
CHIHKAI LIN
*Tatung University*

1 Introduction

This paper reexamines accent patterns in Sino-Japanese trimoraic nouns. As Table 1 shows, there are four accent patterns in Tokyo Japanese trimoraic nouns (H = high, L = low).

<table>
<thead>
<tr>
<th>Position of accent</th>
<th>Accent pattern</th>
<th>Example</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antepenultimate</td>
<td>HLL</td>
<td>inochi</td>
<td>‘life’</td>
</tr>
<tr>
<td>Penultimate</td>
<td>LHL</td>
<td>kokoro</td>
<td>‘heart’</td>
</tr>
<tr>
<td>Final</td>
<td>LHH(L)</td>
<td>otoko</td>
<td>‘man’</td>
</tr>
<tr>
<td>Unaccented</td>
<td>LHH(H)</td>
<td>sakana</td>
<td>‘fish’</td>
</tr>
</tbody>
</table>

Table 1: Accent patterns in Tokyo Japanese
The four accent patterns in Table 1 are divided into accented and unaccented. There are three accent positions in accented trimoraic nouns. Accents can fall on the antepenultimate syllable (HLL), as in does in inochi ‘life’, on the penultimate syllable (LHL), as ko does in kokoro ‘heart’, or on the final syllable, as ko does in otoko ‘man’ LHH. In Tokyo Japanese, trimoraic nouns, like sakana ‘fish’ (LHH), can also be unaccented. Although Table 1 displays four accent patterns, at the surface level, nouns with final accents and unaccented nouns share the same pattern—that is, LHH. Only the addition of an extra component distinguishes them: final accent LHH(L) vs. unaccented LHH(H). The subject marker -ga, for example, does not carry an accent. When following otoko ‘man’ and sakana ‘fish’, the subject marker -ga has different realizations of pitch accent. After otoko ‘man’ LHH(L), the subject marker is associated with a low pitch, and the accentuation of otoko-ga is LHH-L. On the other hand, when the subject marker -ga is added to sakana ‘fish’ LHH(H), which lacks an accent, the accentuation of sakana-ga is LHH-H.

The four accent patterns in trimoraic nouns are not evenly distributed. According to Tanaka and Kubozono (1999: 59), trimoraic word distributions are biased: approximately half are LHH(H), and about forty percent have the HLL pattern. Roughly ten percent have medial accents LHL, and about five percent have final accents LHH(L). As Japanese includes lexical strata, native Japanese, Sino-Japanese and foreign loans, Kubozono (2006, 2008, 2011) have investigated the accent patterns in different strata. Kubozono (2008: 167) has demonstrated that in trimoraic nouns, seventy-one percent of native words are LHH(H), and ninety-three percent of foreign loans are HLL. Sino-Japanese does not favor any accent pattern. Forty-nine percent of trimoraic nouns are accented, and fifty-one are unaccented. Accented trimoraic nouns tend to be antepenultimate accents (Tashiro 1975, Ogawa 2004, Kubozono 2008). Table 2 shows the distributions of trimoraic accented nouns.

<table>
<thead>
<tr>
<th>Lexical strata</th>
<th>Initial accent (± antepenultimate)</th>
<th>Medial accent (± penultimate)</th>
<th>Final accent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>59%</td>
<td>33%</td>
<td>9%</td>
</tr>
<tr>
<td>Sino-Japanese</td>
<td>95%</td>
<td>2%</td>
<td>3%</td>
</tr>
<tr>
<td>Loan</td>
<td>96%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Total</td>
<td>89%</td>
<td>7%</td>
<td>4%</td>
</tr>
</tbody>
</table>

Table 2: Lexical strata and accent patterns in trimoraic accented nouns

Fifty-nine percent of native words have initial accents, thirty-three percent have medial accents and nine percent have final accents. Meanwhile,
ninety-five percent of Sino-Japanese words have initial accents, two percent have medial accents, and three percent have final accents. Lastly, ninety-six percent of foreign loan words have initial accents, two percent have medial accents, and two percent have final accents.

In Japanese phonology, special moras and high vowel devoicing trigger accent shifts. Special moras are moraic nasal, long vowel and geminate, and they cannot carry accents. When there are special moras, accents must shift to adjacent full CV syllables. In tenki ‘weather’, there are three moras, and the second mora is a moraic nasal (tenki). The moraic nasal cannot bear an accent, making the medial accent prohibited; consequently, the accent of tenki falls on the antepenultimate syllable HLL or on the final syllable LHH(L).

Besides special moras, high vowel devoicing triggers accent shifts. High vowels /i/ or /ɯ/ are devoiced when they are located between two voiceless consonants, takusan [takusaN] ‘a lot’ (Tsujimura 2007: 24), or when they are preceded by voiceless consonants in a word-final position, muki [mɯki] ‘direction’ (Tsujimura 2007: 26). When high vowel devoicing occurs, accents shift to other CV syllables (Kitahara 1996, Akinaga 1998, Maekawa 1990, Kawahara 2015). According to Kawahara (2015: 479), the accent of kākusu ‘to hide’ falls on the penultimate syllable ku, but high vowel devoicing triggers a leftward accent shift from the ku syllable to the ka syllable, LHL > HLL.

Sino-Japanese undergoes accent shifts when there are special moras and high vowel devoicing. Tashiro (1975: 150) has suggested that accents in Sino-Japanese tend to fall on the antepenultimate mora, as se does in genshu ‘player’. If antepenultimate moras are special moras or undergo high vowel devoicing, the accents shift to adjacent moras. For example, kessin ‘determination’ has four moras, and the antepenultimate mora is geminate, which fails to carry an accent. The accent must shift to the first mora, and therefore kessin ‘determination’ is HLLL. In kisen ‘steamship’, the antepenultimate mora ki undergoes high vowel devoicing. The accent shifts to the penultimate mora se, and LHL becomes its accent pattern.

Although Kuzobono (2008) has suggested that Sino-Japanese has equal distributions of accented and unaccented trimoraic nouns, the fact that his dataset did not exclude instances of special moras and high vowel devoicing calls this finding into question. Special moras are pervasive in Japanese, and high vowel devoicing is a common phonological process. This makes it necessary to take the two factors into account when recalculating distributions of Japanese accent patterns. To generate accurate distributions of accent patterns in Sino-Japanese trimoraic nouns, I adopt a corpus-based approach. I discuss the steps of establishing the corpus and the data-selection criteria in section 2. The corpus data only include trimoraic nouns
without special moras and high vowel devoicing in non-final positions. I report the distributions in section 3 where I adopt alternate perspectives based on the number of kanji and the internal branching patterns. Sino-Japanese trimoraic nouns can include two or three kanji, and three potential internal branching patterns—left (AB+C), right (A+BC), and no branching (A+B+C). I discuss the interaction of internal branching, accentuation and the number of kanji in section 4, where I also consider gaps in distributions in accent variants. Section 5 concludes this paper.

2 Corpus and Data-selection Criteria

This paper adopts a corpus-based approach to accent patterns of Sino-Japanese trimoraic nouns. I collected data from The Japanese Language Pronunciation and Accent Dictionary (NHK 1998). I sorted proper data as follows. I collected all trimoraic nouns from this dictionary and classified all instances into four groups based on the accent positions listed in Table 1. If a trimoraic noun has more than one accent, I marked it as ‘accent variant’ in the corpus. For instance, ganka ‘ophthalmology’ can be LHH(H) or HLL, and I, therefore, marked it as a variant. In this paper, I separated accent variants from those with single accent patterns. In section 4, I discuss whether accent variants show accent patterns.

After sorting out single accent nouns, I eliminated trimoraic nouns with long vowels (gyoryō ‘fish and hunting’), moraic nasal (gyorū ‘fish scales’) or geminate (tekki ‘ironware’), which this paper does not consider. I also excluded nouns with high vowel devoicing in the first and/or second mora. According to the phonological notation in the dictionary, high vowels /i/ or /ɯ/ in final syllables are not devoiced, meaning the accent does not shift in such instances. For instance, when a morpheme geki ‘excited’ forms a compound with another morpheme, the accent pattern differs depending on where it appears. When geki ‘excited’ precedes sho ‘summer’, the compound is gekisho ‘extremely hot summer’ and the accent is HLL. When geki ‘excited’ follows ka ‘over’, the compound is kageki ‘overexcited’ and the accent is LH. As the accent of geki ‘excited’ is LH, the accent does shift in gekisho ‘extremely hot summer’ from the ki syllable to the ge syllable, whereas the accent does not shift in kageki ‘overexcited’.

I analyzed the selected data from two perspectives: the corresponding number of kanji and the internal branching patterns of the trimoraic nouns. Sino-Japanese trimoraic nouns contain either two or three kanji. The mapping of kanji and moras pertains to a given trimoraic noun’s internal branching pattern. Branching patterns include left branching AB+C, right branching A+BC, and no branching A+B+C. The mapping of kanji and moras along with internal branching is shown in (1) for two kanji and in (2) for three kanji (K = kanji).
Branching and Accentuation in Sino-Japanese

(1) a. Branching
   Two kanji
   Example
   akumu
   ‘nightmare’

   b. Branching
   Two kanji
   Example
   wabaku
   ‘harmony’

(1) displays two mappings of kanji and moras. In (1a), the first two moras correspond to the first kanji, AB → K1, and the third mora to the second kanji. In (1b), the first mora is associated with the first kanji, and the other two moras with the second kanji, BC → K2.

(2) a. Branching
   Three kanji
   Example
   ‘maintenance fee’

b. Branching
   Three kanji
   Example
   kahogo
   ‘overprotected’

c. Branching
   Three kanji
   Example
   ‘Amitabha’

(2) contains three mappings of kanji and moras. Each mora corresponds to a kanji. In (2a), K1 and K2 form a domain, and K3 belongs to the other domain. In (2b), the first domain includes K1, and the second domain consists of K2 and K3. In (2c), there is no internal branching, and one mora corresponds to one kanji.

Together, the four accent patterns in Table 1 and the branching in (1) and (2) produce twenty possible combinations. Section 3 reports the distributions of the twenty combinations. In addition to instances of single accents in the corpus, section 3 also includes the distributions of accent variants in the corpus.

3 Results

This section is divided into two subsections. 3.1 reports distributions of instances with single accents and 3.2 shows distributions of instances with accent variants.

3.1 Distribution of Instances with Single Accent

As Table 3 (below) shows, there are 740 single accent instances. This distribution is biased. Unaccented nouns LHH(H) make up the majority (sixty-one percent). The distribution conforms to Tashiro’s (1975)
observation. About half of Sino-Japanese trimoraic nouns with two or three kanji are unaccented. When the number of kanji is taken into consideration, 705 corpus instances (approximately ninety-five percent) include two kanji. Only thirty-five instances (less than five percent) include three kanji. Four cells exceed fifty instances in the corpus. 381 instances (fifty-two percent) are unaccented in A+BC, as in ikoku ‘foreign country’. Also in A+BC, 122 instances (seventeen percent) have an initial accent—kakyoku ‘songs’ for example. 107 instances (fifteen percent) in AB+C have initial accents, as in kokudo ‘country land’. Sixty-six instances (nine percent) are unaccented in AB+C, as in mokuji ‘index’.

<table>
<thead>
<tr>
<th>Number of kanji</th>
<th>Internal branching</th>
<th>Unaccented</th>
<th>Accented</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LHH(H)</td>
<td>HLL</td>
<td>LHL</td>
</tr>
<tr>
<td>Two</td>
<td>A+BC</td>
<td>381</td>
<td>122</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>AB+C</td>
<td>66</td>
<td>107</td>
<td>24</td>
</tr>
<tr>
<td>Three</td>
<td>A+BC</td>
<td>1</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>AB+C</td>
<td>4</td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>A+B+C</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>454</td>
<td>229</td>
<td>52</td>
</tr>
</tbody>
</table>

Table 3: Distribution of single accent in Sino-Japanese trimoraic nouns

Other cells in Table 3 show relatively low frequencies. The distribution shows twenty-four instances with medial accents in AB+C, as in rokuji ‘six o’clock’, and five unaccented corpus instances. It shows no instances of medial accents in A+BC.

The three kanji distribution shows only thirty-five unaccented and medial accent instances. Nineteen instances in three kanji include medial accents with left branching, as in gekai ‘surgeon’. The three kanji distribution includes eight instances of unaccented nouns with right branching, such as kahogo ‘overprotected’. The remaining instances are sporadic—less than five instances in the corpus. The distribution shows no initial accent and final accent instances.

3.2 Distributions of Accent Variants

Accent variants are trimoraic nouns with more than one accent pattern. Table 4 shows the distribution of 112 instances in A+BC with two kanji.¹

¹The first accent pattern shown in the NHK dictionary is treated as the primary accent, and the second accent pattern is treated as the secondary.
**Table 4: Distribution of accent variants in A+BC (two kanji)**

Table 4 shows seventy instances (sixty-three percent) that alternate between LHH(H) and HLL, as in *iheki* ‘gastric wall’. The distribution includes thirty-nine instances (thirty-five percent) of alternation between HLL and LHH(H), as in *iseki* ‘remains, relic’. The remaining three cells have only one instance.

As Table 5 show, the AB+C with two kanji distribution includes fifty-five instances of accent variants.

**Table 5: Distribution of accent variants in AB+C (two kanji)**

Twenty-three instances alternate between HLL and LHH(H), such as *bakuro* ‘exposure’. Fourteen instances alternate between LHH(H) and HLL, as in *dokugo* ‘after reading’. Twelve instances alternate between HLL and LHL, as in *dokuga* ‘fangs’. In AB+C with two kanji, *richigi* ‘rules and regulations’ has three patterns of accentuation, HLL, LHH(H) or LHH(L).

In the corpus, there are five instances in three kanji with more than one accent pattern. The five instances are *ikiji* ‘spirit’ HLL or LHH(H), *shuisho* ‘a prospectus’ LHL or LHH(L), *segaki* ‘Segaki, feeding the hungry ghosts’ LHH(H) or HLL, *darani* ‘Dharani’ LHL or LHH(L), and *hugiri* ‘ungrateful’, which has all the four accent patterns.

**4 Discussion**

In section 3, I report the distributions of instances of single accents and accent variants in the corpus. This section discusses three issues: (a) the interaction of internal branching, accentuation and kanji, (b) gaps in distributions, and (c) patterns in accent variants.
4.1 Interaction of Internal Branching, Accentuation and Kanji

Kubozono’s (2008) data have shown that unaccented and accented take half of his corpus instances. The distribution in Table 3 differs from Kubozono’s results. Sixty-one percent of the corpus instances (454 in 740) in this distribution are unaccented. Kubozono (2008) also points out that ninety-five percent of his instances are initially accented HLL. The data in my Table 3 conform to the data in Kubozono’s Table 2, indicating that the majority of instances have initial accents. My data, however, show different distributions. About eighty percent of the accented nouns (229 in 286) are initially accented. Eighteen percent of the accented nouns have medial accents, and two percent have final accents.

I conducted a Chi-square test to check whether the data in Table 3 show a significant difference between branching (AB+C vs. A+BC) and accentuation (unaccented vs. accented) in two kanji. I added instances of HLL, LHL, and LHH(L), and conducted the Chi-square test with Yates’s correction. The result shows a significant difference between branching and accentuation in two kanji ($\chi^2 = 109.4$, d.f. = 1, $p < .01$). A+BC unaccented nouns (318 instances) have three times as many instances as accented nouns (124 instances). On the other hand, AB+C accented nouns (134 instances) have two times as many instances as unaccented nouns (66 instances).

In addition to the interaction of internal branching and accentuation, the number of kanji also plays a crucial role. The majority of trimoraic nouns with two kanji manifest different accent patterns from nouns with three kanji. In two kanji, right branching A+BC is highly associated with unaccented nouns (seventy-five percent, 381 in 505), whereas left branching is associated with initial accents (fifty-four percent, 107 in 200). Although the total number of three kanji instances in the corpus is low, more than half of the corpus instances (fifty-four percent, 19 in 35) manifest left branching patterns with medial accents LHL.

4.2 Gaps in Distributions

Table 3 does not include instances for every cell. In two kanji distribution, the right branching, medial accent cell is left blank. Although the blank cell in two kanji could be considered an accidental gap, it is also possible to look deeper into the data. Given that there are three moras in two kanji, one kanji must be heavier than the other. The accentuation is sensitive to the position of a heavy kanji in a trimoraic noun. The fact that Table 3 contains no instances of LHL for A+BC with two kanji suggests that the second and third moras in A+BC should have identical heights, either as HH or LL (B = C). On the other hand, the first two moras in AB+C cannot have identical heights. They have to be HL or LH (A ≠ B).
In three *kanji*, the blank cells are systematic gaps. All the corpus instances manifest in LHH(H) and LHL, and the majority of the corpus instances (eighty percent, 28 in 35) have medial accents, especially in AB+C branching. The blank cells in HLL reveal that high pitch in the first mora is prohibited. As for LHH(L), the fact that Table 3 shows only five instances in the corpus (less than one percent) suggests that low frequency is the main cause of the blank cells in three *kanji*.

### 4.3 Patterns in Accent Variants

The accent variants in section 3.2 have different distributions in A+BC and AB+C. These results have directed to two alternations. The distributions include 146 instances of the LHH(H) and HLL alternation, eighty-four instances of the alternation between LHH(H) as the primary and HLL as the secondary, and sixty-two instances of the alternation between HLL as the primary and LHH(H) as the secondary. The 146 instances take up to eighty-eight percent of the variants (146 in 166).

Nouns without accents largely alternate with nouns with initial accents in the corpus. This fact suggests that the accent variants in Sino-Japanese trimoraic nouns are not random. The alternation takes place in the antepenultimate mora of trimoraic nouns. Suppose that there is an accent at the left edge of LHH(H), namely before the word, (1)LHH(H). The fact that LHH(H) alternates with HLL becomes straightforward. When a trimoraic noun is unaccented, it is expressed as (1)LHH(H). Its accent variant has an accent that shifts rightward to the next interval, (1)LHH(H) → HLL, and the output is the initial accent. When a trimoraic noun has an initial accent, HLL, its accent variant shifts the accent leftward to the left edge of the word, HLL → (1)LHH(H). The accent variant becomes unaccented.

### 5 Conclusion

In this paper, I adopt a corpus-based approach to exploring the accent patterns in Sino-Japanese trimoraic nouns. I analyze instances without special moras and high vowel devoicing. The results reveal the following three facts. First, as suggested by the statistical analyses, internal branching interacts with accent patterns in two *kanji*. Right branching tends to be associated with LHH(H) and left branching with HLL. Second, the instances in three *kanji* consistently appear in LHH(H) and LHL. Third, alternations in the accent variants have a pattern as well. LHH(H) alternates with HLL.

The corpus-based analyses of the accent patterns in Sino-Japanese trimoraic nouns have two implications for Japanese phonology. One implication concerns the number of *kanji*—an issue seldom discussed in the literature. Absent special mora and high vowel devoicing, the accentuation
of Sino-Japanese trimoraic nouns is not only associated with internal branching but also related to the number of kanji. The results open up a new area of inquiry for Japanese phonology. Researchers must take into account how kanji are mapped onto moras.

The other implication for Japanese phonology is that the left edge is more dominant than the right edge in phonological words. The fact that the majority of the variants in Sino-Japanese trimoraic nouns manifest in the alternation of LHH(H) and HLL indicates that accent shifts prefer the left edge. More data—regarding native Japanese trimoraic nouns, for example—are needed to determine whether there is a similar tendency in the alternation.

References


