

The recognition of Japanese-accented and unaccented English words by Japanese listeners

Kiyoko Yoneyama

Department of English
Daito Bunka University
yoneyama@ic.daito.ac.jp

Abstract

This study investigated whether Japanese listeners learning English employ two types of lexical information (word frequency and neighborhood density) when they recognize English words. English words recorded by a native speaker of English and a native speaker of Japanese were presented to Japanese university students in a noise condition. The results of word recognition scores showed that Japanese listeners employed both lexical and pre-lexical levels of information in English word recognition. They were sensitive to both probabilistic phonotactics (bottom-up acoustic information) and word frequency (lexical information). A strong correlation between probabilistic phonotactics and neighborhood density still predict Japanese listeners are influenced by neighborhood density in English word recognition.

1. Introduction

Japanese listeners are influenced by at least two types of lexical information simultaneously in Japanese word recognition: word frequency and neighborhood density [9]. These two types of lexical information are also employed by English listeners [5]. Therefore, there might be a possibility that effects of neighborhood density and lexical frequency are universal across languages.

Interestingly, non-native listeners are also affected by word frequency and neighborhood density when they recognize English. Easy English words (high frequency words of low neighborhood density) were recognized more accurately than difficult English words (low frequency words of high neighborhood density) by non-native listeners [1]. Imai, Flege and Walley found that native Spanish listeners showed a larger neighborhood density effect for native English words than Spanish-accented English words, whereas native English listeners showed a larger neighborhood density effect for Spanish-accented than native words [2][3].

Based on the previous studies of neighborhood density by non-native listeners ([1][2][3]), we hypothesize that Japanese listeners would similarly employ two types of lexical information not only in Japanese, but also in English. In order to test this hypothesis, an English word recognition experiment was conducted with Japanese university students who have been learning English only through a Japanese education system. As in the studies of Imai, Flege and Walley ([2][3]), two types of English stimuli (native-produced stimuli and Japanese-accented stimuli) were prepared.

Our prediction is as follows. If the effects of neighborhood density and lexical frequency are both universal, the Japanese listeners would show these effects even in English. Further, if they are also sensitive to the acoustic

properties of the stimuli, we would expect that the effects would vary depending on the types of stimuli.

2. Experiment

2.1. Methods

2.1.1. Materials

Eighty words were selected for the experiment. These words were familiar to Japanese college students. All words were considered as familiar words by American college students ([8]). Four sets of 20 words that differed orthogonally in text word frequency ([4]) and neighborhood density ([5]) were selected. As shown in Table 1, the four sets are as follows: (a) high-frequency words from dense neighborhoods; (b) high-frequency words from sparse neighborhoods; (c) low-frequency words from dense neighborhoods; (d) low-frequency words from sparse neighborhoods.

Table 1: *Mean word frequency, neighborhood density and word familiarity of test words. Note: WF: word frequency, or occurrence per million words (Kučera and Francis, 1967); ND: neighborhood density, or number of words that differ by a one segment, addition, deletion or substitution (Luce and Pisoni, 1998).*

ND	High WF		Low WF	
	Dense	Sparse	Dense	Sparse
Example	“wait”	“month”	“pen”	“gift”
WF	3.29	3.28	2.74	2.23
ND	26.15	4.25	26.3	4.1
Fam	6.96	6.96	6.99	6.91

The 80 test words were recorded by a male speaker of American English who was born and raised in the Midwest but had lived in Japan for more than 10 years. The stimuli produced by this speaker are referred to as the “native” stimuli. The test words were also recorded by a male junior at Daito Bunka University who was born and raised in the Tokyo area and was judged by the author to speak a Japanese-accented English. The stimuli produced by this speaker are referred to as the “Japanese-accented” stimuli.

The 160 stimuli produced by the two speakers were first digitized at 22.05 kHz and normalized for peak root-mean-square (RMS) amplitude values.

The 160 stimuli (80 test words x 2 speakers) were mixed with a white noise to bring word recognition scores off ceiling. The yielding stimuli having S/N ratios of approximately 0 dB.

The stimulus presentation was controlled by two conditions. First, participants hear 80 test words only one time during the test session. Second, they hear native and

Japanese-accented stimuli equally during the test session. In order to fulfill these conditions, the stimuli were presented as follows.

Two blocks of 40 test words, designated “A” and “B” (with 10 items randomly selected from each of the four lexical sets where neighborhood density and word frequency varied), were formed. The order of words was randomized within each block. Blocks A and B were both presented as native and Japanese-accented stimuli, and the presentation order of Stimulus Type (native vs. Japanese-accented) was counterbalanced. Participants were randomly assigned to one of the four counterbalanced conditions as in Table 2.

Table 2: Four counterbalanced conditions for the stimulus presentations.

	Native Stimuli	Japanese-accented Stimuli
Condition 1	Block A	Block B
Condition 2	Block B	Block A
Condition 3	Block B	Block A
Condition 4	Block A	Block B

2.1.2. Procedures

Participants of each condition were tested as a group in a language laboratory. They completed two primary tasks in the following order: a word recognition task and a written lexical knowledge test.

In the word recognition task, the test words were presented via loudspeakers and were asked to write down each word on a prepared answer sheet. They were given a short demonstration of the task, and a 10-item practice session preceded each of the two blocks. None of the practice words was among the 80 words. The participants were also encouraged to write their responses very clearly.

In the written lexical knowledge test, participants were asked to indicate how well they know about the words. They were given 160 words, half of them were test words and half of them were not. They were instructed to tell which words they learn at school. Five words must be the foils so that all the participants had to report “unknown” for these words.

The written responses by participants in the word recognition task were scored “correct” when they exactly match the target words. However, we anticipated that the Japanese participants might misspell some words. Given that our interest was in spoke word recognition rather than in spelling ability, the experimenter examined each participant’s written responses at the end of the test session.

2.1.3. Participants

Participants were 44 freshmen at Daito Bunka University (Tokyo, Japan). They were Japanese native speakers who had never stayed in English-speaking countries except for short travel visits. None of the participants had any hearing impairment. All the participants knew all test words and correctly reported five foils as “unknown” words.

2.2. Results

Percent correct scores for the 10 items in each of the eight conditions of the stimulus design were obtained for each participant and submitted to a 2 (Stimulus Type: native, Japanese-accented) x 2 (Word Frequency: high, low) x 2

(Neighborhood Density: dense neighborhood, sparse neighborhood) ANOVA.

The ANOVA yielded significant main effects of Stimulus Type, Neighborhood Density and Word Frequency. Native stimuli were recognized more accurately than Japanese-accented stimuli ($F(1, 43) = 103.468, p < 0.001$). The words from dense neighborhood were recognized more accurately than the words from sparse neighborhood ($F(1, 43) = 27.591, p < 0.001$). High frequency words were recognized more accurately than low frequency words ($F(1, 43) = 94.884, p < 0.001$).

The two-way interaction of Stimulus Type and Lexical Frequency was significant ($F(1, 43) = 59.481, p < 0.001$). In native stimuli, high frequency words were more recognized than low frequency words (0.778 vs. 0.549), whereas in Japanese-accented stimuli, they were equally accurately recognized (0.458 vs. 0.423). The two-way interaction of Lexical Frequency and Neighborhood Density was also significant ($F(1, 43) = 7.708, p < 0.01$). In low frequency words, words from dense neighborhood were recognized more accurately than words from sparse neighborhood (0.73 vs. 0.625), whereas in high frequency words, they were equally accurately recognized (0.466 vs. 0.415).

The three-way interaction of Stimulus Type, Lexical Frequency and Neighborhood Density was significant ($F(1, 43) = 10.548, p < 0.01$). Further analyses were conducted for each stimulus type. As shown in Figure 1, in native words, words from dense neighborhood were recognized more accurately than words from sparse neighborhood ($F(1, 43) = 15.617, p < 0.001$). High frequency words were recognized more accurately than low frequency words ($F(1, 43) = 175.95, p < 0.001$). Further, in each word frequency condition, words from dense neighborhood were always recognized more accurately than words from sparse neighborhood (High frequency words: 0.818 and 0.739, $t(1, 43) = 2.872, p < 0.01$; Low frequency words: 0.589 and 0.509, $t(1, 43) = 2.912, p < 0.01$).

As Figure 2 shows, in Japanese-accented words, words from dense neighborhood were recognized more accurately than words from sparse neighborhood ($F(1, 43) = 4.985, p < 0.05$). An interaction between Word Frequency and Neighborhood Density was also significant ($F(1, 43) = 15.654, p < 0.001$). A further analysis revealed that only in low frequency words, words from dense neighborhood were recognized more accurately than words from sparse neighborhood (0.491 and 0.355, $t(1, 43) = 4.574, p < 0.001$).

3. Discussion

The two types of lexical information, word frequency and neighborhood density, simultaneously influence on Japanese word recognition [9]. The results of this experiment provided new findings regarding how Japanese listeners as second-language learners of English process word recognition in English. The results showed that Japanese listeners employ both bottom-up and top-down information when they process English words.

Japanese listeners were influenced by top-down lexical information when they listened to English. Word frequency exerted an important influence on spoken word recognition. As shown in Japanese [9], high frequency words were recognized more accurately than low frequency words in English.

Japanese listeners were also sensitive to bottom-up

acoustic information. Acoustic properties of words influenced on listeners' English word recognition. Especially, Japanese listeners had a difficulty listening to Japanese-accented English words. They performed better for native English words than Japanese-accented English words.

This interpretation is based on the fact that neighborhood density and probabilistic phonotactics are highly correlated: Words from dense neighborhoods generally consist of frequent sounds that yield higher transitional probabilities. The recent studies have shown that transitional probabilities help listeners to access to the lexicon in spoken word recognition [6]. Vitevich and Luce have claimed, a positive effect of neighborhood density is interpreted as an effect of probabilistic phonotactics at a pre-lexical level whereas a negative effect of Neighborhood Density is interpreted as an effect of neighborhood density at a lexical level [7]. A positive correlation between recognition scores and Neighborhood Density may indicate that Japanese listeners may have processed words not at lexical level but at pre-lexical level when they were performing the task.

Why did they perform the task at a pre-lexical level, not at a lexical level? One possible explanation is that the Japanese participants might have needed to focus more on the acoustic information in order to understand the English stimuli presented at a noisy condition. Generally, listening to a foreign language is not an easy task for the second-language learners. This is even harder for those whose second-language communicative abilities are limited. In fact, the Japanese participants in this study were university freshmen whose English experience is not sufficient. Their sensitivity to acoustic information might have been indispensable in order to compensate their lack of sufficient English experience. In order to perform the task better, they needed to focus more on bottom-up acoustic information in a harder condition, such as in a noisy environment.

If this explanation is correct, the Japanese listeners should have needed more bottom-up information for Japanese-accented words in a harder hearing condition. In other words, as the results of the correct word recognition task showed, they should have relied on acoustic information more for Japanese-accented stimuli than for unaccented stimuli.

A comparison of the magnitude differences between the dense neighborhood density in each word frequency condition for the two stimulus types is one way to measure how greatly the acoustic information contributed the correct word recognition. As shown in Figures 1 and 2, three out of four conditions showed that the value of the correct word scores was greater for the words from dense neighborhoods than from sparse neighborhoods. Interestingly, a magnitude difference between the neighborhood density conditions in the low-frequency condition for Japanese-accented words (0.136) is greater than the ones in the both word frequency conditions for native words (high frequency words: 0.079; low frequency words: 0.08). This may indicate that sound probabilistic information at a pre-lexical level was greatly needed in order for the Japanese listeners to perceive sounds correctly during their second-language word recognition. The higher correct word scores for the native stimuli than for the Japanese-accented stimuli might also be a reflection of the Japanese listeners' sensitivity to the salience of the stimuli: the native stimuli were significantly longer than the Japanese-accented stimuli.

Let us briefly consider the possibility of the neighborhood density effect with the Japanese listeners in English word

recognition. As mentioned before, the Japanese listeners desperately needed the acoustic information of the stimuli while they performed the task: the sound probabilistic information helped the listeners to recognize the words. Since transitional probability as an effect at a pre-lexical level is highly correlated with neighborhood density as an effect at a lexical level, there is a high possibility that the Japanese listeners' data show a neighborhood density effect during the English word recognition.

The data of the Japanese-accented high-frequency stimuli might show that Japanese listeners might have been affected by neighborhood density, not by probabilistic phonotactics: Neighborhood density information hindered the recognition of English words. If the Japanese listeners had performed the task in this condition similarly to the other three conditions, the correct word recognition scores would be always higher for the words from dense neighborhoods than for the words from sparse neighborhoods. The Japanese listeners might have needed less acoustic information in this condition than in other three conditions because all words were frequent words so that neighborhood density was in effect. we further need to investigate whether neighborhood density affects Japanese listeners' recognition of English words.

4. Conclusion

This study investigated whether two types of lexical information, word frequency and neighborhood density, are involved in English word recognition by Japanese listeners. The results indicated a word frequency effect, but not a neighborhood density effect. Rather the results showed that Japanese listeners were sensitive to a pre-lexical level of word processing and employed phonotactic probabilities from the bottom-up acoustic information. Because phonotactic probabilities and neighborhood density are highly correlated, there is still a possibility that a neighborhood density effect would be observed in a course of English word recognition by Japanese native listeners if they focus more at a lexical level when they perform a task. Further studies are needed.

5. References

- [1] Bradlow, A.R.; Pisoni, D.B. 1999. Recognition of spoken words by native and non-native listeners: Talker-, listener-, item-related factors. *Journal of the Acoustical Society of America*, 106(4), 2074-2085.
- [2] Imai, S.; Flege, J.E.; Walley, A. 2003. Spoken word recognition of accented and unaccented speech: Lexical factors affecting native and non-native listeners. *Proceedings of International Congress on Phonetic Science 2003*, 845-848.
- [3] Imai, S.; Flege, J.E.; Walley, A. 2005. The recognition of foreign-accented and unaccented English words by native and non-native listeners. *Journal of Acoustical Society of America*, 117(2), 896-907.
- [4] Kučera, F.; Francis, W. 1967. *Computational Analysis of Present Day American English*. Brown University Press: Providence.
- [5] Luce, P.A.; Pisoni, D.B. 1998. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36.
- [6] McQueen, J.M. 1998. Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, 39, 21-46.

- [7] Vitevitch, M.S.; Luce, P.A. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374-408.
- [8] Washington University in St. Louis Speech and Hearing Lab Neighborhood Database as a web-based implementation of the 20,000-word Hoosier Mental Lexicon (2004, January; accessed). <http://128.252.27.56/neighborhood/Home.asp>
- [9] Yoneyama, K. 2002. *Phonological neighborhood and phonetic similarity in spoken word recognition in*

Japanese. Unpublished doctoral dissertation, Department of Linguistics, The Ohio State University.

6. Acknowledgments

This research is supported by Grant-in Aid for Scientific Research (Grant no. 16720092). An earlier version of the paper was presented at the 147th meeting of the Acoustical Society of America (New York, May 24 and May 28, 2004). All errors are, of course, mine.

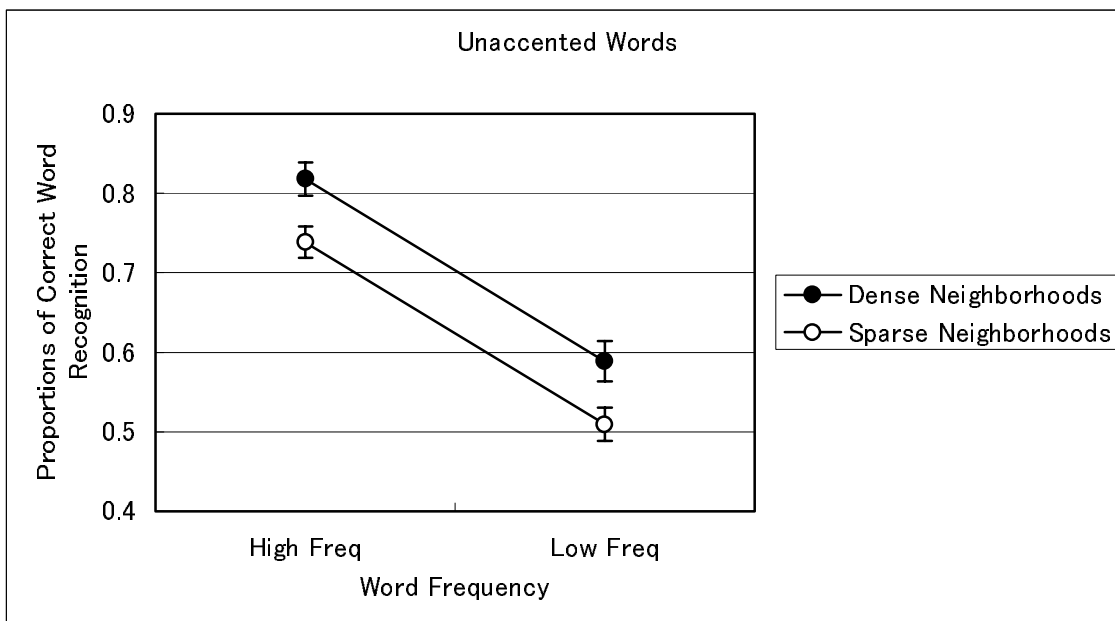


Figure 1: Mean proportions of correct word recognition scores for unaccented stimuli as a function of word frequency and neighborhood density.

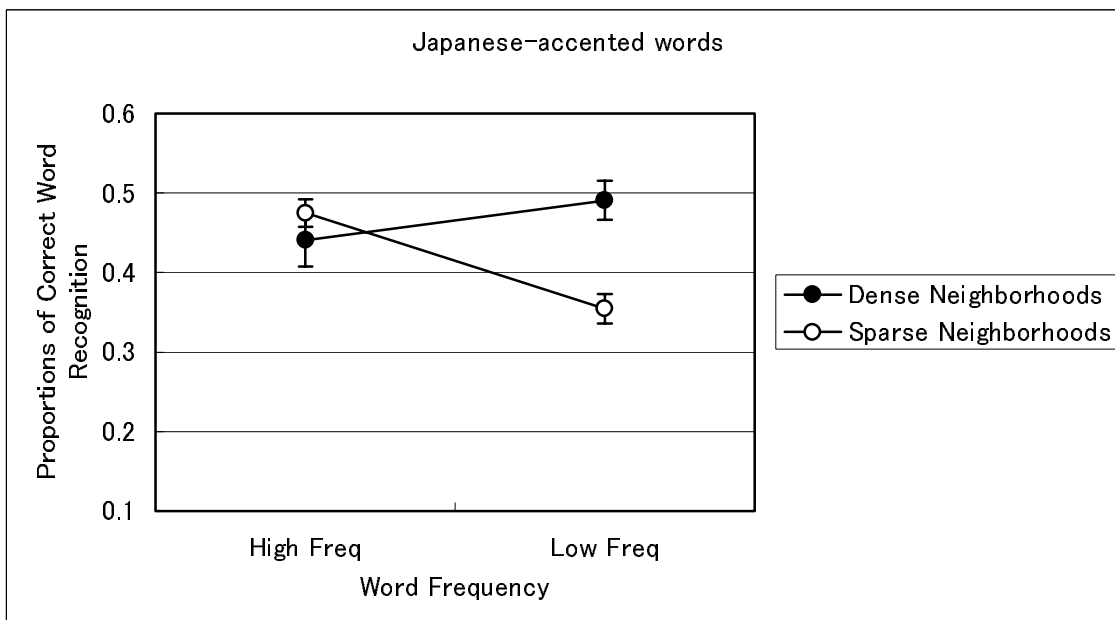


Figure 2: Mean Proportions of correct word recognition scores for Japanese-accented stimuli as a function of word frequency and neighborhood density.