Perception and production of consonant clusters in Japanese-English bilingual and Japanese monolingual speakers

Hinako Masuda, Takayuki Arai

Department of Science and Technology, Sophia University, Tokyo, Japan

`h-masuda@sophia.ac.jp, arai@sophia.ac.jp`

Abstract

Previous research has revealed that Japanese native speakers are more likely to perceive an 'illusory vowel' within consonant clusters compared to French native speakers (Dupoux et al. 1998). The aim of this research is to investigate the differences of perception and production of consonant clusters in Japanese-English bilinguals and Japanese monolinguals. The stimuli of the two experiments consist of 36 pseudo-words that contain the two sequences, VCCV and VCVCV. Results of the perception and the production experiments on the two groups of participants revealed that bilinguals were more likely to achieve high scores in both experiments.

Index Terms: consonant clusters, bilingual, speech perception, speech production

1. Introduction

Various researches on the perception and production of foreign sounds have revealed that the perception and production abilities of foreign sounds are influenced by the speaker's native language [1, 4-6]. As Dupoux et al. claimed, it is difficult for Japanese native speakers to distinguish the difference between VCCV and VCVCV when compared to French native speakers, and that Japanese are likely to perceive an "illusory vowel" between the two consonants. This is due to the difference in the syllable structure of Japanese and French. The syllabic structure of French allows consonant clusters, but Japanese does not. Thus, the Japanese are more likely to perceive and produce consonant clusters as CVCV, which is the main structure in Japanese. For example, "strike" in English is pronounced as [sutoraiku] in Japanese [3]. While Dupoux et al. researched on the perception of consonant clusters in Japanese and French native speakers, the present study researched the perception and production of consonant clusters in two groups of Japanese native speakers with different language learning backgrounds: highly fluent Japanese-English bilinguals and Japanese monolinguals. By dividing Japanese native speakers into two groups with different language learning backgrounds, we will be able to investigate the degree of native language influence on the perception and production of foreign sounds. The aim of this study is to research how different English learning backgrounds affect both perception and production of Japanese native speaker groups.

2. Experiment 1: Perception

Peperkamp et al. has revealed in their research on the perception of stress that highly fluent French-Spanish bilinguals' ability to perceive stress was either Spanish-like or French-like. Although there is stress contrast in Spanish and not in French, results differed according to the country in which the participants were born.

In Experiment 1, we aim to investigate the identification ability of consonant clusters in two groups of Japanese native speakers: Japanese-English bilinguals and Japanese monolinguals. We also investigate whether place of birth has an effect on the results. An AXB forced-choice identification test is assigned to the participants with 192 trials after 8 practice trials. The stimuli used for this experiment are natural utterances of 36 pseudo words that consist of the two sequences VCCV and VCVCV. If the experience of exposure to English has a positive influence on the identification of a phonotactic structure that is not in Japanese, Japanese-English bilinguals should have a higher identification ability compared to Japanese monolinguals.

2.1. Participants

2.1.1. Japanese-English bilingual speakers

17 participants, 12 women and 5 men, participated as Japanese-English bilinguals. Participants' ages ranged from 19 to 25 years of age (median 23.4 years). All participants were born in Japan, and have experience of living in an English-speaking country for 2 to 8.5 years. All participants' first language is Japanese, and their second language is English. All participants are highly fluent in English. None of the participants reported any hearing problems.

2.1.2. Japanese monolingual speakers

22 participants, 8 women and 14 men, participated as Japanese monolinguals. Participants' ages ranged from 18 to 25 years of age (median 20.8 years). The participants do not have experience of living abroad for more than one month. All participants' first language is Japanese. None of the participants reported any hearing problems.

2.2. Materials

The materials used for Experiment 1 are natural utterances of 36 pseudo-words that consist of the two sequences VCCV and VCVCV, spoken by a Japanese-French bilingual speaker. None of the words have meanings in Japanese, English, nor French. All pseudo-words are the same sets of words used in the previous research by Dupoux et al. The first consonants are chosen from [b, g, k, sh], the initial and the finial vowels are chosen from [a, i, u, o], and the second consonants are chosen from [d, g, m, n, sh, t, z]. The list of the pseudo-words used in Experiment 1 is shown in Table 1. All A, X, and B used in the AXB identification experiment are three different tokens of the recorded stimuli.

September 22−26, Brisbane Australia

Table 1: List of 36 pseudo-words.

VCCV	VCVCV
Abmo	Abumo
Ebzo	Ebuzo
Abge	Abuge
Agmi	Agumi
Akmo	Akumo
Ashmi	Ashumi
Ebza	Ebuza
Egdo	Egudo
Ekshi	Ekushi
Eshmo	Eshumo
Ibdo	Ibudo
Igna	Iguna
Ikma	Ikuma
Ishto	Ishuto
Obni	Obuni
Ogza	Oguza
Okna	Okuna
Oshta	Oshuta

2.2.1. Speaker of model sounds

The speaker of the model sounds used in Experiment 1 is a female Japanese-French bilingual speaker. The speaker is a native speaker of Japanese, and is highly fluent in her second language, French. She has experienced living in France for nine years, and is also an experienced phonetician.

2.2.2. Recording of model sounds

The recording of the model sounds took place in a sound-proof room. The sounds were recorded by using Marantz PROFESSIONAL SOLID STATE RECORDER PMD 671 and SONY ECM-959DT microphone. The sounds recorded were then saved onto a personal computer as wave files. The speaker was asked to produce the 36 pseudo-words embedded in the carrier phrase "Dites (*pseudo-word*) deux fois." [Say (pseudo-word) two times.] The set of pseudo words were recorded three times each.

2.3. Procedure

Experiment 1 was performed in a sound-proof room. Participants were instructed to listen to the three sounds (A, X, and B) through the headphones and to judge whether A or B was more similar to X. A personal computer and SONY MDR-CD900ST headphones were used for this experiment. Participants were first given 8 practice trials, which they received no feedback on, and then proceeded to the 192 trials. The practice trials are not valid for this experiment. The 192 trials were divided into 3 sections, and participants could take short breaks between the sections if needed. This experiment took roughly 20 minutes for each participant, including the time to fill out the consent form and the questionnaire.

2.4. Results

The Mann-Whitney test found a significant difference between the two groups' average percentage of correct answers ($p < 0.05$). The average percentage of correct answers in the bilingual speaker group was 98.01%, and 93.18% for monolingual group. The average number of errors made by Japanese-English bilinguals was 3.82 out of 192 trials, whereas those made by Japanese monolinguals was 13.09. The range of the number of errors made by bilinguals was 0 to 12, whereas the number was 2 to 35 by monolinguals.

3. Experiment 2: Production

The aim of Experiment 2 is to investigate the production ability of consonant clusters in Japanese-English bilinguals and Japanese monolinguals. The participants were instructed to read the pseudo-words with the VCVCV sequence written on paper. If the exposure to English has a positive effect on the production of consonant clusters, the Japanese-English bilinguals are speculated to produce consonant clusters without a vowel inserted between the consonants.

3.1. Participants

The same 17 Japanese-English bilinguals and 22 Japanese monolinguals participated in Experiment 2.

3.2. Procedure

After Experiment 1, participants were asked to produce the 18 pseudo-words with the sequence of VCCV (see Table 1). A carrier phrase was not used in order to avoid confusion. The recording environment of the participants' productions is the same as the recording environment of the model sounds in Experiment 1. The sounds recorded were saved onto a personal computer as wave files. Experiment 2 took roughly five minutes for each participant.

3.3. Analysis

3.3.1. Object of analysis

Out of the 18 pseudo-words uttered by the participants, four were randomly selected for analysis. The selection of the pseudo-words was made by categorizing the pseudo-words by the manner of articulation of the two consonants. Since the vowel inserted between two voiceless consonants is likely to be devoiced [2], "Ekshi", "Ishto" and "Oshta" were removed from selection. The selected four pseudo-words are "Abmo", "Ashmi", "Ebza" and "Ibdo".

3.3.2. Acoustical analysis

The criterion for regarding the presence of the vowel [u] inserted between the two consonants is to see whether there is a steady state of the first and second formants between the two consonants that have the characteristics of the vowel [u]. The transitions of the formants are not measured. Therefore, if there is no steady state of the first and the second formants between the two consonants, the consonant cluster is regarded as being produced as a consonant cluster without a vowel. If there is a steady state, even if its duration is short, it is considered as vowel epenthesis. The criterion for measuring the presence of the vowel [u] is shown in Figure 1.

Table 2: Average duration of all words, Abmo, Ashmi, Ebza and Ibdo, the duration of inserted [u], and the percentages of inserted [u] of bilingual and monolingual participants.

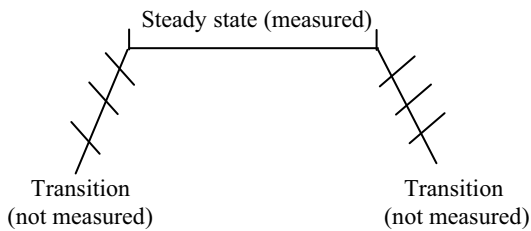|	Duration of word		Duration of inserted [u]		% of inserted [u]	
	Bilingual	Monolingual	Bilingual	Monolingual	Bilingual	Monolingual
All	463.4 ms	452.6 ms	5.8 ms	50.4 ms	1.2%	11.5%
Abmo	454.1 ms	447.6 ms	2.9 ms	56.5 ms	0.7%	13.1%
Ashmi	487.2 ms	479.3 ms	6.4 ms	60.2 ms	1.3%	13.1%
Ebza	439.8 ms	446.5 ms	7.7 ms	40.5 ms	1.8%	9.3%
Ibdo	472.5 ms	437.0 ms	4.8 ms	44.4 ms	0.8%	10.7%

Figure 1: Criterion for measuring the presence of vowel [u].

3.4. Results

The Mann-Whitney test found a significant difference between the two groups' average percentage of the inserted vowel [u] ($p<0.05$). The results of Experiment 2 showed that while Japanese monolinguals were highly likely to insert a vowel [u] between consonant clusters, Japanese-English bilinguals were not likely to do so. The average percentage of the inserted vowel [u] in all four pseudo-words is greater in Japanese monolinguals than Japanese-English bilinguals. The average duration of each pseudo-word, the duration of [u], and the percentages of [u] in each pseudo-word is shown in Table 2.

4. Discussion

4.1. Experiment 1: Perception

In Experiment 1, Japanese-English bilinguals and Japanese monolinguals were tested to determine whether they are able to distinguish the difference between the two sequences VCCV and VCVCV, by using a forced-choice AXB identification task. The experiment consisted of 8 practice trials and 192 experiment trials. The results showed that there was a significant difference in both the average number of errors and the average percentage of correct answers of both participant groups. This result indicates that the bilingual and monolingual participants' ability to identify VCCV and VVCVC differs according to whether the participants have the experience of living abroad or not. The place of birth of participants did not have an influence on the results, as was suggested by Peperkamp et al. However, both groups of participants' percentages of correct answers exceeded 90%. This may have been because of the vowel quality of the material used in Experiment 1. The materials used were natural utterances of VCCV and VCVCV; therefore, digital

editing of the material such as normalization of pitch and duration should be considered in future research.

4.2. Experiment 2: Production

In Experiment 2, the same participants who participated in Experiment 1 were asked to produce the 18 pseudo words that consist of the sequence VCCV (see Table 1). A carrier phrase was not used in order to prevent confusion.

Results from the analysis revealed that Japanese-English bilinguals had a lower tendency to insert an [u] between consonants than Japanese monolinguals. Out of the 68 utterances in bilinguals (17 speakers x 4 pseudo-words), only 8 utterances, or 11.8%, had an [u] inserted between the consonant clusters. On the other hand, out of the 88 utterances in monolinguals (22 speakers x 4 pseudo-words), 70 utterances, or 79.5%, had an [u] inserted between the consonant clusters. Although vowels inserted between unvoiced-voiced consonant set is said to become devoiced [2], it did not influence the results of this experiment. There was no tendency to insert a vowel in any specific pseudo-word in either group of participants; the tendency to insert a vowel between the consonant clusters was more at an individual level. The numbers of errors made in each pseudo-word by both groups of participants are shown in Table 3.

	Bilingual	Monolingual
Abmo	1	19
Ashmi	2	19
Ebza	3	18
Ibdo	2	18

Table 3: The numbers of errors made of each pseudo-word by bilingual and monolingual participant groups.

5. Conclusion

The aim of this study was to investigate how different language learning backgrounds affect the perception and production of consonant clusters in Japanese native speakers. Two groups of participants within Japanese native speakers were selected as participants: Japanese-English bilinguals and Japanese monolinguals.

The results of Experiment 1 revealed that there was a significant difference between bilinguals and monolinguals in the ability to discriminate consonant clusters from vowel-inserted consonant clusters. Although statistical analysis

found that there is a significant difference in the average percentages of correct answers, the difference in the identification ability was not as outstanding as we had expected: both groups' percentage of correct answers exceeded 90%. Although the results of bilinguals were not influenced by the place of birth, as Peperkamp et al. had suggested in previous research, gender may have an influence on the results. Such factors will be considered in future research.

The results of Experiment 2 revealed that there was a significant difference between bilinguals and monolinguals in the percentage of [u] inserted in the pseudo-words. Overall, monolinguals were highly likely to insert the vowel [u] between the two consonants in the consonant clusters compared to bilinguals. However, measuring the epenthetic vowel with the steady state of [u] may have not been the best method of analysis. An alternative method to measure the epenthetic vowel is also another factor that should be considered for future research.

From the results of the two experiments conducted in the present study, it can be concluded that statistically, Japanese-English bilingual and Japanese monolingual participants' perception and production abilities are different. From this conclusion, it can be said that bilinguals have an advantage of perceiving and producing foreign language phonotactics compared to monolinguals.

6. Acknowledgements

7. References

[1] Dupoux, E., Kakehi, K., Hirose, Y., and Pallier, C., "Epenthetic Vowels in Japanese: a Perceptual Illusion?", Journal of Experimental Psychology: Human Perception and Performance, Volume 25, Number 6, 1568-1578, 1999.

[2] Fujimoto, M. and Funatsu, S., "Vowel Epenthesis in Consonant Clusters by Japanese speakers", IEICE Technical Report, SP2007-204 (2008-3), 105-109, March 2008.

[3] Kubozono, H., "Nichieigo taishouniyoru eigogakuenshu series 1, Onseigaku Oninron", Kuroshio Publishers, Tokyo, 1998.

[4] Peperkamp, S., Dupoux, E. and Sebastian-Galles, N., "Perception of stress by French, Spanish, and bilingual speakers", Eurospeech '99 Proceedings, ESCA 7th European Conference on Speech Communication and Technology.

[5] Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J. and Fujimura, O., "An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English.", Perception and Psychophysics, 18, 331-340, 1975.

[6] Takagi, N., "The limits of training Japanese listerners to identify English /r/ and /l/: Eight case studies.", Journal of the Acoustical Society of America, 111(6), 2887-2896, June 2002.