Production variation of English schwa and Japanese listeners' perceptual assimilation pattern of English schwa

Kanako Tomaru・Takayuki Arai (Sophia UNIV)

1. Introduction

Schwa in a word-final position is said to show more variation in F1 than in F2, whereas that in a word-medial position has more variation in F2 than in F1 (Flemming and Johnson, 2007). Although more recent study suggests that schwa in a word-final position is relatively consistent in vowel quality compared schwa in a word-medial position (Flemming, 2009), schwa is kind of a vowel which is considered to show formant variations depending on contexts. The first part of the present report attempts to examine variation of schwa depending on word position (Experiment 1).

Secondly, the present research attempts to show how Japanese listeners perceive English schwa, especially, how their perception reflects schwa's formant variation, based on Perceptual Assimilation Model (PAM) suggested by Best (1994) (Experiment 2). The examination of Japanese listeners' "perceptual assimilation" (Best, 1994; Best and Tyler, 2007) patterns of English vowels has already been reported by Strange *et al.* (1998, 2001); however, which of the five Japanese vowels native speakers of Japanese hear English schwa is not yet clarified. It is intriguing to investigate Japanese listeners' perception of schwa vowels not only because we can follow up the previous research, but also because we can examine how perceptual pattern of schwa reflect its formant variation depending on contexts.

2. Experiment 1: Variation of English Schwa

2.1. Background

English schwa vowels, which are generally considered as weak vowels or reduced vowels that occur in non-stressed positions, are said to have variation depending on word positions (Flemming and Johnson, 2007; Flemming, 2009). In word-final position, schwa is said to have more variation in F1 than in F2 (Flemming and Johnson, 2007), although more resent study suggests word-final schwa vowels are much stable in formant qualities (Flemming, 2009). Word-medial schwa vowels, on the other hand, show greater variation in F2 than in F1 (Flemming and Johnson, 2007). Flemming and Johnson (2007) first made this point clear through examination of schwa in word-medial and word-final positions, using real English words. The first section presents examination of English schwa spoken by native speakers of English to see if we could obtain formant variation of schwa as reported in the previous studies. Unlike Flemming and Johnson (2007), the present research employed nonsense words for examination, where schwa appeared in the following positions: 1) word-initial, 2) word-medial, 3) word-final without coda, and 4) word-final with coda.

2.2. Recording

2.2.1. Speakers

One female Canadian speaker and one male American speaker participated in recording.

2.2.2. Materials

Test materials were nonsense words that contained schwa in various positions. Schwa was surrounded by five different vowels, i.e. [æ/ai], [i], [u], [e], and [oʊ], and four different consonants, i.e. [b], [d], [g], and [z] (see Table 1). The test nonsense words were presented to speakers by using English alphabets, e.g. *tagagite* for [ˈtægəˌgait]. In addition to the test materials, real English words and other nonsense words were also recorded as non-test words. The nonsense test words looked unfamiliar in spellings, so written instructions as well as oral instructions for pronunciation were given. The written instruction gave the speakers examples of words which had the same vowel as in the nonsense test words, e.g. *tagagite*: *ta* as in "tabby", *ga* as in "soda", *gite* as "bite".

2.2.3. Recording condition

Recording of each speaker took place on different days in a quiet room. Test nonsense words and non-test words were recorded on SONY DAT walkman recorder at 48 kHz. The recorded speech was re-digitized using Adobe Audition 1.5 with sampling rate of 44.1 kHz and 16-bit resolution.

2.3. Formant analysis

2.3.1. Schwa in different positions

Formant values at the midpoint of schwa were measured using the Praat software. Standard Deviations (SD) of formants showed that variation was greater in F2 than in F1 in all word positions

Table1: List of test materials. " × " indicates where the test nonsense words are missing.

	Surrounding Vowel Context				
Position	**[æ / ai]**	**[i]**	**[u]**	**[e]**	**[oʊ]**
Initial	[əˈbaiv],[əˈdaiz], [əˈgai], [əˈzaid]	[əˈbip],[əˈdit], [əˈgik],[əˈzi]	[əˈbuf],[əˈdus], [əˈguk],[əˈ_zut]	×	×
Medial	[ˈtæbəˌbait], [ˈtædəˌdait], [ˈtægəˌgait], [ˈtæzəˌzait]	[ˈtibəˌbit], [ˈtidəˌdit], [ˈtigəˌgit], [ˈtizˌzit]	[ˈkubəˌbuk], [ˈkudəˌduk], [ˈkugəˌguk] [ˈkuzəˌzuk]	[ˈtebəˌbet], [ˈtedəˌdet], [ˈtegəˌget], [ˈtezəˌzet]	[ˈkoʊbəˌboʊk], [ˈkoʊdəˌdoʊk], [ˈkoʊgəˌgoʊk], [ˈkoʊzəˌzoʊk]
Final without coda	[ˈsæbə],[ˈsædə], [ˈsægə], [ˈsæzə]	[ˈkibə],[ˈkidə] [ˈkigə],[ˈkizə]	[ˈpubə],[ˈpudə] [ˈpugə], [ˈpuzə]	×	×
Final with coda	[ˈkæbəb], [ˈkædəd], [ˈkægəg], [ˈkæzəz]	[ˈpibəb], [ˈpidəd], [ˈpigəg], [ˈpizəz]	[ˈbubəb], [ˈbudəd], [ˈbugəg], [ˈbuzəz]	×	×

Table 2: Standard Deviation (SD) of schwa produced by the Canadian speaker (left) and American speaker (right) in ERB (Equivalent Rectangular Bandwidth).

	Canadian speaker		American speaker	
Word Position	**F1**	**F2**	**F1**	**F2**
Initial	.50	.80	.71	.86
Medial	.89	.88	.61	.76
Final	.32	.50	.73	.99
Final with coda	.35	.60	1.05	.78

Table 3: Correlation coefficients between formant frequencies (ERB) of schwa and those of preceding and following vowels.

		Canadian speaker	American speaker
F1	**schwa & preceding vowel**	.769 ($p < .001$)	.532 (p=.016)
	schwa & following vowel	.903 ($p < .001$)	-.059 (p=.803)
F2	**schwa & preceding vowel**	.935 ($p < .001$)	.842 ($p < .001$)
	schwa & following vowel	.944 ($p < .001$)	.918 ($p < .001$)

for both speakers (Table 2). To assess the significance, the value of F-distribution for obtained data was calculated, and compared to the value with the significance level of 5%. When the former is greater than the latter, distribution patterns of the two different groups are significantly different from each other. Results showed that the distributional variation of F1 was significantly greater than that of F2 for the Canadian speaker (F= SD of F1 / SD of F2 = 1.71, F (0.05) =1.56). On the other hand, the value of F-distribution for the American speaker showed that the distributional variation of F1 and that of F2 are not significantly different (F=1.28, F (0.05) =1.56). Figure 1 shows formant distribution in different positions in a scale of Equivalent Rectangular Bandwidth (ERB): (a) word-initial, (b) word-medial, (c) word-final without coda, and (d) word-final with coda.

2.3.2. Variation conditioned by surrounding vowels

Only word-medial schwa vowels were subject to the analysis. Formant values of the surrounding vowels as well as those of schwas were measured at the midpoint of the vowels' steady state. For both speakers, positive correlation was obtained between F2 values of schwa and those of surrounding vowels (Table 3); correlation between F1 of schwa and that of preceding vowels were also obtained for both speakers. However, correlation between F1 of schwa and that of following vowels were obtained only for the Canadian speaker.

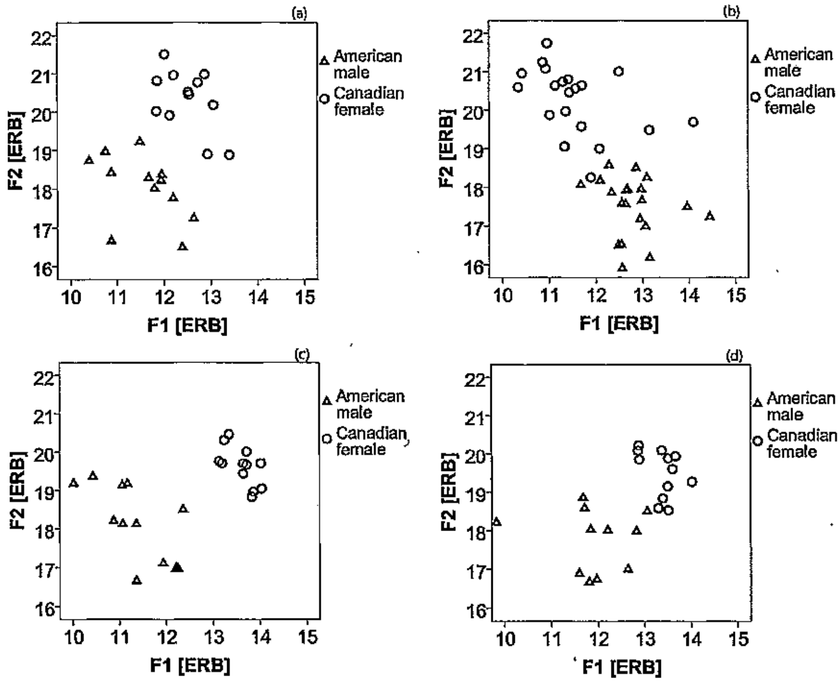

Figure 1: Formant frequencies of schwa produced in different positions: (a) word-initial, (b) word-medial, (c) word-final without coda, and (d) word-final with coda. Formant frequencies are converted from Hz to ERB by using the following formula: ERB = 21.4log10(0.00437f+1).

2.4. Summary

In this section, we demonstrated formant variation of schwa vowels. In the present research, F2 variation was greater than F1 variation in all positions. The results are slightly different from those of the previous study, which suggests F1 variation is greater than F2 variation in word-final position (Flemming and Johnson, 2007). In Flemming (2009), however, word-final schwa is said to be stable in quality; as Fig. 1-c and Fig. 1-d tell you, it can be said that schwa vowels in final positions are more stable than those in the other positions in the present research as well, at least for the Canadian speaker. This finding is parallel to the suggestion of Flemming (2009). In addition, effects of surrounding vowels were observed for both speakers; but correlation between F1 of schwa and that of following vowels were obtained only for the Canadian speaker.

The reason why our results differed from previous findings may be because our test materials were nonsense words, while test materials of the previous study were real words (Flemming and Johnson, 2007). We need further experiments to examine the effects of meaningfulness on schwa variation. Another possibility is that the results reflect dialectal difference of speakers. The effects of dialects on schwa variation are also an open question.

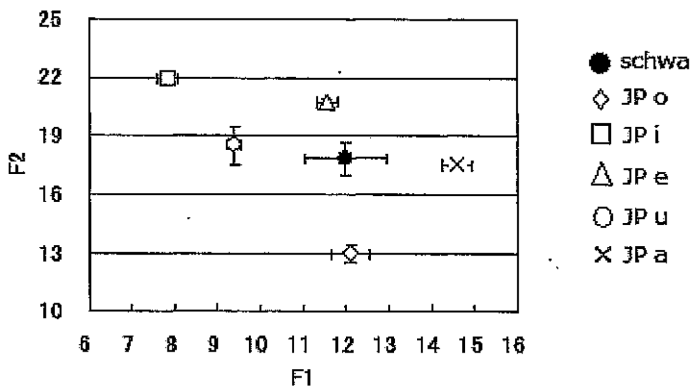

Figure 2: Formant frequencies of schwa and Japanese vowels in ERB. Error bars show SD.

3. Experiment 2: Perception of English Schwa by Japanese listeners

3.1. Background

This part of the report attempts to show perceptual patterns of English schwa by Japanese listeners based on PAM (Best, 1994; Best and Tyler, 2007). PAM predicts that listeners perceptually assimilate foreign speech sound to native speech sound. The word "perceptual assimilation" refers to the process of perceiving an auditory presented non-native speech sound as a native speech sound. For example, if the English vowel /ɪ/ is perceived as equivalent to the Japanese vowel /i/ by a Japanese-speaking listener, the English /ɪ/ is assimilated to the Japanese /i/ by the speaker. Strange *et al.* (1998, 2001) concerned with Japanese listeners' perceptual assimilation of English vowels; but they did not deal with schwa vowels. The present report examines perceptual assimilation patterns of English schwa, especially, how they reflect formant variation of schwa obtained in Experiment 1.

3.2. Listeners

Twenty native speakers of Japanese (3 males, 17 females) ranged in age from 20 to 24 years participated in a perceptual experiment.

3.3. Procedure

Nonsense test words recorded for Experiment 1 were used as stimuli. For a perceptual experiment, participants were divided into four groups, and each group was presented with the same stimuli in different order. The participants wore headphones (SONY Dynamic Stereo Headphones MDR-Z500) connected to EPSON laptops, and they transcribed the stimuli in katakana, as if they were adapting new loanwords to Japanese. They had a practice session before an experimental session.

3.4. Results

Assimilation responses were dominantly Japanese /a/, i.e. 96%. Other responses such as Japanese /u/ and /o/ were obtained, but they were rare, i.e. less than 4%. /u/ and /o/ responses were mainly obtained for the contexts [ˈkægəg] and [ˈkæzəz]. Japanese /e/ and /i/ response were not reported at all.

3.5. Summary and Discussion

The experiment revealed that English schwa vowels were most likely heard as the Japanese vowel /a/ by native speakers of Japanese. That is, listeners assimilated English schwa to Japanese /a/ regardless of formant variation of schwa vowels. The reason why Japanese listeners assimilated English schwa to Japanese /a/ may be the formant similarity between English schwa and Japanese /a/. So, we tested the hypothesis by recording Japanese vowels produced by a male Japanese speaker and compared the formant values of Japanese vowels to those of schwa (Figure 2). Figure 2 shows that schwa does not seem to be any closer to Japanese /a/ than the other Japanese vowels. Thus it requires further examination to reveal why English schwa is perceptually assimilated to Japanese /a/.

4. Conclusion

We first examined formant frequencies of English schwa vowels in various positions, and showed formant variation of schwa. Secondly, we examined Japanese listeners' perception of schwa based on PAM to see if the perception pattern reflects schwa's formant variation. The results showed that schwa vowels were perceptually assimilated to Japanese /a/ predominantly.

References

Best. C. T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). MA: MIT Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: commonalities and complementarities. In O. –S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning: in honor of James Emil Flege* (pp. 13-34). PA: J. Benjamins Pub.

Boersma, Paul & Weenink, David (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.20, retrieved 28 June 2012 from http://www.praat.org/

Flemming, E. (2009). The phonetics of schwa vowels. In D. Minkova (Ed.), *Phonological weakness in English from old to present-day English* (pp.78-95). NY: Palgrave Macmillan.

Flemming, E., & Johnson, S. (2007). Rosa's roses: reduced vowels in American English. *Journal of the International Phonetic Association, 37*, 83-96.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., & Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America, 109,* 1691-1704.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics, 26,* 311-344.