Children’s Use of Orthographic Cues in Language Processing
Takayo Sugimoto (takayosg@vega.aichi-u.ac.jp)
Aichi University
1-1 Machihatakemachi, Toyohashi,
Aichi, 441-8522 Japan

Abstract

Rendaku, or sequential voicing, is a morphophonemic process in Japanese in which the voiceless word-initial consonant of the second element (=E2) of a compound word becomes voiced (e.g., ori + kami → origami, ‘folding’ + ‘paper’ → ‘paper folding’, /k/ becomes /g/). In adult grammar, rendaku is subject to two conditions: It applies if and only if (a) E2 is a Yamato word (native vocabulary) in the lexicon and (b) it contains no voiced consonant (e.g., b, d, & g). Recent psycholinguistic studies have revealed that Japanese-speaking preschoolers do not follow adult’s grammar; they develop their original prosodically-based rendaku processing strategy (preschooler-specific rendaku strategy). Their strategies qualitatively change in the early middle childhood to be adult-like rendaku, creating a discontinuity in children’s word-processing strategies. This study investigated factors responsible for this developmental discontinuity. We conducted an experiment using cross-modal linguistic stimuli (prosody & orthography) to see whether children’s orthographic knowledge affects their rendaku strategy or not. Our results showed that orthographic cues affected literate children’s rendaku processing. They were aware the correspondence between types of orthography and word categories in Japanese.

Keywords: preschoolers; rendaku; orthography; word category; pitch accent;

Introduction

Rendaku is a morphophonemic process in Japanese that the initial voiceless consonant of the second element (E2) of a compound becomes voiced (Vance, 2015).

(1) Compound Word Formation and Rendaku

E1 + E2 → Compound Word
ori + kami → origami  /k/ → /g/ ‘folding’ ‘paper’ ‘paper folding’

Rendaku has long been studied based on the data from adult subjects (Labrune, 2013). Neuropsychological studies have shown that rendaku is not just a matter of pronunciation such as English consonantal assimilation (e.g., transcript [s] vs. transcribed /s/→[z]); rendaku has functions of signaling the syntactic (grammatical), semantic (=meaning) aspects of E2 within a compound (Ogata, E., Hayshi, R., Imaizumi, S., Hirata, N., & Mori, K., 2000).

Thus, studying children’s rendaku acquisition should lead to our deeper and comprehensive understanding of language development and cultural influences on our language faculty.

In adult grammar, rendaku is generally subject to 2 conditions (Ito & Mester, 1986, 1995, Fukuda & Fukuda, 1999, Fukuda, 2002) 1:

(2) Rendaku conditions

Rendaku applies if

(a) E2 is a Yamato morpheme (native vocabulary) and
(b) E2 contains no voiced obstruent in it (Lyman’s Law).

The rendaku conditions seems complex and difficult for children to acquire since in order to be able to apply rendaku properly, children must know the lexical strata and Lyman’s Law. Lexical strata are word categories stratified in the Japanese lexicon: Yamato, Sino-Japanese, foreign loans, and onomatopoeia (Ito & Mester, 1995, McCawley, 1968). Lyman’s Law (Ito & Mester, 1995, Vance, 2015, and many others), which prohibits more than one voiced consonant within a morpheme (an element or unit of a word), is active only in the native vocabulary. This means that children must know which word belongs to the native vocabulary.

Children also have to parse the E2 and decide the applicability of rendaku immediately. How do children acquire this seemingly complex knowledge about rendaku? Do they know these two conditions from the beginning? Or do they learn them in some order?

Rendaku acquisition studies have assumed that children learn the adult’s rendaku conditions (Fukuda & Fukuda, 1999, Fukuda, 2002). However, recent psycholinguistic studies have revealed that children do not acquire rendaku along with the adult’s grammar. Instead, Japanese-speaking preschoolers develop a prosodically-based rendaku strategy (preschooler-specific rendaku strategy, Sugimoto, 2013a).

(3) Preschooler-specific rendaku strategy

Apply rendaku if E2 is an unaccented word.

The preschooler-specific rendaku strategy is also observed in English-Japanese simultaneous bilinguals (Sugimoto, 2015a). Sugimoto (2015b) reported that

1 These conditions are not strict restrictions and there are some exceptional cases. For the purpose of conducting experiments, we assume and start from these conditions.
preliterate blind children also show the prosodically-based preschooler-specific rendaku strategy.

Japanese is a language with a pitch accent system. Pitch accent is a prosodic feature of a word and it differentiates the meaning of each Japanese word. Pitch accent can be divided into two types: Accented and unaccented. Accented words have a tonal rise followed by a sudden fall. Unaccented words have no such a tonal (rise & fall) contour.

Preschooler-specific rendaku processing reflects the distribution of pitch accent types of native vocabulary, to which rendaku applies (Table 1). Children are aware that rendaku does not apply every word; it applies a particular category of words. They actively construct their own Rendaku rule, making best use of their knowledge: Apply rendaku if E2 is an unaccented native word.

Table 1: Word categories and pitch accent

<table>
<thead>
<tr>
<th>Lexical strata</th>
<th>Pitch accent</th>
<th>Accented words’ (examples)</th>
<th>Unaccented words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yamato words (=native vocabulary)</td>
<td>29%</td>
<td>(ka’rasu ‘crow’)</td>
<td>71% (sakana ‘fish’)</td>
</tr>
<tr>
<td>Sino-Japanese words (vocabulary of Chinese origin)</td>
<td>49%</td>
<td>(hu’dan ‘daily’)</td>
<td>51% (saihu ‘purse’)</td>
</tr>
<tr>
<td>Loan words (excluding Sino-Japanese)</td>
<td>93%</td>
<td>(ke’eki ‘cake’)</td>
<td>7% (pijano ‘piano’)</td>
</tr>
</tbody>
</table>

(Sugimoto, 2016, data originally from Kubozono, 2006, p180)

The commonly observed children’s prosodically-based rendaku processing strategy changes in quality around in the early middle childhood to be adult-like rendaku, creating a discontinuity in children’s rendaku strategies (Sugimoto, 2015b). School age children no longer depend on prosodic information of E2. Their rendaku patterns become similar to those of adults being independent from the prosody of E2. These qualitative change in children’s rendaku processing strategy create a developmental discontinuity.

It is easy to imagine that some developmental changes and learning outcomes during these years might motivate children’s rendaku strategy at around their entry into elementary school.

What kind of development or learning motivates and determines such changes or a discontinuity?

One possibility would be the influence of literacy. Since three types of Japanese characters represent lexical strata, it could be likely that children become well aware of Japanese lexical strata and word categories along with usages of Japanese hiragana, katakana, and kanji characters. If literacy affects children’s rendaku processing strategy, how and why?

The purpose of this study was to investigate factors responsible for this developmental discontinuity. We conducted an experiment using cross-modal linguistic stimuli (prosody & orthography) to see whether literacy affects children’s rendaku processing strategy or not.

(4) Research questions
Q1. Do literate preschoolers know the correspondence between types of orthography and word category?
Q2. How do they use their orthographic knowledge in their rendaku processing?

(5) Working Hypothesis
If literate preschoolers use their orthographic knowledge in rendaku processing, then their rendaku patterns should differ, depending upon types of orthography given.

Method
We used a compound noun formation task (Nicoladis, 2003, Sugimoto, 2013a&b, 2016) to see Japanese-speaking preschoolers’ language processing strategy described below.

Participants and ethical considerations
The total number of 73 six-year olds with hiragana & katakana literacy living in the Tokyo dialect area participated in our study. 40 children were assigned to the no orthography condition (control group); 16 children were in the hiragana condition; 17 children were in the katakana condition.

Prior to the study, we obtained written form of parental permission from all the participants.

Procedure
Children were tested individually in a quiet room. We went through three trials: 4 warm-up trials, 4 comprehension trials, and finally the 16 test trials (E1=hima’wari). In each trial, children were shown 3 types of pictures on a laptop computer, E1 & E2 in a random order. The experimenter read aloud E1 and E2 and then asked the children to name picture C (compound noun).

(6) Compound noun formation task (Sugimoto, 2016)

\[
\text{E1} + \text{E2} \rightarrow \text{C (Compound noun)} \\
\text{hima’wari} + \text{ka’rasu} \rightarrow \text{himawara’rasu} \\
\text{‘sunflower’} + \text{‘crow’} \rightarrow \text{‘sunflower crow’}
\]

The experimenter read aloud the following statements followed by a question as in (7).

(7) Instructions in the experiment
E1: Kore-wa himawari-desu. ‘Here’s a hima’wari’
E2: Kore-wa karasu-desu. ‘Here’s a ka’rasu.’
C: Koreni name-e tuketekudasai.
‘How would you name it?’

---

\(^2\) The apostrophe in Japanese accented words (e.g., ka’rasu (=crow)) indicates the position of pitch accent.
Compound noun formation task with cross-modal linguistic stimuli: Visual(orthography) & auditory (speech) information

We used a modified version of compound noun formation task, cross-modal stimuli, which used two types of visual stimuli (picture and orthography) and auditory stimuli (experimenter read aloud the stimuli).

Design and material

The experiment used a 3 factor inter-subject design: 2(pitch accent)×2(word)×2(orthography). First, the pitch accent of E2 was controlled to two types: unaccented vs. accented (Kubozono, 2006). We used known words and novel words. As for novel words, we counterbalance poss

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaccented</td>
<td>6.2</td>
<td>1.244</td>
<td>40</td>
</tr>
<tr>
<td>Hiragana</td>
<td>3.38</td>
<td>2.391</td>
<td>16</td>
</tr>
<tr>
<td>Katakana</td>
<td>1.82</td>
<td>1.286</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>4.56</td>
<td>2.444</td>
<td>73</td>
</tr>
</tbody>
</table>

In order to counterbalance possible phonotactic effects, we divided each condition group into two subgroups and switched types of pitch accent assignment between the subgroups of each condition (see Tables 2 & 3 in Appendix).

We also used three types of Japanese orthography conditions (visual stimuli) as in (8).

<table>
<thead>
<tr>
<th>(8) 3 orthography conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. No orthography (control)</td>
</tr>
<tr>
<td>B. Hiragana condition</td>
</tr>
<tr>
<td>C. Katakana condition</td>
</tr>
</tbody>
</table>

Visual stimuli for each of these three conditions are illustrated in Figures 1a-c, respectively.

Measures

During the experiment, we used a SONY IC recorder (ICD SX-1000) and recorded the children’s utterances. After the experiment, two people, one of whom was the author, listened separately to the recordings and transcribed them, judging whether or not the children voiced the target consonant. The reliability of the judgments (Cohen’s κ) was calculated. The agreement ratio was κ = .96, which is considered highly reliable.

Our scoring was the following. All of the 16 compound nouns in the test trial are subject to rendaku (see Table 2 in Appendix). When a subject voiced the morpheme-initial obstruent of E2, then we judged that he/she had correctly applied rendaku; for each compound, one point was added to the score. We calculated the total scores and subtotals by condition (pitch accent, orthography, and word type).

Results

Descriptive statistics is shown in Table 4. A three-way ANOVA found a significant interaction of pitch accent and orthography \[F(2,70)=15.061, p<.001, \eta^2 = .301\]. The simple effect of pitch accent was significant in the control group and in the hiragana condition \[F(1,70)=93.092, p<.001, \eta^2 = .571\], \[F(1,70)=4.750, p=.033, \eta^2 = .064\], respectively. But the simple effect of pitch accent was not significant in the katakana condition \[F(1,70)=.20, p=.888, \eta^2 < .0001\].

Multiple comparisons of within-condition show significant differences between unaccented words and accented words in the control and the hiragana conditions (Figure 2). In the two conditions, children used the prosodically-based preschooler-specific rendaku strategy. On the other hand, the katakana condition show no significant difference between the two pitch accent types \(p<.001, 95\% IC: .891-.774\). Thus, children in the katakana condition did not use the preschooer-specific rendaku strategy. Compared to the other two conditions, their rendaku processing rates in the katakana conditions were inhibited, but when we look at the within-group difference, children in the katakana condition applied rendaku to accented words more often than unaccented words.

Table 4: Descriptive statistics of rendaku score

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaccented</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>6.2</td>
<td>1.244</td>
<td>40</td>
</tr>
<tr>
<td>Hiragana</td>
<td>3.38</td>
<td>2.391</td>
<td>16</td>
</tr>
<tr>
<td>Katakana</td>
<td>1.82</td>
<td>1.286</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>4.56</td>
<td>2.444</td>
<td>73</td>
</tr>
</tbody>
</table>

(8 pts. are the maximum score for each pitch-accent type)
Discussion

The significant interaction of pitch accent and orthography conditions, which indicates that literate preschoolers were affected by orthographic cues and showed different rendaku processing. Children in the no orthography condition (control) and in the hiragana condition showed similar rendaku patterns: apply rendaku to unaccented E2s. But children in the katakana condition showed a different rendaku pattern: they did not apply rendaku to unaccented words, contrary to the other two groups. When katakana is given, which represents foreign loan words in Japanese, children’s rendaku processing was inhibited. Children in (c) did not to apply rendaku to non-native vocabulary as much as those in the other conditions did.

It follows from this that the katakana condition made children assume the word stimuli were nonnative vocabulary. Katakana is used for foreign loan words in the Japanese writing system. Children seem to know katakana is used for non-native vocabulary, and judged that the stimuli (both existent native vocabulary and novel words were given) were non-native vocabulary, to which rendaku does not apply.

Our results suggest that literate preschoolers used orthographic cues in rendaku processing. Preschoolers used different strategies, depending on types of orthography provided in the stimuli.

Preschoolers in the ‘no orthography’ and the hiragana conditions showed the similar tendency: they applied rendaku to unaccented words more often than accent words while those in the katakana condition seemed reluctant to apply it. It follows from these that children attend to orthographic cues but children still use prosodically-based rendaku strategy.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Application of rendaku</th>
<th>Use of preschooler-specific strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>No orthography (pictures only)</td>
<td>Yes</td>
<td>Yes, U &gt; A</td>
</tr>
<tr>
<td>Hiragana</td>
<td>Yes</td>
<td>Yes, U &gt; A</td>
</tr>
<tr>
<td>Katakana</td>
<td>Yes, but highly inhibited</td>
<td>No, U ≈ A</td>
</tr>
</tbody>
</table>

(U = unaccented words, A= accented words)

Why, then, did preliterate preschoolers in the katakana condition were not willing to apply rendaku to unaccented words more often than accent words? We can think of the following possibilities. Preliterate children may be aware of some correspondence between types of orthography and word categories. That is, they may know hiragana is used for a certain category of words and katakana for another. They may also be aware that katakana is used to represent non-native Japanese vocabulary, to which rendaku does not apply.

Preschoolers first define the rendaku word category based on prosodic information, that is, pitch-accent (preschooler-specific rendaku strategy). We only found that our literate preschoolers can differentiate rendaku strategies based on a rough katakana vs. non-katakana distinction, not hiragana vs. katakana distinction. We need their longitudinal data to consider orthographic effects on the developmental changes in children’s language processing.

Conclusion

Japanese speaking preschoolers, when they have no orthographic knowledge, first define the rendaku category based on pitch-accent, i.e., preschooler-specific rendaku strategy. After acquiring orthography, children seem to gradually change their rendaku processing strategy, using information such as the relationship between orthography and word category. Something more is needed for their redefinition of the rendaku category, which may cause the developmental change in their rendaku processing strategy. We need developmental paths of individual children.

Acknowledgments

The author is grateful for all the participants of this study. This work was financially supported in part by JSPS KAKENHI Grant #26580080 and “The Japanese Lexicon: A Rendaku Encyclopedia” (Project leader: Prof. Timothy J. Vance, National Institute for National Language and Japanese Linguistics).

Figure 2: Rendaku Score by Condition (8pts.)

Table 5: Orthographic cues and children’s application of rendaku

886
References


Appendix

Table 2: List of stimuli for 16 E2s used in the test trial

<table>
<thead>
<tr>
<th>8 known words</th>
<th>8 novel words (Old Japanese)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accented</td>
<td>Unaccented</td>
</tr>
<tr>
<td>1 ta’nuki ‘raccoon dog’</td>
<td>sakura ‘cherry blossom’</td>
</tr>
<tr>
<td>2 ka’rasu ‘crow’</td>
<td>tsukue ‘desk’</td>
</tr>
<tr>
<td>3 ho’uki ‘bloom’</td>
<td>hatake ‘field’</td>
</tr>
<tr>
<td>4 ho’taru ‘light bug’</td>
<td>kuruma ‘car’</td>
</tr>
</tbody>
</table>

(E1: himawari ‘sunflower’;  E2: 16 words listed in Table 2)

# For 8 novel words, we created two types of pitch accent assignment patterns to counterbalance possible phonotactic effects (see Table 3 below). In particular, we divided children in each of three conditions into two groups and used different pitch accent assignment for each group in each condition.

Table 3: Two types of pitch accent assignment

<table>
<thead>
<tr>
<th>Novel word stimuli (E2)</th>
<th>Groups</th>
<th>Group A of each condition</th>
<th>Group B of each condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>tekona, hokai, tatara, hokora</td>
<td>Accented e.g., te’kona (antipenulte)</td>
<td>Unaccented e.g., tekona (no pitch accent)</td>
<td></td>
</tr>
<tr>
<td>tokama, hikime, sasara, koromo</td>
<td>Unaccented e.g., tokama (no pitch accent)</td>
<td>Accented e.g., to’kama (antipenulte)</td>
<td></td>
</tr>
</tbody>
</table>