A Rendaku Database for Old Japanese

Timothy J. Vance
NINJAL

Mark Irwin
Yamagata Univ.
Rendaku and Lyman’s Law
**Rendaku** is a morphophonemic phenomenon which affects non-initial elements of Japanese compounds.

If such an element begins with a voiceless obstruent when it occurs as a word on its own, the voiced counterpart of this obstruent may appear in a compound.
For example:

/k/ → /g/

/hana/ + /kami/ → /hanagami/
‘nose’ ‘paper’ ‘tissue’

/t/ → /d/

/uširo/ + /tate/ → /uširodate/
‘behind’ ‘shield’ ‘backer’
Rendaku was already well established in Old Japanese, the language recorded in texts originally written in the 8th century CE.

At that time, however, the consonants corresponding to modern Japanese voiced obstruents were prenasalized.
Rendaku was pervasively irregular in Old Japanese, and it remains so today.

That is, there are many cases where rendaku might be expected to occur but does not.
Many of these exceptions are systematic, at least to some degree, since they involve a known inhibiting factor.

The best known inhibiting factor is Lyman’s Law.
Lyman’s Law says that if a non-initial element in a compound already contains a voiced obstruent as an independent word, then *rendaku* is blocked.
For example:

/hitori/+ /tabi/ → /hitori\text{tabi}/
‘1 person’ ‘trip’ ‘solitary journey’

/doku/+ /tokage/ → /doku\text{tokage}/
‘poison’ ‘lizard’ ‘poisonous lizard’
In Old Japanese, there is clear evidence for a “strong version” of Lyman’s Law.

In this version, *rendaku* is blocked not just by a (prenasalized) voiced obstruent in the non-initial element itself but also by a voiced obstruent in the initial element.
For example:

\[
{\text{OJ}} /\text{midu}/ + {\text{OJ}} /\text{tori}/ \rightarrow {\text{OJ}} /\text{mizutori}/
\]

‘water’ ‘bird’ ‘waterfowl’

Dictionaries list the modern Tokyo Japanese descendant of this word as /mizutori/, but some speakers have rendaku: /mizudori/.
In general, the strong version of Lyman’s Law does not constrain *rendaku* in modern Tokyo Japanese compounds.

For example:

/\text{tabi}/ + /\text{hito}/ \rightarrow /\text{tabibito}/

‘journey’ ‘person’ ‘traveler’
A Rendaku Database for Old Japanese
A dictionary published in 1967 contains essentially the entire attested Old Japanese lexicon. Our “Rendaku Database for Old Japanese” is based on the headwords in this dictionary.
時代別　国語大辞典
上代編
三省堂
Many of these vocabulary items are attested in phonograms, i.e., Chinese characters used to represent Japanese words syllable-by-syllable. Many others, however, are attested only logographically.
Our database contains all compounds attested in phonograms and ending in an element (E2) that begins with a voiceless obstruent as an independent word.
For example:

$OJ/\text{sit}a/ + OJ/\text{pa}/ \rightarrow OJ/\text{sit}a\text{ba}/$

‘bottom’ ‘leaf’ ‘lower leaves’

written 之多婆  in the *Man’yōshū*

In compiling this database, we have had to confront a variety of philological problems that we will not discuss in any detail here.
The difficulties include:

[1] compounds not attested in a “reliable” 8th-century source

[2] ambiguous phonogram/logogram spellings

[3] inconsistent spellings

[4] phonograms interpolated later
A preliminary version of the database will be available in the near future at this URL:

http://www.ninjal.ac.jp/rendaku/en/

Inevitably, given the mind-numbingly tedious nature of the work involved, there will be mistakes.
The website will include a detailed explanation of the process we followed and a plea for help finding and correcting errors.
The database is just an Excel spreadsheet.

This small portion shows the kind of information it provides.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>PH</td>
<td>E1</td>
<td>K1</td>
<td>E2</td>
<td>K2</td>
<td>C(-4)</td>
<td>V(-3)</td>
<td>C(-2)</td>
<td>V(-1)</td>
<td>C(+1)</td>
<td>V(+2)</td>
<td>C(+3)</td>
<td>V(+4)</td>
<td>R</td>
<td>F-W ROM</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>kapa</td>
<td>kami</td>
<td>上</td>
<td>k</td>
<td>a</td>
<td>p</td>
<td>a</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>1A</td>
<td>0</td>
<td>kap+ami</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>ko</td>
<td>kami</td>
<td>上</td>
<td>k</td>
<td>o</td>
<td>t</td>
<td>o</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>1A</td>
<td>1</td>
<td>to+ami</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>sa</td>
<td>kami</td>
<td>そ</td>
<td>s</td>
<td>a</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>1A</td>
<td>0</td>
<td>a+ami</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>sume</td>
<td>kami</td>
<td>神</td>
<td>s</td>
<td>u</td>
<td>m</td>
<td>e</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>1B</td>
<td>0</td>
<td>sumy+ami</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>asi</td>
<td>kami</td>
<td>神</td>
<td>a</td>
<td>s</td>
<td>i</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>o</td>
<td>1A</td>
<td>1</td>
<td>si+amiO</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>ma</td>
<td>kami</td>
<td>鳥</td>
<td>m</td>
<td>a</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>0</td>
<td>0</td>
<td>ma+amiO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>ku</td>
<td>kane</td>
<td>金</td>
<td>k</td>
<td>u</td>
<td>k</td>
<td>a</td>
<td>n</td>
<td>e</td>
<td>1A</td>
<td>1</td>
<td>ku+ame</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>ma</td>
<td>kane</td>
<td>金</td>
<td>m</td>
<td>a</td>
<td>k</td>
<td>a</td>
<td>n</td>
<td>e</td>
<td>0</td>
<td>ma+kane</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
It is easy to test how well a generalization holds by sorting the data appropriately.

For instance, to test the strong version of Lyman’s Law, the first step is to sort by column H, i.e., the last consonant in the first element (E1) of each compound.
Almost all the compounds of interest are those that have a voiced obstruent in this column (romanized as \(b, d, g, \) or \(z\)).

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PH</td>
<td>E1</td>
<td>K1</td>
<td>E2</td>
<td>K2</td>
<td>(-4)</td>
<td>(-3)</td>
<td>(-2)</td>
<td>(-1)</td>
<td>(+1)</td>
<td>(+2)</td>
<td>(+3)</td>
<td>(+4)</td>
<td>R</td>
<td>F-W ROM</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>kappa</td>
<td>kamı́</td>
<td>2</td>
<td>k</td>
<td>a</td>
<td>p</td>
<td>a</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>i^A</td>
<td>0</td>
<td>kappa+kami</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>koto</td>
<td>kamı́</td>
<td>2</td>
<td>k</td>
<td>o^B</td>
<td>t</td>
<td>o^B</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>i^A</td>
<td>1</td>
<td>koto+gami</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>sa</td>
<td>kamı́</td>
<td>2</td>
<td>s</td>
<td>a</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>i^A</td>
<td>0</td>
<td>sa+gami</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>sume</td>
<td>kamı́</td>
<td>2</td>
<td>s</td>
<td>u</td>
<td>m</td>
<td>e^A</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>i^B</td>
<td>0</td>
<td>sume+kamwi</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>asi</td>
<td>kamO</td>
<td>2</td>
<td>a</td>
<td>s</td>
<td>i</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>o^A</td>
<td>1</td>
<td>asi+gamoO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>ma</td>
<td>kamO</td>
<td>2</td>
<td>m</td>
<td>a</td>
<td>k</td>
<td>a</td>
<td>m</td>
<td>o^B</td>
<td>0</td>
<td>ma+kamoO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>ku</td>
<td>kane</td>
<td>2</td>
<td>k</td>
<td>u</td>
<td>a</td>
<td>n</td>
<td>e</td>
<td>1</td>
<td>ku+gane</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>ma</td>
<td>kane</td>
<td>2</td>
<td>m</td>
<td>a</td>
<td>k</td>
<td>a</td>
<td>n</td>
<td>e</td>
<td>0</td>
<td>ma+kane</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(sume = \) 形状音 (honorific)
There are only 39 such examples in the database.

One or two are philologically problematic, and a few would be expected to resist *rendaku* for other reasons.

But none of the 39 has *rendaku*. 
There are only four examples with a voiced obstruent in an earlier syllable of E1 (in the penultimate syllable in all four cases).

One of these four examples has *rendaku*. 
This is very little to go on, but maybe only a voiced obstruent in the final syllable of E1 inhibited *rendaku*. 
The Ultra-Strong Version of Lyman’s Law
Roy Andrew Miller published an article in 1984 in which he argued that an “ultra-strong” version of Lyman’s Law held in Old Japanese.

Miller’s claim is based on his interpretation of two obscure passages in an 11th-century glossary to the Chinese Suvarnapabhrasā Sutra.
We will not consider whether or not Miller’s interpretation is correct.

We will simply show how our database can be used to test the putative ultra-strong version of Lyman’s Law.
The following passage is Miller’s clearest statement of the claim.

... whenever the first morph in [a] compound [in Old Japanese] contained a sonant [i.e., voiced] consonant whether stop or continuant in its final syllable, this prior sonant inhibited the implementation of the otherwise regular voicing rule [i.e., *rendaku*].
In other words, if the final syllable of E1 in an Old Japanese compound began with a voiced consonant (whether obstruent or sonorant), then *rendaku* was blocked.
Our database enables us to test Miller’s ultra-strong version of Lyman’s Law.

We have already shown that the database contains no examples of *rendaku* immediately preceded by a syllable beginning with a voiced obstruent.
We now consider examples with a potential *rendaku* site immediately preceded by a sonorant consonant.

Once again, the first step is to sort by the last consonant in each E1.
The compounds of interest are those that have a sonorant in this column (romanized as \textit{m, n, r, w, or y}). We further restricted this subset by removing examples that are likely to have resisted \textit{rendaku} for some other reason.
That is, we discarded compounds of the following types:

[1] medial voiced obstruent in E2
[2] compound not a noun
[3] coordinate (dvandva) meaning
[4] numeral E1
[5] honorific 0J/mi/ as E1
We also discarded reduplicated words, since mimetic reduplicated words resist *rendaku*, while non-mimetic reduplicated words favor it. These reductions left 220 candidate compounds with a sonorant+vowel syllable immediately preceding the potential *rendaku* site.
Of these 220, 88 (40%) have *rendaku*.

Clearly, the proposed ultra-strong version of Lyman’s Law did not hold in Old Japanese.
Miller actually adds some provisos that effectively insulate his claim from disconfirmation, but we have chosen to treat it as an empirical statement.
Rendaku
Following /r/V
Syllables
It is often claimed that an /r/V syllable immediately preceding a potential *rendaku* site made *rendaku* less likely in Old Japanese.

Our database allows us to group compounds according to the consonant in this syllable and compare the *rendaku* rates.
Here again, we exclude compounds of the types that we excluded in our test of the ultra-strong version of Lyman’s Law.

The graph on the following slide shows the results of this procedure.
overall mean: 41%
obstruent mean: 45% /r/ mean: 26%
non-/r/ sonorant mean: 43%
Our sample is fairly small and certainly not random, but the *rendaku* rate is noticeably low for Old Japanese compounds with OJ/r/ in the syllable immediately preceding the potential *rendaku* site.
Comparisons with Modern Tokyo Japanese
Irwin, with assistance from Mizuki Miyashita of the University of Montana, has compiled a dictionary-based database for investigating *rendaku* in modern Tokyo (“standard”) Japanese.
This “Rendaku Database” is, of course, much larger than our database for Old Japanese, and version 2.0 is now available at this URL:

http://www-h.yamagata-u.ac.jp/~irwin/site/Rendaku_Database.html
We will close today’s presentation with a simple comparison between modern Japanese and Old Japanese. Specifically, we will look at the relationship between *rendaku* rate and the consonant in the syllable immediately preceding the potential *rendaku* site.
The modern sample was restricted in much that same way as the Old Japanese sample and also by imposing a few additional restrictions that were not relevant in the Old Japanese case (e.g., limiting E2 to native Japanese items).
As the graph on the following slide shows, the rate following /r/V is slightly below average.
The results of a one-way ANOVA show no statistically significant difference among the *rendaku* rates for different consonants in the last syllable before the boundary.

Even if the apparent inhibiting effect of /r/ in Old Japanese was genuine, it has not persisted.
Thank you for your kind attention.