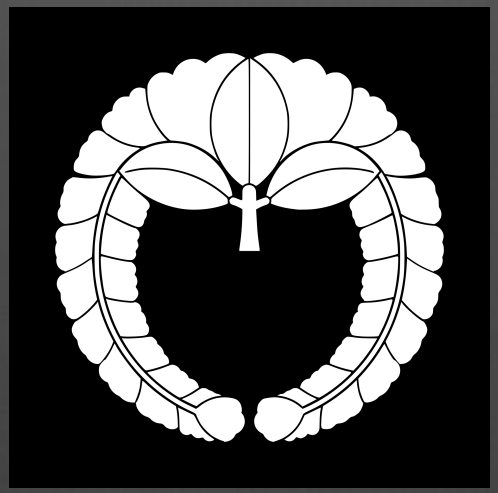
Rendaku & Labial Lenition



Mark Irwin,
Yamagata University

Paul Lyddon,
Osaka Jogakuin College

EAJS, Kobe University
24 September, 2016

What is Rendaku?

- Rendaku – or sequential voicing – is a productive morphophonemic process found in Japanese
- It is a process found only in compounds, and restricted to non-initial elements which begin in a voiceless obstruent
- The initial voiceless obstruent of a non-initial element may voice

kawa	+	*soko*	>	*kawa.**z**oko*
‘river’		‘bottom’		‘riverbed’
tuke	+	*tume*	>	*cuke.**d**ume*
‘attach’		‘fingernail’		‘false nail’
hana	+	*kami*	>	*hana.**g**ami*
‘nose’		‘paper’		‘tissue’
usiro	+	*tate*	>	*usiro.**d**ate*
‘behind’		‘shield’		‘backer, supporter’

Rules & Constraints

- Rendaku occurs with markedly different frequencies across different lexical strata – it is essentially a phenomenon limited to native Japanese words
- Rendaku is not restricted to Modern Tokyo Japanese
- Synchronically, it is found in all modern Japanese dialects, though research has been insufficient for determining to what extent behaviour differs
- Diachronically, it is found in texts throughout all stages of the history of the language

Rules & Constraints

- Rendaku is governed by a number of systematic rules and constraints which control the triggering of the phenomenon
- The most well-known of these is Motoori/Lyman's Law [1, 2]
- Also important in terms of today's presentation is the right-branch condition [3]

Motoori/Lyman's Law

- If a non-initial element already contains a voiced obstruent, then rendaku is blocked:

asi	+	*kubi*	>	*asi.**k**ubi (*asi.**g**ubi)*
'leg'		'neck'		'ankle'
mamizu	+	*kurage*	>	*mamizu.**k**urage (*mamizu.**g**urage)*
'freshwater'		'jellyfish'		'freshwater jellyfish'

Right Branch Condition

- Rendaku is restricted to elements on the right branch of a constituent structure
- Thus, *[nuri.**g**asa].ire* ‘case for lacquered umbrellas’ but *nuri.[**k**asa.ire]* ‘umbrella case which is lacquered’
- In other words, rendaku is blocked when a compound is itself part of a larger compound: *huyoo.**t**eate* ‘family allowance’ *zidoo.**h**urikomi* ‘automatic transfer’

Research Background

- Despite the rules and constraints governing rendaku mentioned earlier, rendaku does not always occur where it should… and sometimes occurs where it should not
- In other words, explanations are not always available for its presence or absence
- Our proposal received a grant-in-aid from the Japan Society for the Promotion of Science and it is on one of our research questions which we will report today

The Rendaku Database

- The rendaku database [4] contains 34,464 entries, all of which are compounds possessing a potential rendaku site
- Data are taken from two major dictionaries [5, 6] and tagged for a range of criteria, including vocabulary stratum, length, accent and part of speech
- Although further dictionaries and tags are planned in the future, the rendaku database is already the largest ever compiled
- The rendaku databases has already spawned research on prefixes [7], duplicate moras [8], 'loving and hating' [9] and homophony [10]
- A detailed description of the database can be found in [11]

Morphemes Beginning in *h*

- During construction of the database and subsequent work in it, it came to our attention that non-initial elements beginning in *h* appeared to undergo rendaku more frequently
- In other words, we hypothesized that words such as *hata* 'flag', *hukuro* 'bag' or *hi* 'fire', when appearing as the non-initial element in compounds, appear to be more susceptible to rendaku than non-initial elements beginning in any of the other three voiceless obstruent, *k*, *s* or *t*
- Today's presentation will outline our method for testing this hypothesis, as well as proposing a possible reason for why this should be

Pre-Analysis

- It is well known that verbs dampen rendaku [12]
- It is also well known that rendaku operates sporadically in the Sino-Japanese and hardly ever in the foreign stratum [12, 13, 14]
- In order to obtain an accurate analysis of rendaku rates, only Japanese (*wago*) non-verbs were considered
- In addition, in order to create an environment devoid of systematicity, the following compounds were also excluded:
  - those which are personal or place names
  - whose non-initial elements are abbreviations (e.g. *sasi* < *sashimi*)
  - whose non-initial elements are potentially analyzable as suffixes (e.g. *–sama)*
  - whose non-initial elements possess an unbound voiced allomorph (e.g. *kawa* ~ *gawa*)

Pre-Analysis

- Finally, only those compounds which appear in both the dictionaries employed in the Rendaku Database were considered
- In this way, ‘dictionary fossils’ or obsolete compounds did not come under analysis
- The remaining ‘commonplace’ compounds are arguably more likely an active part of a Japanese native speaker’s lexicon

Pre-Analysis

- Each compound was then assigned a binary status: whether or not it exhibits rendaku
- Comparison across dictionaries yielded 3 possibilities:

  i. A rendaku reading only (= exclusively rendaku)
  ii. A non-rendaku reading only (= exclusively non-rendaku)
  iii. Both rendaku and non-rendaku readings (= ambivalent)

- As the frequency of rendaku in ambivalent cases could not be determined, the 223 compounds fitting this description were also excluded

The DATABASE

- This left us with a residue, which will henceforth be referred to as the DATABASE
- The DATABASE consists of 6917 compounds, 20% the size of the original rendaku database

Rendaku Occurrence

non-initial element	*n*	never	always	
all	6,917	25.2%	74.8%	**100%**
begins in *k*	2864	25.9%	74.1%	**100%**
begins in *s*	969	29.3%	70.7%	**100%**
begins in *t*	1567	32.7%	67.3%	**100%**
begins in *h*	1517	13.6%	86.4%	**100%**

Statistical Analysis

- A Somers' d test was conducted to determine whether rendakuable elements beginning in *h* incurred rendaku more frequently than those starting with *k*, *s* or *t*
- The results showed a statistically significant correlation of moderate strength ($\Delta = .135$, $p < .0005$)
- In other words, knowing whether a rendakuable element begins in *h* increases by 13.5% the likelihood of being able to predict whether that element will undergo rendaku
- Interestingly, the tendency for rendaku to occur less frequently as E2 initial phoneme place of articulation moves progressively away from the back of the mouth (i.e. /h/ > /k/ > /s/ > /t/) is also statistically significant ($\rho = -.136$, $p < .005$)

Labial Lenition

- For the remainder of this presentation, we will explore one possible reason why elements beginning in *h* (hereafter '*h*-elements') should show statistically significant elevated rendaku rates
- Labial lenition (hereafter 'LL') is a sound change that began around the 8th century and is, arguably, still ongoing – and thus incomplete – today
- 1,300 years is a long time and LL is well documented [15, 16, 17]
- LL can be divided into two types: initial LL and internal LL

Initial LL

- Initial LL was restricted to word-initial environment
- Here *p* > *f* > *h* (*f* = [ɸ])
- The first stage *p* > *f* was complete in most dialects by around the 10th century, while the second stage *f* > *h* was largely complete by around the 17th

pana > *fana* > *hana*	'flower'	
potaru > *fotaru* > *hotaru*	'firefly'	etc…

- Note, however, that *h* before *u* is still pronounced [ɸ] in the standard language today

Internal LL

- Internal LL was restricted to word-internal environment
- Here, in a highly simplified form *p* > *w* > *ø*
- The first stage *p* > *w* was complete in most dialects by around the 10th century, while the second stage *w* > *ø*, which occurred before all vowels except *a*, was largely complete by around the 18th

kapa > *kawa*	'river'
kopori > *kowori* > *koori*	'ice'
ipye > *ife* > *iwe* > *iye* > *ie*	'house' etc…

Phonological Word

- Two interesting LL pathways are:

  pa > *fa* > *wa* (は) 'topic particle'

  pye > *fe* > *e* (へ) 'to (motion)'

- Both are clitics or agglutinative particles: the domain of 'initial' and 'internal' refers to the phonological word

  kapa.pe > *kawa.we* > *kawa.e* 'to the river'

Rendaku Compounds: Chids

- Not all of what are analysed as compounds are perceived as such by speakers
- Rendaku is found with, for example:

kimono	(= wear + thing)	*sakasa.**g**imono*
koya	(= small + house)	*inu.**g**oya*
siai	(= do + together)	*syooka.**z**iai*

- This is despite the fact that the right branch condition dictates that rendaku is blocked when a compound is itself part of a larger compound
- We will call such compounds 'chids' and let Q = the proportion of compounds which are chids
- We assume that chids are evenly spread throughout the speaker lexicon,and that Q is not subject to linguistic constraints

- When a rendakuable *k-*, *s-* or *t*-element fails to undergo rendaku, it may be because of (i) reasons unknown or because (ii) the compound is a chid

tosi	+	*kasa*	>	*tosi.**k**asa (*tosi.**g**asa)*
'year'		'bulk'		'the older person'

- Despite the fact that *kasa* 'bulk' regularly undergoes rendaku, and the absence of constraints such as Motoori/Lyman's law preventing it from undergoing rendaku, in the case of *tosikasa,* it fails to do so

- This could be for a reason we are as yet not aware of, or because *kasa* in this compound has been bleached of its meaning and *tosikasa* is perceived by speakers as a chid: as a non-compound

- There is no way for us to ascertain which is the case

- With *h*-elements, however, the historical operation of internal LL means that we have no such problem

ura	+	*hara*	>	*ura.**h**ara (*ura.**b**ara)*
'back'		'belly'		'opposite, reverse'

- Despite the fact that *hara* 'belly' regularly undergoes rendaku, and and the absence of constraints such as Motoori/Lyman's law preventing it from undergoing rendaku, in the case of *urahara*, it fails to do so

- This cannot be because *urahara* is perceived by speakers as a chid, since if it were, internal LL would have caused *h* > *w*, e.g. *urahara* > *urawara*

- In this case, *hara* resists rendaku for reasons unknown

- The point of the foregoing discussion is that compounds such as *tosikasa* appear in the database and are tagged as not undergoing rendaku, while on the other hand a hypothetical compound such as **urawara* would not appear in the database (*wara* is not considered as an allomorph of *hara~bara* for database purposes)
- As we are assuming that 'chidness' occurs evenly throughout the lexicon, the absence of non-rendaku forms such as **urawara* in the database mean the overall rendaku rate for *h*-elements is not being dragged down
- Could this be the reason why *h*-elements show significantly higher rendaku rates?

w-allomorphs

- To prove this, we need to show that internal LL has occurred with sufficient frequency to account for the statistically significant higher rendaku rates found in *h*-elements
- Recall that, with internal LL, the first stage $p > w$ was complete in most dialects by around the 10th century: any '*w*-allomorphs' will be old
- *W*-allomorphs in the context of rendaku had been commented upon as long ago as 1910 in a paper by Ogura [18]
- He noted the following:

siranu	+	*hi*	>	*siranu.**i***
‘unknown’		‘fire’		‘St Elmo’s fire’
siri	+	*he*	>	*siri.**e***
‘hind’		‘vicinity’		‘rearwards’
i	+	*ho*	>	*i.**o***
‘5’		‘100’		‘500’

- All other known examples include the *h*-element *hara* ‘field’, and all except the first are family or place names

kusa	+	*hara*	>	*kusa.**w**ara*
‘grass’		‘field’		‘grassy plain’
huzi	+	*hara*	>	*huzi.**w**ara*
‘wisteria’		‘field’		‘Fujiwara (name)’
hagi	+	*hara*	>	*hagi.**w**ara*
‘clover’		‘field’		‘Hagiwara (name)’

In some cases, a family name has both *w-* and *b-* (*Kuwahara~Kuwabara*) or *w-* and *h-* allomorphs (*Sugawara~Sugahara*), indicating variation in chidness perception through time and across dialects, combined with as yet undiagnosed regional rendaku triggers

- The following two examples can also be cited but in neither of these cases is the non-initial element rendakuable:

yuku	+	*he*	>	*yuku.**e***
‘go’		‘area’		‘whereabouts’
hi	+	*hada*	>	*ki.**w**ada*
‘cypress’		‘skin’		‘cypress bark’

Conclusions

- Rendaku occurs statistically significantly more often in cases where the E2 starts with *h*
- The absence of chids involving labial lenition in the Rendaku Database does not provide an adequate explanation for this finding
- An alternate explanation might be that rendaku in these cases serves to maintain compound transparency by reinforcing E2 identity as *h* otherwise progressively becomes /ø/
- A linear trend of decreasing rendaku frequency, $h > k > s > t$, provides additional evidence of a phonemic rather than phonetic phenomenon, but further analysis is necessary before any definitive explanation can be offered

Conclusions (cont'd.)

- Rendaku is conditioned not only phonotactically but also morpho-syntactically, and the interaction between all the various governing factors is complex, for even well-attested constraints such as Mootori/Lymans's Law exhibit notable exceptions that cannot easily be accounted for
- The Rendaku Database provides a tool that allows researchers to identify, examine, and hopefully one day unravel the relevant relationships underlying rendaku

(1) Motoori, Norinaga. 1822. *Kojiki den.* Edo: Maegawa Rokuzaemon; Osaka: Matsumura Kyūbei; Kyoto: Imai Kihei; Owari: Katano Tōshirō.
(2) Lyman, Benjamin. 1894. Change from surd to sonant in Japanese compounds. *Oriental Club of Philadelphia*, 1-17.
(3) Otsu, Yukio. 1980. Some aspects of rendaku in Japanese and related problems. In *Theoretical issues in Japanese linguistics*, ed. by Yukio Otsu and Ann Farmer, 207-227. MIT Working Papers in Linguistics.
(4) Irwin, Mark; Miyashita, Mizuki & Russell, Kerri. 2016. The rendaku database version 3.0. Available online at http://www-h.yamagata-u.ac.jp/~irwin/site/Rendaku_Database.html.
(5) Shinmura, Izuru. 2008. ed. *Kōjien dai-6 han*. Tokyo: Iwanami Shoten.
(6) Watanabe, Toshirō, Edmund Skrzypczak & Paul Snowden (eds.). 2008. *Kenkyūsha's new Japanese-English dictionary digital supplemented edition*. Tokyo: Kenkyūsha.
(7) Irwin, Mark. 2012. Rendaku dampening and prefixes. *NINJAL Research Papers* 4: 27-36.
(8) Irwin, Mark. 2014. Rendaku across duplicate moras. *NINJAL Research Papers* 7:93–109.
(9) Irwin, Mark. 2012. Rendaku lovers, rendaku haters and the logistic curve. Paper presented at the 22nd Japanese/Korean Linguistics Conference, Tokyo.
(10) Irwin, Mark & Lyddon, Paul. 2016. Rendaku and Homophony. *Phonological Studies* 19:11-18.
(11) Irwin, Mark. 2016. The rendaku database. In *Perspectives on rendaku: sequential voicing in Japanese compounds*, ed. Vance, Timothy and Irwin, Mark.
(12) Vance, Timothy. 2015. Rendaku. In *Handbook of Japanese Phonetics and Phonology*, Haruo Kubozono (ed.), 397–441. Berlin: De Gruyter Mouton.
(13) Irwin, Mark. 2011. *Loanwords in Japanese*. Amsterdam: John Benjamins. pp. 150-153.
(14) Vance, Timothy. 1996. Sequential voicing in Sino-Japanese. *Journal of the Association of Teachers of Japanese* 30: 22-43.
(15) Martin, Samuel. 1987. *The Japanese language through time.* New Haven: Yale University Press.
(16) Kiyose, Gisaburō. 1985. 'Heianchō hagyōshiin-p onron [Evidence that Modern Japanese *h-* was *p-* in the Heian Period],' *Onsei no Kenkyū* 21, 73-87.
(17) Wenck, Günther. 1957. *Japanische Phonetik Vol. 3: die Phonetik des Sinojapanischen*. Otto Harrassowitz, Wiesbaden.
(18) Ogura Shinpei. 1910. Raiman-shi no rendakuron (Lyman's Rendaku Theory). *Kokugakuin Zasshi* 16: 9-23, 31-45.

Acknowledgments

- This research was supported by a Grant-in-Aid for Scientific Research (C) no. 26370436 from the Japan Society for the Promotion of Science
- The authors wish to extend their gratitude Timothy Vance at NINJAL

- A pdf of this paper is available on my website:
  http://www-h.yamagata-u.ac.jp/~irwin/site/Home.html
  or search Mark Irwin Yamagata University