Impact of talker variability on L2 word recognition among Japanese EFL learners

Noriko Matsuda

Aino University

doi: <http://dx.doi.org/10.7820/vli.v06.2.Matsuda>

Abstract

This article is a pilot study investigating auditory word priming in 40 Japanese learners of English using speeded repetition tasks to measure the impact of talker changes on second language (L2) word recognition. The results showed that by focusing more on the perceptual dimension, in single-talker conditions, word recognition time was statistically significantly shorter and a perceptual learning effect was seen. However, with talker changes, word recognition time significantly increased and the repetition effects were nullified. The results indicate that Japanese English as a foreign language (EFL) learners have high sensitivity to individual attributes of speech and seem to need some variations in auditory input for some period of time without any need for comprehension in order to form robust representations of L2 words.

1 Background

When listening to the news on TV in one's first language (L1), one seldom experiences sudden difficulty in understanding what is being said when newscasters change. However, many L2 learners have trouble understanding a new talker. Talker variability affects low-level speech-perception skills of L2 learners because the skill is automatized in L1 but not in L2. During language learning and acquisition, one cannot properly comprehend what is being said if one is unable to ignore the variations in prosody and pronunciation between different talkers, identify and retain vocabulary patterns, and associate meanings with those patterns. In spoken language, listeners cannot separate linguistic information or content from acoustic elements (paralinguistic information) such as differences between talkers' voices and emotional inflections. Therefore, exploring how such variations are processed is an important topic in considering the mechanisms of spoken language learning and acquisition. The cognitive process of L2 speech perception is complicated and has many unknown parts as both the ability to recognize sounds in the target language as well as vocabulary knowledge are limited and differ according to learners' proficiency levels. To understand this complexity, the effects of talker variability from the view point of cognitive development and vocabulary learning are mentioned below.

1.1 Talker Variability and Cognitive Development

Based on perception and cognitive research on L1 speakers, it is widely believed that the capacity to identify common linguistic information from

different talkers is present in early infancy, before infants acquire the ability to link words with meaning. Research on L1 speech perception development in infants has shown that at two months of age, the learning of syllables does not proceed very well when infants are exposed to multiple talkers (Jusczyk, Pisoni, & Mullennix, 1992). According to Houston and Jusczyk (2000), common linguistic information can be recognized by infants at seven and a half months of age when talkers are of the same gender, but not when talkers are of different genders. Processing information with no influence of talker variability becomes possible at ten and a half months of age. At this stage, infants begin to process linguistic information independently of the acoustic features of individual talkers' voices. At 12 months of age, they are generally able to do this quite well.

At this point, an infant's attention to information that is relatively less important to understanding spoken language, such as acoustic features in utterances, is inhibited for some time. A study of native Japanese speakers showed that at the age of two, they have difficulty learning new words similar to words they already know but differing in pitch accentual patterns. In comparison, by age three, they can learn such words easily (Yamamoto & Haryu, 2016). In addition, in early infancy, children may be sensitive to emotional prosody in utterances. From infancy to later childhood, the phenomenon of lexical bias comes into play: children prioritize the linguistic content of utterances over the manner of speaking to infer a talker's feelings (Friend & Bryant, 2000). A talker's way of speaking is given more importance as children grow older. This overlaps precisely with the period of development of the executive function, which is believed to control human attention (Chevalier, 2015; Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2010; Jerger, Martin, & Pirozzolo, 1988).

While this research field continues to be subject to many debates concerning L1 speech perception development, as mentioned in this overview, some important findings have been made. As children's knowledge of their L1 develops, they become able to independently process paralinguistic information (modularity of processing). Furthermore, children appear to allocate limited cognitive resources to more important linguistic information as they develop the ability to switch attention.

Various L1 studies on adults have demonstrated the robustness of their ability to cope with talker variability. One study has shown that even in a slightly noisy environment, speaker adaptation occurs after listening to only five syllables spoken by a single talker (Kato & Kakei, 1988). Moreover, adults' capacity to cope with various environmental changes in their L1 is said to be due to perceiving speech hierarchically (top-down processing) by concentrating on not only its acoustic aspects, but also on a variety of information from memory, experience, and knowledge.

1.2 Talker Variability and Vocabulary Learning

Several types of acoustic variability were known to be effective regarding vocabulary learning. According to Barcroft (2015), the high variability conditions of phonetically relevant acoustic features were more effective for lexical development than the low variability conditions. In his study, the effects of five different sources of acoustic variability were summarized using the results of several

different studies: talker variability, speaking style variability, speaking rate variability, amplitude variability, and F0 variability. The effects of the variability were measured by word identification in noise for L1 studies and recall tests for L2 studies. The results showed that the first three sources have negative effects on L1 word identification but positive effects on L2 vocabulary learning (and only talker variability has positive effects on L1 vocabulary learning). Neither amplitude variability nor F0 variability (except for tone-language speakers) affect L1 and L2 vocabulary learning. Barcroft insisted that some sources of acoustic variability “can facilitate L2 vocabulary learning by producing more robust or more distributed representations of the L2 words in the mind.” (Barcroft, 2015, p. 153) In addition, a number of studies have shown that for both adults and children, more robust representations can be formed by being exposed to different L2 talkers during the early stages of learning (Lively, Logan, & Pisoni, 1993; Kingston, 2003; Rost & McMurray, 2009, 2010).

Regarding the negative results of L1 word identification tasks, Sommer, Nygaard, and Pisoni (1994) stated that talker variability could be a factor that produces some cognitive load (or cost), thus hindering L1 word identification in noise. Similarly, talker variability is also known to be a factor that imposes cognitive loads on L2 learners’ speech processing (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Goldinger, 1996; Trofimovich, 2005). Some L2 priming studies have measured word recognition speed and accuracy, and they have also shown similar results: the priming effects (perceptual learning effects) were greatly reduced if a word that the participants had processed once was repeated by a different voice (e.g., Trofimovich, 2005). Priming effects are “facilitative effects of an encounter with a stimulus on subsequent processing of the same stimulus (direct priming) or a related stimulus (indirect priming)” (Tulving, Schacter, & Stark, 1982, p.336). However, talker variability seems to reduce these effects. Considering that auditory priming is said to be a mechanism that supports spoken language learning and acquisition, it would be worthwhile to see the effects of talker variability on L2 vocabulary learning from the perspective of auditory priming.

Based on research on L1 speech perception development in children, it is possible that L2 speech containing paralinguistic information may be processed quite differently, depending on the linguistic information that the cognitive resources are allocated to. Cognitive resource allocation requires the development of the attention switching function of the executive function, which can be assumed to be adequately developed by adulthood. Therefore, for adults, it is highly likely that cognitive resource allocation in L2 processing is determined by the focus of the learner’s attention. In L1 research, there were no differences in priming effects when participants listening to words focused their attention on the perceptual dimension or the semantic dimension. According to Tulving and Schacter (1990), the nonsemantic nature of priming suggests the existence of a presemantic perceptual system in human memory. In contrast, L2 studies have suggested that focusing on the semantic dimension results in negative priming effects, depending on the learner’s proficiency.

A priming experiment using talker variability and attention focus as independent variables showed no effects of talker variability on priming effects in L1. In L2, talker variability caused negative effects and no priming effect was seen, regardless of whether attention was focused on the perceptual or semantic

dimension (Trofimovich, 2005). A similar L2 study used length of stay in the L2 country as well as attention focus and talker variability as independent variables. A priming effect was seen only in the longer-stay group when attention was focused on the perceptual dimension (Trofimovich, 2008). These two studies used audio recordings of the speech of six native speakers. When the speaker was changed in the test phase, a voice of the other gender was used. Tasks focusing on the perceptual dimension asked the participants to rate the sound clarity of each word, while meaning-focused tasks asked them to rate the pleasantness of word meaning (i.e., to rate how fun the meaning of each word was). These studies showed that when the participants focused on the perceptual dimension (the sounds of the words), talker change did not affect speech processing in participants with a long exposure to the L2; however, when participants focused on the semantic dimension (the meanings of the words), talker change greatly reduced the priming effect. We can predict that for Japanese EFL learners with little exposure to English speech, talker changes will reduce the priming effect, regardless of the attention focus. Furthermore, the combined effect of talker variability and focusing attention on meaning is likely to produce a significant decrease in the priming effect.

An auditory priming effect has previously been demonstrated in Japanese EFL learners (Matsuda, 2013; Sugiura & Hori, 2012). However, as far as the researcher is aware, no studies taking into account talker variability and cognitive resource allocation (attention focus while listening to vocabulary words) have been conducted. The goal of this study is to examine the effect that talker variability has on L2 word recognition among Japanese EFL learners. The researcher posed the following research questions based on the outcomes of previous studies:

- (1) Does talker variability negatively affect L2 word recognition in Japanese EFL learners?
- (2) Does the combination of talker variability and cognitive resource allocation affect L2 word recognition in Japanese EFL learners?

The following hypotheses were formulated based on previous research:

- (1) The L2 priming effect will decrease with talker variability.
- (2) If attention is focused on the semantic dimension while listening to words in the presence of talker variability, the L2 priming effect will decrease.

2 Method

2.1 Participants

The participants in this study were 40 Japanese learners of English (undergraduate and graduate students) at two private universities in the Kansai region of Japan. There is an overlap with the participants of Matsuda (2017). The Oxford Quick Placement Test with a 60 point maximum (Oxford University Press, University of Cambridge & Association of Language Testers in Europe, 2001) and the Listening Test of the Oxford Placement Test with a 100 point maximum (Allan, 2004) were used to assess their proficiency in English. Table 1 shows the participants' background as captured by questionnaire.

Table 1. Participants' Background Revised from Matsuda (2017)

	<i>M</i>	<i>SD</i>
The Oxford Quick Placement Test	30.95	7.14
The Oxford Placement Test (Listening Test)	74.67	8.35
Age starting English study	11.41	1.72
Years of formal English education	11.92	1.33
Years of residence in English-speaking countries	0.20	0.96
Age	20.90	2.67
Self-ratings ^a	4.51	1.22
Listening	4.56	1.69
Speaking	3.92	1.38
Reading	4.97	1.39
Writing	4.56	1.48

Note. *SD* = standard deviation. ^aRatings scored on a 10-point scale with 1 = minimum proficiency and 10 = near-native proficiency. *N* = 40.

2.2 Materials

Vocabulary groups, revised from Matsuda (2012), were created based on controlled familiarity of spoken words (Yokokawa et al., 2009; $F[7, 136] = 0.33$, $p = .94$, $\eta_p^2 = .02$ [Spoken]), familiarity of written words (Yokokawa et al., 2006; $F[7, 136] = 0.81$, $p = .58$, $\eta_p^2 = .04$), frequency (British National Corpus or BNC; $F[7, 136] = 0.67$, $p = .70$, $\eta_p^2 = .03$), number of syllables ($F[7, 136] = 0.06$, $p = .99$, $\eta_p^2 = 0.003$), duration ($F[7, 136] = 0.06$, $p = .99$, $\eta_p^2 = .003$), and initial consonants (initial consonants of six out of 36 pairs were uncontrolled) (see Appendices A to D).

The recorded voices of two native English speakers were used for this study. The voices were monaurally digitally recorded in a sound booth at a sampling rate of 48 Hz and quantization at 16 bits. Praat (Boersma & Weenink, 2011) was used to standardize the volume and speed of speech, a major factor affecting intelligibility and comprehensibility (Jones, Berry, & Stevens, 2007).

2.3 Procedure

All the participants underwent two phases of the experiment: a study phase and a test phase. The study phase included a rhyme judgment task to focus the attention of participants on the perceptual dimension of words, and a synonym judgment task focusing on the semantic dimension of words. Eighteen pairs of words were used for each task. The participants were asked to determine whether or not two consecutive words they heard rhymed (rhyme judgment task) and whether or not the two words had similar meanings (synonym judgment task). In the test phase, a speeded repetition task was used, where participants were asked to repeat the words they heard as soon as possible. New vocabulary groups were added (unrepeated words) to compare the reaction time (RT) of the repetition. To see the effects of talker variability, the voice was changed from the study to the test phase (Female → Male and Male → Female), as shown in Appendices C and D. Two evaluators, including the experimenter, analyzed the RT data for errors

(the concordance rate in error evaluations was 94.66%). Incorrect data, which accounted for 5.10% of the results, were excluded from the analyses.

2.4 Data Analyses

RT was used as the dependent variable in the current experiment as in Matsuda (2017), and an analysis of variance (ANOVA) was conducted to investigate the priming effects.

3 Results and Discussion

Table 2 and Figure 1 show the RT data and the priming effects. The RT data of the repetition task were submitted to 2 (Task) $\times$ 3 (Voice Variability) repeated measures ANOVA, with Task (rhyme judgment task or synonym judgment task in the study phase) and Voice Variability (repeated words in the same voice or different voice in the test phase, or unrepeatd words) as within-subjects factors. Post-hoc comparisons were conducted using the procedure from Benjamini and Hochberg (1995). This analysis revealed a significant interaction between Task and Voice Variability, $F(2, 78) = 30.42, p < .001, \eta_p^2 = 0.44, power = 1.00$.

The simple main effect of Task was significant for the same voice, $F(1, 39) = 8.16$, adjusted $p = 0.02, \eta_p^2 = 0.17$, but not for different voices, $F(1, 39) = 1.34$, adjusted $p = 0.32, \eta_p^2 = 0.03$, nor for unrepeatd words, $F(1, 39) = 0.19$, adjusted $p = 0.66, \eta_p^2 = 0.005$. The participants could repeat the words significantly faster when rhyme judgment was conducted in the study phase and the words were presented in the same voice.

The simple main effect of Voice Variability was significant for both the rhyme judgment task, $F(2, 78) = 50.83$, adjusted $p < 0.001, \eta_p^2 = 0.57$, and the synonym judgment task, $F(2, 78) = 4.87$, adjusted $p = 0.02, \eta_p^2 = 0.11$.

Under the rhyme judgment task conditions, the post-hoc test of Voice Variability showed that the RT data of the same voice and a different voice, as well as the same voice and unrepeatd words, were statistically significantly different (adjusted $ps < 0.001, < 0.001$, respectively). However, a different voice and unrepeatd words did not show statistically significant differences (adjusted $p = 0.58$). Under the synonym judgment task conditions, the post-hoc test of Voice Variability revealed that the RT data of a different voice and unrepeatd words were significantly different (adjusted $p = 0.02$), while the same voice and a different voice, as well as the same voice and unrepeatd words, did not show statistically significant differences

Table 2. Mean Reaction Times (ms)

	Rhyme		Synonym	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Unrepeatd	836.44	106.70	840.26	104.50
Repeatd				
Same voice	806.70	112.70	832.53	109.53
Different voice	838.29	105.93	828.07	107.60

Figure 1. Mean priming effects (ms).
 Note: Vertical lines indicate standard error.

(adjusted p s = 0.28, = 0.06, respectively). The rhyme judgment task in the study phase revealed a priming effect only for words repeated in the same voice. However, the synonym judgment task conducted in the study phase indicated a priming effect only for words repeated in a different voice. Importantly, there were no statistically significant differences between the RT data for the same and different voices.

4 Discussion

The results of this experiment only partially supported Hypothesis 1 (The L2 priming effect will decrease with talker variability). When participants listened to words focusing on the perceptual dimension, there was a statistically significant effect of talker variability and a negative priming effect was seen. This appears to be a reasonable result considering the fact that most learners of English in Japan learn in environments with little exposure to spoken English and that previous studies have had similar results. However, when participants focused on the semantic dimension, there were no statistically significant differences in RTs, with or without talker variability. Furthermore, contrary to Hypothesis 2 (If attention is focused on the semantic dimension while listening to words in the presence of talker variability, L2 priming effect will decrease), when participants focused on the semantic dimension of words with talker variability, a statistically significant priming effect was seen; whereas, with no talker variability, there was only a trend toward statistical significance (adjusted p = .06).

One might conclude that the reason for these results is that this experiment had less variation in speech than previous studies (six speakers in the previous studies compared with two in the present study). The participants may have grown accustomed to the talkers' voices, resulting in a priming effect. However, if this were the case, the same results would have occurred when the participants focused on the perceptual dimension. Furthermore, a close examination of the mean RTs for each item showed that there was no particular item to which the participants responded with notable speed. Considering the fact that no statistically significant differences in RTs were seen when the focus was on the semantic dimension, regardless of whether or not the talker remained the same, the likelihood is high that the reason must lie in the contents of the tasks.

In the previous studies, a meaning-focused task asked participants to rate how fun the words were. However, this study used a synonym judgment task. One can imagine that there may have been some L1 interference. While there is also a high likelihood that L1 translation was involved in order to rate the pleasantness of a word, comparing two words requires greater cognitive effort, which may have strengthened the related memory traces. However, this alone does not explain why a priming effect was seen, regardless of talker variability. The experimenter would like to propose the possibility that attention to paralinguistic information (i.e., the emotions, age, gender, and dialect of the talker) may have been inhibited, as it was not essential to processing the language. This is similar to what is seen in L1 speech perception development, as shown in previous research. Since the cognitive load related to the meaning-focused task in this study was greater than that in the previous studies, it likely required the allocation of more cognitive resources to process semantic information. This in turn may have reduced attention to paralinguistic information. As a result, participants may have been less susceptible to any effects from variations in paralinguistic information. Furthermore, if participants were less susceptible to the acoustic aspects of speech, there is a possibility that top-down speech perception occurred during the test period.

5 Conclusion

The limitations of this study must be noted. This study used a limited range of voices and does not contain information on the longitudinal effects of paralinguistic information. This is something that the researcher hopes to address through future research.

Looking strictly at the priming effect when participants focused on the perceptual dimension of words in the presence of talker variability, one could suggest that Japanese EFL learners were unable to process paralinguistic information independently of linguistic information. However, the perceptual learning effect was quite high when talkers were the same, suggesting the potential presence of a talker adaptation mechanism. The results of this experiment indicate the possibility that attention to L2 paralinguistic information is inhibited in situations where it is necessary to allocate large amounts of cognitive resources to semantic information. This is a phenomenon seen during the development of L1 speech perception in children. This implies the existence of a common system across L1 and L2 for the allocation of the majority of cognitive resources to the essential part of spoken language: meaning. Sufficient exposure to a variety of voices (exemplars)

could lighten the cognitive load of L2 learners, allowing for the possibility that L2 speech perception can be changed. In this study, we could see the perceptual learning effect when L2 learners focused on the perceptual dimension.

This suggests that it is essential to allow L2 learners some time to listen without comprehending to become accustomed to the perceptual dimension and that some variations in auditory input should also be provided.

The pedagogical implications are that, simply put, in order to retain vocabulary information that can be processed quickly without being affected by various perceptual factors, the number of exemplars should be increased. Moreover, the kinds of exemplars and their presentation in the classroom should be carefully chosen because the vocabulary knowledge of L2 learners is dependent on the quality of these exemplars.

References

- Allan, D. (2004). *Oxford placement test 2*. Oxford, England: Oxford University Press.
- Barcroft, J. (2015). Effects of acoustically varied input. In J. Barcroft (Ed.), *Lexical input processing and vocabulary learning* (pp. 149–162). Amsterdam: John Benjamins.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B*, 57, 289–300.
- Boersma P., & Weenink D. J. M. (2011). Praat: doing phonetics by computer (Version 5.2.13) [Computer program]. Retrieved from <http://www.praat.org/>
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299–2310. doi:10.1121/1.418276
- Chevalier, N. (2015). The development of executive function: Toward more optimal coordination of control with age. *Child Development Perspectives*, 9, 239–244. doi:10.1111/cdep.12138
- Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science*, 13(1), 120–133. doi:10.1111/j.1467-7687.2009.00864.x
- Friend, M., & Bryant, J. B. (2000). A developmental lexical bias in the interpretation of discrepant messages. *Merrill-Palmer Quarterly*, 46(2), 342–369.
- Goldinger, S.D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1166–1183. doi:10.1037/0278-7393.22.5.1166
- Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1570–1582. doi:10.1037/0096-1523.26.5.1570
- Jerger, S., Martin, R. C., & Pirozzolo, F. J. (1988). A developmental study of the auditory Stroop effect. *Brain and Language*, 35(1), 86–104. doi:10.1016/0093-934X(88)90102-2

- Jones, C., Berry, L., & Stevens, C. (2007). Synthesized speech intelligibility and persuasion: Speech rate and non-native listeners. *Computer Speech and Language*, 21, 641–651. doi:10.1016/j.csl.2007.03.001
- Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, 43, 253–291. doi:10.1016/0010-0277(92)90014-9
- Kato, K., & Kakei, K. (1988). *Onseichikaku ni okeru washa eno tekiousei no kentou* [Listener adaptability to individual speaker differences in monosyllabic speech perception]. *Nihon Onkyougaku Kaishi* [Acoustical Science and Technology], 44, 180–186.
- Kingston, J. (2003). Learning foreign vowels. *Language and Speech*, 46(2–3), 295–348. doi:10.1177/00238309030460020201
- Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, 94, 1242–1255. doi:10.1121/1.408177
- Matsuda, N. (2012). Effects of auditory word repetition on speech processing of Japanese EFL learners, *Language Education & Technology*, 49, 143–172.
- Matsuda, N. (2013). Second-language speech processing: Auditory word priming in Japanese EFL learners and native English speakers, *Journal of the Japan Society for Speech Sciences*, 14, 43–62.
- Matsuda, N. (2017). Evidence of the effects of text-to-speech synthetic speech to improve second language learning, *JACET Journal*, 61, 149–164.
- Oxford University Press., University of Cambridge., & Association of Language Testers in Europe. (2001). *Quick placement test*. Oxford, England: Oxford University Press.
- Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, 12, 339–349. doi:10.1111/j.1467-7687.2008.00786.x
- Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, 15, 608–635. doi:10.1111/j.1532-7078.2010.00033.x
- Sommers, M. S., Nygaard, L. C., & Pisoni, D. B. (1994). Stimulus variability and spoken word recognition. I. Effects of variability in speaking rate and overall amplitude. *The Journal of the Acoustical Society of America*, 96(3), 1314–1324. doi:10.1121/1.411453
- Sugiura, K., & Hori, T. (2012). Auditory priming effect in Japanese learners of English: An investigation using a repetition task of spoken words. *KATE Journal*, 26, 39–51.
- Trofimovich, P. (2005). Spoken-word processing in native and second languages: An investigation of auditory word priming. *Applied Psycholinguistics*, 26, 479–504. doi: 10.1017.S0142716405050265

- Trofimovich, P. (2008). What do second language listeners know about spoken words? Effects of experience and attention in spoken word processing. *Journal of Psycholinguistic Research*, 37, 309–329. doi: 10.1007/s10936-008-9069-z
- Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, 19, 301–306. doi:10.1126/science.2296719
- Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 336–342. doi:10.1037/0278-7393.8.4.336
- Yamamoto, H. & Haryu, E. (2016). *Youji no tangogakushu ni okeru akusento patan riyō no hattatsu katei* [Japanese children's learning of homophones with different accentual patterns]. *Ninchi Kagaku* [Cognitive Studies], 23(1), 22–36.
- Yokokawa, H., Ikemura, D., Kadota, S., Sato, H., Shimamoto, T., Sugiura, K., Tanimura, M., ...Yoshimura, M. (2009). *Kyōikukenkyū no tameno daiginengo detabesu: Nihonjin eigogakushusha no eitangoshinmitsudo onseihen* [The data base of a second language for teaching and research: Spoken English word familiarity of Japanese EFL learners]. Tokyo, Japan: Kuroshioshuppan.
- Yokokawa, H., Ikemura, D., Kadota, S., Sakai, M., Shimamoto, T., Tanimura, M., Nakanishi, Y., ...Yoshimura, M. (2006). *Kyōikukenkyū no tameno daiginengo detabesu: Nihonjin eigogakushusha no eitangoshinmitsudo moji-hen* [The data base of a second language for teaching and research: Written English word familiarity of Japanese EFL learners]. Tokyo, Japan: Kuroshioshuppan.

Appendix A

Words for Repetition Task after Rhyme Judgment Task (Same Voice)

Word	Voice (Study)	Voice (Test)	Familiarity (Spoken)	Familiarity (Written)	Frequency (BNC)	Syllable Number	Duration
Repeated Word							
contrast	Female	Female	5.49	5.11	8172	2	857.78
energy	Female	Female	5.94	5.42	13083	3	671.31
husband	Female	Female	5.60	5.08	12263	2	634.01
influence	Female	Female	5.21	5.23	15130	3	790.65
myself	Female	Female	6.08	5.91	12444	2	611.64
open	Female	Female	6.13	6.77	46095	2	566.88
remember	Female	Female	6.79	6.07	26748	3	663.85
sentence	Female	Female	5.54	5.89	10127	2	880.16
water	Female	Female	6.70	6.77	35767	2	604.18
addition	Male	Male	5.97	4.70	10664	3	693.69
character	Male	Male	5.34	6.21	12511	3	677.27
design	Male	Male	5.27	5.75	26375	2	708.60
feature	Male	Male	6.37	5.28	17219	2	559.42
however	Male	Male	6.25	5.73	60498	3	589.26
listen	Male	Male	6.34	6.17	12080	2	618.66
present	Male	Male	6.40	5.85	36806	2	686.23
probably	Male	Male	6.05	5.63	27303	3	716.06
recent	Male	Male	5.79	5.25	15812	2	648.93
Average			5.96	5.71	22172.06	2.39	676.59
Unrepeated Word							
address	Female	Female	5.09	6.41	11984	2	745.90
basic	Female	Female	6.04	6.05	10860	2	604.18
career	Female	Female	5.36	5.00	9441	2	723.50
difference	Female	Female	5.15	6.48	19138	3	790.65
island	Female	Female	6.45	5.85	7649	2	604.18
movement	Female	Female	5.94	5.05	17880	2	708.60
October	Female	Female	5.61	5.88	10600	3	775.73
party	Female	Female	6.51	6.48	52979	2	566.88
welcome	Female	Female	6.67	6.30	9570	2	675.59
anyone	Male	Male	6.30	5.68	14956	3	555.22
area	Male	Male	6.17	6.00	58449	3	589.26
couple	Male	Male	5.64	5.54	15330	2	522.13
easy	Male	Male	6.57	6.80	21480	2	512.68
expression	Male	Male	5.18	4.80	8756	3	835.41
money	Male	Male	6.72	6.61	37892	2	560.05
restaurant	Male	Male	6.82	5.90	5100	3	709.83
Sunday	Male	Male	6.51	6.53	10100	2	753.36
suppose	Male	Male	5.63	5.25	14482	2	835.41
Average			6.02	5.92	18702.56	2.33	670.47

Appendix B

Words for Repetition Task after Synonym Judgment Task (Same Voice)

Word	Voice (Study)	Voice (Test)	Familiarity (Spoken)	Familiarity (Written)	Frequency (BNC)	Syllable Number	Duration
Repeated Word							
control	Female	Female	6.13	6.49	38281	2	673.37
decide	Female	Female	5.10	5.31	24380	2	694.41
father	Female	Female	6.33	6.29	23216	2	652.33
history	Female	Female	6.11	6.13	20064	3	596.21
measure	Female	Female	6.07	4.48	17443	2	645.31
official	Female	Female	5.49	6.22	15931	3	638.30
paper	Female	Female	6.71	6.50	23694	2	575.17
response	Female	Female	5.20	5.08	14627	2	848.73
window	Female	Female	6.78	6.18	19340	2	631.29
computer	Male	Male	6.40	6.80	16976	3	687.40
concept	Male	Male	5.93	5.18	9093	2	743.51
ever	Male	Male	6.55	6.16	27195	2	512.04
exercise	Male	Male	6.52	5.78	12721	3	960.96
image	Male	Male	6.53	6.54	11024	2	603.23
over	Male	Male	5.73	6.55	135170	2	568.16
police	Male	Male	5.97	6.27	27508	2	708.44
possible	Male	Male	5.59	6.27	34178	3	687.40
realize	Male	Male	6.05	5.50	15575	3	813.66
Average			6.07	5.99	27023.11	2.33	680.00
Unrepeated Word							
accident	Female	Female	6.36	6.02	8374	3	736.50
brother	Female	Female	5.87	6.15	11757	2	540.10
council	Female	Female	5.78	3.07	34496	2	715.46
doctor	Female	Female	5.62	6.38	13684	2	610.24
even	Female	Female	5.61	6.14	90473	2	568.16
happen	Female	Female	6.08	6.02	32075	2	610.24
relation	Female	Female	5.82	5.17	19628	3	764.56
sample	Female	Female	5.16	5.92	8182	2	715.46
worry	Female	Female	6.19	6.16	9006	2	589.20
already	Male	Male	6.24	5.73	34292	3	785.60
anything	Male	Male	6.62	6.31	28321	3	701.43
example	Male	Male	6.87	6.11	43402	3	694.41
important	Male	Male	5.56	6.67	39265	3	708.44
little	Male	Male	5.70	6.47	63383	2	512.04
machine	Male	Male	6.09	6.37	13518	2	715.46
real	Male	Male	5.27	6.15	22982	2	596.21
suggest	Male	Male	5.71	5.30	28665	2	932.90
summer	Male	Male	6.71	6.73	11563	2	603.23
Average			5.96	5.94	28503.67	2.33	672.20

Appendix C

Words for Repetition Task after Rhyme Judgment Task (Different Voice)

Word	Voice (Study)	Voice (Test)	Familiarity (Spoken)	Familiarity (Written)	Frequency (BNC)	Syllable Number	Duration
Repeated Word							
copy	Female	Male	6.59	5.79	10562	2	589.94
difficult	Female	Male	6.21	6.63	22033	3	714.87
either	Female	Male	5.16	5.39	27766	2	548.30
into	Female	Male	5.45	5.82	163469	2	610.76
mother	Female	Male	6.60	6.60	27784	2	610.76
opinion	Female	Male	5.21	5.58	9213	3	694.05
reduce	Female	Male	5.32	4.61	17226	2	687.11
science	Female	Male	5.93	5.74	12644	2	812.03
wonder	Female	Male	6.17	5.87	14375	2	596.88
collection	Male	Female	6.67	5.91	9639	3	749.57
cover	Male	Female	6.11	5.75	24698	2	513.59
depend	Male	Female	5.65	5.34	10125	2	652.40
exactly	Male	Female	6.46	5.43	10729	3	791.21
heavy	Male	Female	5.11	5.92	10439	2	471.95
language	Male	Female	6.29	6.35	22117	2	909.20
okay	Male	Female	6.40	4.98	12190	2	569.12
position	Male	Female	5.48	6.27	28071	3	659.34
arrangement	Male	Female	5.25	4.24	9054	3	867.56
Average			5.89	5.68	24563.00	2.33	669.37
Unrepeated Word							
animal	Female	Male	6.16	6.25	15250	3	617.70
attention	Female	Male	6.34	5.57	13968	3	784.27
body	Female	Male	5.00	6.35	32231	2	645.46
finish	Female	Male	6.77	6.53	13902	2	596.88
hospital	Female	Male	5.91	6.15	16898	3	728.75
market	Female	Male	5.45	6.11	36905	2	596.88
practice	Female	Male	6.25	5.92	24019	2	784.27
reaction	Female	Male	5.73	5.76	7565	3	839.80
support	Female	Male	5.87	6.13	40248	2	714.87
culture	Male	Female	6.07	6.55	10196	2	666.29
everything	Male	Female	6.27	6.44	18675	3	742.63
imagine	Male	Female	6.64	5.65	8300	3	687.11
morning	Male	Female	6.68	6.55	21845	2	645.46
prevent	Male	Female	5.64	4.66	10286	2	638.52
radio	Male	Female	6.50	6.38	9066	3	756.51
ready	Male	Female	6.31	5.89	10110	2	569.12
surface	Male	Female	5.11	4.79	10361	2	749.57
worker	Male	Female	5.23	5.89	18247	2	596.88
Average			6.00	5.98	17670.67	2.39	686.72

Appendix D

Words for Repetition Task after Synonym Judgment Task (Different Voice)

Word	Voice (Study)	Voice (Test)	Familiarity (Spoken)	Familiarity (Written)	Frequency (BNC)	Syllable Number	Duration
Repeated Word							
advantage	Female	Male	5.76	5.40	10285	3	877.13
degree	Female	Male	5.22	4.43	12996	2	604.25
happy	Female	Male	6.42	6.75	12854	2	545.77
holiday	Female	Male	6.30	6.50	9731	3	656.23
interview	Female	Male	5.38	6.30	9008	3	708.20
major	Female	Male	6.04	6.23	23629	2	643.23
message	Female	Male	6.47	6.58	8938	2	773.18
subject	Female	Male	5.03	6.25	32392	2	708.20
woman	Female	Male	5.05	6.45	63087	2	578.26
another	Male	Female	5.38	6.40	60182	3	643.23
argument	Male	Female	5.35	3.98	12125	3	688.71
better	Male	Female	6.27	6.47	15626	2	558.77
country	Male	Female	6.37	6.35	48177	2	649.73
document	Male	Female	6.54	5.29	10498	3	773.18
include	Male	Female	5.50	5.37	34753	2	727.70
occasion	Male	Female	5.75	4.22	9152	3	734.19
only	Male	Female	6.10	6.62	152903	2	558.77
supply	Male	Female	6.20	5.50	16892	2	695.21
Average			5.84	5.84	30179.33	2.39	673.55
Unrepeated Word							
company	Female	Male	5.20	6.30	57754	3	597.75
contact	Female	Male	5.46	5.84	13867	2	753.68
county	Female	Male	5.36	4.52	9745	2	565.26
encourage	Female	Male	5.63	4.98	10664	3	930.45
leader	Female	Male	6.30	5.63	15903	2	623.74
maybe	Female	Male	5.70	5.67	10472	2	636.73
production	Female	Male	6.00	5.87	15837	3	701.71
receive	Female	Male	5.86	4.96	24111	2	688.71
research	Female	Male	6.30	5.72	27663	2	714.70
everyone	Male	Female	6.71	6.15	13337	3	695.21
follow	Male	Female	5.60	5.46	46145	2	558.77
music	Male	Female	6.98	6.88	15024	2	552.27
prepare	Male	Female	5.20	4.86	14961	2	636.73
program	Male	Female	6.48	6.14	32068	2	753.68
recognize	Male	Female	5.57	4.35	15203	3	892.21
recover	Male	Female	5.99	5.10	4932	3	662.72
station	Male	Female	6.31	6.75	12328	2	721.20
weather	Male	Female	6.54	5.92	42042	2	493.79
Average			5.96	5.62	21225.33	2.33	676.63