

A Role of Pitch Accent in Spoken-Word Recognition in Accentless Japanese Dialects: Evidence from Fukushima Listeners

Takashi Otake & Marii Higuchi

Department of English
Dokkyo University, Japan
{otake@dokkyo.ac.jp}

Abstract

An experiment was conducted to examine whether word initial pitch accent information could be exploited to reduce possible word candidates by speakers of an accentless dialect in Japan. 40 native high school students from Fukushima were presented with Tokyo Japanese materials used in an earlier study, employing a gating task. Results show that the subjects performed significantly above chance, but their responses showed less sensitivity to the information in the input, and greater bias toward vocabulary distribution frequencies, than had been observed with Tokyo Japanese listeners. The whole response pattern was identical to that of speakers of Kumamoto dialect which is another accentless dialect in our earlier study. The results suggest that one of the main features of accentless dialects may be characterized by less effective use of word initial pitch information to reduce possible word candidates.

1. Introduction

New research in human spoken-word recognition has shown that listeners exploit relevant phonetic information as soon as it becomes available [1] [2]; any inter-word distinctions in the initial portions of words could therefore be of great use to listeners, enabling drastic reductions in the set of possible candidate words for recognition. It is reported that this function maybe manifested not only at the segmental level but also at the suprasegmental level [3].

Pitch-accent information in Japanese may have this function. Although there are complexities in the pitch-accent system of Tokyo Japanese, there are in general only two possible ways to label the first two morae of the citation form of a word: HL- or LH- [4]. In other words, important pitch-accent distinctions may be expressed in the initial portion of Tokyo Japanese words to reduce possible word candidates. Thus, it is highly desirable to investigate whether the exploitation of pitch information is actually involved with the reduction of possible word candidates.

Recent studies provide strong evidence that Tokyo Japanese listeners make early and effective use of pitch-accent information in recognizing spoken words [5] [6]. Minematsu and Hirose found that mismatched words of Tokyo Japanese, presented in isolation, were harder to recognize than the same words with correct accent [5]. They found that Type 1 four mora words (HLLL), presented incrementally in successively larger fragments, could be recognized on the basis of less information than Type 0 (LHHH) or Type 2 (LHLL) words. As the four-mora vocabulary of Tokyo Japanese contains less than 10% Type 1 words, this result suggests that listeners were effectively using accent to narrow down the set of potential candidate words.

One of the experiments examined in Cutler and Otake (1999) presented listeners with initial fragments of words and asked them to guess the words; the listeners' guesses overwhelmingly had the same initial accent structure (HL- or LH-) as the presented word even when all that the listener had heard of the word was the first consonant and a portion of the vowel that followed it (e.g., *na-* from *nagasa* HLL or *nagashi* LHH) [6]. Thus, listeners could effectively exploit the cues to accent pattern available even in the first syllable of a word. In other words, the suprasegmental structure of Tokyo Japanese words is, like the segmental structure, of use in narrowing down the set of potential candidates for lexical recognition.

However, many varieties of Japanese differ from Tokyo Japanese in respect of accentual structure, and not all varieties make use of accent to convey distinctions between words at all. The so-called "accentless" dialects are spoken in two principal areas of Japan: a large area spreading from the northern Kanto area (Tochigi/Ibaraki Prefectures) to the southern part of the Tohoku area (Fukushima and Miyagi Prefectures) on Honshu Island, and the central area (Kumamoto and Miyazaki Prefectures) of Kyushu Island [7].

The words in the accentless dialects are characterized by flat pitch contours, a pattern which is very different from that of other dialects. The accentless dialects make use of pitch contours to express other linguistic distinctions at the sentence level [8]. But they do not make lexical distinctions with pitch accent. Therefore, for users of such dialects there would be little point in exploiting the suprasegmental features of the initial portions of spoken words in word recognition, in the way Tokyo Japanese listeners do: word recognition in accentless Japanese could not benefit from attention to suprasegmental information.

In our previous study, both Tochigi/Ibaraki and Kumamoto dialect listeners were investigated with the same task and linguistic materials to see whether they were equally sensitive, like Tokyo listeners, or less sensitive to the pitch accent information [9]. The results showed that both dialect groups were less sensitive to the pitch information in recognizing Tokyo Japanese words. Thus, it is likely that one of the main features of accentless dialect speakers may be characterized by a way of exploiting a word initial pitch accent information to reduce possible word candidates.

In order to test the validity of this hypothesis, we have decided to investigate another accentless dialect, Fukushima dialect which is spoken in the northern part of Tochigi/Ibaraki dialects and to make a direct comparison with Kumamoto dialect which was investigated in our earlier study [9]. If the findings in our earlier study represent the main characteristics of accentless dialect in spoken-word recognition in Japanese, Fukushima speakers, just like Kumamoto listeners must show the evidence that their early use of word initial pitch information should be less effective than Tokyo listeners.

2. Experiment

2.1 Experiment

2.1.1 Materials

The materials were those used by Cutler and Otake (1999; Experiment 2)[6]. They consisted of 24 pairs of Japanese words, with each pair having a common initial bimoraic CVCV sequence, but differing in segmental structure from the fifth segment on. The accent pattern of the two words in each pair differed, in that one word the initial CVCV sequence was HL, in the other LH. Thus, *nagasa* and *nagashi* formed a pair; both begin *naga-*; the accent pattern of *nagasa* is HLL, while *nagashi* is LHH. Twenty-two pairs had three morae, the remaining two pairs four morae, and no words contained moraic nasals, geminate consonants or long vowels. The complete set of pairs (in HL-/LH-order) is shown in Table 1.

Table 1 Stimuli used in the experiment

<i>bakuhi/bakuchi</i>	<i>hanabi/hanawa</i>	<i>hokubu.hokuro</i>	<i>kamotsu/kamome</i>
<i>karasu/karada</i>	<i>karahuru/karamatsu</i>	<i>karuteru/karudera</i>	
<i>kasegi/kasetsu</i>	<i>kokugi/kokuge</i>	<i>maguchi/maguro</i>	<i>moguri/mogura</i>
<i>mokuba/mokuji</i>	<i>nagasa/nagashi</i>	<i>namida/namiki</i>	<i>nimotsu/nimono</i>
<i>nomichi/nomiya</i>	<i>sashizu/sashiki</i>	<i>sekiri/sekiryu</i>	<i>tachiba.tachiki</i>
<i>tomato/tomari</i>	<i>wakaba/wakate</i>	<i>warabi/waraji</i>	<i>yomichi/yomise</i>

Twenty-four further words served as practice and warm up items. Some of these contained moraic nasals, geminate consonants, or long vowels. Twelve were three-mora words and 12 were four-mora words.

All experimental and filler words, each preceded by a short carrier phrase *Sore wa ...* ("It is...") appeared in gated form on a tape made from an original recording by a male native speaker of Tokyo Japanese. In the gated version, each word, always preceded by the carrier phrase, was presented in increasingly large fragments, incrementing in each case by a portion of the word up to the midpoint of the following phoneme. This was achieved by determining as near as possible on the basis of visual and auditory information, the boundaries of each phonetic segment, following which a marker was placed at the midpoint of each such demarcated region. The first fragment then presented the carrier plus the target word up to the first marker, i.e., presented no more than part of the word's initial consonant; the next fragment added further material up to the second marker, i.e., the midpoint of the first vowel, and so on. The cuts producing the fragments were made in such a way as to avoid abrupt amplitude changes which might lead to the perception of illusory clicks. The gating procedure, and acoustic analyses of the materials, are described in greater details in Cutler and Otake (1999)[6].

There were two experimental tapes, each containing all filler words and one member of each experimental pair. Accent pattern was counterbalanced across tapes; each tape contained 12 HL- and 12 LH- experimental words, and the members of any pair occurred at the same position on both tapes.

2.1.2 Subjects

Forty second-year high school students (five were first year) at the northern part of Fukushima where the variety of Japanese spoken in the area is characterized as an accentless dialect, took part in this experiment. All had grown up in Fukushima prefecture and had received all their schooling in their home area. They were tested as a group at a language laboratory room at Fukushima Minami High School. A native speaker of Fukushima Japanese, the second author of this paper, conducted the experiment using this dialect.

2.1.3 Procedure

There were two subject groups, of 20 listeners each, and each group heard a different one of the two tapes. The subjects were tested in groups of up to 18. The tape was played over headphones from a cassette tape recorder which was built in the language laboratory system. Presentation was given with 8 second pause after each fragment to allow time for the subject to record a guess as to the word's identity. The guesses were written on a response sheet in their normal Japanese orthographic form.

The earlier studies concentrated their analyses on the candidate word guessed at fragments 1-4, and in particular on the comparison of the accent pattern of these candidate words for target words beginning HL- and LH- [6][9]. From fragment 5 onwards, segmental information could distinguish between the members of the word pairs; for the first four fragments, however, the only distinguishing information was accentual. The word guesses for the first four fragments thus most directly address the question of interest: could listeners make use of the accentual cues available in the initial bimoraic portion of each stimulus pair (such as *naga-* in *nagasa* HLL and *nagashi* LHH), a portion which was segmentally matched but accentually different? In the present analyses we examine these initial guesses.

2.1 Result

2.2.1 Analysis of Fukushima listeners

The word guesses were scored by hand and the accent pattern of each guess in Tokyo Japanese determined. As in the study of Cutler and Otake (1999)[6], only the first two morae of the guessed words were taken into consideration, which effectively resulted in a two-way classification of accent patterns into HL- (Type 1 accent) vs. LH- (Types 0, 2, 3, etc.).

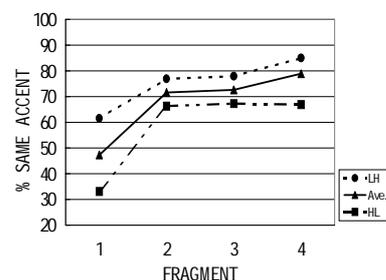


Figure 1: FUKUSHIMA LISTENERS: proportion of guessed words with the same initial accent pattern.

Fig. 1 shows the proportion of guesses which had the same initial accent pattern as the spoken word, for each of the first four fragments, separately for HL- and LH- words and averaged across these. Percent correct for each fragment was subjected to a sign test of statistical significance of the difference from chance, and analyses of variances were carried out on the accuracy scores across word types.

The guesses provided for fragment 1 matched the initial accent pattern of the target word in 47.3% of cases, which is not significantly different from chance. The guesses provided at fragment 2 matched the initial accent pattern of the target word in 71.6% of cases, which is significantly above chance performance ($z=10.18$, $p<0.001$), and again accuracy continued to rise across fragments 3 and 4. Accuracy was again greater for LH- words than for HL- words ($F [1, 39]=92.773$, $p<0.001$), and this difference was again separately significant at every fragment (position 1: $F [1, 78] = 65.26$, $p<0.001$; position 2: $F [1, 78] = 12.50$, $p<0.001$; position 3: $F [1, 78] = 10.40$, $p<0.01$; position 4: $F [1, 78] = 38.91$, $p<0.001$).

2.2.2 Joint analysis with Kumamoto listeners

First, let us to look at the result of Kumamoto listeners in the earlier study [9], which is shown in Fig. 2.

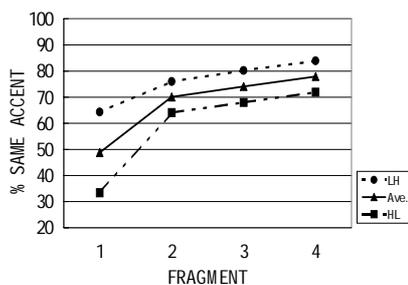


Figure 2: KUMAMOTO LISTENERS: proportion of guessed words with the same initial accent pattern.

The guesses provided for fragment 1 matched the initial accent pattern of the target word in 47.4% of cases, which was not significantly different from chance. The guesses provided at fragment 2 matched the initial accent pattern of the target word in 69.3% of cases, which was significantly above the chance performance ($z=10.64$, $p<0.001$), and again accuracy continued to rise across fragments 3 and 4. Accuracy was again significantly greater for LH- words than for HL- words ($F [1, 31]= 55.67$, $p<0.001$), and this difference was separately significant at every fragment. As can be seen from these figures, it is easily noticed that the whole pattern in the statistical analysis of Fukushima listeners was almost identical.

In order to make a direct comparison between Fukushima listeners with Kumamoto listeners, a joint analysis of the accuracy scores was undertaken. Fig. 3 shows the mean accuracy scores between Fukushima listeners and Kumamoto listeners. As can be seen from this figure, the mean scores at the four positions were almost identical. The joint analysis revealed that there was no difference between the two groups ($F [1, 142] = 0.145$, $p = 704$). In this figure, the result of Tokyo listeners [6] is also included. The joint

analysis of the accuracy scores between Fukushima listeners and Tokyo listeners show the significant difference between the two groups ($F [1, 150] = 18.91$, $p<0.001$). Separate comparisons between the subject groups revealed that Tokyo listeners had higher accuracy than Fukushima listeners at all of fragments 2-4 (fragment 2: $F [1, 150] = 5.37$, $p<0.05$; fragment 3: $F [1, 150] = 28.00$, $p<0.001$; fragment 4: $F [1, 150] = 20.44$, $p<0.001$). The results suggest that the sensitivity to pitch-accent by Tokyo listeners was higher than Fukushima listeners.

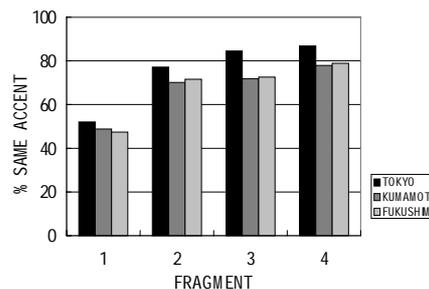


Figure 3: Proportion of guessed words with the same initial accent pattern as the spoken word (Average).

Fig. 4 shows the accuracy scores for LH-words between Fukushima listeners and Kumamoto listeners. As can be seen from this figure, the scores at the four positions were almost identical. The joint analysis revealed that there was no difference between the two groups ($F [1, 142]=0.145$, $p<0.704$). The joint analysis of the accuracy scores between Fukushima listeners and Tokyo listeners show the significant difference between the two groups ($F [1, 74] = 4.06$, $p<0.05$). Separate comparisons between the subject groups revealed that Tokyo listeners had higher accuracy than Fukushima listeners at fragment 3 (fragment 2: $F [1, 74] = 0.20$, $p=0.888$; fragment 3: $F [1, 74] = 7.2$, $p<0.05$; fragment 4: $F [1, 74] = 1.39$, $p=0.243$).

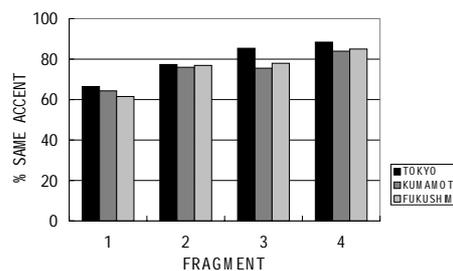


Figure 4: Proportion of guessed words with the same initial accent pattern as the spoken word (LH-words).

Fig. 5 shows the accuracy scores for HL-words between Fukushima listeners and Kumamoto listeners. As can be seen from this figure, the scores at the four positions were almost identical. The joint analysis revealed that there was no difference between the two groups ($F [1, 142]=0.145$, $p<0.704$). The joint analysis of the accuracy scores between Fukushima listeners and Tokyo listeners show the significant difference between the two groups ($F [1, 74] = 3045.8$, $p<0.001$). Separate comparisons between the subject groups revealed that Tokyo listeners had higher accuracy than Fukushima listeners at all of fragments 2-4 (fragment 2: $F [1, 74] = 9.61$, $p<0.05$; fragment 3: $F [1, 74] = 24.11$, $p<0.001$; fragment 4: $F [1, 74] = 30.60$, $p<0.001$).

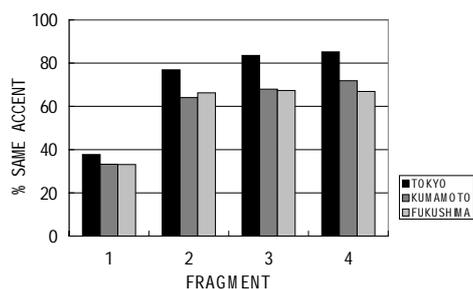


Figure 5: Proportion of guessed words with the same initial accent pattern as the spoken word (HL-words).

3. Discussion

This study has investigated a role of pitch accent information by Fukushima listeners in terms of spoken-word recognition. Our main question raised in the study was whether the word initial pitch information was involved with the reduction of possible word candidates in accentless speakers. Since the dialect tested in this study was Fukushima, another accentless dialect, our prediction based upon our earlier studies [6] [9] was that Fukushima listeners should show less sensitivity to the word initial pitch information. The results in the experiment using the gating task with the same materials in the earlier studies have shown that there was not much difference with Kumamoto listeners but a clear difference with Tokyo listeners. Thus, our prediction was sustained.

What this result suggests is that so called “accentless speakers” which is mainly been defined in terms of speech production can be also characterized in terms of speech perception. Since accentless speakers are not solely surrounded by their local dialect, it is reported that precise estimates of actual dialect use in speech production could be difficult to establish. Thus far, the study of accentless dialects has been mainly evaluated on the basis of speech production with a special reference to Tokyo dialect [10]. However, the findings in this study may suggest that a methodology based upon speech perception which is involved with spoken-word recognition may be more suitable to examine because it may reflect the role of the pitch accent information stored in the mental lexicon in a more clear manner.

4. Conclusion

The present study investigated a role pitch accent in spoken-word recognition in Fukushima listeners, which is classified as an accentless Japanese dialect. The analyses have shown that Fukushima listeners perform very similar to that of Kumamoto listeners, which is another accentless dialect. The sensitivity to the information in the input is lower than that of the Tokyo listeners and their bias is higher. The evidence obtained from this study suggests that accentless dialect listeners may in general show the same tendency. Thus, the examination of the role of pitch accent may provide us with a new methodology.

5. Acknowledgement

This paper is a revised version of BA thesis written by the second author. This research was supported by a grant from the Japan Society for the Promotion of Science (Grant-in-Aid for Scientific Research (C) 11610566) which was given to the first author. We thank to Mr. Koichi Nihei and Masatake Watanabe of Fukushima Minami High School for the arrangement of subjects.

6. References

- [1] Martin, J. G. and Bunnell, H. T., 1981. Perception of anticipatory coarticulation effects. *Journal of Acoustical Society of America*. 69:559-567.
- [2] Lahiri, A. and Marslen-Wilson, W. D., 1991. The mental representation of lexical form: A phonological approach to the recognition of lexicon. *Cognition*. 38, 245-294.
- [3] Cutler, A., Dahan, D., and van Donselaar, W., 1997. Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40-2, 141-201.
- [4] Haraguchi, S., 1977. *The Tone Pattern of Japanese: An Autosegmental Theory of Tonology*. Tokyo: Kaitakusha.
- [5] Minematsu, N. and Hirose, K., 1995. Role of prosodic features in the human process of perceiving spoken words and sentences in Japanese. *Journal of Acoustical Society of Japan* 16, 311-320.
- [6] Cutler, A. and Otake, T., 1999. Pitch accent in spoken-word recognition in Japanese. *Journal of Acoustical Society of America* 105, 1877-1888.
- [7] Shibatani, M., 1990. *The Language of Japan*. Cambridge: Cambridge University Press.
- [8] Maekawa, K., 1997. Akusento to intoneishon: Akusento no nai chiiki (Accent and intonation in accentless dialects). In *Nihongo Onsei 1: Shohougen no akusento to intoneishon (Japanese speech 1: Accent and intonation in various dialects)*, 97-122. Tokyo: Sanseido.
- [9] Otake, T. and Cutler, A., 1999. Perception of suprasegmental structure in non-native dialect. *Journal of Phonetics* 27, 229-253.
- [10] Inoue, F., 1984. Akusento no seisei to chikaku: Kantou ni okeru chiikisa to nenreisa (Production and perception of (Japanese pitch) accent: area and age differences in the Kanto area). In *Kindaich Haruhiko Hakase Kinen Ronbunshuu (Papers in celebration of Dr. Haruhiko Kindaichi's 70th birthday)*, Vol. 2, 119-153. Tokyo: Sanseido.