Perception of English consonant clusters in quiet by Japanese native listeners with advanced and intermediate level English proficiency *

☆Hinako Masuda and Takayuki Arai
(Graduate School of Science and Technology, Sophia University)

1 Introduction

Foreign language acquisition is more or less difficult for everyone, and is caused by reasons such as first language interference, age of acquisition, the length and quality of exposure to the target language, and many more. Furthermore, one of the difficulties that foreign language learners face during language learning is the perception and production of sounds that are not in their first language. The difference in the phonemic inventories of the learner's first and target languages is of the possible reasons that cause such difficulties.

Cutler *et al*. [1] performed an English vowel and consonant identification experiment, and found that English native listeners performed significantly higher than non-native (Dutch) listeners in both quiet and noisy listening conditions. In the case of Japanese native listeners' perception of English consonants, Ueda *et al*. [2] stated that Japanese native listeners could not reach native-like performance in the identification of English /r/ and /l/ in noisy listening conditions, even after going through 15 days of training sessions.

Many experiments regarding perception of single consonants exist; however, the perception of consonant clusters still remains unknown in many aspects. Japanese native listeners are well-known for perceiving an "illusory vowel" between consonants. Dupoux *et al*. [3] found that Japanese native listeners were unable to distinguish "ebzo" and "ebuzo" (/VCCV/ and /VCVCV/). In other words, Japanese native listeners perceive "ebzo" as "ebuzo" due to first language interference. However, the influence of foreign language proficiency is still unknown, but is a factor worth considering. Japanese native speakers also produce vowel epenthesis within consonant clusters [4, 5]. This is due to the difference in the syllable structures of Japanese and English. Japanese is an open-syllable language, opposed to English being a closed-syllable language. Therefore, Japanese are likely to insert a vowel between consonants to avoid consonant clusters, which are phonologically illegal in the Japanese syllable structure.

The present paper focuses on the perception of English consonant clusters by Japanese native listeners with advanced and intermediate level English proficiency. While many studies on the perception of singleton consonants exist, many aspects of the perception of consonant clusters by Japanese with different English proficiencies are still unknown.

The present study aims to 1) investigate the difference in the perception of consonant clusters by Japanese native listeners with advanced and intermediate level English proficiency, 2) to compare their results with the English native control group, and 3) to compare the differences in the consonant confusions among the three participant groups.

2 Experiment

2.1 Participants

Twenty-three Japanese native listeners participated in the present study. Twelve participants (4 male, 8 female) were advanced learners of English, who had achieved higher than 850 on TOEIC® examination [6] or achieved equivalent scores on TOEFL® examination [7], and/or were placed in advanced English classes at university.

The remaining eleven participants (4 male, 7 female) were intermediate level learners of

* Perception of English consonant clusters in quiet by Japanese native listeners with advanced and intermediate level English proficiency, by MASUDA, Hinako and ARAI, Takayuki (Sophia University).

English, who had achieved lower than 700 on TOEIC® examination. Participants who do not have experience of living abroad received English education from age twelve at junior high schools in Japan.

Two native listeners of American English (2 male) also participated as a control group. None of the participants reported any hearing problems. The data of participants are shown in Table 1.

Table 1 Data of participants

	Intermediate learners	Advanced learners	English native listeners
Number of Participants	N = 11	N = 12	N = 2
Mean age (range)	22.9 (19 – 29)	22.9 (19 – 30)	27, 41

2.2 Stimuli

Twenty-three English legal consonant clusters were presented to participants in the C_1C_2V sequence. C_1 consists of one of the ten consonants (/s/ /b/ /d/ /g/ /p/ /t/ /k/ /f/ /th/ /sh/), C_2 consists of one of the nine consonants (/r/ /l/ /w/ /m/ /n/ /p/ /t/ /k/ /f/), and the vowel /a/ is used for V. C_1 stands for the first consonant in the consonant cluster (example: /s/ in /spa/), and C_2 stands for the second consonant in the consonant cluster (example: /p/ in /spa/). All consonant clusters were followed by the vowel /a/ to form C_1C_2V sequences. The vowel /a/ (V in C_1C_2V) was used in order to reduce coarticulatory differences among the stimuli [1]. The list of the C_1C_2V words is shown in Table 2.

The speaker of the stimuli is a female, Japanese-English bilingual speaker. The recording of the stimuli took place in a sound-proof room, using a digital sound recorder (Marantz PMD 660) and a microphone (SONY ECM-23F5) at a sampling frequency of 48 kHz. The stimuli were later down-sampled to 16 kHz.

Table 2 List of 23 phonologically legal English initial consonant clusters

CCV
bra, dra, gra, pra, tra, kra, fra, thra, shra, pla, kla, fla, sla, twa, kwa, thwa, swa, sma, sna, spa, sta, ska, sfa

2.3 Procedure

A laptop computer was used to present the stimuli and to record the participants' response. All procedure of the experiment was conducted by using Praat [8]. Participants were presented with the stimuli through a digital audio amplifier (ONKYO MA-500U) and headphones (STAX SR-303 and STAX SRM-323A). The laptop computer and the digital audio amplifier were digitally connected via USB interface.

All participants were given a total of 23 practice trials (six in quiet, nine in multi-speaker babble noise, and eight in white noise) before proceeding to the main experiment. The practice trials were not scored. After the practice trials, participants proceeded to the main experiment where 138 trials (23 consonant clusters x 2 repetitions x 3 listening conditions) were presented. The three listening conditions were 1) quiet, 2) multi-speaker babble noise at 0 dB, and 3) white noise at 0 dB. They were asked to listen to each stimulus and to write down what they heard. The participants were told that they would hear a "consonant – consonant – 'a'" sequence, and were asked to type the letters they heard on an Excel worksheet. The present paper reports only the results of the quiet listening condition. Thus a total of 1150 consonant clusters (25 participants x 23 consonant clusters x 2 repetitions) were analyzed.

3 Results

3.1 General results

The approximate average percentages of correct responses were 74% for intermediate level learners, 85% for advanced learners, and 84% for English native listeners (See Figure 1). Intermediate level learners made a total of 132 trial errors (approximately 26%) out of 506 trials

(11 participants x 23 consonants x 2 repetitions), advanced learners made a total of 81 trial errors (approximately 15%) out of 552 trials (12 participants x 23 consonants x 2 repetitions), and English native listeners made a total of 15 trial errors (approximately 16%) out of 92 trials (2 participants x 23 consonants x 2 repetitions). Chi-square test found a significant difference between intermediate level learners and advanced learners, and intermediate level learners and English native listeners, but not between advanced learners and English native listeners (p<0.05).

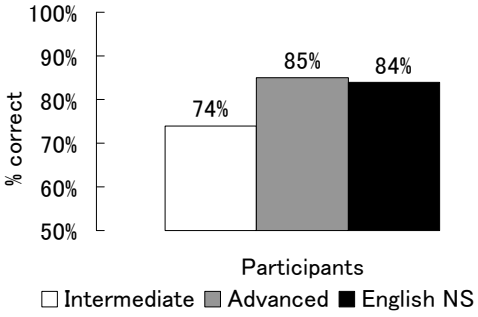

Figure 1 Average percentage of correct responses by the three listener groups

The average number of trial errors made by intermediate level learners, advanced learners, and English native listeners are shown in Figure 2. Intermediate level learners made more errors than both advanced learners and English native listeners, and advanced learners made similar amount of errors as English native listeners. The average number of consonantal errors (example: perceiving /spa/ as /thfa/ is considered as two consonantal errors) made by the three listener groups is shown in Figure 3.

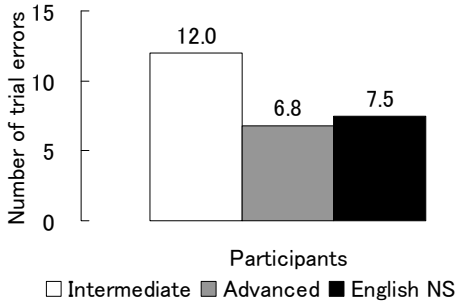

Figure 2 Average numbers of trial errors made by the three listener groups

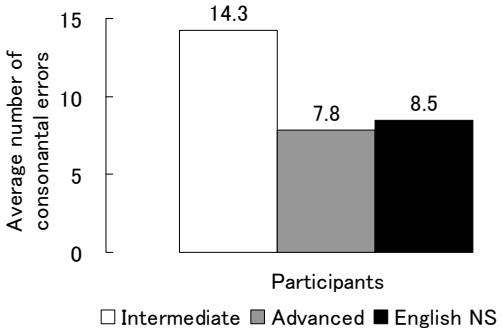

Figure 3 Average numbers of consonantal errors made by the three listener groups

3.2 Confusion matrices

Consonant confusion matrices by the three listener groups are calculated into percentages and are shown in Tables 3 to 5, respectively. Percentages may not add up to 100 due to rounding. Rows represent the stimuli presented to the participants, and columns represent the participants' responses.

Table 3 Consonant confusion matrix in quiet for intermediate learners (%)

	s	b	d	g	p	t	k	f	θ	ʃ	r	l	w	m	n
s	74							3	17	6					
b		100													
d			86	5					9						
g				100											
p					97	3									
t						88			12						
k							100								
f					2			87	6				5		
θ						9		11	78	2					
ʃ									5	95					
r											82	16	1		1
l											49	51			
w								2			1		97		
m														100	
n															100

Table 4 Consonant confusion matrix in quiet for advanced learners (%)

	s	b	d	g	p	t	k	f	θ	ʃ	r	l	w	m	n	others
s	89							1	10							
b		96														v (4)
d			67			4			25							v (4)
g				100												
p					96	3		1								
t						99			1							
k							98	1		1						
f								88	8				4			
θ						2		2	96							
ʃ									1	99						
r											95	4	1			
l											35	65				
w											1		99			
m								4						96		
n															100	

Table 5 Consonant confusion matrix in quiet for English native listeners (%)

	s	b	d	g	p	t	k	f	θ	ʃ	r	l	w	m	n	others
s	88							9	3							
b		100														
d			100													
g				100												
p		8			84			8								
t						100										
k				13			87									
f								67								v (33)
θ						25		13	62							
ʃ										100						
r											94	6				
l												100				
w													100			
m														100		
n															100	

3.3 Consonant positions

The breakdown of errors regarding consonant position by the three listener groups is shown in Figure 4. The percentage of C1 errors increased as English proficiency got higher, whereas the percentage of C_2 errors decreased as English proficiency got higher. Chi-square test found a significant trend between the breakdown of errors in C_1 and C_2 by the three listener groups (p=0.08).

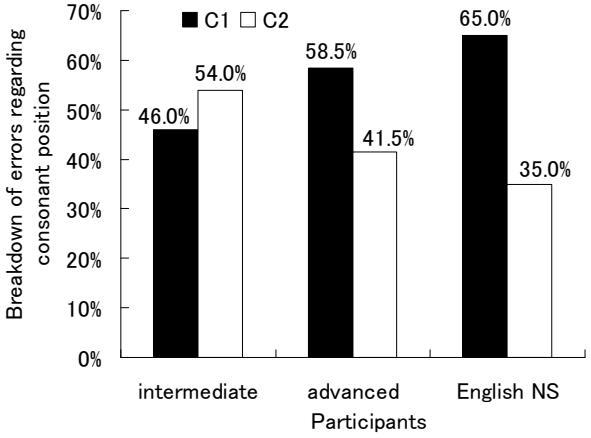

Figure 4 Breakdown of errors regarding consonant position by the three listener groups

4 Discussion

The present study reported the results of the perception of English consonant clusters in quiet by twenty-three Japanese native listeners and two American English native listeners. The results of the perception experiment revealed that Japanese native listeners with intermediate level English proficiency made more trial errors as well as consonantal errors than Japanese native listeners with advanced level English proficiency and English native listeners. Advanced learners, on the other hand, made similar numbers of errors as English native listeners.

The confusion matrices show that /s/ - /th/ and /t/ - /th/ are confused by both Japanese and English native listeners. On the other hand, /r/ - /l/ confusion, a common confusion for Japanese native listeners, was observed in the two Japanese listener groups, but not in English listeners. However, the confusions of /r/ and /l/ were asymmetric among the Japanese native listeners.

5 Conclusion

The present experiment revealed that 1) intermediate level learners made more trial and consonantal errors compared to advanced learners, 2) advanced learners made similar numbers of trial and consonantal errors as English native listeners, and 3) all three listener groups confused /s, t, th/, while /r/ - /l/ confusion was unique to Japanese native listeners, especially intermediate learners. However, further research is necessary with an increased number of trials as well as English native participants.

Acknowledgments

This study work was partially supported by Sophia University Open Research Center from MEXT.

References

[1] Cutler *et al.*, J. Acoust. Soc. Am., 116(6), 3668-3678, 2004.

[2] Ueda *et al.*, Acoust. Sci. & Tech., 28(4), 251-259, 2007.

[3] Dupoux *et al.*, J. Experimental Psychology: Human Perception and Performance, 25(6), 1568-1578, 1998.

[4] Masuda and Arai, Acoust. Sci. Tech., 2010, to appear.

[5] Fujimoto and Funatsu, IEICE Technical Report, SP2007-204 (2008-3), 105-109, 2008.

[6] TOEIC® (Test of English for International Communication), provided by ETS (Educational Testing Service).

[7] TOEFL® (Test of English as a Foreign Language), provided by ETS (Educational Testing Service).

[8] Boersma and Weenink, Praat: doing phonetics by computer (Version 5.0.09) [Computer program].