79.2
/l/**	98.3	74.2	/l/**	86.7	65.8

For stops and fricatives most errors happened between the voiced and voiceless contrast, especially when the consonants appeared in syllable-final position. This suggests that for both Chinese speakers and Japanese speakers, voicing characteristics are the most difficult to identify. Also, note from Fig. 1, that in the perception of voicing in syllable-final consonants, Chinese speakers made more mistakes than Japanese speakers.

Figure 1: Percentage of errors occurring within voiced-voiceless contrast in syllable-final position.



3. EXPERIMENT 2

Experiment 2 focuses on how Chinese and Japanese speakers make use of the preceding vowel duration cue to perceive voicing in syllable-final consonants. Hillenbrand [3] and Lawrence [4] discussed the importance of the preceding vowel duration in the perception of voicing characteristics of consonants, and especially for stops compared with fricatives. In experiment 1, most errors occurred between the word pairs with voiced and voiceless sounds with Chinese speakers making more mistakes than Japanese speakers. When a stop consonant is reduced, the cues for listeners to use to perceive the voicing characteristics change, preceding vowel duration supposedly plays an important role. Experiment 2 is conducted to verify the hypothesis that Chinese speakers may be less sensitive than Japanese speakers to vowel duration cues.

3.1. Subjects

The subjects were the same ten Chinese and ten Japanese native speakers who participated in Experiment 1.

3.2. Materials

Six English words with a simple CVC structure (cop, cob, bet, bed, pick and pig) were recorded by an American speaker. For each word, the vowel durations were measured. Stimuli were made by editing the words ending in voiceless stops: the release burst was eliminated and the duration of the

preceding vowel was lengthened in five steps. In the final step, the duration is equal to that of the preceding vowel of the words ending with voiced consonants. Table 2 shows the vowel duration of each stimulus. 18 stimuli were made and randomized so that each word appeared two times.

Table 2: Vowel durations of stimuli used in experiment 2 (ms).

	1	2	3	4	5	6
/bet/-/bed/	193	215	237	259	281	303
/sɑp/-/sɑb/	193	216	237	258	279	300
/pɪk/-/pɪg/	173	185	197	209	221	233

3.3. Procedures

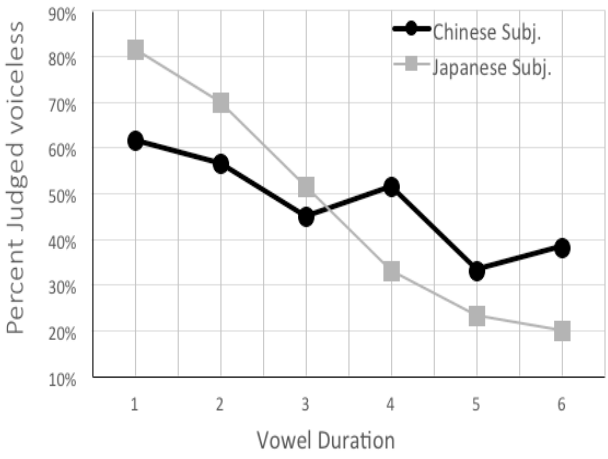
Subjects were told to listen to some English words, for which the last consonant was reduced. They were asked to identify if the word ended with a voiced or a voiceless consonant.

3.4. Result

The results showed a considerable difference between Chinese and Japanese speakers (Fig. 2).

In Figure 2,the percentage judged as voiceless by Japanese speakers gets lower if the preceding vowel duration gets longer. There was a statistically significant difference between groups of different vowel durations (ANOVA test, $p<0.001$). This means for Japanese speakers, the consonant following the vowel was likely to be perceived as voiceless when the preceding vowel duration is shorter. This is the same result as reported for English native speakers by Lawrence [4]. Thus Japanese speakers show a tendency to use vowel duration as a voicing cue in way similar to English speakers.

Figure 2: Percentage judged as voiceless was a function of vowel duration.



In the case of Chinese speakers, not significant differences in identification of voicing of consonants were found as a function of vowel duration (ANOVA test, $p=0.26$). This indicates that Chinese speakers do not make use of vowel duration for perceiving voiced-voiceless final stops.

4. SUMMARY AND FURTHER STUDY

In this study, two experiments were conducted. Experiment 1 showed that Chinese speakers performed less well than Japanese speakers in the perception of voiced-voiceless final consonant contrasts. Experiment 2 showed that Chinese speakers are not sensitive to differences in vowel duration as a cue to voicing of final stops. This finding may account for their relatively poor performance in Experiment 1 for identifying voicing features of syllable-final stops. This is considered to be a negative transfer from their first language. As was discussed above, vowel duration is considered to be the strongest cue for the voicing of reduced final stops, but many other acoustic cues also have potential influence on the perception of voicing features. Further study is needed to examine other acoustic parameters such as VOT, release burst, and formant transitions.

5. REFERENCES

[1] Enomoto, A. 2005. Perception of coda consonants by Japanese learners of English. *Journal of the Phonetic Society of Japan*. Vol.9 No.1: 73-84.

[2] Flege, J.E. 1989. The perception of word-final /t/ and /d/ by native and Chinese listeners, *J. Acoust. Soc. Am*. 86(5): 1684-1697.

[3] Hillenbrand, J., Ingrisano, D.R., Smith, B.L. and Flege, J.E. 1984. Perception of the voiced-voiceless contrast in syllable-final stops. *J. Acoust. Soc. Am*. 76(1): 18-26.

[4] Lawrence J.R. 1971. Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *J. Acoust. Soc. Am*. 51: 1296-1303.

[5] Rodgers, B., Purnell, T., Salmons, J. 2008. Expanding the set of acoustic features of the post-vocalic voicing contrast in English. *J. Acoust. Soc. Am*. 123(5): 38-88.

[6] Recasens, Daniel. 2004. The effect of syllable position on consonant reduction (evidence from Catalan consonant clusters). *Journal of Phonetics*. 32: 435-453.

[7] Treiman, R. 1984. On the status of final consonant clusters in English syllables. Journal of verbal learning and verbal behavior. 23: 343-356.