Left-dominant Accentuation: The Case of Sino-Japanese Numeral Compounds

Chuyu HUANG

Abstract

本研究は 2 字漢数詞複合語（例：一語、一気）のアクセントを考察したものである。日本語の漢数詞（例：いち、に、さん）は他の漢語形態素と共に漢語を形成する。漢数詞複合語（Sino-Japanese numeral compound）は、アクセントに関して非漢数詞複合語と異なることが指摘されている。本研究は漢数詞と漢数詞複合語のアクセントの観察を通じて、漢数詞複合語はどのように異なる性質を持っているかを探求する。本研究の意義は記述的側面と理論的側面がある。記述的側面では漢数詞と漢数詞複合語のアクセントの関係および音節・形態構造との関係を明らかにすることであり、理論的側面では従来の複合語アクセント規則と違い、これらの語が左側支配型アクセント（left-dominant accent）であることを示唆し、更に既存の一般的な制約によってそれを解釈することである。

Key Words: compound accent, Sino-Japanese, numeral, left-dominant accent

1. Introduction

As a pitch-accent language, Japanese uses the location of lexical prominence to distinguish between words. The existence and the position of an accent in a word are lexically distinctive and non-predictable in simple words. For example, one can never predict the accentual contrast of [a'me] ('rain' 雨; ''' stands for the position where accent falls) and [ame-] ('candy' 飴; '-' at the word end stands for unaccentedness) merely by its semantic information. However, when two words combine into a compound, the accent placement can be predicted to some extent. The accent of a compound is not a simple one-plus-one addition, but is calculated by a number of phonological processes.

But if the process is not decided by simple addition, the question remains which component decides the accent of the compound. Previous research has tended to focus on the second component rather than the first. This is called the "right-dominant" accent. However, it fails to explain some cases, including compounds containing numerals. The present study focuses on the numeral compounds in Sino-Japanese, illustrating some cases where the compound accent rule proposed by previous research cannot account for the placement of the accent, and introducing a

new possible analysis for the accent placement in these cases. It will be shown that in these cases, the first component is the provider of the accent information. Thus it will be called the "left-dominant" accent.

This study aims to investigate the left-dominant accentuation in Sino-Japanese two-character words, especially compounds whose first element is a numeral. This study is constructed in the following way: Section 2 introduces some basic assumptions and accentual facts about Sino-Japanese along with previous studies about the compound accent rule. A hypothesis where the accent of the first component is decisive in compound accent will be proposed and proved by empirical evidence in Section 3. Section 4 proposes a theoretical analysis within the framework of Optimality Theory, and Section 5 summarizes the study.

2. Sino-Japanese accent in Japanese

2.1 Sino-Japanese accent in Tokyo-Japanese

Japanese has three lexical strata: Yamato (*Wago*), which is the native lexical stratum, Sino-Japanese (*Kango*; henceforth, SJ) words borrowed from Chinese over different periods of history, and loanwords (*Gairaigo*), mainly words borrowed from western languages. SJ words form a separate lexical stratum with unique morpheme-structural, prosodic, and segmental characteristics (McCawley 1968, Ito and Mester 2015, among others).

The accent of two-character Sino-Japanese words differs among various dialects. In Kagoshima dialect, for instance, the accent type of N1, the first component, will be inherited and reflected in the combined form (Uwano 1997). In Tokyo Japanese, on the other hand, the accent type of the second component will affect the combined form. Moreover, as for two-character SJ words, it is said that the accentuation depends on the phonological structure (Ogawa 2006).

As mentioned by Ogawa (2006), the syllabic and moraic structure have an impact on the accentuation. For instance, when the first component has one heavy (H) syllable or two light (L) syllables, the accent tends to be located on the first mora from the left. Also, prosodic length plays a role in accentuation. Compared to three-mora words, four-mora words show an extremely high unaccented ratio regardless of morpheme boundary and syllable weight. The dominant accentuation of different prosodic length and structure of two-character SJ words is outlined in (1).

(1) Accent and its prosodic and morphological structure

H#L(LL#L): [ha'ku#si[1]] ('doctor' 博士), [bu'n#ka] ('culture' 文化)

Four-mora words: [dai#gaku-] ('university' 大学), [syus#seki-] ('presence' 出席)

2.2 Sino-Japanese words' status as compounds

It is still controversial whether SJ two-character words are compounds or not. Some argue that SJ two-character-words are simplex words rather than complex because many of them are constructed by bound morphemes that are seldom used independently. However, a Chinese character (*Kanji)* has a meaning of its own and a clear boundary in orthography, thus it might be appropriate to see each character as independent morphemes. From this point of view, each two-character-word can be analyzed as a compound, consisting of by two components: the first component (henceforth N1) and the second (henceforth N2). In addition, a number of previous studies indicated that the accent of these words can be predicted by the accent type of N2 (Kubozono 1997, Tanaka 2004, among others). Due to the complex characteristics of *Kanji*, Labrune (2012) described these words as *fixed compounds* to differentiate from the compounds made of free morphemes, also highlighting the dependent characteristic of many of them.

Despite this dependent morphological characteristic, these words show a high predictability in respect to the accent. The accent of two-character SJ words can be accounted for by the compound accent rule proposed in previous studies. The rules of compound accent have been hotly discussed. The most well-known analysis is McCawley (1968), in which the second components in compounds are divided into three kinds, having different accentual status as follows:

(2) Three kinds of short N2

a. N2= deaccenting morpheme, *e.g.* mura'saki+iro' → [murasaki#iro-] ('purple color' 紫色)

b. N2= preaccenting morpheme, *e.g.* imo'+musi- → [imo'#musi] ('caterpillar' 芋虫)

c. N2= initial-accenting morpheme, *e.g.* tumugi-+i'to → [tumugi#i'to] ('silk thread' 紬糸)

In (2a), a deaccenting morpheme *iro* serves as the second morpheme and causes deaccentuation on the compound, while morphemes in (2b) result in a pre-N2 accent. In (2c), these morphemes trigger an accent on the initial syllable of N2.

The accentuation of many SJ two-character words can be accounted for by this categorization. *Sei*, as a deaccenting morpheme, frequently causes unaccentedness of compounds (*e.g.* [akusei-], [isei-]). Similarly, preacccenting morphemes and initial-accenting morphemes can also be found in Sino-Japanese. This provides direct evidence that two-character SJ words phonologically belong to compounds and their accent can be predicted as those compounds in other lexical strata. A summary of these rules and some examples can be found in (3).

(3) Three kinds of short N2 (Sino-Japanese)

a. N2= deaccenting morpheme,

e.g. se'i → [aku#sei-] ('malignant' 悪性), [ryoo#sei-] ('benign' 良性),
[i#sei-] ('the opposite sex' 異性)

b. N2= pre-accenting morpheme,

e.g ke'n → [akita'#ken] ('Akita prefecture' 秋田県),
[aomori'#ken] ('Aomori prefecture' 青森県)

c. N2= initial-accenting morpheme,

e.g. su'u → [kan#su'u] ('function' 関数), [ten#su'u] ('point' 点数),
[hen#su'u] ('variable' 変数)

Another analysis of compound accent by Kubozono was that when two words are combined, a juncture accent occurs (Kubozono 1995, 1997). For example, [pi'tti] ('pitch') and [a'kusento] ('accent') will be [pitti#a'kusento] ('pitch accent'). This rule predicts that the accent falls on the boundary between two components when compounding together. However, on which side of the boundary the accent appears remained unsolved. Kubozono (1995) proposed an analysis based on Optimality Theory. He used a faithfulness constraint which requires that the output of N2 be faithful to the input accent of N2. By putting another constraint which requires the accent not to be on the final prosodic units in a higher ranking, most cases of short N2 compounds can be correctly predicted. In other words, the accent of N2 would be preserved as long as it is not on the final prosodic unit.

Note that both of the accent rules have to access the accent information of N2 instead of N1. In other words, the accent of N1 is not predicted to affect the compound accent. The accent of compounds is predicted solely by N2. However, some cases do not seem to be accounted for by these rules. For example, the compounds containing a numeral, such as *iti* ('one'), as seen in (4):

(4) Words that compound accent rules can and cannot predict:

a. N2=go ('word' 語) (deaccenting morpheme)

[zyuku#go-] ('idiom' 熟語), [kei#go-] ('honorific language' 敬語), [nihon#go-] ('Japanese' 日本語)

cf. exception: [iti'#go] ('one word' 一語)

b. N2=ri ('principle' 理) (word initial accenting morpheme)[2]

[ga'ku#ri] ('scientific principle' 学理), [ka'n#ri] ('management' 管理),
[go'#ori] ('reasonability' 合理)

cf. exception: [iti'#ri] ('some truth' 一理)

Go has been regarded as a deaccenting morpheme. When it appears as N2, compounds will be deaccented as in [zyuku#go-]. However, when N1 is a numeral, it becomes accented and seems to inherit the accent position of N1. Similarly, in (b), the morpheme *ri* seems to cause a word-initial accent in two-character compounds except in the case when N1 is a numeral.

Compound accent rules proposed in the literature still fail to explain compounds with a numeral as N1 as shown above, thus the present study proposes a hypothesis to account for this on the assumption that N1 instead of N2 decides the accent of compound when N1 is a numeral.

Some previous studies already indicated that numerals might have an impact on the compound accent, making numeral compounds different from non-numeral words (Uwano 1997, Akinaga 1999, Labrune 2012), but lacked a thorough investigation. This study aims to be the first investigation of the accent of SJ numeral compounds. The hypothesis is that the accent of N1 will be preserved in compounds. Disyllabic numerals were used because these numerals have an accent on the second syllable from the left and when it is preserved, it can be distinguished from the dominant accent type in Sino-Japanese. This reasoning will be illustrated in the next section.

3. Accent of Sino-Japanese numeral compounds

In this section, all words with a disyllabic numeral N1 (*iti, roku, siti, hati*) were selected from *Sanseido Daijirin*, a dictionary where the accent is noted by a number to show the position or presence of accent. 453 words were collected.

The reason why these disyllabic numerals were selected was that they are composed of two light syllables and possess an accent on the second syllable from the left. If their accent is preserved in trisyllabic words, an accent type (LL'#L) different from the most dominant type (*e.g.* L'L#L) should occur. If the alternative hypothesis that denies the N1 dominant analysis is true, it would predict that the accent of these compounds appears in the position where the accent of most SJ words do.[3] All words are divided into two groups by their prosodic length: three-mora words and four-mora words.

3.1 Three-mora words

Disyllabic numerals in SJ include *iti* ('one' 一), *roku* ('six' 六), *siti* ('seven' 七), and *hati* ('eight' 八). Because all of them have two moras, in three-mora words with a disyllabic numeral as N1, the N2 is inevitably one mora long. Before seeing the numeral words, it is important to note where the accent of SJ words generally falls. As Ogawa (2006) pointed out, the accent in three-mora words is classified by its prosodic length and morpheme boundary as follows:

Table 1. Three-mora Two-character SJ Words (Ogawa 2006).

	μ'μ μ	μ μ' μ	μ μ μ'	μ μ μ
H#L	1750(78%)		13(1%)	494(22%)
L#H	230(15%)	44(3%)		1298(83%)
LL#L	187(50%)	48(13%)	6(2%)	134(36%)
L#LL	138(19%)	7(1%)	4(1%)	586(80%)
total	2305(47%)	99(2%)	23(0%)	2512(51%)

Since Tokyo Japanese does not allow an accent on the second mora of a heavy syllable, μ μ' μ is not allowed in H#L structure. Likewise, μ μ μ' is not allowed in L#H structure, either.

SJ numeral compounds, however, differ in several ways. About 51% of general SJ three-mora compounds are unaccented in total, in that H#L and LL#LL are 22% and 36% respectively. Take *iti*, for example, which contrasts with general SJ three-mora compounds in that the unaccented type can nonetheless be hardly found in numeral compounds:

Table 2. Accent of Numeral Compounds containing *Iti*.

	Accented	Unaccented	both	Total
three-mora	60(95.2%)	2(3.2%)	1(1.6%)	63(100%)

The prosodic structure also seems to play a role in accent. As shown in Table 3, all LLL words have a middle accent, while all HL words have an initial accent. Note that *iti* will be realized as a heavy syllable *it* when followed by a voiceless obstruent.

Table 3. The Accent Position and Syllabic Structure of Three-mora Words[4] *iti*

	μ'μμ	μμ'μ	μμμ'	total
LLL	0(0%)	20(100%)	0(0%)	20(100%)
HL	41(100%)		0(0%)	41(100%)
total	41(67.2%)	20(32.8%)	0(0%)	61(100%)

Not only *iti*, but other SJ disyllabic numerals also show a similar tendency. The syllabic structure seems to be correlated with the accent: LLL with a penultimate accent and HL with an accent on the first syllable from the left, which is the same as *iti*.

Table 4. Accent of Numeral Compounds with Three Moras (All disyllabic numerals)

	μ'μμ	μμ'μ	μμμ'	total
LLL	0(0%)	50(100%)	0(0%)	50(100%)
HL	53(98.1%)		1[5](1.9%)	54(100%)
total	53(51%)	50(48%)	1(1%)	104(100%)

The middle accent in LLL (namely LL'L) can be predicted simply by the accent position of N1. The accent in HL is not on the second mora from the left as expected. An accent shift occurs in this case and the accent moves onto the leftmost mora. This can be also accounted for by a phonological constraint in Tokyo Japanese, because a geminate cannot bear an accent.

Although the accent of numeral compounds can be predicted by simply saying the accent position of N1 is preserved, the observations that have been made so far do not exclude the possibility that all N2s in the data are pre-accenting morphemes by coincidence. However, there are also some deaccenting morphemes in the data as will be shown below. Numeral compound whose N2 elements are these deaccenting morphemes are predicted to be accentless by the right-dominant compound accent rules, but this prediction is incorrect.

(5) Unpredictable cases by right-dominant compound accent rules

a. N2= go ('word' 語)

[iti'#go] ('one word' 一語), [roku'#go] ('six words' 六語), [hati'#go] ('eight words' 八語)

cf. [zyuku#go-] ('idiom' 熟語), [kei#go-] ('honorific language' 敬語), [nihon#go-] ('Japanese' 日本語)

b. N2= za ('seat; constellation' 座)

[iti'#za] ('one status' 一座), [siti'#za] ('seven traditional exclusive distributors' 七座)

cf. [sei#za-] ('sitting formally' 正座), [sisi#za-] ('Leo' 獅子座), [zyooruri#za-] ('joruri drama theater' 浄瑠璃座)

The accentuation of three-mora words can be summarized as follows:
(6) Generalization of three-mora SJ numeral compounds

LL+L→LL'#L

H+L → H'#L

3.2 Four-mora words

This subsection will consider compounds with four moras. Ogawa (2006) indicated that nearly 90% of four-mora two-character SJ words are unaccented regardless of their different prosodic structures, which can be seen as the dominant accent type:

Table 5. The Relation between Four-mora Two-character Accent (Ogawa 2006)

	μ'μ μ μ	μ μ'μ μ	μ μ μ'μ	μ μ μ μ'	μ μ μ μ
H#H	326(8%)		169(4%)		**3832(89%)**
H#LL	156(10%)		2(0%)	10(1%)	**1460(90%)**
LL#H	6(1%)	41(7%)	13(2%)		**488(89%)**
LL#LL	0(0%)	16(4%)	0(0%)	1(0%)	**381(96%)**
Total	488(7%)	57(1%)	184(3%)	11(0%)	**6161(89%)**

SJ two-character numeral compounds, however, have a lower deaccentuation ratio than other general SJ two-character compounds. This is apparent because only 55.2% were unaccented:

Table 6. SJ Numeral Words with *Iti* as N1.

	accented	unaccented	both	total
four-mora words	42(21.9%)	**106(55.2%)**	44(22.9%)	192(100%)

This is observed not only in the words with *iti*, but also other disyllabic SJ numeral compounds.

Table 7. LL Two-character Four-mora SJ Numerals Deaccentuation Ratio

	accented	unaccented	both	total
iti	42(21.9%)	**106(55.2%)**	44(22.9%)	192(100%)
roku	14(29.2%)	**28(58.3%)**	6(12.5%)	48(100%)
siti	15(31.3%)	**22(45.8%)**	11(22.9%)	48(100%)
hati	18(31.6%)	**28(49.1%)**	11(19.3%)	57(100%)
total	89(25.8%)	**184(53.3%)**	72(20.9%)	345(100%)

It is known that four-mora words are more likely to be unaccented than other mora length. For example, foreign words in Japanese are generally accented, but many four-mora words are unaccented (Kubozono 2006). However, if we focus only on the accented words, we can still find that the position of accent in these accented compounds is coordinated to the accent position of the first component. In LL#LL and LL#H, which start with two light syllables, have a high ratio of LL'#LL and LL'#H. Both of LL'#LL and LL'#H preserve the accent of N1, which is the same as in LL#L.

(7) Four-mora words which begin with LL

a. LL+LL → LL'#LL

[iti'#oku] ('one hundred million' 一億), [iti'#iti] ('one by one' 一々)

b. LL+H → LL'#H

[iti'#gyoo] ('one row' 一行), [iti'#zyoo] ('one whip' 一条)

As for the compounds beginning with a heavy syllable, according to Ogawa (2006), most cases of SJ words with HLL or HH structure tend to be accented initially when they have an accent which is shown as follows:

Table 8. SJ Four-mora Word Accent and Prosodic Structure (Ogawa 2006).

	μ'μ μ μ	μ μ'μ μ	μ μ μ'μ	μ μ μ μ'	μ μ μ μ
H#H	326(8%)		169(4%)		3832(89%)
H#LL	156(10%)		2(0%)	10(1%)	1460(90%)

From the table above, it is clear that almost 90% of four-mora words which begin with a heavy syllable are unaccented.

Returning to the hypothesis, one would reasonably expect that the accent of N1 in SJ numeral compounds would be preserved if possible, but when these disyllabic numbers become a heavy syllable, it means that the second mora undergoes gemination and cannot bear the accent anymore. In this case, the accent should shift in componds with HL structure as mentioned. However, the way the accent shift behaves in these compounds is different from HL: the accent position tends to fall on the final syllable both in HH and HLL structure, instead of on the first syllable.

(8) Accent of four-mora words starting with a heavy syllable

a. H+H → H#H'
[it#te'n] ('one point' 一点), [ik#ka'i] ('once' 一回),
[is#su'n] ('one inch' 一寸), [is#si'n] ('wholeheartedness' 一心)

b. H+LL → H#LL'
[ik#kyoku'] ('one song' 一曲), [ip#patu'] ('one shot' 一発),
[is#situ'] ('one room' 一室), [is#seki'] ('one boat' 一隻)

With the generalization above, the accentuation pattern of the four-mora SJ compounds can be summarized as below:

(9) <u>The generalization of the accent of four-mora SJ numeral compounds</u>

a. LL+LL Put the accent on the second syllable from the left
e.g. LL+LL → LL'#LL [iti'#oku] ('one hundred million' 一億)

b. LL+H Put the accent on the second syllable from the left
e.g. LL+H → LL'#H [iti'#gyoo] ('one row' 一行)

c. H+H Put the accent on the final syllable
e.g. H+H → H#H' [it#te'n] ('one point' 一点)

d. H+LL Put the accent on the final syllable
e.g. H+LL → H#LL' [ik#koku'] ('one moment' 一刻)

Thus far we have generalized the accentuation of SJ two-character numeral compounds. The most remarkble result to emerge from the data is that N1 is responsible for the accent when it is a numeral. In the next section, the left-dominant accentation will be analyzed using Optimality Theory (OT) and a possible phonological explanation will also be presented.

4. Analysis

This section proposes an analysis of the generation of (8) in Section 3 using Optimality Theory (OT). The Optimality Theory used here is based on the version developed in Prince and Smolensky (1993/2004).

4.1 Constraints on Japanese Syllabic Structure

This sub-section reviews the constraints on the Japanese phonological system that are proposed in previous studies. First, it is important to introduce the foot structure of Japanese. Poser (1990) claims that a foot is bimoraic in Japanese with evidence such as hypocoristic formation. This can be demonstrated by the constraint **FT-BIN(μ)**. The second constraint is **ALIGN-L**. This constraint ensures the left edge of foot agrees with the left edge of the stem. A *kanji,* a single Chinese character, is defined as a stem in this study. Third, **PARSE σ** is a constraint which ensures all syllables are exhaustively parsed into feet:

(10) **FT-BIN(μ)**: Feet are bimoraic (Poser 1990, Prince and Smolensky 1993/2004).

e.g. μ μ → ✓(μ μ),[6] *(μ) μ

(11) **ALIGN-L(Ft, Stem):** Align the left edge of foot to the left edge of the stem.[7]

e.g. [μ μ]STEM[μ]STEM → ✓(μμ) (μ), *(μ) (μμ)

(12) **PARSE σ**: Syllables are parsed into feet (Prince and Smolensky 1993/2004).

The ranking of the three constraints is hypothesized as follows:

(13) **FT-BIN(μ) >> ALIGN-L(Ft, Stem) >> PARSE σ**

	μ μ # μ	FT-BIN(μ)	ALIGN-L	PARSE σ
A	☞ (μ μ)#μ			*
B	μ (μ# μ)		*!	*
C	(μ μ)# (μ)	*!		

With this hypothesized ranking, the optimal foot structure can be decided. For three-mora words, because **ALIGN-L(Ft, Stem)** is dominated by **FT-BIN(μ)**, candidates A and B are favored over candidate C, which is completely parsed into feet. Since we have assumed that **FT-BIN(μ)** outranks other two constraints here, a foot should be bimoraic and any degenerate foot will be relatively disfavored. However, this is not a problem in quadrimoraic words because four moras can be parsed into two full feet. For example, in LLLL structure, the optimal structure would be as follows:

(14) Foot Structure of LLLL

	LL # LL	FT-BIN(μ)	ALIGN-L	PARSE σ
A	☞ (L L)# (L L)			
B	L (L# L) L		*!	**
C	(L L)#L L			*!*

4.2 Constraints on the Accent

In this sub-section, the three general constraints over accentuation will be introduced. In this study, the SJ numerals, *e.g. iti, roku, siti, hati*, are assumed to carry an accent nucleus on the second mora from the left in the underlying form. Another important assumption in this study is that the second vowel of these numerals does not exist in the underlying form.[8] It will be shown in this section that the accentuation can be analyzed with these assumptions (*i.e.* the underlying forms are /it'/, /rok'/, /sit'/, /hat'/). This can be ensured by other constraints (*cf.* Ito and Mester 1994). However, these constraints are assumed to be in a higher ranking and thus will not be discussed in this study.

As for the constraints on accent position, the following general constraints are used for this analsysis:

(15) **Peak-Prominence (PK-PROM)**: Accent is on the most prominent nucleus of the syllable (*cf.* Prince and Smolensky 1993/2004).
(16) **NOFOLP(N1)**: Corresponding prominences have corresponding sponsors and links (*i.e.* No accent shift of N1).
(17) **RIGHTMOST(σ)**: The accented syllable is rightmost of the word.

First, as mentioned above, a non-vowel mora cannot bear an accent in surface form in Tokyo Japanese. This prosodic characteristic can be promised by the constraint **Peak-Prominence (PK-PROM)** which assigns a violation mark when the accent nucleus does not fall on the most prominent segment in a syllable.

The second constraint is a faithfulness one: **NOFOLP**. It is a faithfulness constraint to ensure that the accent in the input corresponds to the accent in the output. **NOFOLP(N1)** is violated when the accent position in the output is not in agreement with the one in the input (*cf.* Alderete 1999). When the accent is located on another mora, the constraint will be violated. On the other hand, the constraint will not be violated even if there is a vowel epenthesis as long as the second mora is accented. (*e.g.* it'+ki → [ik'-ki], *[i'k-ki]. it'+ri' → [iti'-ri], *[i'ti-ri])

The third constraint is a markedness constraint. **RIGHTMOST(σ)** is a general constraint that has also been observed in Japanese foreign words, and has been used for the analysis of compound

accent (=Edgemost(σ), Kubozono 1997).

As was stated above in the beginning of this section, when a voiceless obstruent (such as /p/, /t/, /k/) follows another voiceless obstruent, a geminate will appear in surface form (/pp/, /tt/, /kk/). Since the Sino-Japanese numerals used in this study, *iti, roku, siti, hati*, are assumed to have the underlying forms /it'/, /rok'/, /sit'/, and /hat'/, when they are compounded, the surface form contains a geminate mora. These numeral compounds have the same characteristics as all other lexical strata in that a non-vowel mora cannot bear an accent. Therefore, **PK-PROM** is never violated. On the other hand, although the accent of N1 is preserved in many cases, the accent shift sill occurs when the accented mora is geminated. From the observation above, it can be deduced that **PK-PROM** outranks **NOFOLP(N1)**.

(18) **PK-PROM** >> **NOFOLP(N1)**

一気 it'+ki	**PK-PROM**	**NOFOLP(N1)**
(ik')ki	*!	
☞(i'k)ki		*

When N1 undergoes gemination in four-mora words, the ultimate accent type appears. When **NOFOLP(N1)** is violated, the output with a rightmost accent becomes optimal. On the other hand, in LLLL, /iti'+oku/ violates **RIGHTMOST(σ)** but not **NOFOLP(N1)** and wins in the end. Thus **RIGHTMOST(σ)** is outranked by **NOFOLP(N1)**.

(19) **NOFOLP(N1)** >> **RIGHTMOST(σ)**

一点 it'+ten	**NOFOLP(N1)**	**RIGHTMOST(σ)**
(i't)(ten)	*	*!
☞ (it)(te'n)	*	

The constraint ranking is as below:

(20) **PK-PROM >> NOFOLP(N1) >> RIGHTMOST(σ)**

With the ranking in (19), the generalization in Section 3 can be predicted correctly as (20)-(24). The reason why the accent does not move to the right edge in HL structure is due to the foot structure given in (12). Since **FT-BIN(μ)** ranks above **PARSE σ**, any accent on a degenerate foot is avoided.

(21) LL+L

一理 it'+ri'	PK-PROM	NOFOLP(N1)	RIGHTMOST(σ)
☞ (iti')ri			*
(i'ti)ri		*!	**

(22) H+L

一気 it'+ki	PK-PROM	NOFOLP(N1)	RIGHTMOST(σ)
(ik')ki	*!		*
☞(i'k)ki		*	*

(23) LL+LL (LLH)

一億 it'+o'ku	PK-PROM	NOFOLP(N1)	RIGHTMOST(σ)
☞ (iti')(oku)			**
(iti)(o'ku)		*!	*
(i'ti)(oku)		*!	***
(iti)(oku')		*!	

(24) H+H

一点 it'+ten	PK-PROM	NOFOLP(N1)	RIGHTMOST(σ)
(it')(ten)	*!		*
(i't)(ten)		*	*!
☞ (it)(te'n)		*	
(it)(ten')	*!	*	

(25) H+LL

一刻 it'+ko'ku	PK-PROM	NOFOLP(N1)	RIGHTMOST(σ)
(ik')(koku)	*!		**
(i'k)(koku)		*	*!*
(ik)(ko'ku)		*	*!
☞(ik)(koku')		*	

4.3 Why the Left?

Even with the analysis put forward by this study, the question still remains why the left component rather than the right component can affect the compound accent. Thus far, the motivation of left-dominant accent is still not clear. The left-dominant accent needs an explanation indeed about why both left and right dominant accent rules exist in a language. However, the left-

dominant compound accent is widely observed in other Japanese dialects (Uwano 1997, among others). Therefore, the left-dominant phenomenon is typologically supported.

Moreover, the left-dominant accent is not incompatible with the right-dominant accent rule. Kubozono (forthcoming) indicated that the left-dominant rule still exists along with the right-dominant rule with some examples in native words. The present study provides further evidence for this view. In other words, there are two kinds of compound accent rules that synchronically exist in Tokyo Japanese: one is the right-dominant rule and the other is the left-dominant rule proposed in this study. As a result, this study implies that this left-dominant rule still exists in compounds containing a numeral.

With the constraints that are generally used in the analysis of right-dominant compound accent in the framework of OT, the left-dominant examples provided in this study can be accounted for as stated in Sub-Section 4.2. However, the question remains what linguistic motivation the left-dominant cases have. Additionally, why re-ranking of constraints occurs in left-dominant accent when compared to the analyses of right-dominant compound accent is a vital issue for future studies.

5. Summary

This paper has examined the case of the left-dominant accent. This research targeted compounds with a disyllabic numeral as N1. The results indicated that the accent of N1 is preserved in SJ compounds when N1 is a numeral. In addition, accent shift occurs as the accented mora is geminated: the accent nucleus shifts to the first mora from the left in three-mora words, while it shifts to the ultimate syllable in four-mora words. The present study generalized this special accentuation and analyzed it with general constraints used in Optimality Theory. Finally, this study implies that both left-dominant compound accent rules and right-dominant compound accent rules exist in Tokyo Japanese.

Notes

* An earlier version of this study was presented at the 11th meeting of Phonology Festa on March 11, 2016. I deeply appreciate all the comments from the audience. Also, I would like to thank the two anonymous reviewers, Professor Shin-ichi Tanaka, Daiki Hashimoto, Alina Kordesch, Itsuki Minemi, for giving me a lot of useful suggestions, and Professor Haruo Kubozono, for sharing his latest opinion toward left-dominant accent with me. All errors are my own.

[1] *Kunreisiki* transcription is used in this study for principle, except that long vowels are transcribed with two vowels instead of macron for the sake of convenience (*i.e.* [tookyoo] instead of [tôkyô]).

[2] These morphemes assign an accent on the first syllable of the word, thus they are called word initial accenting morphemes in this study. This kind of accentuation was not discussed in McCawley (1968) and other

previous studies, but it is still productive in Tokyo Japanese. Another example of word initial accenting morphemes is /zi/ (‘temple’ 寺), assigning an accent on the very first syllable of word, *e.g.* [to’odai#zi] (‘Toodai Temple’ 東大寺), [ko’oen#zi] (‘Kooen Temple’ 高円寺). These morphemes await future studies.

[3] *San* (‘three’ 三) and *yon* (‘four’ 四) are not included in the present study. Since *san* is unaccented in Tokyo Japanese, it is difficult to identify whether the accent type is “preserved” or not, depending on the definition of “preserve.” *Yon* is not a Sino-Japanese word in a strict sense, so it will not be discussed in the present study. The Sino-Japanese counterpart of *yon* is *si* (‘four’ 四).

[4] In this table and Table 4, words with both accented and unaccented types are excluded.

[5] The only word with an ultimately accent in our data is [hakka] ‘八卦,’ which has a special reading and is seldom used in modern Japanese.

[6] The domain of a foot is indicated by ‘( )’ in this study.

[7] It is still debatable that a Chinese character (*kanji*) is a stem or a root. This will not be discussed in this study.

[8] In Japanese, disyllabic morphemes (CVCV) are hypothesized to have an underlying form CVC, because the second vowel is predictable in most cases (Itô and Mester 2015, for instance). Further evidence comes from the gemination of disyllabic Chinese character. For instance, /gak/(‘learn’ 学) and /gakkoo-/ (‘school’ 学校). If we assume the first Chinese character as /gaku/ in underlying form, the question will remain why /gakukoo/ is not acceptable. In the present study, it is further hypothesized that disyllabic SJ numerals also have CVC in underlying form but carry an accent on the second mora.

References

Akinaga, Kazue. 1999. *Tôkyôben-akusento-no Henyô*. The accent change in Tokyo dialect. Tokyo: Kasamashoin.

Alderete, John. 1999. *Morphologically Governed Accent in Optimality Theory*. Doctoral dissertation. University of Massachusetts Amherst. Online: http://roa.rutgers.edu/files/309-0399.

Itô, Junko and Armin Mester. 1994. Reflections on CodaCond and alignment. *Phonology at Santa Cruz.* 3.

Itô, Junko and Armin Mester. 2015. Sino-Japanese Phonology. *Handbook of Japanese Phonetics and Phonology*. 289-312. The Hague: Mouton de Gruyter.

Kubozono, Haruo. 1995. Constraint interaction in Japanese phonology: Evidence from compound accent. *Phonology at Santa Cruz*. 4. 21-38. University of California at Santa Cruz.

Kubozono, Haruo. 1997. Lexical markedness and variation: A nonderivational account of Japanese compound accent. *WCCFL*. 15. 273-287.

Kubozono, Haruo. 2006. Where does loanword prosody come from?: A case study of Japanese loanword accent. *Lingua*. 116(7). 1140-1170.

Kubozono, Haruo. forthcoming. Pitch Accent. ed. by Yoko Hasegawa. *The Cambridge Handbook of Japanese Linguistics*. Cambridge: Cambridge University Press.

Labrune, Laurence. 2012. *The Phonology of Japanese*. Oxford: Oxford University Press.

Martin, Samuel E. 1952. Morphophonemics of standard colloquial Japanese. *Language*. 28(3). 7-115.

Matsumura, Akira. 1999. *Daijirin*, ed. by Matsumura, Akira et al. Tokyo: Sanseido.

McCarthy, John J and Alan Prince. 1993. Prosodic morphology I: Constraint interaction and satisfaction. ms. University of Massachusetts, Amherst and Rutgers University.

McCawley, James David. 1968. *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.

Mutsukawa, Masahiko. 2009. *Japanese Loanword Phonology: The Nature of Inputs and The Loanword Sublexicon*. Tokyo: Hituzi Syobo.

Ogawa, Shinji. 2006. *Nihongo Shohôgen-no Nizi-kango-akusento*. Master thesis, Graduate School of Humanities and Faculty of Kobe Universisty. Online: http://researchmap.jp/?action=cv_download_main&upload_id=39301.

Prince, Alan and Paul Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction In Generative Grammar*. Blackwell Publishing.

Poser, Williams J. 1984. *The Phonetics and Phonology of Tone and Intonation in Japanese*. Doctoral dissertation, Massachusetts Institute of Technology. Online: www.ai.mit.edu/projects/dm/theses/poser84.pdf.

Poser, Williams J. 1990. Evidence for foot structure in Japanese. *Language*. 78-105.

Tanaka, Shin-ichi. 2004. Variation and Change in the Accentual System of Japanese Compounds: Implications for the Debate over Indexing vs. Reranking. *Tsukuba English Studies*. 22. 359-369.

Uwano, Zendô. 1997. Fukugô-meisi-kara mita nihongo-shohôgen-no akusento. *Akusento, Intoneisyon, Rizumu-to-pôzu*. 231-270. Tokyo: Sanseido.