

A LOGISTIC REGRESSION APPROACH TO ACCENT CLASS DIVISION IN JAPANESE DIALECTS: WITH SPECIAL REFERENCE TO THE KEIHAN-TYPE ACCENT SYSTEM IN PERIPHERAL KINKI REGIONS

Murata, Mami

Osaka University
murata@cjlc.osaka-u.ac.jp

ABSTRACT

The accent system of Modern Japanese is characterized by patterns of high and low tones which are conventionally specified for each word, and words are grouped into several classes according to their accent patterns. Put differently, words belonging to the same class exhibit the same accent pattern. Some dialects, however, permit more than one accent patterns in the same class, which will be referred to as “accent class division” (ACD). In an attempt to explain this phenomenon, this paper takes a logistic regression approach, with the hypothesis that the features of consonants and vowels and their combinations are the possible factors. The data are from the Keihan-type accent in peripheral Kinki regions, especially Tokushima and Mie prefectures, where ACD is widely observed. It will be shown that the composition of segments in a word affects the choice of accent patterns, resulting in ACD.¹

Keywords: Pitch accent, Japanese dialects, the Keihan-type accent, logistic regression

1. INTRODUCTION

Historical studies on Japanese accent (especially those dealing with the Kyoto accent system) have shown that it is not individual words that have undergone accent changes but an accent class as a whole, and that words belonging to the same class have changed their accent patterns almost at the same time. In some dialects, however, there are more than one accent patterns in the same class, which will be referred to as “accent class division” (ACD). This phenomenon is often observed in regions where different accent systems are in contact with each other.

A typical case is seen in the Toyama dialect (Figure 1), which has the Tarui-type accent system (this system is normally found in regions bordering on

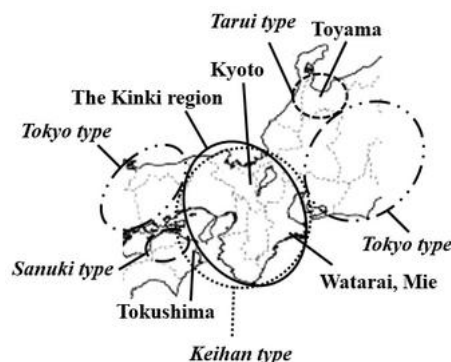


Figure 1: Accent types in and around the Kinki region, Japan

Table 1: Accent patterns of 2-mora nouns in the Toyama dialect

	Accent pattern	Examples
Class 1	LH-H	<i>kao-ga, kaze-ga</i>
Class 2	LH-L	<i>ishi-ga, yuki-ga</i>
Class 3	LH-L or HL-L	<i>ashi-ga, mame-ga</i>
Class 4	LH-H	<i>ito-ga, ine-ga</i>
Class 5	HL-L	<i>ame-ga, saru-ga</i>

Table 2: ACD in the Toyama dialect (Class-3 nouns)

	Examples
HL	<i>ashi, mimi, oni, hiru, inu, doku</i>
LH	<i>iwa, uta, mame, hito, imo</i>

the Keihan-type and Tokyo-type regions). In this dialect, 2-mora nouns in Class 3 show ACD. The data is given in Table 1, where L and H indicate a low/high tone respectively. Each noun is followed by the nominative particle *ga*.

Table 1 shows that in contrast with the other four classes, Class 3 has two different accent patterns, a typical case of ACD. This phenomenon is straightforward since it can be explained as a set of simple phonological rules. According to [2], 2-mora nouns

¹This work is supported by JSPS KAKENHI Grant Number 26370449 (‘A Study on the Relation between Voicedness of Plosives and Prosody produced and/or perceived by the Learners of Japanese of various non-native Speakers’).

belonging to Class 3 fall into either the HL or LH patterns (Table 2), which is adequately explained by the following rules. The HL pattern is chosen when the vowel in the second mora is a close vowel or when the second mora is a special mora; the LH pattern is chosen when the vowel in the second mora is an open vowel.

As the Tokushima and Mie dialects are spoken in the peripheries of the Keihan-type region (Figure 1), it is natural that they exhibit ACD, too. However, ACD in these dialects is so complicated that it cannot be explained by simple rules. This paper thus attempts to tackle this issue by making use of logistic regression, with the hypothesis that the features of consonants and vowels and their combinations are the possible factors. More specifically, it is hypothesized that the choice of accent patterns is affected not only by openness or closedness of vowels but also by the following factors: the presence or absence of a consonant, whether a consonant, if any, is voiced or voiceless, and the consonant-vowel combination in the first and second moras. The statistical analysis in this study will demonstrate the adequacy of this hypothesis.

The rest of the paper is organized as follows. Section 2 reviews the previous literature on ACD in the Tokushima and Mie dialects and summarizes the phenomenon. Section 3 introduces the data for the analysis. Section 4 describes the methodology employed. Section 5 shows and interprets the results. Finally Section 6 concludes the study.

2. PREVIOUS STUDIES

Unlike the Toyama dialect outlined in Section 1, it has been reported that ACD in the Tokushima and Mie dialects is complicated and no study has succeeded in finding out factors behind the complexity.

2.1. The Tokushima dialect

[1] is a report of his fieldwork in 1968 in the middle stream area of the Yoshino River, the boundary between the Sanuki-type and Keihan-type accent systems (Figure 1). Table 3 is the result of his survey of accent patterns of Class 3, 2-mora nouns. As shown, ACD is observable within the same class. According to [1], this division cannot be explained phonologically or grammatically.

In 2009, about 40 years after [1], [3] conducted a survey on the same topic in the same area. [3] is intended to be more comprehensive than [1] by including more informants and more nouns in the survey. The result indicates that there is still ACD in the area just like 40 years ago, without any change regarding

Table 3: ACD in the Tokushima dialect (Class 3 nouns)

		Examples
HH		<i>ashi, ie, inu, ono, hone, nawa, kame</i>
HL		<i>ami, oni, kumo, sao, hana, yama, kutsu</i>

accent patterns or grouping of nouns. [3], however, also fails to provide an adequate explanation for the phenomenon.

2.2. The Mie dialect

[4] reports ACD observed in Watarai, Mie prefecture (Figure 1). As summarized in Table 4, ACD occurs in 2-mora nouns in Classes 4 and 5, resulting in the HL and LF patterns (F indicates tonal step-down). [4] says that he could find out no phonological rule that explains the division.

Table 4: ACD in the Mie dialect

		Examples
Class 4	HL	<i>ato, iki, kasa, tane, matsu</i>
	LF	<i>obi, hari, kinu, kama</i>
Class 5	HL	<i>asa, kage, kumo, oke</i>
	LF	<i>aka, ase, koi, saru, funa</i>

3. DATA

3.1. The Tokushima dialect

The data for the Tokushima dialect comes from the survey by [3]. The informants included were 16 native speakers of the dialect aged more than 60 years old, for their language was assumed to exhibit traditional accent patterns. The data was collected by having them read aloud a list of 38 2-mora nouns in Class 3. Each informant read the list twice, and the number of the data collected amounted to 1,216 instances (38*2*16).

3.2. The Mie dialect

The data in [4] was used. Due to its insufficiency, however, the author will do an additional survey in June, 2015.

4. METHODOLOGY

The data was analyzed using R. In order to test the hypothesis proposed in Section 1, logistic regression was applied.

5. RESULTS

5.1. The Tokushima dialect

The result shows that the features of the vowel and consonant and their combination in a word are the factors behind ACD in the Tokushima dialect, a result which suggests the adequacy of the hypothesis suggested in Section 1. The results in detail are:

- Openness or closedness of vowels in the first and second moras plays a significant role.
- When the first mora consists of a close vowel (without a consonant) and the second mora of a voiced consonant and an open vowel, words tend to exhibit the HH pattern. Nouns such as *ude*, *iro*, and *uma* are likely to follow this pattern.
- When the first mora consists of a consonant and an open vowel and the second mora of a close vowel (without a consonant), words tend to exhibit the HL pattern. Nouns such as *kai* and *koi* are likely to follow this pattern.

These results indicate that on the basis of simple vowel rules like those in the Toyama dialect, the Tokushima dialect has developed complex consonant rules which led to ACD.

5.2. The Mie dialect

Due to the small data size in [4], a logistic regression approach could not be taken and the author is planning to collect more data by an additional survey. By doing so, it is expected that the factors behind ACD in the Mie dialect will also be revealed.

6. CONCLUSION

This paper has shown that ACD in Japanese dialects is explainable in terms of the features of consonants and vowels and their combination. Although it has explored and clarified internal factors behind ACD, further studies are needed to examine the effect of external factors like dialect contact.

7. REFERENCES

- [1] Kato, N. 1968. *Kyokaichitai ni okeru akusento no mondai: yoshinogawaryuiki wo chushin to shite* [Issues concerning accent in boundary regions: with special reference to the middle stream area of the Yoshino River]. Paper read at the 6th meeting of Dialectological Circle of Japan.
- [2] Matsumori, A., Nitta, T., Kibe, N., Nakai, Y. 2012. *Nihongo Akusento Nyumon [An Introduction to Japanese Accent]*. Tokyo: Sanseido.
- [3] Murata, M. 2012. Tokushimaken yoshinogawaryuiki no akusento: Keihan shiki akusento to Sanuki shiki akusento no kyokai [Accent in the Yoshino River basin, Tokushima prefecture: the boundary between the Keihan-type accent and the Sanuki-type accent]. *Bunri Synergy* 16, 119–124.
- [4] Uwano, Z. 1987. Nihon hondo syohougen akusento no keifu to bumpu (2) [Genealogy and distribution of accent in the dialects of mainland Japan (2)]. *The Japan Academy Kiyo* 42, 15–70.