Triple operations of rendaku processing: Native Chinese and Korean speakers learning Japanese *

Katsuo Tamaoka[1], Kyoko Hayakawa[2], and Timothy J. Vance[3]

[1]Nagoya University, [2]Kwansei Gakuin University, and [3]National Institute for Japanese Language and Linguistics

The present study investigated a hypothesis proposing the involvement of three operations in processing Japanese sequential voicing (rendaku): a lexical-specific operation, an etymology-specific operation, and a rule-based operation. Second language (L2) learners of Japanese are in the process of constructing an L2 mental lexicon of Japanese, and this lexicon is assumed to display a clear contrast between rule-based and etymology-specific occurrences of rendaku in early-stage learning and lexical-specific rendaku at later stages as a result of memory-based lexical learning. Native Chinese (*N*=32) and Korean (*N*=32) speakers learning Japanese, matched for their lexical and grammatical knowledge, avoided applying rendaku in compounds with a medial voiced obstruent in the second element, indicating that Lyman's Law is an active principle even in L2 acquisition. Both L2 learner groups also showed sensitivity to lexical strata by distinguishing Japanese-origin words (*wago*) from Sino-Japanese words (*kango*) and alphabetic loanwords (*gairaigo*). Thus, as factors of rendaku processing, Lyman's Law is considered rule-based while lexical stratum is considered etymology-specific. In contrast, both L2 learner groups showed a low occurrence of rendaku both for Lyman's Law exceptions (i.e., X+*basigo*) and for infrequent or rare words (i.e., X+*zyootyuu*). These instances can be considered memory-based, lexical-specific rendaku, which L2 learners must acquire as individual lexical items. This study indicated that all three described operations were used by L2 Japanese speakers to process rendaku.

Areas of interest: rendaku, triple operations, Lyman's Law, lexical strata, L2 Japanese learners

1. Introduction

In Japanese, when two words or morphemes are compounded, the initial consonant of the second element may be voiced; the compound of *ao* 'blue' and *sora* 'sky' becomes *aozora* 'blue sky'. This morphophonological phenomenon is called sequential voicing or rendaku. Rendaku refers to the voicing of the initial

voiceless obstruent in the second element of a compound word (e.g., Haraguchi, 2001; Irwin, 2005; Itô & Mester, 1986, 2003; Kubozono, 2005; McCawley, 1968; Otsu, 1980; Rosen, 2001; Vance, 1979, 1987, 2005, 2008, 2014a, 2014b). More than 30,000 rendaku compounds are listed in the rendaku database (Irwin, 2012; Irwin & Miyashita, 2016). Basically, the following initial voiceless consonants in second elements turn into the voiced consonants (Vance, 2016): /f/~/b/ *fue* 'flute' to *bue*, /h/~/b/ *hati* 'bee' to *bati*, /t/~/d/ *tori* 'bird' to *dori*, /k/~/g/ *kame* 'turtle' to *game*, /c/~/z/ *cuna* or *tuna* 'rope' to *zuna*, /s/~/z/ *sara* 'dish' to *zara*, /č/~/ǰ/ *tikara* /čikara/ 'power' to *zikara*, and /š/~/ǰ/ *sina* /šina/ 'item' to *zina* /ǰina/. Yet, it is not the case that all second elements of compounds begin with a voiced obstruent. It is also not obvious when rendaku should or should not apply in compounds. In the compound *kotori* 'a small bird,' the first element *ko* 'small' is combined with the second element *tori* 'bird,' and the initial consonant /t/ of the second element 'bird' is not voiced to /d/. In contrast, when the same morpheme *ko* 'small' is combined with *hako* 'box' to produce a compound, the initial consonant /h/ is voiced to /b/, producing *kobako* 'a small box.' As this pair of examples indicates, rules for rendaku seem to be very difficult to identify. However, this study shows that L2 learners take advantage of multiple operations when processing rendaku.

2. Triple operations in rendaku
This paper proposes that three possible operations can be used to process rendaku. The seemingly random distribution of rendaku suggests that some compound words are lexically specified as words that will undergo rendaku, while others are not. This can be understood as a lexical-specific process. For instance, when the first element *mugi* 'wheat' and the second element *hatake* 'field' are combined, the initial consonant of the second element /h/ is voiced to /b/, resulting in *mugibatake* 'wheat field.' However, when *naga* 'long' and *sikaku* 'square' are combined, the first consonant /s/ of the second element is not voiced as /z/, but remains voiceless as *nagasikaku* 'rectangle.' Once native Japanese speakers have learned these compounds, such as voiced *mugibatake* 'wheat field' and voiceless *nagasikaku* 'rectangle,' they are stored in the mental lexicon independent from their elements (*mugi* 'wheat' and *hatake* 'field,' and *naga* 'long' and *sikaku* 'square'). The irregularity of rendaku occurrences leads to the proposal that lexical-specific memory plays a role in understanding the distribution of rendaku, and that compound words are stored with phonological information about whether or not they co-occur with rendaku in the mental lexicon (i.e., phonological representations) as a result of learning each compound.

Although rendaku seems to occur randomly in compounds, there are cases where the distribution of rendaku is clearer. Lyman' Law is proposed as at least one consistently-applicable rule governing rendaku. Lyman's Law, named after Benjamin Smith Lyman (for details see Itô & Mester, 1986, 2003; Kawahara & Sano, 2014; Kubozono, 2005; Kubozono & Ota, 1998; Otsu, 1980; Vance, 2005,

2007, 2008), stipulates that the existence of a voiced obstruent in a second element prohibits rendaku. For instance, the second element *kaze* 'wind' is combined with *haru* 'spring' in *harukaze* 'spring wind,' and with *kami* 'divine' in *kamikaze* 'divine wind'. The word *kaze* already contains a voiced obstruent, so the initial consonant /k/ is not voiced to /g/. There is, however, a well-known exception to Lyman's Law; the compound of *nawa* 'rope' and *hasigo* 'ladder' appears with rendaku as *nawabasigo* 'rope ladder' (Kindaichi, 1976).

Lyman's Law is discussed as a particular case of the Obligatory Contour Principle (OCP). The OCP is a melodic level principle that prevents similar or identical phonological features from being repeated (e.g., Leben, 1973; Goldsmith, 1976; McCarthy, 1986). In case of Lyman's Law, when the second element of a compound word contains a voiced obstruent, rendaku is blocked to avoid having two voiced obstruent in a single element. As such, Lyman's Law can be considered rule-based.

Another possible operation used in governing the distribution of rendaku one that is etymology-specific. Rendaku most commonly occurs among Japanese origin words *wago*, in only a few cases in Sino-Japanese words *kango* (Okumura, 1955; Vance, 1996, 2007), and very seldom in alphabetic loanwords *gairaigo* (Irwin, 2011; Takayama, 1999). Some *kango* do exhibit rendaku in compounds. For instance, a compound of *kabusiki* 'stock' and *kaisya* 'company' is pronounced *kabusikigaisya* 'joint-stock company' as the initial consonant /k/ of the second element becomes voiced /g/. According to Nakagawa (1966), rendaku among *kango* is a barometer for 'nativization.' Irwin (2005) says that the proportion of words exhibiting rendaku is approximately 90% for *wago*, 20% for *kango* mononoms, 10% for *kango* binoms, and negligible for *gairaigo*. Thus, rendaku is restricted by Japanese lexical strata (e.g., Irwin, 2005; Itô and Mester, 1986, 2003; Kubozono, 1995). Native Japanese speakers formulate the rendaku etymology-specific operation based on the lexical type of the elements included in the compound (*wago*, *kango*, and *gairaigo*). Because rendaku is most frequent when combining *wago* elements, this process can be understood as an etymology-specific operation.

The present study proposes three different types of operations for rendaku processing. The triple operations can be also understood in the framework of phonological processing. Max Coltheart and his colleagues (e.g., M. Coltheart, Curtis, Atkins & Haller, 1993; M. Coltheart, & Rastle, 1994) proposed a dual-route model for phonological processing of visually-presented stimuli. This model explains that real words are processed by addressed phonology as whole word units stored in the mental lexicon. Given that compounds including phonological information (with/without rendaku) could be possibly stored as whole words in memory, this addressed phonology route might be applied to produce compounds with rendaku. In contrast, the dual-route model also suggests that pronounceable English nonwords are processed by assembled phonology applying grapheme-to-phoneme conversion rules. Other groups of

researchers utilizing the dual route models (e.g., V. Coltheart & Laxon, 1990; Patterson, 1986) suggests a sub-lexical or sub-word assembly of phonology, which could be by a group of letters or morphemic unit.

For rendaku, this paper proposes that the process of assembled phonology could be two-fold. Lyman's Law could be regarded as one aspect of assembled phonology when gathering phonological units, which functions like grapheme-to-phoneme conversion rules as proposed by the original dual route model (e.g., M. Coltheart et al., 1993, 1994). When native Japanese speakers perceive the inclusion of a voiced obstruent in the second element, they stop the voicing of the second element's initial obstruent. Since rendaku is produced when combining two morphemes or words, lexical strata can be regarded as another type of assembled phonology (e.g., V. Coltheart & Laxon, 1990; Patterson, 1986). When combining lexical elements, information about the specific lexical element, such as whether that element is *wago* or *kango*, or even *gairaigo*, affects voicing. As such, rendaku can be understood as an extension of the dual route model, that is, addressed and assembled phonology. However, in the case of rendaku processing, assembled phonology should be interpreted as two different operations, the purely phonological rule-based assembly (e.g., Lyman's Law) and the etymological-specific phonological-unit assembly (i.e., lexical strata).

The study of children with various deficiencies has contributed to a better understanding of rendaku processing. Fukuda and Fukuda (1994), for example, investigated rendaku in Japanese specifically language-impaired (SLI) children. Their study found that these SLI children were unlikely to voice the second elements of non-frequent and novel compounds. Fukuda and Fukuda (1994) suggested that the SLI children appeared to rely heavily on item-specific memory. Furthermore, Fukuda and Fukuda (1999) investigated the production of rendaku in six Japanese SLI children and age-matched non-SLI children. Their study indicated that SLI children did not voice most of the initial obstruents of the second elements in non-frequent and novel compounds. In contrast, the age-matched non-SLI children did voice the appropriate obstruents of all compounds. Fukuda and Fukuda (1994, 1999) concluded that SLI children either did not have or were unable to construct a productive rule of rendaku. SLI children seem to use only their lexical-specific memory or addressed phonology when processing rendaku. They also seem to be hindered in rule-based assembled rendaku production. In any case, as SLI children appear to use only a single type of processing, the lexical-specific approach, no further investigation is possible regarding rendaku in these participants.

An event-related potential (ERP) study by Kobayashi, Sugioka, and Ito (2014) clearly showed the triple operations of rendaku processing, although they proposed a Dual Mechanism Model of rule-based and memory-based processing. The ERP study by Kobayashi et al. (2014) contained three rendaku types, each of which showed a different EEG pattern. The existence of three ERP patterns

seems to support a theory of three operations, rather than a theory of bipartite operations.

The first pattern was observed in rendaku-immune words. Some words and morphemes of Japanese origin (i.e., *wago*) are never voiced. These words are called rendaku-immune (Irwin, 2009). For example, the Japanese-origin word, *hime* 'princess' is not voiced (i.e., no rendaku) no matter what it is combined with. Examples of this are *kaguyahime* 'the Moon Princess' from the Tale of the Bamboo Cutter, *sirayukihime* 'Snow White,' *utahime* 'songstress' etc. (taken from Asai, 2014). These rendaku-immune words were presented to native Japanese speakers in the unvoiced/non-rendaku (i.e., *hime*) and voiced/rendaku (i.e., **bime*) conditions embedded in a sentence. It should be noted that the issue of rendaku immunity only arises for words etymologically classified as *wago*, since *kango* and *gairaigo* words are typically rendaku-immune. Compared to the rendaku condition, the non-rendaku condition for rendaku-immune words elicited a negativity at the time range of 300-600 ms (interpreted as a LAN component, but it could also be interpreted as a N400 component) and a positivity around 400-800 ms (interpreted as P600). The N400 is the negative electrical brain activity observed around 400 ms in response to meaningful and potentially meaningful stimuli (e.g., Kutas & Federmeier, 2000) while the P600 is the positive electrical brain activity elicited by rule-based syntactic re-analysis phenomena (e.g., Friederici, 2002; Hagoort, 2003). Since both N400 and P600 were elicited for the condition of rendaku-immune words in *wago* lexical stratum, these words could be conceptually activated first (N400), and resist rendaku (P600) when making a compound because of the rendaku-immune property. This can be considered as an instance of etymology-specific processing.

The second pattern was observed in rendaku avoidance due to Lyman's Law. Kobayashi et al. (2014) presented a compound word with and without rendaku (e.g., *sunakabe* with no rendaku, and **sunagabe*; cf. *kabe* 'wall'). Lyman's Law systematically constrained the distribution of rendaku in all Kobayashi et al.'s stimuli – the initial obstruent of the second element could not be voiced. Because this resistance to rendaku is rule-based, we expect a P600 to be elicited. The ERP result indeed showed a positivity in the time range of 400 – 800 ms (interpreted as P600). Taking this as an evidence for implementation of Lyman's Law, Kobayashi et al. (2014) suggested that rule-based processing of rendaku is also possible.

The third pattern was observed in a condition involving neither Lyman's Law nor rendaku-immunity. Kobayashi et al. (2014) again presented a second element of a compound word with and without rendaku. For example, the compound of *mame* 'small' and *kai* 'shellfish' were combined as **mamekai* 'pea shellfish' without rendaku, and as *mamegai* with rendaku. In this pair, the voiced condition (rendaku) is correct while the voiceless condition (non-rendaku) is incorrect. Since the voiceless condition is incorrect, E400 is expected to be observed. In fact, the voiceless (non-rendaku) condition elicited a broad

negativity in the time range of 300 – 800 ms (interpreted as N400), when compared with the voiced (rendaku) condition. This result suggests that the voiceless condition was understood as containing a semantic violation. Thus, native Japanese speakers presumably understand these voiced (rendaku) compound words as an activation of a whole word using a lexical-specific operation.

Kobayashi et al. (2014) proposed the Dual Mechanism Model on the basis of the model proposed by Pinker (1991, 1994, 1999). Pinker's model hypothesized that the regular past tense is accomplished by rule-based regularity, just adding the morphemic inflection -*ed*, whereas the irregular past tense is processed with analogical memory. The ERP studies displayed the different ERP patterns for regular and irregular past tense in English (Morris & Holcomb, 2005) and in German (Penke, Weyerts, Gross, Zander, Münte, & Clahsen, 1997; Weyerts, Penke, Dohrn, Clahsen, & Münte, 1997). However, Pinker's model of verbal processing for past tense is not directly related to phonological processing. The ERPs of Kobayashi et al. (2014) displayed the three ERP patterns depending on the three different types of voiced/voiceless compounds (Kobayashi et al., 2014). We rather consider that these patterns support the triple operation processing model for rendaku; the pattern of both the N400 (or LAN) and P600 components for the etymology-specific operation applied to compounds with rendaku immune second elements, the pattern of only the P600 component for the rule-based operation applied to compounds involving Lyman's Law, and the pattern of only the N400 component for the lexical-specific operation applied to common rendaku compounds.

3. Five predictions about rendaku and L2 learners of Japanese
In L2 Japanese language education, rendaku seems not to have received much attention. Nakazawa, Vance, Irwin and Lyddon (2016) gave a rendaku survey to 234 Chinese-speaking students studying the Japanese language at Ming Chuan University in Taipei, Taiwan. Regarding the knowledge of rendaku, only 18 students responded that they were 'very familiar with it' (7.69%). 190 students had 'heard of it' (81.20%), 24 students had 'never heard of it' (10.26%), and 2 students gave no response (0.85%). When asked how they deal with learning rendaku, 77 students responded 'I just memorize' (32.9 %), 50 students said 'I use strategies to some extent'(21.37%), 102 students said 'I am not much interested' (43.59%), and 5 students did not respond (2.14%). A majority of these L2 students had heard of rendaku, but did not pay much attention to it. Given these responses, it might expected that, since the mental lexicon of these L2 learners is not clearly established during the learning process, they would display a clear contrast at least between rule-based and lexical-specific rendaku. L2 learners might implicitly perceive rules of Lyman's Law and the tendencies in lexical strata (e.g., rendaku mostly occurs in *wago*) in the early-stages of Japanese learning. In contrast, since they simply memorized compounds not

covered by these regularities, relatively rare compounds with rendaku will be acquired in the later stages as a result of memory-based lexical-specific learning.

By carrying out three experiments, the present study investigated five predictions involving the triple operations of rendaku: (1) the rule-based operation for Lyman's Law, (2) the lexical-specific operation for the Lyman's Law exception X+*basigo*, involving, to some degree, an etymology-specific operation for the first element, and (3) lexical-specific and etymology-specific for X+*zyootyuu*. The present study also examined an additional factor that might influence rendaku by comparing two groups of learners with different first (L1) languages: Chinese and Korean.

Prediction #1 involves Lyman's Law (for details, see Itô & Mester, 1986, 2003; Kawahara & Sano, 2014; Kubozono, 2005; Kubozono & Ota, 1998; Otsu, 1980; Vance, 2005, 2007, 2008), which states that rendaku does not occur when the second element of a compound word contains a medial voiced obstruent. Nakazawa et al. (2016) also questioned students on their knowledge of Lyman's Law. Only 5 students responded themselves as 'very familiar with it' (2.14%), 50 students had 'heard of it' (21.37%), 174 students had 'never heard of it' (74.36%), and 5 students did not respond (2.14%). Since L2 learners of Japanese had little knowledge of Lyman's Law according to the survey by Nakazawa et al. (2016), we might expect the L2 learners in our study not to apply this rule. However, on the other hand, if L2 learners can implicitly internalize Lyman's Law, they would show sensitivity to this rule. The first experiment tests whether L2 learners demonstrate such a sensitivity to Lyman's Law by measuring L2 learner's judgments for compounds containing second elements with/without voiced obstruents and first elements of different etymologies.

Prediction #2 also involves Lyman's Law. A well-known exception to Lyman's Law (Kindaichi, 1976; Otsu, 1980) is *nawabasigo* 'rope ladder,' a compound of *nawa* 'rope' and *hasigo* 'ladder'. The word *hasigo* contains a voiced obstruent, but native Japanese speakers frequently voice its initial consonant in compounds: X+*basigo* (Ihara & Murata, 2006). Although L2 learners with advanced Japanese proficiency most probably know *nawa* and *hasigo*, they would be unlikely to know that the second element of *hashigo* should be voiced as *basigo* when forming a compound. In such a case, following Lyman's Law, L2 learners will not voice the initial consonant of the second element: *nawahasigo*. The second experiment tests whether L2 learners treat X+*basigo* compounds as exceptions to Lyman's Law.

Prediction #3 involves lexical strata, i.e., the etymological categories *wago*, *kango* and *gairaigo*. Rendaku occurs more frequently when the second element is a *wago*, whereas it rarely occurs when the second element is a *kango* or *gairaigo*. Itô and Mester (2003) stated that "the class membership of the first member plays no role in the realization of linking morphemes" (p. 147). They provided three different examples of etymological types for first elements combined with the second elements *tayori* 'news' and *hanasi* 'a story':

ume-dayori 'news of plum blossoms' (*ume* 'plum'+ *tayori*) for a *wago* first element, *kankyoo-banasi* 'a talk about environment' (*kankyoo* 'environment' + *hanasi*) for a *kango* first element, and *supootsu-dayori* (*supootu* 'sports' + *tayori*) for a *gairaigo* first element (for more details see Itô & Mester, 2003, p. 147). Based on these examples, where rendaku occurs in compounds with first elements of three different etymological types, Itô and Mester concluded that characteristics of first elements have no influence on the rendaku of a second element. However, an experimental study by Ihara and Murata (2006) indicated that native Japanese speakers voice the initial consonant of *hasigo* 'ladder' more frequently when the first element is *wago* (e.g., *naga* 'long': *nagabasigo*) than when the first element is *kango* (e.g., *tetusei* 'iron') or *gairaigo* (e.g., *sutenresu* 'stainless steel'). Thus, the etymological type of the first element does seem to affect rendaku in the second element, and this characteristic is considered as a type of the etymology-specific operation. Thus, it is predicted that L2 learners of Japanese may develop sensitivity to the lexical stratum of the first element. The second experiment also investigates whether L2 learners demonstrate such sensitivity.

Prediction #4 is derived from a previous study by Tamaoka and Ikeda (2010) regarding X+*zyootyuu*. *Syootyuu* is a traditional Japanese alcoholic drink which is usually made from barley, sweet potato, rice, buckwheat, or brown sugar. The word *syootyuu* can be combined with the name of the particular main ingredient to produce a compound word. As a result, the initial consonant in *syootyuu* (sy = [ɕ]) frequently becomes voiced (zy = [(d)z]). In fact, native Japanese speakers displayed high rendaku ratios in these X+*zyootyuu* compounds (Tamaoka & Ikeda, 2010): (1) a *wago imo* 'sweet potato'+*shoochuu* showed the highest frequency of voicing at 93.83%, (2) a *wago kome* 'rice'+*shoochuu* (88.89%) and a *wago soba* 'buckwheat'+*shoochuu* (84.69%) showed similar percentages, (3) a *wago mugi* 'wheat'+*shoochuu* (72.59%) showed a significantly lower rate of voicing than *imo*, *kome* or *soba*, and (4) a *kango kokutoo* 'brown sugar'+*shoochuu*' (56.44%) was the lowest. These *syootyuu* compounds are lexically learned in each case by native Japanese speakers. However, L2 learners will not be so familiar to these X+*zyootyuu* compounds, so they may not be able to apply rendaku to the second element *syootyuu* in a manner similar to native Japanese speakers. Therefore, it is predicted that L2 learners simply combine *syootyuu* without rendaku, resulting in lower rendaku ratios for all X+*zyootyuu* compounds.

Prediction #5 involves the influence of the learner's first language (L1) on rendaku. Korean has a phenomenon called *sai-sori* that resembles rendaku (Labrune, 2013). The result of combining *mul* 'water' and *koki* 'meat' is *mulk'oki* 'fish' (where *k* is lax and *k'* is tense). In contrast, Chinese does not provide its native speakers with anything similar. Although third-tone sandhi results in a pitch change, it does not create a segmental change. Thus, Prediction #5 is that native speakers of Korean will be more sensitive to Japanese rendaku

than native speakers of Chinese. Two L2 groups (Chinese and Korean) were compared in all three experiments.

In order to investigate these five predictions, the present study conducted three rendaku experiments with two different groups of Chinese and Korean L1 speakers learning L2 Japanese. This study aims to show that L2 learners make use of three separate operations when dealing with rendaku: rule-based, etymology-specific and lexical-specific operations.

4. Experiments

4.1 Participants

Native Chinese (N=32; 20 females and 12 males) and Korean (N=32; 17 females and 15 males) speakers, who had been living in Japan and majoring in various subjects at Japanese universities, participated in the three experiments reported here. The two groups were similar in age at the time of the experiments. The 32 Chinese students had an average age of 25 years and one month with a standard deviation of three years and two months, while the 32 Korean students had an average age of 24 years and six months with a standard deviation of three years and 10 months, $t(62)$=0.70, *ns*.

A questionnaire regarding Japanese learning was administered to all participants. The two groups were similar with respect to how long they had been studying Japanese (six students did not report, but these responses were treated as missing values for this calculation, and their data was still included in the experiments): the 30 Chinese-speaking learners who responded had studied an average of six years and zero months with a standard deviation of two years and four months, while the 28 Korean-speaking learners who responded had studied an average of five years and nine months with a standard deviation of two years and 11 months. There was no statistically significant difference in length of Japanese study between the two groups, $t(56)$=0.34, *ns*. The two groups also did not differ with respect to length of residence in Japan. The average for the 31 Chinese participants who responded was two years and 3 months, and the average for the 30 Korean participants who responded was two years and four months, $t(59)$=0.611, *ns*. Self-reported Japanese usage (hours per week, speaking and listening) was also very similar, with 31 Chinese speakers reporting an average of 10.6 hours and 30 Korean speakers reporting and average of 10.4 hours. Again, there was no significant difference, $t(59)$=0.140, *ns*.

More importantly, the Japanese proficiency levels of all participants were measured by a lexical knowledge test (Miyaoka, Tamaoka, & Sakai, 2011) and a grammatical knowledge test (Miyaoka, Tamaoka, & Sakai, 2014). The lexical knowledge test consisted of 48 questions (maximum points = 48) sub-divided into 12 *wago*, 12 *kango*, 12 *gairaigo* and 12 function words. This test has been used in many previous studies, and displays a very high Cronbach's alpha reliability: $N = 278$, $M = 22.06$, $α = 0.87$ for native Chinese speakers learning Japanese in China (Chu & Tamaoka, 2013), $N = 127$, $M = 32.32$, $α = 0.85$ for

native Chinese speakers learning Japanese in China (Yamato, Tamaoka, & Chu, 2013), $N = 238$, $M = 27.34$, $\alpha = 0.85$ for native Chinese speakers learning Japanese in China (Komori, Tamaoka, Saito, & Miyaoka, 2014), and $N = 78$, $M = 28.42$, $\alpha = 0.88$ for native Korean speakers learning Japanese in Korea (Tamaoka, Miyaoka, Kim, & Lim, 2011). In the present study, no statistically significant difference in lexical knowledge test scores was found between the 32 native Chinese speakers ($M = 38.63$, $SD = 4.67$) and the 32 native Koreans speakers ($M=39.00$, $SD=5.66$), $t(62)=0.29$, ns. In addition, their mean scores, showing about 80 percent accuracy, were much higher than those in the previous studies conducted on students learning Japanese at universities in China and Korea. This suggests that the participants in the present study had rich Japanese vocabularies.

The grammatical knowledge test consisted of 36 questions (maximum points = 36) sub-divided into 12 questions regarding morphological inflections, 12 questions regarding local dependency, and 12 questions regarding complex structure. This test also showed a very high Cronbach's alpha reliability of $N = 278$, $M = 23.43$, $\alpha = 0.88$ for native Chinese speakers learning Japanese in China (Chu & Tamaoka, 2013), but a relatively lower reliability of $N = 127$, $M = 27.86$, $\alpha = 0.69$ for native Chinese speakers learning Japanese in China (Yamato, Tamaoka, & Chu, 2013). No significant difference in scores on the grammatical knowledge test was found between the 32 native Chinese speakers ($M = 31.97$, $SD = 2.13$) and the 32 native Koreans speakers ($M = 32.69$, $SD = 2.26$), $t(62)= 1.31$, ns. The mean scores of both L1 groups were about 90 percent accuracy, much higher than the previous studies conducted on students at universities in China. This suggests that the participants in the present study had an advanced understanding of Japanese grammar.

The Chinese and Korean learners of Japanese in the present study were similar in multiple respects and had attained highly advanced lexical and grammatical knowledge, which makes it possible to directly compare the responses of these two groups in the three rendaku experiments.

4.2 Experiment 1 – Rule-based operation of Lyman's Law
The phonological process of rendaku is not introduced systematically in Japanese language education for foreign students at Chinese and Korean universities. As reported by Nakazawa et al. (2016), Chinese L1 speakers learning L2 Japanese lack conscious awareness of rendaku, although they are somewhat aware of rendaku. Therefore, the participants in the present study were expected to have little knowledge of Lyman's Law.

4.2.1 Stimuli
Experiment 1 used eight compounds: four with and four without a voiced obstruent in the second element. The compounds with a voiced obstruent in the second element were 黒羊 *kuro+hituzi* 'black sheep,' 親雀 *oya+suzume* 'parent

sparrow,' 鉄釘 *tetu+kugi* 'iron nail,' and 合鍵 *ai+kagi* 'duplicate key'. Lexical stratum (*wago*, *kango*, or *gairaigo*) was determined for each element. Word frequency was also recorded, based on the Mainichi Newspaper corpus (containing all articles published between 2000 and 2010). Finally, each item's level on the Japanese Language Proficiency Test (JLPT) created by the Japan Foundation and Association of International Education, Japan (2004) was determined (1st to 4th, or 0 for unlisted items). The lexical stratum, frequency, and JLPT level of each item are as follows: *oya* (*wago*, 24166, 2nd), *suzume* (*wago*, 1339, 0), *tetu* (*kango*, 5128, 2nd), *kugi* (*wago*, 2014, 2nd), *a(u)* (*wago*, 9331, 3rd), *kagi* (*wago*, 10653, 4th), *kuro* (*wago*, 37153, 4th), *hituzi* (*wago*, 2629, 0), *huru(i)* (*wago*, 4101, 4th), and *tanuki* (*wago*, 865, 0). The compounds with no voiced obstruent in the second element were 歯車 *ha+guruma* 'gear,' 生魚 *nama+zakana* 'raw fish,' 蜜蜂 *mitu+bati* 'honeybee' and 古狸 *huru+danuki* 'old raccoon'. The lexical stratum, frequency, and JLPT level of these elements are: *ha* (*wago*, 4591, 4th), *kuruma* (*wago*, 29746, 4th), *nama* (*wago*, 7382, 2nd), *sakana* (*wago*, 7716, 4th), *mitu* (*kango*, 714, 2nd), and *hati* (*wago*, 670, 0). Although four of the elements were not included in the JLPT, these elements are considered familiar, high-frequency words or morphemes.

4.2.2 Procedure
All participants were presented with two versions of the compound and asked to select which was correct: the compound with or without voicing of the initial consonant of the second member of the compound. For example, 生魚 *nama+zakana* was written in kanji and hiragana as 生 (なま) ＋ 魚 (さかな). The two choices were presented in hiragana:なまさかな *namasakana* (without rendaku) and なまざかな *namazakana* (with rendaku). Participants were asked to check one pronunciation or the other for each item. This same method was employed in all three experiments. Table 1 shows the number of students who selected each pronunciation for each item. The label "voiced" denotes the pronunciation with rendaku, i.e., with a voiced consonant at the beginning of the second element, and the label "voiceless" label denotes the pronunciation without rendaku, i.e., with a voiceless consonant at the beginning of the second element.

4.2.3 Analysis and results
In order to determine the strength of the factors influencing the voiced/voiceless decisions, a decision tree analysis was used in the present study. A decision tree analysis depicts the results in a tree-like graph referred to as a dendrogram. Decision trees are commonly used in operations research to help identify a strategy most likely to reach a goal. In the present study, a decision tree analysis (more specifically, a classification tree analysis for predicting a categorical variable; IBM SPSS Statistics Version 21.0) with the chi-squared automatic

interaction detector (CHAID) algorithm was used to predict voiced/voiceless (i.e., ±rendaku) decisions by three variables: (1) first language (e.g., L1 language), (2) presence or absence of a medial voiced obstruent in the second element of the compound, and (3) the identity of the four different words of each type, the two types being with or without a voiced obstruent in the second element. The decision tree analysis also provided an overall estimate of relative risk. In this analysis, it was 24.02% with a standard error of 1.89%, suggesting that 78.98% of voiced/voiceless decisions were correctly classified by the three variables.

Table 1
Frequencies of voiced/voiceless decisions by Chinese and Koreans in Experiment 1

Condition	Compounds in sounds	Chinese (N=32)			Koreans (N=32)		
		Voiced	Voiceless	% voiced	Voiced	Voiceless	% voiced
Voiced obstruent	*kuro+hituzi*	4	28	12.50%	3	29	9.38%
	oya+suzume	5	27	15.63%	6	26	18.75%
	tetu+kugi	12	20	37.50%	24	8	75.00%
	ai+kagi	10	22	31.25%	10	22	31.25%
Sum and voiced rate (%)		31	97	24.22%	43	85	33.59%
Voiced obstruent <u>not</u> included	*ha+guruma*	32	0	100.00%	32	0	100.00%
	nama+zakana	23	9	71.88%	28	4	87.50%
	mitu+bati	23	9	71.88%	22	10	68.75%
	huru+danuki	18	14	56.25%	21	11	65.63%
Sum and voiced rate (%)		96	32	75.00%	103	25	80.47%

Note: Gray-colored columns indicate correct responses. White-colored columns indicate incorrect responses

The results of the decision tree analysis are depicted in the decision tree dendrogram in Figure 1. The dendrogram clearly indicates that, regardless of L1 (automatically excluded in the tree), all participants were likely to avoid applying rendaku in compounds containing a voiced obstruent in the second element: 28.91% for compounds with a voiced obstruent in the second element (Node #1), versus 77.73% for compounds without (Node #2) [χ^2(1)=122.61, *p*<.001]. The difference in the rendaku percentage between the two categories was very large (48.82%), supporting the conclusion that Lyman's Law is a rule-based principle adhered to even by L2 learners of Japanese. This result was not related to a speaker's L1.

The factor of stimulus differences ranked below the robust factor of Lyman's Law in the dendrogram. The detailed differences among stimulus compounds are shown in Figure 1 for both types of compound (i.e., those with and those without a medial voiced obstruent in the second element). In the case of compounds involving Lyman's Law (inclusion of a voiced obstruent in the second element), the compound with the *kango* first element *tetu* (56.25%) was voiced more frequently than the compounds with the *wago* first elements (see Node #5 in Figure 1). This pattern indicates that the *wago* first elements were

recognized and favored the voiceless form in accordance with Lyman's Law. In contrast, in the case of compounds without a voiced obstruent in the second element, the *kango* first element *mitu* (and the *wago huru*) showed a lower rendaku ratio of 65.62%, compared to the high ratios for the *wago* first elements, *ha* (100%) and *nama* (79.69%). These results partly support sensitivity to etymology.

4.2.4 Discussion

Regardless of the participant's L1, Experiment 1 showed a clear contrast between the second elements of the compounds with/without voiced obstruents. Those compounds without voiced obstruents displayed a high rendaku ratio, whereas rendaku was avoided in those with voiced obstruents. Therefore, the rule-based operation of Lyman's Law was well-perceived by both L1 Chinese and L1 Korean participants. This result was different from Nakazawa et al. (2016), which reported that L1 Chinese speakers displayed an equally high rendaku rate for both compounds with and without voiced obstruents in second elements. The different results between the present study and Nakazawa et al. (2016) could be caused by Japanese language proficiency. The participants in the present study were highly advanced in both their lexical and grammatical knowledge of Japanese, and had been studying at a university in Japan. Nakazawa et al. (2016) investigated students living in Taiwan, who may have not had enough lexical knowledge to implicitly learn Lyman's Law. In addition, Experiment 1 partially showed a partial sensitivity to the etymology of the first element of the compound, suggesting that participants had some degree of awareness of the etymology-specific operation

4.3 Experiment 2 – Lexical-item-specific rendaku: An exception to Lyman's Law

Experiment 2 examined the occurrence of rendaku in L2 learners' Japanese by using first elements from three different lexical strata (*wago*, *kango* and *gairaigo*) in compound words with the second elements *hasigo* and *hasami*.

4.3.1 Stimuli

The first elements, *nawa*, *tetusei*, and *sutiiru*, were used with *hasigo*, and *hagane*, and *tetusei*, and *sutiiru* were used with *hasami*. The lexical stratum, frequency, and JLPT level of these elements are: *nawa* (*wago*, 461, 2nd), *tetusei* (*kango*, 552, *tetu*=2nd, *sei*=3rd), *sutiiru* (*gairaigo*, 693, 0), *hagane* (*wago*, 304, 0), *hasigo* (*wago*, 847, 2nd), and *hasami* (*wago*, 954, 2nd).
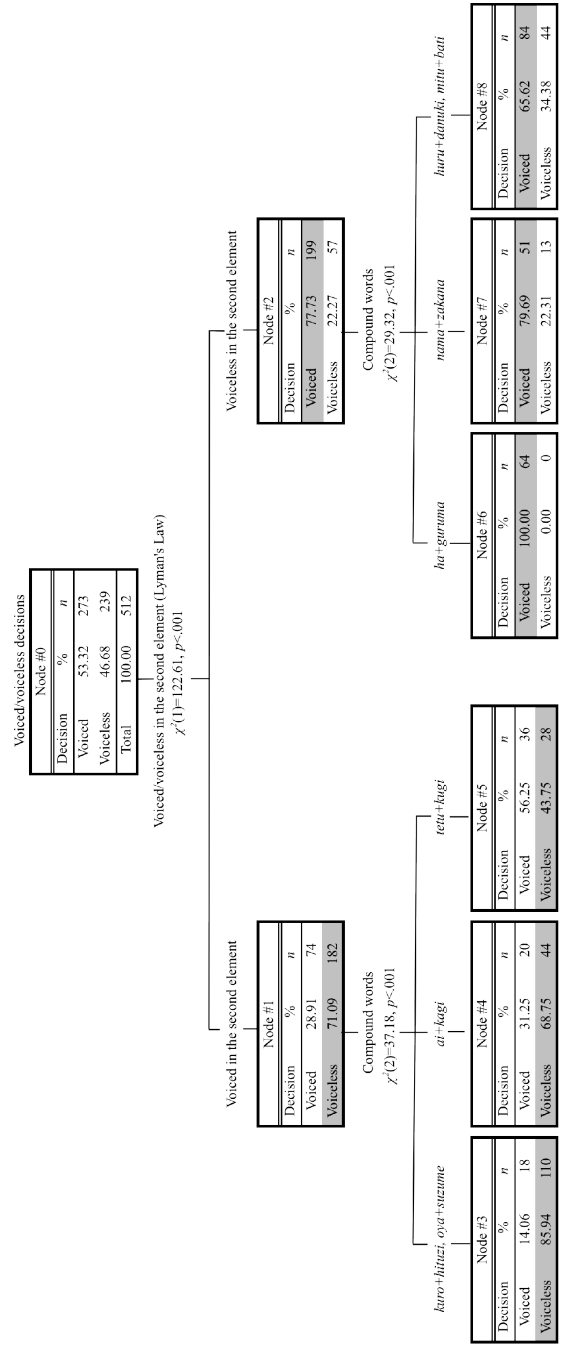
Voiced/voiceless decisions

Node #0

Decision	%	n
Voiced	53.32	273
Voiceless	46.68	239
Total	100.00	512

Voiced/voiceless in the second element (Lyman's Law)
$\chi^2(1)=122.61, p<.001$

Voiced in the second element

Node #1

Decision	%	n
Voiced	28.91	74
Voiceless	71.09	182

Voiceless in the second element

Node #2

Decision	%	n
Voiced	77.73	199
Voiceless	22.27	57

Compound words
$\chi^2(2)=37.18, p<.001$

kuro+hituzi, oya+suzume

Node #3

Decision	%	n
Voiced	14.06	18
Voiceless	85.94	110

ati+kagi

Node #4

Decision	%	n
Voiced	31.25	20
Voiceless	68.75	44

tetu+kugi

Node #5

Decision	%	n
Voiced	56.25	36
Voiceless	43.75	28

Compound words
$\chi^2(2)=29.32, p<.001$

ha+guruma

Node #6

Decision	%	n
Voiced	100.00	64
Voiceless	0.00	0

nama+zakana

Node #7

Decision	%	n
Voiced	79.69	51
Voiceless	22.31	13

huru+danuki, mitu+bati

Node #8

Decision	%	n
Voiced	65.62	84
Voiceless	34.38	44

Figure 1. Influence of Lyman's Law for voiced/voiceless decisions by native Chinese and Korean speakers

Note 1: N=64 (32 Chinese and 32 Koreans).

Note 2: Gray-colored columns indicate more frequently selected choices.

4.3.2 Procedure
Experiment 2 utilized the same procedure as in Experiment 1, where participants were asked to choose a reading with or without rendaku.

4.3.3 Analysis and results
The numbers of Chinese and Korean students who selected the rendaku pronunciation (voiced) and the non-rendaku pronunciation (voiceless) in each case are reported in Table 2. A decision tree analysis was conducted to predict the voiced/voiceless decisions by (1) native language (Chinese and Korean), (2) lexical stratum of the first element, and (3) second element (*hasigo* and *hasami*). The overall estimate of relative risk in the analysis was 39.58% with a standard error of 2.50%. The result of the decision tree analysis is drawn as a dendrogram in Figure 2.

Table 2
*Voiced/voiceless decisions for X-*hasigo *and X-*hasami *by native Chinese and Korean speakers*

X-*hashigo and* X-*hasami*	Lexical strata	Chinese (*N*=32)			Korean (*N*=32)		
		Voiced	Voiceless	% voiced	Voiced	Voiceless	% voiced
nawa+hasigo	wago	15	17	46.88%	19	13	59.38%
tetusei+hasigo	kango	8	24	25.00%	15	17	46.88%
sutiiru+hasigo	gairaigo	8	24	25.00%	12	20	37.50%
hagane+hasami	wago	17	15	53.13%	20	12	62.50%
tetusei+hasami	kango	11	21	34.38%	18	14	56.25%
sutiiru+hasami	gairaigo	12	20	37.50%	11	21	34.38%

As shown in the decision tree dendrogram of Figure 2, lexical strata elicited differences in that *wago* (55.47%) in Node #1 had a higher rendaku frequency than both *kango* and *gairaigo* together (37.11%) in Node #2 [$\chi^2(1)$=11.72, *p*<.01]. The decision tree automatically combined both *kango* and *gairago*, indicating that these two categories showed a similar ratio trend. No difference was indicated in *wago* (no tree is grown from Node #1), so neither the Chinese nor the Korean L1 group differed in rendaku decisions between the exceptional X+*basigo* compound (*nawa+hasigo*) and non-exceptional X+*basami* compound (*hagane+hasami*). Neither L1 group had acquired *nawabasigo* 'rope ladder', an exceptional case that must be learned as a special lexical item. Additionally, as seen in Node #3 and Node #4 in Figure 2, Chinese learners were less likely than Korean learners (by a margin of 13.28%) to apply rendaku in compounds with the *kango* first element *tetusei*. This could be due to their L1, since Chinese speakers have a clear sense of *kango*, which are unlikely to undergo rendaku.
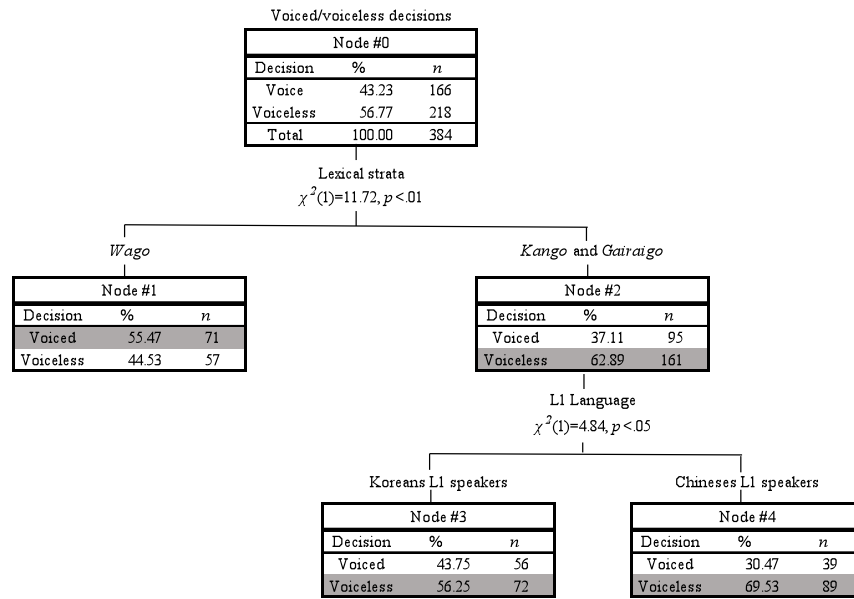
Figure 2. Lexical strata and Chinese/Korean L1 Languages for voiced/voiceless decisions
Note: *N*=64 (32 Chinese and 32 Koreans).

4.3.4 Discussion

As Tamaoka, Ihara, Murata, and Lim (2009) demonstrated, native Japanese speakers are sensitive to the lexical strata of both the first and the second element when determining whether a specific second element will have rendaku. The present study has also demonstrated sensitivity to lexical strata by showing that both Chinese and Korean L1 speakers learning Japanese differentiated *wago* from *kango* and *gairaigo* when determining the distribution of rendaku. Thus, the behavior of L2 Japanese learners is consistent with the idea that rendaku in different lexical strata must involve an etymology-specific operation. In addition, Experiment 2 indicated that L1 Chinese learners of Japanese were sensitive to the lexical stratum of *kango*, whereas L1 Korean learners were not.

4.4 Experiment 3 – Lexical-item specific rendaku in *syootyuu*

The word *syootyuu* can be combined with the name of a particular main ingredient to produce a compound word. *Syootyuu* is one of the few *kango* words that undergoes rendaku. Native Japanese speakers voice the initial consonant in *syootyuu* frequently in such compounds (Tamaoka & Ikeda, 2010). Given this behavior, compounds with *syootyuu* as the second element are assumed to be

memorized as specific lexical items. Experiment 3 examined how L2 learners of Japanese have acquired this type of lexical-specific items.

4.4.1 Stimuli

The five ingredients *imo* 'sweet potato,' *kome* 'rice,' *soba* 'buckwheat,' *mugi* 'barley' and *kokutoo* 'brown sugar' were compounded with *syootyuu* to investigate whether L1 Chinese and L1 Korean learners of L2 Japanese voice the initial consonant of the second element *syootyuu* in the same way as native Japanese speakers do. The lexical stratum, frequency, and JLPT level of each word were: *soba* (*wago*, 5947, 4th), *imo* (*wago*, 1432, 0), *mugi* (*wago*, 1379, 0), *kome* (*wago*, 5439, 3rd), *kokutoo* (*kango*, 71, 0), and *syootyuu* (*kango*, 1521, 0). *Kokutoo* is less frequently used than the other words.

4.4.2 Procedure

Experiment 3 used the same procedure as with Experiments 1 and 2, where participants were required to select either a rendaku (voiced) or non-rendaku (voiceless) pronunciation of each compound.

4.4.3 Analysis and results

The voiced/voiceless frequencies for X-*zyootyuu* rendaku are shown in Table 3. Table 3 shows a clear overall trend for L1 Chinese and L1 Korean learners. As with Experiments 1 and 2, Experiment 3 also used a decision tree analysis, to predict voiced/voiceless decisions by (1) the first elements of the five ingredients, i.e., *imo*, *kome*, *soba*, *mugi* and *kokutoo*, and (2) the L1 (Chinese and Korean) of the Japanese learners. The decision tree analysis drew no dendrogram (a total voiced frequency of 79 or 24.69%, and a total voiceless frequency of 241, or 75.31%), suggesting neither the ingredients (the first elements) nor the native language affected rendaku. In other words, no difference in rendaku ratios was found for the first element of the compounds, and this trend did not differ between L1 Chinese and L1 Korean speakers. To make sure this null difference was real, a 5 (five ingredients) $\times$ 2 (voiced or voiceless) chi-square test of independence was separately conducted for each language group. There were no differences among the five ingredients for either the L1 Chinese speakers [$\chi^2(4)=1.93$, *ns*] or the L1 Korean speakers [$\chi^2(4)=1.77$, *ns*].

Table 3
Voiced/voiceless decisions for X-zyootyuu by L1 Chinese and L1 Korean learners of Japanese

Type of *syootyuu*	Chinese (*N*=32)			Koreans (*N*=32)		
	Voiced	Voiceless	% voiced	Voiced	Voiceless	% voiced
imo + *syootyuu*	8	24	25.00%	9	23	28.13%
kome + *syootyuu*	6	26	18.75%	10	22	31.25%
soba + *syootyuu*	8	24	25.00%	9	23	28.13%
mugi + *syootyuu*	10	22	31.25%	6	26	18.75%
kokutoo + *syootyuu*	6	26	18.75%	7	25	21.88%

Unlike native Japanese speakers (Tamaoka & Ikeda, 2010), the voiced (rendaku) ratios are very low for both L1 Chinese and L1 Korean speakers. Tamaoka and Ikeda (2010) reported a rendaku trend among 405 native Japanese speakers from six different prefectures (or regions): Kagoshima, Oita, Fukuoka, Yamaguchi, Hiroshima, and Shizuoka. Regardless of regional differences native speakers showed a very high rendaku rate in *imo+zyootyuu* at 93.83%, *kome+zyootyuu* at 88.89%, *soba+zyootyuu* at 84.69% and *mugi+zyootyuu* at 72.59%, and a relatively low rate in *kokutoo+zyootyuu* at 56.44%. In order to make the differences between learners of Japanese and native Japanese speakers visually clear, voiced (rendaku) ratios are depicted in Figure 3. In contrast to the high frequency of rendaku in *syootyuu* for native Japanese speakers (Tamaoka & Ikeda, 2010), rendaku rates for L1 Chinese and L1 Korean speakers learning Japanese were very low, ranging from 18.75% to 31.25% (see also Table 3**).**

4.4.4 Discussion
Experiment 3 showed very low voiced (rendaku) ratios for L1 Chinese and L1 Korean speakers. We interpret these results as indicating that these Chinese and Korean speakers learning Japanese are unaware that *syootyuu* can undergo rendaku. This trend of low voiced ratios in the responses of the L2 learners supports the theory that X-*zyootyuu* compounds are a part of the lexical-specific operation. The absence of such a tendency in the L1 Chinese and L1 Korean learners of Japanese also supports the idea that *syootyuu* is not marked by the lexical-specific operation in the L2 learner's mental lexicon. It should be noted that *kokutoo* is *kango*, so one might wish to draw the conclusion that *kango* are less prone to rendaku than *wago*. In fact, native Japanese speakers applied rendaku following *kokutoo* less frequently than following other ingredient names (Tamaoka & Ikeda, 2010). However, since the lengths of the first and second elements affect rendaku frequency (Tamaoka et al., 2009), four-mora CV+CV+CV+V *kokutoo* cannot be directly compared to the other first elements, which all have two-mora (C)V+CV structure (*imo*, *kome*, *soba*, *mugi*).
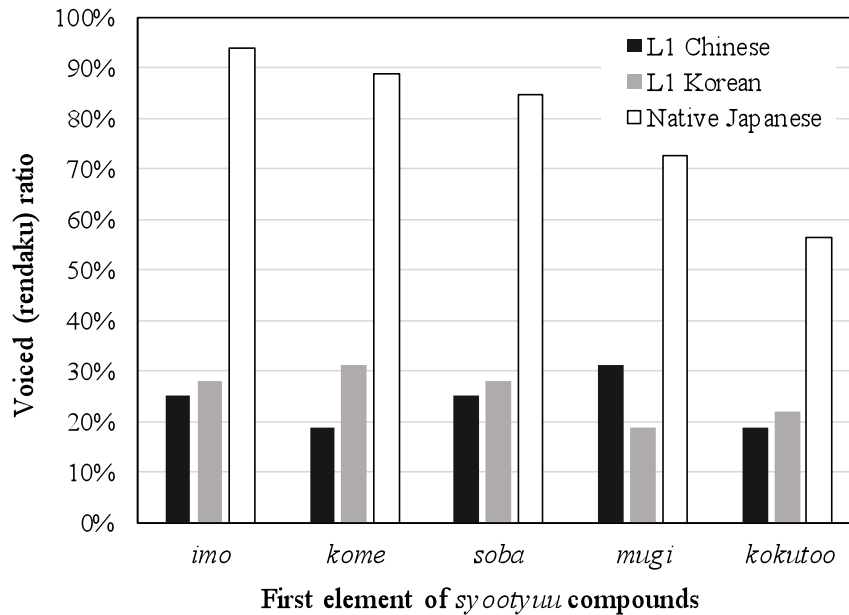
Figure 3. Voiced ratios for X-*zyootyuu* by L1 Chinese and L1 Korean speakers and native Japanese speakers
Note: Thirty-two Chinese, 32 Koreans, and 405 native Japanese. The frequencies of voiced ratios among native Japanese speakers were taken from Tamaoka and Ikeda (2010).

5. Overall discussion

Using L1 Chinese and L1 Korean learners of L2 Japanese, the present study attempted to demonstrate the triple operations of rule-based, etymology-specific and lexical-specific rendaku operations (see Section 2), an assumption drawn from Kobayashi et al. (2014). This notion of triple operations is also consistent with the phonological processing of the dual route model, addressed and assembled phonology (e.g., M. Coltheart, Curtis, Atkins, & Haller, 1993; M. Coltheart & Rastle, 1994). With the addition of sub-lexical or sub-word assembly of phonology (e.g., V. Coltheart & Laxon, 1990; Patterson, 1986), assembled phonology is interpreted as two different operations, the rule-based (e.g., Lyman's Law) and the etymological-specific (i.e., lexical strata). Kobayashi et al. (2014) clearly displayed three different ERP patterns for the three types of rendaku operations: the single P600 pattern for the rule-based operation (Lyman's Law), the double N400 (or LAN) and P600 pattern for the etymology-specific operation, and the single N400 pattern for the lexical-specific operation. Within the framework of the triple rendaku operation theory, five

predictions (see Section 3) for L1 Chinese and L1 Korean learners of Japanese were evaluated using the results of the three experiments (see Section 4) reported in the present study.

The rule-based operation of Lyman's Law was consistently observed in both L1 Chinese and L1 Korean speakers learning L2 Japanese. Given Lyman's Law (see Itô & Mester, 1986, 2003; Kawahara & Sano, 2014; Kubozono, 2005; Kubozono & Ota, 1998; Otsu, 1980; Vance, 2005, 2007, 2008), L1 Chinese and L1 Korean speakers were expected to avoid rendaku in compounds containing a voiced obstruent in the second element (Prediction #1). Experiment 1 indicated that learners' L1 background was not a factor, since both L1 Chinese and L1 Korean speakers were likely to avoid applying rendaku when it would violate Lyman's Law. Although differences among stimulus compounds were found (see Figure 1), these were observed within the overall robust trend governed by Lyman's Law. Thus, combined with results from native Japanese speakers (Ihara & Murata, 2006; Tamaoka et al., 2009), these results are consistent with the idea that Lyman's Law reflects a principle as a sub-case of the OCP (e.g., Goldsmith, 1976; Leben, 1973; McCarthy, 1986), which refers to the tendency to avoid repetition of the same or similar phonological features.

The compound word *nawabasigo* 'rope ladder' is a well-known exception to Lyman's Law (Kindaichi, 1976; Otsu, 1980). Murata (1984) reported a high rendaku rate for native Japanese speakers in compounds with *hasigo* as the second element. Of 197 native Japanese speakers, 190 (96.45%) applied rendaku in a compound of *naga* 'long' and *hasigo* 'ladder,' pronouncing it as *nagabasigo*. However, since this word is rare, appearing only 43 times in 11 years in the Mainichi Newspaper (the authors' original calculation), compounds of the form X+*basigo* will generally not be stored in the mental lexicons of L2 Japanese learners. In other words, the participants in this study were expected not to choose X+*basigo* (Prediction #2). As predicted, Experiment 2 showed that only 46.88% of Chinese learners and 59.38% of Korean learners chose the rendaku form *nawabasigo* instead of *nawahasigo*. The behavior of the second element *hasigo* in compounds must be learned as a particular fact, i.e., it is lexical-specific rendaku. On the other hand, L2 learners of Japanese did exhibit sensitivity to lexical strata (e.g., Irwin, 2005; Itô and Mester, 1986, 2003; Kubozono, 1995) by distinguishing the first elements of *wago* from *kango* and *gairaigo* (Prediction #3 and see Figure 2). Contrary to Itô and Mester's (2003) suggestion that the lexical stratum of the first element has no effect on rendaku, L2 learners of Japanese, like native Japanese speakers (Tamaoka et al., 2009), exhibited sensitivity to the lexical stratum as the etymology-specific operation.

Compounds of X-*zyootyuu* are rather rare words for L1 Chinese and L1 Korean speakers learning L2 Japanese. In addition, the element *syootyuu* is a *kango* word. This lexical category shows rendaku voicing less frequently than *wago* (Irwin, 2005). Therefore, despite the high rate of rendaku in *syootyuu* for native Japanese speakers (Tamaoka & Ikeda, 2010), these learners of Japanese

were not expected to apply rendaku to the element *syootyuu* (Prediction #4). Experiment 3 showed that the voiced (rendaku) rates for L1 Chinese and L1 Korean speakers were very low, ranging from 18.75% to 31.25% (see Table 3 and Figure 3). Thus, it appears that *syootyuu* compounds are lexical-specific items which must be added to the L2 mental lexicon at a very late stage of Japanese learning.

The only influence of L1 (Chinese vs. Korean) (Prediction #5) was a difference between *kango* and *gairaigo* in compounds ending in *hasigo~basigo* (see Figure 2). L1 Chinese speakers were better than L1 Korean speakers in terms of not applying rendaku in compounds ending in *hasami~basami* when the first element is *kango* (see Figure 2). Native Chinese speakers are familiar with kanji and *kango*. Among the 2,060 basic two-kanji compound words in levels 4 to 2 of the Japanese proficiency test (Japan Foundation and Association of International Education, Japan, 2004), 1,509 words (73.25%) are orthographically and conceptually similar across Chinese and Japanese (Park, Xiong, & Tamaoka, 2014; Xiong & Tamaoka, 2014). This database is available on the Web-site with a search engine (http://kanjigodb.herokuapp.com; for the usage of the search engine, see Yu & Tamaoka, 2015). Native Chinese speakers learning Japanese may be able to detect *kango* more accurately than native Korean speakers. If, as Nakagawa (1966) suggested, rendaku in *kango* is a barometer of 'nativization,' familiarity with *kango* may be the reason that L1 Chinese learners had a lower rendaku ratio than L1 Korean learners for X-*basigo* compounds with the *kango* first element.

In summary, L1 Chinese and L1 Korean learners of Japanese displayed a pattern of responses that support a clear contrast: consistent appearance of rule-based (Lyman's Law) and etymology-specific (lexical strata) rendaku operations in the early stages of learning, followed by lexical-specific rendaku operations (X-*basigo* and X-*zyootyuu*) at later stages, as a result of memory-based lexical learning.

References

Asai, A. (2014). *Rendaku seiki-no keikoo to teichaku-ka* [The productivity of rendaku sequential voicing and its fixity in frequent occurrences]. *NINJAL Research Papers, 7,* 27–44.

Chu, X., & Tamaoka, K. (2013). *Chuugokujin nihongo gakushuusha-niyoru goi/bunpoo chishiki kara jutsubu koozoo-no rikai-o chuukai-toshita dokkai-eno ingakankei moderu-no kenshoo* [An investigation of causal relation model from lexical/grammatical knowledge via understanding of predicate structure to reading comprehension by native Chinese speakers learning Japanese]. *Kotoba-no Kagaku* [*Studia Linguistica*]*, 26,* 5–24.

Coltheart, M., Curtis, B., Atkins, P. & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed processing approaches. *Psychological Review, 100(4),* 589–608.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance, 20(6),* 1197–1211.

Coltheart, V., & Laxon, V. J. (1990). The development of phonological mediation in reading comprehension. *Language and Cognitive Processes, 5(2),* 81–104.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6(2),* 78–84.

Fukuda, S. E., & Fukuda, S. (1994). To voice or not to voice: The operation of rendaku in the Japanese developmentally language-impaired. *McGill Working Papers in Linguistics, 10(1-2),* 178–193.

Fukuda, S. E., & Fukuda, S. (1999). The operation of rendaku in the Japanese specifically language-impaired: A preliminary investigation. *Folia phoniatrica et logopaedica, [International Journal of Phoniatrics, Speech Therapy and Communication Pathology], 51(1-2),* 36–54.

Goldsmith, J. A. (1976). *Autosegmental phonology.* Doctoral dissertation, Massachusetts Institute of Technology.

Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *Neuroimage, 20, Supplement 1,* 18–29.

Haraguchi, S. (2001). *On rendaku. On'in Kenkyuu* [Phonological Studies], 4, 9–32.

Ihara, M., & Murata, T. (2006). *Nihongo-no rendaku ni kansuru ikutsuka-no jikken* [Some experiments on sequential voicing]. *On'in Kenkyuu* [*Phonological Studies*]*, 9,* 17–24.

Irwin, M. (2005). Rendaku-based lexical hierarchies in Japanese: The behavior of Sino-Japanese mononoms in hybrid non compounds. *Journal of East Asian Linguistics, 14(2),* 121–153.

Irwin, M. (2009). Prosodic size and rendaku immunity. *Journal of East Asian Linguistics, 18(3),* 179–196.

Irwin, M. (2011). *Loanwords in Japanese.* Amsterdam: John Benjamins.

Irwin, M. (2012). Rendaku dampening and prefixes. *NINJAL Research Papers, 4,* 27–36.

Irwin, M., & Miyashita, M. (2016). *The rendaku database v2.7.* Downloadable from the Web-site, http://www-h.yamagata-u.ac.jp/~irwin/site/Rendaku_Database.html

Itô, J., & Mester, A. (1986). The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. *Linguistic Inquiry, 17(1),* 49–73.

Itô, J., & Mester, A. (2003). *Japanese morphophonemics: Markedness and word structure* (*Linguistics Inquiry Monograph Forty-one*). Cambridge, MA: MIT Press.

Japan Foundation and Association of International Education, Japan (2004). *Nihongo nooryoku shiken shutsudai kijun* (*kaitei-ban*) [*Japanese language proficiency test: Test content specifications* (revised edition)]. Tokyo: Bonjinsha.

Kawahara, S., & Sano, S. (2014). Testing Rosen's Rule and strong Lyman's Law. *NINJAL Research Papers, 7,* 111–120.

Kindaichi, H. (1976). Rendaku-*no kai* [Solution of sequential voicing]. *Sophia linguistics* (*published by Sophia University, Tokyo*)*, 2,* 1–22.

Kobayashi, Y., Sugioka, Y., & Ito, T. (2014). Rendaku (Japanese sequential voicing) as rule application: An ERP study. *NeuroReport, 25(16),* 1296–1301.

Komori, K., Tamaoka, K., Saito, N., & Miyaoka, Y. (2014). *Daini genngo-toshite Nihongo-o manabu chuugokugo washa-no nihongo-no kanjigo no shuutoku-ni kansuru koosatsu* [The acquisition of Japanese kanji compound words by Chinese native speakers learning Japanese as a second language]. *Chuugokugo*

Washa-notameno Nihongo Kyooiku Kennkyuu [*Research of Japanese Education for Native Chinese Speakers Learning Japanese*]*, 5,* 1–16.

Kubozono, H. (1995). *Gokeisei to on'inkoozoo* [*Lexical formation and phonological structure*]. Tokyo: Kurosio Publishers.

Kubozono, H. (2005). Rendaku: Its domain and linguistic conditions. In J. van de Weijer, K. Nanjo, & T. Nishihara (Eds.), *Voicing in Japanese* (pp. 5–24). Berlin, Germany: Mouton de Gruyter.

Kubozono, H., & Ota, S. (1998). *On'in koozoo-to akusento* [*Phonological structure and accents*]. Tokyo: Kenkyūsha.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4(12),* 463–470.

Labrune, Laurence (2013). *Rendaku in cross-linguistic perspective.* Paper presented at the 3rd International Conference on Phonetics and Phonology. NINJAL, Tokyo, 22 December 2013.

Leben, W. R. (1973). *Suprasegmental phonology*. Doctoral dissertation, Massachusetts Institute of Technology.

McCarthy, J. J. (1986). OCP effects: Gemination and antigemination. *Linguistic Inquiry, 17(2),* 207–263.

McCawley, J. D. (1968). *The phonological component of a grammar of Japanese*. The Hague: Mouton.

Miyaoka, Y., Tamaoka, K., & Sakai, H. (2011). *Nihongo goi tesuto-no kaihatsu-to shinraisei: Chuugokugo-o bogo-tosuru nihongo gakushuusha-no deeta-niyoru tesuto hyooka* [Developing Japanese lexical knowledge test and its reliability: Test evaluation by data of native Chinese speakers learning Japanese]. *Hiroshima Keizai Daigaku Kenkyuu Ronbunshuu* [*Bulletin of Hiroshima University of Economics*]*, 34(1),* 1–18.

Miyaoka, Y., Tamaoka, K., & Sakai, H. (2014). *Nihongo-no bunpoo nooryoku tesuto-no kaihatsu-to shinraisei: Nihongo gakushuusha-no deeta-niyoru tesuto hyooka* [Developing Japanese grammatical ability test and its reliability: Test evaluation by data of Japanese learners]. *Hiroshima Keizai Daigaku Kenkyuu Ronbunshuu* [*Bulletin of Hiroshima University of Economics*]*, 36(4),* 33–46.

Morris, J., & Holcomb, P. J. (2005). Event-related potentials to violations of inflectional verb morphology in English. *Cognitive Brain Research, 25(3),* 963–981.

Murata, T. (1984). *Nihongo-no rendaku-ni tsuite* [*On rendaku in the Japanese language*]. A handout of the talk presented at the meeting of Fukuoka Linguistic Society.

Nakagawa, Y. (1966). *Rendaku rensei (kashoo) -no keihu* [A genealogy of sequential voicing and sequential non-voicing (working label)]. *Kokugo Kokubun* [*Japanese Language and Literature*]*, 35(6),* 302–314.

Nakazawa, N., Vance, T. J., Irwin, M., & Lyddon, P. A. (2016). Rendaku awareness of Japanese learners in Taiwan: Students at Ming Chuan University. In T. J. Vance & M. Irwin (Eds.), *Sequential voicing in Japanese compounds: Papers from the NINJAL rendaku project* (pp. 43–59). Amsterdam: John Benjamins.

Okumura, M. (1955). Rendaku. In *Kokugogakkai* (Ed.) *Kokugogaku jiten* (pp. 961–962). Tokyo: Tokyodo.

Otsu, Y. (1980). Some aspects of rendaku in Japanese and related problems. In Y. Otsu and A. Farmer (Eds.), *MIT Working Papers in Linguistics 2: Theoretical Issues in Japanese Linguistics,* 207–227.

Park, S., Xiong, K., & Tamaoka, K. (2014). *Dookei niji kanjigo-no hinshisei-ni kansuru nikkanchuu deeta beesu* [A database of grammatical categories for orthographically-similar two-kanji compound words among the Japanese, Korean and Chinese languages]. *Kotoba-no Kagaku* (*Studia Linguistica*)*, 27,* 53–111.

Patterson, K. (1986). Lexical but nonsemantic spelling? *Cognitive Neuropsychology, 3(3),* 341–367.

Penke, M., Weyerts, H., Gross, M., Zander, E., Münte, T. F., & Clahsen, H. (1997). How the brain processes complex words: an event-related potential study of German verb inflections. *Cognitive Brain Research, 6(1),* 37–52.

Pinker, S. (1991). Rules of language. *Science, 253(5019)*, 530–535.

Pinker, S. (1994). *The language instinct.* New York: William Morrow.

Pinker, S. (1999). *Words and rules: The ingredients of language.* New York: Basic Books.

Rosen, E. R. (2001). *Phonological processes interacting with lexicon: Variable and non-regular effects in Japanese phonology.* Doctoral dissertation, University of British Columbia, Canada.

Takayama, T. (1999). *Shakuyoogo-no rendaku/koo'onka-ni tsuite* [Regarding rendaku on alphabetic loanwords]. *Report of the special research project for the typological investigation of languages and cultures of the East and West, 1,* 375–385.

Tamaoka, K., & Ikeda, F. (2010). Whiskey or Bhiskey?: Influence of first-element and dialect region on sequential voicing of *shoochuu. Gengo Kenkyuu, 137,* 65–79.

Tamaoka, K., Ihara, M., Murata, T., & Lim, H. (2009). Effects of first-element phonological-length and etymological-type features on sequential voicing (rendaku) of second elements. *Journal of Japanese Linguistics, 25,* 17–38.

Tamaoka, K., Miyaoka, Y., Kim, S., & Lim, H. (2011). *Kankokugo-o bogo tosuru Nihongo gakushuusha-no goi chishiki-ga onomatope-no shuutoku-ni ataeru eikyoo* [Influences of lexical categories on acquisition of sound-symbolic words by native learning Japanese]. *Gengo Kyooiku Hyooka Kenkyuu* [*Assessment and Evaluation in Language Education*]*, 2,* 36–41.

Vance, T. J. (1979). *Nonsense-word experiments in phonology and their application to 'rendaku' in Japanese.* Doctoral dissertation, University of Chicago.

Vance, T. J. (1987). *An introduction to Japanese phonology.* Albany: State University of New York.

Vance, T. J. (1996). Sequential voicing in Sino-Japanese. *Journal of the Association of Teachers of Japanese, 30(1),* 22–43.

Vance, T. J. (2005). Sequential voicing and Lyman's Law in Old Japanese. *Polymorphous linguistics: Jim McCawley's legacy* (pp. 27–43). Cambridge, MA: MIT Press.

Vance, T. J. (2007). Have we learned anything about rendaku that Lyman didn't already know? In B. Frellesvig, M. Shibatani, & J. C. Smith (Eds.), *Current issues in the history and structure of Japanese* (pp. 153–170). Tokyo: Kurosio Publishers.

Vance, T. J. (2008). *The sounds of Japanese.* Cambridge, UK: Cambridge University Press.

Vance, T. J. (2014a). *Rendaku to onomatope-no joogo* [Rendaku and mimetic reduplication]. *NINJAL Project Review, 5(1),* 32–38.

Vance, T. J. (2014b). If rendaku isn't a rule, what in the world is it? In K. Kabata & T. Ono (Eds.), *Usage-based approaches to Japanese grammar* (pp. 137–152), Amsterdam: John Benjamins.

Vance, T. J. (2016). Introduction. In T. J. Vance & M. Irwin (Eds.), *Sequential voicing in Japanese compounds: Papers from the NINJAL rendaku project* (pp. 1–9). Amsterdam: John Benjamins.

Weyerts, H., Penke, M., Dohrn, U., Clahsen, H., & Münte, T. F. (1997). Brain potentials indicate differences between regular and irregular German plurals. *Neuroreport, 8(4),* 957–962.

Xiong, K., & Tamaoka, K. (2014). *Nittuu dookei niji kanjigo-no hinshisei-no taioo kankei-ni kansuru koosatsu* [A descriptive analysis of Japanese-and-Chinese orthographically-similar two-kanji compound words according to the database of grammatical categories]. *Kotoba-no Kagaku* (*Studia Linguistica*)*, 27,* 25–51.

Yamato, Y., Tamaoka, K., & Chu, X. (2013). *Chuugokujin nihongo gakushuusha-no tekisuto-no on-rain yomi niokeru goi-to bunpoo nooryoku-no eikyoo* [Effects of lexical and grammatical ability for on-line text processing by native Chinese speakers learning Japanese]. *Nihon Kyooka Kyooiku Gakkai-shi* [*The Bulletin of Japanese Curriculum Research and Development*]*, 36(1),* 33–43.

Yu, S., & Tamaoka, K. (2015). *Nikkanchuu dookei niji kanjigo-no hinshisei webu kensaku dezain* [A Web-accessible search engine for grammatical category for orthographically-similar two-kanji compound words between Japanese, Korean and Chinese]. *Kotoba-no Kagaku* (*Studia Linguistica*)*, 29,* 43–61.

Katsuo Tamaoka
Graduate School of Languages and Cultures
Nagoya University
Japan
ktamaoka@lang.nagoya-u.ac.jp

Kyoko Hayakawa
Kwansei Gakuin University
Japan
hayakawa@kwansei.ac.jp

Timothy J. Vance
National Institute for Japanese Language and Linguistics
Japan
tjvance825@gmail.com