Acceleration of Lexical Access Caused by Word Familiarity and by Accent Nucleus Perception in Japanese

Nobuaki MINEMATSU Seiichi NAKAGAWA

Dept. of Information and Computer Sciences, Toyohashi Univ. of Technology, 1-1 Hibarigaoka, Tempaku-chou, Toyohashi-shi, Aichi-ken, 441-8580 JAPAN

Abstract: Two factors, accent nuclei and word familiarity, were focused upon in terms of their capabilities to accelerate lexical access in Japanese. The minimum length of the initial portion of a word utterance required for its correct recognition was examined through perceptual experiments. As a result, the high familiarity and the accent nucleus at the first mora were found to significantly shorten the initial portion of the word required for its recognition. And it was also observed that the accent nucleus at the second mora could accelerate the lexical access only when the word was high familiar. Further, the comparison of the accelerating effects by the above two factors was also done.

1. INTRODUCTION

In Japanese, the prosodic features of content words are fully represented by the accent type, denoted by a high-low pattern of F_0 for each constituent mora. The number of the accent types in actual use is strongly limited and the accent types can be distinguished only by positions of their accent nuclei. Since only one accent nucleus can be found in a word, once a position of the nucleus is known, its accent type can be determined. This means that the early appearance of the nucleus can give a listener the full information of the word's accent type before the completion of recognizing the word. And in most cases, each word is lexically allowed to have one accent type. This fact let us assume that the perception of type-1 accent words, which are with the earliest appearance of the nuclei, would be distinct from that of words of the other accent types. And our previous study[1] showed the singularity of the type-1 accent words. Literatures [2][3] reported the identification of the accent types only by using the first morae of word stimuli and they also found the special effect caused by the accent nuclei at the first morae of the words.

However, the word stimuli used in [1] were limited to 4-mora words only. In the current study, using *the gating task*, 3-/5-mora words were additionally utilized to analyze the accelerating effect caused by the accent nucleus perception. Furthermore, another factor of accelerating the lexical access, the word familiarity, was introduced here. To increase the reliability of the experimental results, the familiarity of each word stimulus was rated subjectively by each subject. And the accelerating effect caused by the word familiarity[4] was also analyzed and the interaction between the effects by the above two factors was investigated.

2. DESIGN OF EXPERIMENTS

2.1. Preparation of Words

2.1.1. 3/4/5-mora words with high familiarity

3/4/5-mora words with the high familiarity were prepared. Here, for each number of the morae, 10 words were prepared separately for each of the accent types 0 to 2[1]. In the preparation, the familiarity of these words was rated by the experimenters in advance. However, the analysis of the results of the gating experiments required the familiarity of each word rated by each subject. Therefore, after all the gating experiments, the word familiarity was rated by the same subjects that joined the gating experiments.

2.1.2. 4-mora words with various levels of familiarity

4-mora words with various levels of the familiarity were further prepared for each of the accent types 0 to 2. The number of the words was 40 for each type. As in Section **2.1.1.**, the *actual* familiarity of each word was rated afterward. 4-mora words with the high familiarity prepared in this section were the same as those prepared inSection **2.1.1.**. Namely in this paper, 180 different words were prepared in total.

2.2. Recording of Word Stimuli

All the words were spoken at a natural rate by a male adult and recorded in a sound-proof room to give word stimuli. The length of the word stimuli was statistically analyzed for each number of morae and each accent type. According to ANOVA, in every number of morae, no significant difference was found between any two

[1] Type-i accent words have their accent nuclei at their i-th morae except for type-0 words, which have no nuclei.

of the accent types. Namely, the word stimuli of a given number of morae can be said to have almost the same duration among the accent types, which is a required condition for the gating experiments.

2.3. Procedures and Task

In a session, the 180 word stimuli were randomly presented to subjects through headphones, where the initial d [msec] of each stimulus was retained and the rest of the stimulus was replaced by silence. The subjects were requested to guess the missing part and write down an entire word on a given sheet after each presentation. The length of the retained portion (d) was defined as in the following equation;

$$d_n^m = D_n + \frac{D_n}{3} \times m, \tag{1}$$

where d_n^m is the length of the word fragment for an n-mora word in the m-th session and D_n is the averaged length of the n-mora words. Namely, in the first session, the initial portion of $1\frac{1}{3}$ morae was presented and the portion was lengthened by $\frac{1}{3}$ morae in the subsequent sessions. The interval between two consecutive stimuli was 2.5 [sec] and immediately before the presentation, a beep sound of 0.1 [sec] was presented for subjects to stop guessing the stimulus. In writing a word, the subjects were allowed to do with *hira-gana*s, phonograms of Japanese. If homonyms existed for the *hira-gana* representation, its meaning was asked afterward. After the writing, the subjects were also requested to check whether the word was a guessed one or just a sequence of morae. In the latter case, no entry of the lexicon is thought to be accessed.

2.4. Minimum Length of the Initial Portion of the Word Required for Its Correct Recognition

With the increase of d_n^m, more similar words were guessed to the target word. Using these results, the minimum length of the initial portion needed for the correct recognition was estimated for each word and each subject. In the estimation, the minimum length was defined as the length of the word fragment with a longer duration than which no other words than the target was guessed in the subsequent sessions.

2.5. Rating Experiments of the Word Familiarity

After the gating experiments, the familiarity of each word was rated by each subject using a seven-degree (1∼7) scale. And the degree of 1 was defined as *unknown*. In the experiment, each of the 180 words was presented on a sheet in two forms of *hira-gana*s and *kanji* characters, which are ideograms of Japanese. The *kanji* representation used here denoted the meaning of the word assumed by the experimenters. Naturally, the subjects had sometimes guessed a homonym of the *kanji* representation in the gating experiments. In this case, they were also asked to report the meaning they had guessed and to rate its familiarity. In the discussions below, the familiarity of the meaning guessed in the gating experiments is used. In this paper, 8 adult Japanese of Tokyo dialect participated in each experiment as subjects.

3. RESULTS AND DISCUSSIONS

3.1. 3/4/5-mora Words with High Familiarity

For 3/5-mora words, 80 (10×8) ratings of the familiarity were obtained for each accent type. And for 4-mora words, 320 (40×8) ratings were obtained for each accent type. *High familiar* stimuli were defined as the stimuli whose familiarity ratings were 6 or 7 and extracted out of all the stimuli. In this section, only the high familiar stimuli were focused upon. **Figure 1** shows the minimum length of the initial portions of the stimuli required for their correct recognition for each number of morae and each accent type. In these figures, the duration of the word fragment is indicated in two manners, a left axis as a function of the session number m and a right axis as a function of the number of morae. ANOVA was conducted on these results, which is shown in **Table 1**. From p**Figure 1** and **Table 1**, we can say that type-1 < type-2 < type-0 in the required minimum duration for 3-mora stimuli, and that type-1 ≈ type-2 < type-0 for 4/5-mora stimuli. Namely, in the case of 3/5-mora words, the accelerating effect by the accent nucleus perception was observed.

3.2. 4-mora Words with Various Levels of Familiarity

High/middle/low familiar stimuli were defined as the stimuli whose familiarity ratings were 6/7, 4/5, and 2/3 respectively and they were extracted out of all the 4-mora stimuli. **Figure 2** shows the the minimum length of the initial portions of the stimuli needed for their correct recognition for each level of the familiarity and each accent type. Results of ANOVA are shown in **Table 2**. From the figures and the table, it can be said that, for high familiar stimuli, type-1 ≈ type-2 < type-0 in the required minimum duration, and that

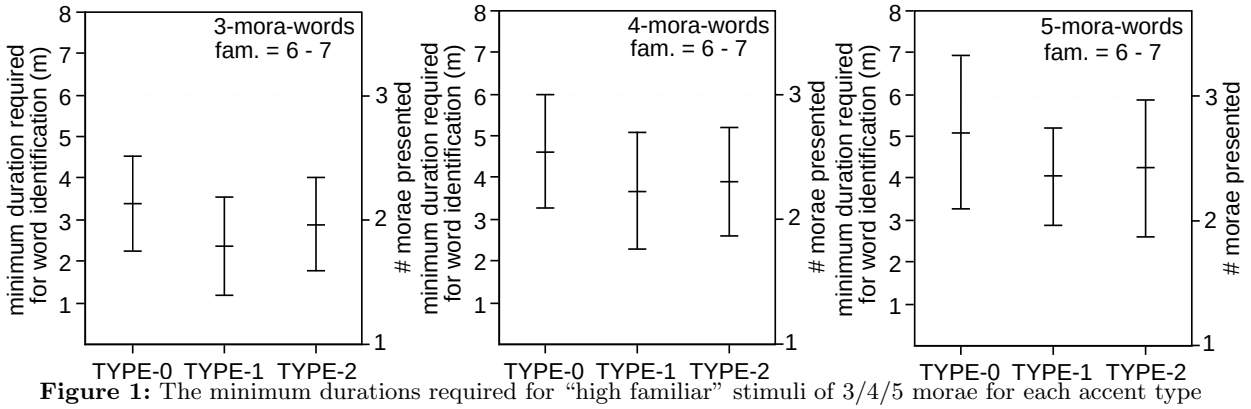

Figure 1: The minimum durations required for "high familiar" stimuli of 3/4/5 morae for each accent type

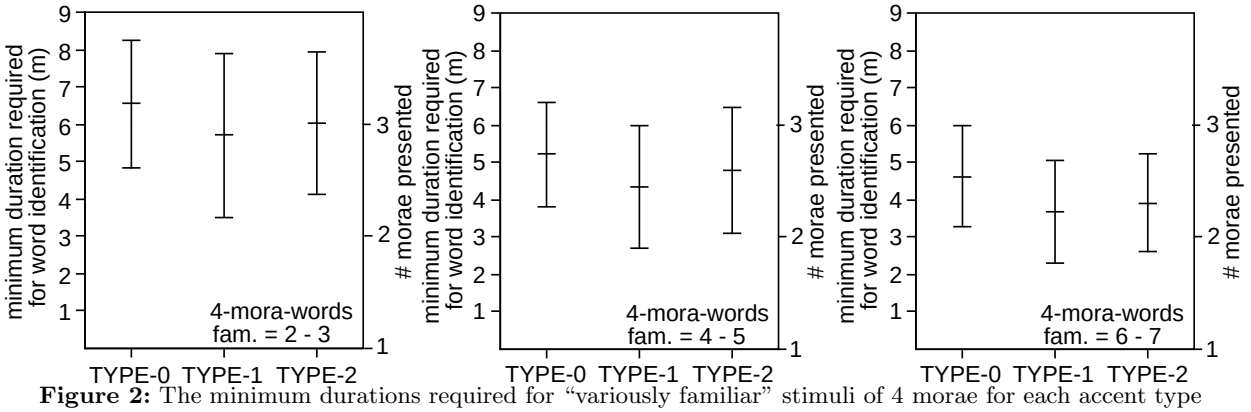

Figure 2: The minimum durations required for "variously familiar" stimuli of 4 morae for each accent type

Table 1: Results of ANOVA conducted on high familiar stimuli of 3/4/5 morae

types	0–1	1–2	0–2
3 morae	<1%	<1%	<1%
4 morae	<1%	—	<1%
5 morae	<1%	—	<5%

Table 2: Results of ANOVA conducted on 4-mora stimuli of various levels of the familiarity

types	0–1	1–2	0–2
fam. = 2∼3 (Low)	<5%	<5%	—
fam. = 4∼5 (Mid.)	<1%	<5%	—
fam. = 6∼7 (High)	<1%	—	<1%

type-1 < type-2 ≈ type-0 for middle/low familiar stimuli. These mean that the accent nucleus at the second mora can accelerate the lexical access only when the stimulus is high familiar. Namely, the dependency of the accelerating effect by the accent nucleus perception on the word familiarity was observed in the nuclei at the second morae. As for the effect by the accent nuclei at the first morae, it was observed independently of the word familiarity. Therefore, the application of the automatic detection of the accent nuclei to speech recognition should be limited to the nuclei at the first morae of the words[5].

3.3. Accelerating Effect by Word Familiarity

The accelerating effect caused by the word familiarity was examined separately for each accent type. The high/middle/low familiar stimuli defined in the previous section were also used and analyzed here. ANOVA was performed on the stimuli and the results are listed on **Table 3**, where it was tested that the required minimum durations for the higher familiar stimuli are shorter than those of the lower familiarity. Clearly shown in the table, for each accent type, the minimum duration is shortened by the higher word familiarity. These results also indicate the importance to control the familiarity of the stimuli used in the experiments where words are guessed only from word fragments in such tasks as gating or lexical decision[4].

3.4. Comparison of the Accelerating Effects by the Two Factors

In this paper, two factors, the accent nucleus and the word familiarity, were considered. By using all the results of the gating experiments, it is possible to compare the accelerating effects caused by the two factors. A stimulus set can be defined using the stimuli of an accent type and of a level of the familiarity. And a set is written as $(a,\ f)$, where a is the accent type and f is the familiarity of the set. From the previous

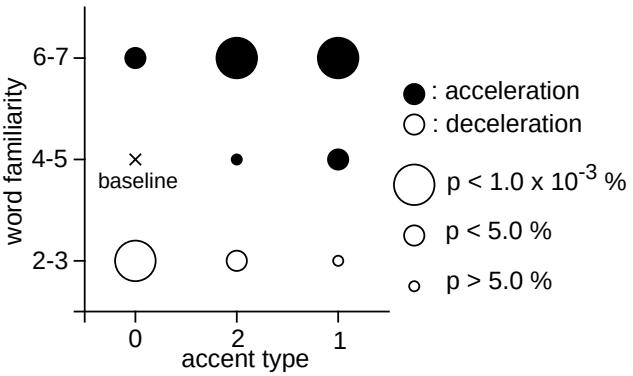

Figure 3: Difference of the minimum duration between stimulus sets using (0, 4–5) as their baseline.

Table 3: Results of ANOVA conducted on 4-mora stimuli of the accent types 0 to 2

fam.	Low–Mid.	Mid.–High
type-0	<1%	<1%
type-1	<1%	<1%
type-2	<1%	<1%

discussions, the required durations of (1, 4–5) are shorter than those of (0, 4–5) and that the minimum durations of (0, 6–7) are shorter than those of (0, 4–5). And the comparison between (1, 4–5) and (0, 6–7) is possible. According to ANOVA, no difference was found between them even at a significance level of 0.15. This result denotes that the accelerating effect caused by increasing the familiarity from 4–5 to 6–7 for type-0 accent stimuli is as strong as that caused by changing the accent type from 0 to 1 for middle familiar stimuli. The same analyses were done in the other stimulus sets, which are drawn in **Figure 3**. In the figure, all the sets except (0, 4–5) are represented by their significance levels (p), which were calculated using the difference from the baseline (0, 4–5), and their directions of the change of the required minimum durations. The former are drawn by size of the circles and the latter are drawn by color of the circles. From the figure, it is globally seen that each of the increase of the familiarity and the change of the accent type (0→2→1) can accelerate the lexical access. And also it can be locally seen that the change of both the familiarity and the accent type from (0, 4–5) to (1, 2–3) produces a small deceleration. It means that the change of the familiarity produces a slightly stronger influence than that of the accent type in this case.

4. CONCLUSIONS

In this paper, two factors, the accent nuclei and the word familiarity, were focused upon in terms of their capabilities to accelerate the lexical access. Using the gating task, the minimum length of the initial portion of a word stimulus required for its correct recognition was examined. As a result, the higher word familiarity and the accent nucleus at the first mora were found to significantly shorten the initial portion required for the correct recognition. And it was also observed that the accent nucleus at the second mora could accelerate the word recognition only when the word was high familiar. Furthermore, using several examples, the comparison between the accelerating effects caused by the above two factors was done, where a stimulus set was defined by two attributes corresponding to the two factors.

REFERENCES

1. N. Minematsu and K. Hirose, "Role of prosodic features in the human process of perceiving spoken words and sentences in Japanese," J. Acoust. Soc. Japan(E), 16, 5, pp.311–320 (1995).
2. A. Cutler and T. Otake, "The processing of word prosody in Japanese," Proc. 6th Australian Internat. Conf. on Speech Science and Technology, pp.599–604 (1996).
3. A. Cutler and T. Otake, "Pitch accent in spoken-word recognition in Japanese," J. Acoust. Soc. Am., Vol.105, No.3, pp.1877–1888 (1999).
4. S. Amano and T. Kondo, "Word familiarity effect on lexical decision for Japanese spoken words," Report of Autumn Meet. Acoust. Soc. Jpn., 1-8-8, pp.363-365, (1998, in Japanese).
5. N. Minematsu and S. Nakagawa, "Automatic Detection of Accent Nuclei at the Head of Words for Speech Recognition," Proc. ICSLP'96, pp.1620–1623 (1996).