FAMILIARITY EFFECT ON SPOKEN WORD RECOGNITION IN JAPANESE

Shigeaki Amano, Tadahisa Kondo, and Kazumi Kato
NTT Communication Science Laboratories

ABSTRACT
Word familiarity effects on lexical decision times and recognition scores were investigated for two sets of 100 spoken Japanese words whose familiarity ratings ranged from middle (4.0) to high (7.0) taken from a recently-developed word familiarity database containing about 80,000 Japanese words. The results showed that the higher the word familiarity rating is, the shorter the reaction time is, and that the higher the word familiarity is, the higher the recognition score is at all signal-to-noise ratios of -5.0, -2.5, 0.0, and 2.5 dB, and also at no noise. The results showed that the word familiarity has a strong effect on spoken word recognition in Japanese as seen in English.

1. INTRODUCTION
Word familiarity is a well-known factor which affects spoken word recognition. For example, Connine, Mullennix, Shernoff, and Yelen [1] showed that high-familiarity spoken words in English have shorter reaction times than low-familiarity spoken words in a lexical decision task and a delayed naming task. They also showed that the high-familiarity spoken words are recognized better than the low-familiarity spoken words in these tasks. Their results clearly indicate that the word familiarity affects speed and accuracy of spoken word recognition. Their results support the idea that controlling the word familiarity is necessary, when choosing experimental stimuli in psycholinguistic experiments.

Hitherto, there has been no study which has shown the word familiarity effect on spoken word recognition of Japanese. This is mainly because there was no word familiarity database which is large and reliable enough for such psycholinguistic studies. Furthermore, lack of the database has caused great difficulties in controlling the word familiarity in psycholinguistic experiments in Japanese.

Recently, Amano and Kondo [2] developed a series of databases (PSYLEX) for about 80,000 Japanese words. One of them is a word familiarity database [3] which contains audio, visual, and audio-visual familiarity ratings (1: familiar - 7: familiar) given by 32 subjects. By using spoken words which were selected from the word familiarity database as stimuli, this study investigated the effects of the word familiarity on spoken word recognition in Japanese. Specifically, the effects was studied on lexical decision times and recognition scores of Japanese words.

2. EXPERIMENT 1
2.1. Objectives
The objective of Experiment 1 is to investigate the relationships between word familiarity and reaction times of lexical decision for Japanese words.

2.2. Method
2.2.1. Subjects.
All possible Japanese participants undertook the kanji-word reading ability test (100 Rakan) which was developed by Kondo and Amano [4], and only those 30 (15 males and 15 females) who scored from 70 to 80% accuracy participated in the experiment. This screening was to ensure that all the participants had similar language competence which is above average. The participants were 18 to 30 years old and they had no hearing impairment.

2.2.2. Stimuli.
Two hundred words were selected from the word familiarity database in PSYLEX [2], according to the following conditions.

W1: Auditory word familiarity is ranged 4.0-4.5, 4.5-5.0, 5.0-5.5, 5.5-6.0, and 6.0-7.0.
W2: Word length is 4 mora.
W3: Accent pattern is low-high-high-high.
W4: No homophone exists.
W5: Difference between auditory and visual word familiarity is less than 0.15.
W6: No homograph exists.
W7: Orthography consists of two kanji characters.

The W5, W6, and W7 were extra conditions for future visual experiments using the same words as the current auditory experiment. Two word sets of 100 words were made by collecting 20 words in each word familiarity range. Nonwords consisted of very low-familiarity words which were selected from the word familiarity database with the following conditions.

N1: Auditory word familiarity is ranged 1.0-2.0.
N2: The conditions of W2 to W7.

The N1 condition assures that the stimuli are not recognized as words and that they do not violate Japanese phonotactic rules (e.g., transitional probability of mora). One nonword set was made by collecting 100 nonwords.

Two stimulus sets of 200 stimuli were made by combining the word set 1 and the nonword set (hereafter called stimulus set 1), and the word set 2 and the nonword set (hereafter called stimulus set 2). Another 10 words and 10 nonwords were selected for practice trials from the word familiarity database according to the same conditions described above.

Speech files of these stimuli were extracted from PSYLEX [2]. They were digitally recorded files (16 bit - 16 kHz sampling rate) pronounced by one Japanese female speaker. All stimuli were digitally adjusted to have the same root-mean-square power.

2.2.3. Procedure.
The stimuli were diotically presented to the subjects through headphones at 74 dB SPL. In each trial, a warning
signal (1 kHz pure tone, 150 ms) was presented first. One second after it, a word or a nonword was presented. The subjects were instructed to push one key with their first finger and another key with their second finger as soon as possible when they thought that they heard a word or a nonword respectively. The subjects were also instructed to try to avoid errors as much as possible. The interval between successive trials was about two seconds. Reaction times were measured from word/nonword stimulus onset.

A half of the subjects heard the stimulus set 1, the other half heard stimulus set 2. Each subject heard 200 experimental trials after three sets of 20 practice trials. There was a short break after every 50 trials. The order of the stimuli was randomized for each subject.

2.3. Results
Responses were determined as errors when one of the following conditions was satisfied.

E1: wrong key pressing
E2: reaction times which is more than 1600 ms
E3: reaction times which is less than 160 ms

The data from the two subjects were excluded from the analyses because their error rates were greater than 25%. Otherwise, the erroneous responses were excluded from the reaction time analyses.

Mean reaction times for lexical decision were plotted against the auditory word familiarity in Figure 1. A two-way analysis of variance was conducted on the reaction times. The main effect of the word familiarity was significant \( [F(4,2367) = 141.2, p < .001] \). The main effect of the stimulus set was not significant. The HSD test revealed that paired comparisons over different word familiarity ranges were all significant \( (p < .05) \).

Error rates for lexical decision were plotted against the auditory word familiarity in Figure 2. A two-way analysis of variance was also conducted on the error rates. The main effects of the word familiarity was significant \( [F(4,134) = 83.8, p < .001] \). The main effect of the stimulus set was not significant. The HSD test revealed that paired comparisons over all different word familiarity ranges were all significant \( (p < .05) \), except for the paired comparison between “5.5-6.0” and “6.0-7.0”.

2.4. Discussion
Figure 1 and the analysis of variance showed a significant effect of word familiarity: the higher the word familiarity is, the shorter the reaction time is. However, it might be argued that this effect of the word familiarity could be an artifact of word duration. That is, the shorter the duration is, the shorter the reaction time is. However, this could not be the case because there was no significant difference in the word duration among different word familiarity ranges.
It might be argued, however, that the effect could be caused by speed-accuracy trade off. That is, the higher the error rate is, the shorter the reaction time is. However, this could not be the case, because Figure 1 and 2 showed a completely opposite relationship between speed and accuracy. That is, the higher the word familiarity is, the smaller the error rate is as shown in Figure 2, and the shorter the reaction time is as shown in Figure 1.

Taken these facts together, it can be argued that the differences shown in Figure 1 were truly due to the word familiarity, and not to other factors. Therefore, the results indicated that the word familiarity has strong effects both on the reaction times and error rates of lexical decisions in Japanese. That is, the higher the word familiarity is, the faster the lexical decision is. Similarly, the higher the word familiarity is, the more accurate the lexical decision is. These two findings coincide with those of Connine et al.’s study [1]. Therefore, it is concluded that word familiarity has the same effect on lexical decisions in Japanese as in English.

3. EXPERIMENT 2

3.1. Objectives
Experiment 2 was conducted to investigate the relationships between word familiarity and recognition score of a word with or without noise. It was thought that this experiment should add a further dimension to the familiarity effects on spoken word recognition in a ‘non-speeded’ task rather than a ‘speeded’ task in Experiment 1.

3.2. Method
3.2.1. Subjects. Forty Japanese adults (20 males and 20 females) aged between 18 and 30 years participated in the experiment. They scored from 70 to 80% accuracy in the kanji-word reading ability test (100 Rakan) [4]. They had no hearing impairment.

3.2.2. Stimuli. The same 200 words in Experiment 1 and 100 Japanese moras were used as stimuli in Experiment 2. The word stimuli were prepared by adding or not adding the Hoth noise to the 200 words. The Hoth noise had a spectral shape with about 5dB/oct decreasing in higher frequency. Frequency range of the noise was between 80 Hz and 8 kHz. The noise started 100 ms before the onset of a word stimulus, and lasted about 100 ms to 200 ms after the end of the word stimulus. The onset of the noise was tapered for 20 ms to prevent an audible click. There were five signal-to-noise ratios (-5.0, -2.5, 0.0, and 2.5 dB, or no noise which is hereafter called ‘org’). The 10 word stimuli used for the practice trial in Experiment 1 were also prepared in the same noise-adding operation for a practice trial in Experiment 2.

The 100 moras were prepared by the same noise-adding operation as the word stimuli. The 100 moras were pronounced by the same female speaker as the word stimuli and they were digitally recorded at 16 bit - 16 kHz sampling rate. All word and mora stimuli were digitally adjusted to have the same root-mean-square power after the noise-adding operation. All the ‘org’ stimuli were also digitally adjusted to have the same root-mean-square power.

3.2.3. Procedure. The word stimuli were presented in the first session and the mora stimuli were presented in the second session. The stimuli were diotically presented to the subjects through headphones at 70 dBSPL. The subjects were instructed to type the presented word or mora in ‘hiragana’ using a computer keyboard with a ‘romaji-to-hiragana’ conversion software. They were also instructed to make a guess and type their answer even if they could not identify the word or mora.

Forty words in each word familiarity range were assigned to different signal-to-noise ratio, so that one subject did not hear a particular word more than once in different noise conditions. The assignment of these was counterbalanced across subjects. One subject heard 200 words in total. One hundred moras were divided into 5 groups and they are assigned to different signal-to-noise ratio in the same way as word stimuli. One subject heard 100 moras in total.

Each subject undertook 200 trials for the word stimuli in the first session after the 10 practice trials using practice stimuli. There was a short break after 100 trials. Each subject undertook 100 trials for the mora stimuli in the second session. The order of the stimuli was randomized for each subject.

Figure 3. Recognition score for words in various auditory word familiarity as a function of signal-to-noise ratio. Recognition score of 100 moras is also plotted.

3.3. Results
Word recognition scores were obtained by dividing the number of correct answers by the total number of answers for each spoken word at each signal-to-noise ratio. Mora recognition scores were obtained by dividing the number of correct answers by the total number of answers for each mora at each signal-to-noise ratio. The word recognition scores were averaged at each auditory word
familiarity range and each signal-to-noise ratio, and they were plotted in Figure 3. The mora recognition scores were averaged at each signal-to-noise ratio, and they were also plotted in Figure 3.

To examine effects of the word familiarity and the signal-to-noise ratio, a two-way analysis of variance (word familiarity x signal-to-noise ratio) was conducted on recognition scores of words. The interaction between the effect of the word familiarity and the effect of the signal-to-noise ratio was significant \[ F(16,975) = 24.0, p < .001 \]. The simple main effect of the word familiarity was significant at all signal-to-noise ratios (at -5.0 dB, \( F(4,195) = 13.2, p < .001 \); at -2.5 dB, \( F(4,195) = 14.2, p < .001 \); at 0.0 dB, \( F(4,195) = 10.6, p < .001 \); at 2.5 dB, \( F(4,195) = 7.70, p < .001 \); at ‘org’, \( F(4,195) = 3.54, p < .001 \)). Also the simple main effect of signal-to-noise ratio was significant at all word familiarity ranges (at 4.0-4.5, \( F(4,195) = 22.2, p < .001 \); at 4.5-5.0, \( F(4,195) = 20.1, p < .001 \); at 5.0-5.5, \( F(4,195) = 19.4, p < .001 \); at 5.5-6.0, \( F(4,195) = 20.8, p < .001 \); at 6.0-7.0, \( F(4,195) = 8.05, p < .001 \)).

To examine an effect of the signal-to-noise ratio on mora recognition, a one-way analysis of variance was conducted on mora recognition scores. The main effect of the signal-to-noise ratio was significant \[ F(4,495) = 112.0, p < .001 \].

To examine differences of recognition scores between the moras and the words in familiarity range of 4.0-4.5, their scores were combined and analyzed with a two-way analysis of variance (word-versus-mora x signal-to-noise ratio). The interaction between the effect of the word-versus-mora and the effect of the signal-to-noise ratio was significant \[ F(9,690) = 7.18, p < .001 \]. The simple main effect of the word-versus-mora was significant at all signal-to-noise ratios (at -5.0 dB, \( F(1,138) = 12.9, p < .001 \); at -2.5 dB, \( F(1,138) = 20.8, p < .001 \); at 0.0 dB, \( F(1,138) = 33.3, p < .001 \); at 2.5 dB, \( F(1,138) = 31.9, p < .001 \)) except for the ‘org’ condition.

3.4. Discussion

Significant effects of word familiarity and signal-to-noise ratio were observed as shown in Figure 3. That is, the higher the word familiarity is or the higher the signal-to-noise ratio is, the better the word recognition score is. This result clearly showed that the word familiarity affects the word recognition in a ‘non-speeded’ task.

Interestingly, the significant interaction was found between the word familiarity effect and the signal-to-noise ratio effect. It indicates that the word familiarity effect is greater at low signal-to-noise ratio than at high signal-to-noise ratio. In other words, high familiarity words are more tolerant to noise degradation than lower familiarity words.

This tolerance is also shown in the study of Sakamoto et al., Suzuki, Amano, Ozawa, Kondo, and Sone [5]. Using the word stimuli selected from PSYLEX [2], they collected word recognition scores as a function of word familiarity and sound pressure level. They found that higher familiarity words are recognized better than the lower familiarity words, and that the difference in the recognition scores between the higher and lower familiarity words is greater when they were degraded by lowering sound pressure level.

Since word familiarity is regarded as part of lexical information, the current as well as Sakamoto et al.’s results lend support to the view that subjects rely more on top-down lexical information than on bottom-up acoustic information when speech input is degraded. Although this reliance could be smaller for the words with familiarity range of 4.0-4.5 than those with higher familiarity ranges, it is still observable for the former, because they were recognized better than the moras to which no lexical information is available.

However, this advantage could also be due to several other factors. One of the factors could be information derived from phonotactic rules in Japanese. That is, transitional probability of moras within a Japanese word might be used for identifying successive moras, which result in facilitating the word recognition. While, such information is not available for mora recognition.

Another possible factor is coarticulatory information. That is, transitional acoustic cues among moras within a word could be used for identifying successive moras, which result in facilitating the word recognition. However, such information is unavailable for mora recognition because the mora was presented alone and there were no successive mora. Maybe these two factors together with the top-down lexical information caused the recognition advantage of the words over the moras.

Both results from Experiments 1 and 2 as well as Sakamoto et al.’s study show that word familiarity is a strong factor for spoken word recognition in Japanese in both a ‘speeded’ task and a ‘non-speeded’ task. Therefore, controlling the word familiarity is important in lexical decision and recognition tasks. Otherwise, the obtained data would be contaminated, and wrong conclusion might be drawn. This would be also true for other experimental tasks such as a naming task or a priming task. It is suggested that controlling the word familiarity is necessary for any psycholinguistic experiments with Japanese spoken words.

4. Conclusion

The present study investigated word familiarity effects on spoken word recognition in Japanese. It is found that the word familiarity has strong effect on lexical decision times and recognition scores of word with or without noise. That is, the higher the word familiarity is, the faster and the better the word recognition is. It is also found that the word familiarity effect is greater at low signal-to-noise ratio than at high signal-to-noise ratio. It is suggested that the word familiarity should be controlled when Japanese words are used as stimuli in psycholinguistic experiments.

REFERENCES


