Japanese Language and Culture[1]

An Introduction to Hiragana

9 January 2016

日本語: Hiragana (ひらがな)

- Several ways to describe *hiragana*, the main *kana* script:
  - Using *rōmaji* (romanization of Japanese script)
  - Using the IPA system to describe sounds
  - Using hiragana itself (metaphysical)
  - Using *kanji*, a more advanced script

Rōmaji	**IPA**	**Hiragana**	**Kanji**
hiragana	[çiɽagana]	ひらがな	平仮名

Table 1: Writing the word *hiragana* in various scripts.

- Hiragana forms the **core** of the Japanese sound system/written system
  - But NOT omnipresent (i.e. see extra sounds in *katakana*)
  - Each symbol in hiragana represents one Japanese *mora*
    - *Mora*: A unit of sound that determines where the syllables/stresses/weights go in the phonology of a language (pl. *morae*)
  - Linguistic constructs used in hiragana: *monographs*, *diacritics*, *digraphs*, *digraphs with diacritics*
    - *Monographs* (hiragana: ごじゅうおん, rōmaji: *gojūon*):
      - ☐ A single **syllabographic** character (one symbol)
    - *Diacritics* (hiragana: ごじゅうおんはんだくてん, rōmaji: *gojūon (han)dakuten*):
      - ☐ Marks added to monographs to indicate different pronunciations than the ones they are usually given
    - *Digraphs* (hiragana: ようおん, rōmaji: *yōon*):
      - ☐ Two juxtaposed syllabographic characters (side-by-side)
    - *Digraphs with diacritics* (hiragana: ようおんはんだくてん, rōmaji: *yōon (han)dakuten*):
      - ☐ Marks added to digraphs to indicate different pronunciations than the ones they are usually given
    - Each digraph (with or without diacritics) generally corresponds to one mora
- Thus, learning hiragana comprises learning **all** the possible Japanese sounds (exceptions: certain katakana sounds)
- Ignore kanji for now: Much more complex *stroke order* (often requiring between 10 to 15 strokes)
  - cf. hiragana (and katakana, actually): No kana requires more than 4 strokes!
- Hiragana characteristics:
  - Smooth, curvy (similar to cursive in English)
  - Fluid, can be written down quickly if stroke order is correct
- Simplicity of the syllabary's sound system
  - Don't even need to know how to pronounce all 46 basic characters
  - Just need to know the pronunciation of 5 vowels (important to get these correct immediately, as it builds)
  - All consonant sounds are based off of the vowel sounds: *consonant* = ○ + *corresponding vowel* (e.g. $ka = k + a$)
  - Unlike English, *a* (あ) is always [a] in hiragana, for example (convenient pronunciation rules arise as a result)

[1]Designed and structured by Chirag Bharadwaj, Cornell University, *B.Sc.* Computer Science, 2017.

- All vowels are **short** unless marked otherwise
  - We will learn marks for long vowels later on
  - *Interlingual differences*: Some sounds may be different than in English
  - *Intralingual differences*: Some sounds may be compressed, others may not be compressed
  - This is why IPA is important!
    - Need to speciy the differences and subtleties in sound voicings
    - Need to show what is actually the same despite different appearances/romanizations (i.e. rōmaji ≠ perfect)
- Approach to learning/practicing hiragana:
  - Study approximately 20 per day (4 sets of 5 hiragana syllabograms each)
  - Memorize using the **mnemonics**, practice by making a chart like this for each set of 5 (notice the stroke order):

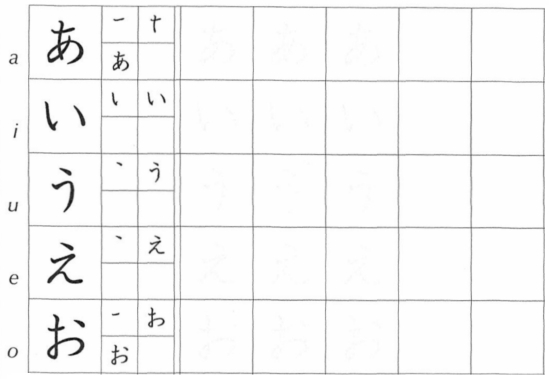


Figure 1: An outline of how we should set up our notebooks for practicing kana stroke order.

  - Make this table, write down the stroke order, reproduce the kana five times while covering up order/appearance
  - Eventually break the associations with the mnemonics/pictures, move to pure recognition/reproducibility
  - Learn the correct stroke order using the **documentation on stroke order**—very important!!
- So without further ado, let's introduce all the characters! A lot of sets of 5 kana will have exceptions in the '*i*' and '*u*' columns (using stronger sounds instead of the expected weaker sounds). Monograph symbols, rōmaji, mnemonics.

Monographs (ごじゅうおん)

(a) H: あ, R: *a*, IPA: [a]
mnemonic: capital '**A**'

(b) H: い, R: *i*, IPA: [i]
mnemonic: two **ee**ls

(c) H: う, R: *u*, IPA: [ɯ̟]
mnemonic: sideways '**u**'

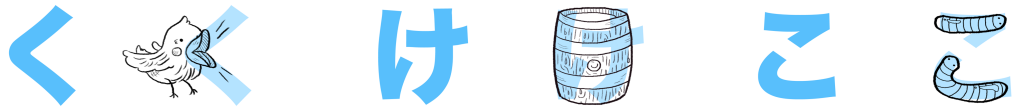
(d) H: え, R: *e*, IPA: [e]
mnemonic: **e**xotic bird

(e) H: お, R: *o*, IPA: [o]
mnemonic: two '**o**'s

Figure 2: The five vowels in hiragana. The pronunciations/order is different from English, however. Note that the '**u**' is a compressed sound, like in the Japanese name *Ts**u**baki-chan* or in the English word *h**oo**d*.

(a) H: か, R: *ka*, IPA: [ka]
mnemonic: **ca**n-can dance

(b) H: き, R: *ki*, IPA: [ki]
mnemonic: **key**



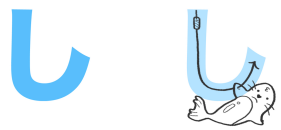
(c) H: く, R: *ku*, IPA: [kɯ̟]
mnemonic: **cu**ckoo bird

(d) H: け, R: *ke*, IPA: [ke]
mnemonic: **ke**g of beer

(e) H: こ, R: *ko*, IPA: [ko]
mnemonic: **co**habitation

Figure 3: The five *k*-* hiragana syllabograms. No exceptions in pronunciation.

(a) H: さ, R: *sa*, IPA: [sa]
mnemonic: **si**gn



(b) H: し, R: *shi*, IPA: [ɕi]
mnemonic: fi**shy** hook

(c) H: す, R: *su*, IPA: [sɯ̟]
mnemonic: **sw**ing

(d) H: せ, R: *se*, IPA: [se]
mnemonic: **sa**bertooth

(e) H: そ, R: *so*, IPA: [so]
mnemonic: **so**ngbird

Figure 4: The five *s*-* hiragana syllabograms. Note that there is an exception for **shi** (not *si*).

(a) H: た, R: *ta*, IPA: [ta]
mnemonic: looks like **ta**

(b) H: ち, R: *chi*, IPA: [t͡ɕi]
mnemonic: no **chi**n

(c) H: つ, R: *tsu*, IPA: [t͡sɯ̟]
mnemonic: **two** fish hook

(d) H: て, R: *te*, IPA: [te]
mnemonic: **te**n

(e) H: と, R: *to*, IPA: [to]
mnemonic: **toe**nail

Figure 5: The five *t*-* hiragana syllabograms. Note that there are exceptions for **chi** and **tsu** (not *ti* or *tu*).

(a) H: な, R: *na*, IPA: [na]
mnemonic: **nu**n praying

(b) H: に, R: *ni*, IPA: [ni]
mnemonic: **nee**dle

(c) H: ぬ, R: *nu*, IPA: [nɯ̟]
mnemonic: **noo**dles

(d) H: ね, R: *ne*, IPA: [ne]
mnemonic: **ne**ko (*cat*)

(e) H: の, R: *no*, IPA: [no]
mnemonic: **no** smoking

Figure 6: The five *n*-* hiragana syllabograms. No exceptions in pronunciation.

(a) H: は, R: *ha*, IPA: [ha]
mnemonic: Looks like **ha**

(b) H: ひ, R: *hi*, IPA: [çi]
mnemonic: **he** looks sad

(c) H: ふ, R: *fu*, IPA: [ɸɯ̟]
mnemonic: **hu**la dancer

(d) H: へ, R: *he*, IPA: [he]
mnemonic: Mt.St.**He**lens

(e) H: ほ, R: *ho*, IPA: [ho]
mnemonic: **he**adless guy!

Figure 7: The five *h*-* hiragana syllabograms. Note that **hi** is *palatalized* and **fu** is an exception (not *hu*).

(a) H: ま, R: *ma*, IPA: [ma]
mnemonic: **mo**nster

(b) H: み, R: *mi*, IPA: [mi]
mnemonic: 21? **Me**? No.

(c) H: む, R: *mu*, IPA: [mɯ̟]
mnemonic: Cows say **moo**

(d) H: め, R: *me*, IPA: [me]
mnemonic: **me** (*eye*)

(e) H: も, R: *mo*, IPA: [mo]
mnemonic: **mo**re fishies

Figure 8: The five *m*-* hiragana syllabograms. No exceptions in pronunciation.

(a) H: や, R: *ya*, IPA: [ja]
mnemonic: **ya**k

(b) H: ゆ, R: *yu*, IPA: [jɯ̟]
mnemonic: **u**nique fish

(c) H: よ, R: *yo*, IPA: [jo]
mnemonic: **yo** wtf!

Figure 9: The three *y*-* hiragana syllabograms. No exceptions in pronunciation. Note that **yi** and **ye** do not exist (nor have they ever), as *i* and *e* are near-perfect approximants of the same sounds (i.e. they have always been used in their place).

(a) H: ら, R: *ra*, IPA: [ɾa]
mnemonic: DJ **ra**pping

(b) H: り, R: *ri*, IPA: [ɾi]
mnemonic: two **ree**ds

(c) H: る, R: *ru*, IPA: [ɾɯ̟]
mnemonic: **rou**te 3

(d) H: れ, R: *re*, IPA: [ɾe]
mnemonic: **re**tching

(e) H: ろ, R: *ro*, IPA: [ɾo]
mnemonic: third **ro**ad

Figure 10: The five *r*-* hiragana syllabograms. The '**r**' sound in Japanese is a tricky one to master. It requires a **flap** of the tongue (hence the name for the corresponding phoneme: *retroflex flap*). The result has aspects of the English '**r**', '**d**', and '**l**'.

(a) H: わ, R: *wa*, IPA: [ɯ̟a]
mnemonic: **wa**sp

(b) H: ゐ, R: *wi/i*, IPA: [ɯ̟i]
mnemonic: **we** are 2gether

(c) H: ゑ, R: *we/e*, IPA: [ɯ̟e]
mnemonic: looks like **we**

(d) H: を, R: *wo*, IPA: [ɯ̟o]
mnemonic: **whoa**!

(e) H: ん, R: *n'*, IPA: [n]
mnemonic: looks like **n**

Figure 11: The four *w*-* and the *n'* hiragana syllabograms. The '**w**' is a compressed sound. Note that **wi** and **we** are very rare, as *modern Japanese* (i.e. the post-World War II dialect) uses *i* and *e* in their place. They are mostly obsolete outside of Okinawan dialects of Japanese (southern islands, near Taiwan). There is no **wu** in Japanese.

	-a	*-i*	*-u*	*-e*	*-o*
ϕ	あ	い	う	え	お
k-	か	き	く	け	こ
s-	さ	し	す	せ	そ
t-	た	ち	つ	て	と
n-	な	に	ぬ	ね	の
h-	は	ひ	ふ	へ	ほ
m-	ま	み	む	め	も
y-	や		ゆ		よ
r-	ら	り	る	れ	ろ
w-	わ	ゐ		ゑ	を

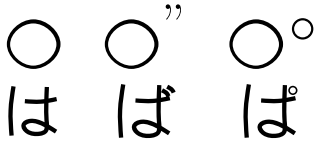
Table 2: Table of monographs (ごじゅうおん, or *gojūon*) in hiragana. This is meant as a summary.

Diacritics (ごじゅうおんはんだくてん)

- The monographs form the basic sounds/written scripts of hiragana
- Can make things slightly more complicated in multiple ways
  - Combining sounds through variational marks (diacritics)
  - Combining symbols through juxtaposition (digraphs)
  - Simple in principle if the gojūon are understood well
- The *dakuten* are **voicing marks** used as diacritic signs
- Indicates that the corresponding modified consonant is **voiced** instead of **unvoiced** (using the vocal chords to vibrate/articulate the sound)
- There are also *handakuten*, for **semi-voiced** consonants (*plosives*)
- We should know how to write these in hiragana by now, as we can write them using gojūon:
  - Dakuten: だくてん
  - Handakuten: はんだくてん
- Denotation of diacritics (○ represents an arbitrary consonant monograph):

dakuten *handakuten*

- Basic idea of how diacritical marks/voicing principles work:



ha **ba** **pa**

Table 3: An example using the diacritical marks. Here, **ba** is with a voiced *b*, and **pa** is with a semi-voiced *p*. The *h*-* kana are the **only** ones with semi-voiced consonants available (handakuten) in addition to the voiced consonants (dakuten).

- So the general idea of diacritics have been laid out:
  - Easy once monographs are known, as these are just modification marks
  - Simple *voice shifts* using standard phonetic phenomenon from linguistics (i.e. IPA knowledge should guide us)
  - Pronounce each consonant through vibration of the mouth rather than air from back of mouth

- We will establish the **official** phonetic shifts as a result of this voicing process now:

	gojūon	**dakuten**	**handakuten**
k-	*k-*	*g-*	NONE
s-	*s-*	*z-*	NONE
t-	*t-*	*d-*	NONE
h-	*h-*	*b-*	*p-*

Table 4: Only some consonants can be voiced (the other ones are already voiced). These are natural phonetic shifts due to voicing. Sometimes, linguists use *ng-* as a handakuten for *k-*, but this is not ever used in the Japanese language itself.

- What about kana with exceptions in pronunciation in the gojūon?
  - e.g. し (*shi*). Is じ pronounced *zi*? Isn't that a "weaker"/semi-voiced sound?
  - **No**: Need to observe IPA for these exceptions and determine how to voice them based on natural vibrations
  - e.g. じ is pronounced *ji*, not *zi* (subtle difference, but is important in understanding voicings of exceptions)
  - ...that is, **exceptions breed exceptions** (no exceptions to this rule!)
- The complete table of (han)dakuten for the gojūon is shown below with the corresponding rōmaji/IPA:

	-a	*-i*	*-u*	*-e*	*-o*
g-	が	ぎ	ぐ	げ	ご
	ga	gi	gu	ge	go
z-	ざ	じ	ず	ぜ	ぞ
	za	ji	zu	ze	zo
d-	だ	ぢ	づ	で	ど
	da	dzi	dzu	de	do
b-	ば	び	ぶ	べ	ぼ
	ba	bi	bu	be	bo
p-	ぱ	ぴ	ぷ	ぺ	ぽ
	pa	pi	pu	pe	po

(a) The official table of dakuten and handakuten, using rōmaji to characterize each sound.

	-a	*-i*	*-u*	*-e*	*-o*
g-	が	ぎ	ぐ	げ	ご
	[ga]	[gi]	[gu̟]	[ge]	[go]
z-	ざ	じ	ず	ぜ	ぞ
	[za]	[d͡ʑi]	[zu̟]	[ze]	[zo]
d-	だ	ぢ	づ	で	ど
	[da]	[d͡ʑi]	[zu̟]	[de]	[do]
b-	ば	び	ぶ	べ	ぼ
	[ba]	[bi]	[bu̟]	[be]	[bo]
p-	ぱ	ぴ	ぷ	ぺ	ぽ
	[pa]	[pi]	[pu̟]	[pe]	[po]

(b) The official table of dakuten and handakuten, using IPA to characterize each sound.

Table 5: The table of dakuten/handakuten for gojūon. Notice that the pronunciations of **ji/dzi** and **zu/dzu** are *similar*.

- Some authors also voice the vowel **u** to **vu** with the dakuten diacritical mark (mostly uncommon)

Digraphs (ようおん)

- We have **changed sounds** through voice shift and semi-voicings
  - ...but can we combine sounds?
  - ...but can we combine symbols?
- The *yōon* are **palatalization marks** as mora signs
  - Similar to how the dakuten are voicing marks used as diacritic signs
- Indicates that the corresponding modified consonant has a *palatalized central phoneme*
  - i.e. contractions, dipthongs, etc. in English
- In hiragana: ようおん (yōon)... we will talk about why it is "youon" later, similar to how gojūon is romanized "gojuuon" from the hiragana ごじゅうおん (these are the foretold long vowels)
- Construction: Standard *i*-column consonant prefix with a **smaller-sized** *ya*, *yu*, or *yo* kana annexed on

- Denotation of digraphs (○ represents an arbitrary consonant monograph):

○や	○ゆ	○よ
-ya	*-yu*	*-yo*

- For comparison in size with the **regular-sized** *ya*, *yu*, or *yo* kana in juxtaposition (these are **not** digraphs!):

○や	○ゆ	○よ

- Basic idea of how diacritical marks/voicing principles work:

○	○や	○ゆ	○よ
き	きゃ	きゅ	きょ
ki	**kya**	**kyu**	**kyo**

Table 6: An example using digraphs for the *k*-*kana. Here, the **ki** is merged with the *a*, *u*, and *o* sounds.

- So the general idea of digraphs have been laid out:
  - Managed to develop a way to combine certain sounds through juxtaposition of symbols (glide/palatalization)
  - More complicated model of our hiragana syllabary—complex *combination kana* based on existing characters
  - However, again easy once monographs are known, as these are just modification marks too
  - Need to be careful: SMALLER versions of the *y*-* kana are used! (i.e. get better at size discrimination)
  - e.g. *kyo* (digraph) is not the same as *kiyo* (def.: skilled): きょ v. きよ (subtle difference in size)
  - Doesn't make sense to have digraphs for *y*-* kana: *ya*, *yu*, *yo* are reproduced anyways
- Again, **exceptions breed exceptions**, so *shi* would create *sha*, *shu*, *sho*, **not** *sya*, *syu*, *syo* (stronger sounds)
  - Similarly for *chi*: digraphs are *cha*, *chu*, *cho*, **not** *tya*, *tyu*, *tyo* (again, stronger sounds, not weaker ones)
- The complete table of yōon is shown below with the corresponding rōmaji/IPA:

	-ya	*-yu*	*-yo*
k-	きゃ	きゅ	きょ
	kya	kyu	kyo
s-	しゃ	しゅ	しょ
	sha	shu	sho
t-	ちゃ	ちゅ	ちょ
	cha	chu	cho
n-	にゃ	にゅ	にょ
	nya	nyu	nyo
h-	ひゃ	ひゅ	ひょ
	hya	hyu	hyo
m-	みゃ	みゅ	みょ
	mya	myu	myo
r-	りゃ	りゅ	りょ
	rya	ryo	ryu

(a) The official table of yōon, using rōmaji to characterize each sound.

	-ya	*-yu*	*-yo*
k-	きゃ	きゅ	きょ
	[kʲa]	[kʲʉ̟]	[kʲo]
s-	しゃ	しゅ	しょ
	[ɕa]	[ɕʉ̟]	[ɕo]
t-	ちゃ	ちゅ	ちょ
	[t͡ɕa]	[t͡ɕʉ̟]	[t͡ɕo]
n-	にゃ	にゅ	にょ
	[nʲa]	[nʲʉ̟]	[nʲo]
h-	ひゃ	ひゅ	ひょ
	[ça]	[çʉ̟]	[ço]
m-	みゃ	みゅ	みょ
	[mʲa]	[mʲʉ̟]	[mʲo]
r-	りゃ	りょ	りゅ
	[ɾʲa]	[ɾʲʉ̟]	[ɾʲo]

(b) The official table of yōon, using IPA to characterize each sound.

Table 7: The table of yōon in hiragana.

- Side note: *myu* is very rare, only occurring naturally in the family name Omamyūda, or おまみゅうだ (小豆生田)

- We have successfully combined symbols and sounds, but there is still a problem:
  - Two systems of changing/modifying the hiragana characters:
    - Diacritics: For voiced/unvoiced consonant shifts
    - Digraphs: For combining sounds into less fundamental units of sound (merging vowels)
  - Can we somehow combine them and cover all of the bases?
    - Yes, by using *digraphs with diacritics*

Digraphs with Diacritics (ようおんはんだくてん)

- Let us consider the relationship between the four types of syllabograms now:

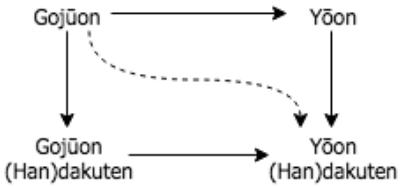


Figure 12: The relationships between the monographs, diacritics, digraphs, and digraphs with diacritics in Japanese. The solid lines represent direct (explicit) relationships, while the dashed line represents an indirect (implicit) relationship.

- We now just need to combined voiced sounds with the merged-vowel sounds (i.e. diacritics and digraphs combined)
- Simple enough: Use the dakuten/handakuten symbol to voice the consonant, then merge with *y-** kana like before
- As usual, **exceptions breed exceptions**, so some kana will be "unusual"
- The complete table of (han)dakuten for the yōon is shown below with the corresponding rōmaji/IPA:

	-ya	*-yu*	*-yo*
g-	ぎゃ gya	ぎゅ gyu	ぎょ gyo
z-	じゃ ja	じゅ ju	じょ jo
d-	ぢゃ dzya	ぢゅ dzyu	ぢょ dzyo
b-	びゃ bya	びゅ byu	びょ byo
p-	ぴゃ pya	ぴゅ pyu	ぴょ pyo

(a) The official table of dakuten and handakuten, using rōmaji to characterize each sound.

	-ya	*-yu*	*-yo*
g-	ぎゃ [gʲa]	ぎゅ [gʲʉ]	ぎょ [gʲo]
z-	じゃ [d͡ʑa]	じゅ [d͡ʑu]	じょ [d͡ʑo]
d-	ぢゃ [d͡ʑa]	ぢゅ [d͡ʑu]	ぢょ [d͡ʑo]
b-	びゃ [bʲa]	びゅ [bʲʉ]	びょ [bʲo]
p-	ぴゃ [pʲa]	ぴゅ [pʲʉ]	ぴょ [pʲo]

(b) The official table of dakuten and handakuten, using IPA to characterize each sound.

Table 8: The table of dakuten/handakuten for yōon. Notice the *similarity* in pronunciation of **ja/dzya**, **ju/dzyu**, **jo/dzyo**.

- Note that the *j-** kana and the *dz-** kana are **not** interchangeable despite same IPA! But it's a bit more complicated than that:
  - *Rendaku voicing*: sequential voicing in Japanese morphophonology that governs how consonants are voiced if they repreat in sequence
  - Example of rendaku voicing: *hito* (person), *hito* + *hito* = *hitobito* (person + person = people): ひと v. ひとびと
  - Another one: *ori* (fold), *kami* (paper), *ori* + *kami* = *origami* (fold + paper = paperfolding): おりがみ
  - In other words, there is an unexpected voicing on the initial syllable of the joined/annexed word
  - Mostly an advanced concept, governed by lexical rules (Lyman's law, etc.)—more on this later
  - If the first two syllables of a word consist of one syllable without dakuten marks followed by the same syllable with dakuten marks, then the same hiragana is used to write the sounds, presumably

    - *chijimeru* (def.: to boil down, to shrink): ちぢめる, **not** ちじめる
    - *tsudzuku* (def.: to continue): つづく, **not** つずく
  - But if one of the syllables has rendaku voicing, then the **original** hiragana is used:
    - *hana* (nose), *chi* (blood), *hanadzi* (nosebleed) by rendaku: はなち, **not** はなじ
  - Adding to the agony: On the other hand, ちゃ/ちゅ/ちょ are possible in rendaku, but rarely used (cf. じゃ/じゅ/じょ)
  - Again, advanced concept, so we will revisit this in the future
  - Bottom line, though: Cannot just claim that we use the *j-** kana and the *dz-** kana interchangeably
- We have successfully combined symbols and sounds, and we have covered all of our bases:
  - **Complete** model of hiragana now, built from the ground-up
  - Also examined from the top-down: Looked at existing words and understood how they fit into these patterns
  - Very complex combination kana based on 46^2 basic hiragana characters
  - Can represent a range/multitude of sounds based off of 5 very simple ones
  - We will use these sounds to represent **classical** Japanese concepts
  - What about foreign loanwords that utilize sounds not present among these 107^3 characters?
    - We need to use katakana for this purpose! Need to introduce additional sounds on top of the existing ones (since foreign languages also use these sounds)
    - Thus, the sounds of hiragana ARE the sounds of Japanese
    - **Pitfalls**: Misusing hiragana/katakana for certain "reborrowed" words: e.g. *waifu* (ワイフ v. わいふ)

Functional Marks and Other Aspects of Hiragana

- Similar to (small) や/ゆ/よ, there exists a っ character (as opposed to つ) in hiragana
- This so-called "small *tsu*", or っ, is used as a *geminate* in hiragana
  - *Geminate*: Consonant elongation where the consonant is spoken for audibly longer than the short version
  - Called *sokuon* in Japanese (そくおん), used for reduplication purposes
  - Similar to '*kk*' in English, e.g. in *bookkeeping* (slightest pause between *k*'s)
  - e.g. a good way to think about it is *book'keeping*, where the apostrophe signifies a pause
- Affects rōmaji through "twinning" of the affected consonant
- Denotation of sokuon (○ represents an arbitrary kana—could be monograph, diacritic, digraph, or digraph+diacritics):

sokuon

- That is, the sokuon mark comes **before** the affected kana:

いた	*ita*	かった	*katta* (won)
いった	(cf.) *itta*	さっか	*sakka* (writer)
けこう	*kekou*	はっぱ	*happa* (leaf)
けっこう	(cf.) *kekkou*	ざっし	*zasshi* (magazine)

Table 9: The left column shows geminated v. ungeminated words, while the right column shows common uses of the sokuon.

- However, to have the geminate effect for the *n-** kana, we need to use the ん instead of the sokuon
- Solution:

-nna: ～んな　*-nni*: ～んに　*-nnu*: ～んぬ　*-nne*: ～んね　*-nno*: ～んの

- Has the same effect as the sokuon:

さんねん　*sannen* (three years)　あんない　*annai* (guide)

[2] If we paid attention closely, we'd see that it is actually 48, but we did not use two of them beyond an introduction.
[3] 109, as noted above.

- For digraphs, only the **first** consonant is duplicated, except for *ch*, which duplicates as *tch*
- These techniques won't work for vowels, though
  - If we could reduplicate vowels, we would end up with a way to develop a long-vowel system
  - Similar to English... is there a way to just juxtapose two vowels with the intention of conveying "long vowel"?
  - Yes, can use a technique similar to how we twinned the *n-** kana
- We will explore the technique via example
  - When to double with the same vowel, when to use other vowels instead to simulate "doubling"
  - Does it affect meaning which way we double? Yes...
- There are some exceptions, as usual (mostly historical reasons, another recurring theme in Japanese language study)
- *aa*
  - おばあさん (*obaasan*, or grandmother) vs. おばさん (*obasan*, or aunt)
- *ii*
  - おじいさん (*ojiisan*, or grandfather) vs. おじさん (*ojisan*, or uncle)
- *uu*
  - すうじ (*sūji*, or number)
- *ee*
  - Hiragana **usually** achieves *ee* with *ei*, as in せんせい, or *sensei* (master)
  - えいが (*eiga*, or movie) uses *ei*, but おねえさん (*oneesan*, or sister) is true *ee*
- *oo*
  - Hiragana **usually** achieves *oo* with *ou*, as in ありがとう, or *arigatou* (thanks)
  - ほうりつ (*hōritsu*, or law) uses *ou*, but とお (*too*, or ten) is true *oo*
- Note that these are not pronounced like in English either, e.g. *ee* is *eh-eh*, not like the letter 'E' in English
- Two schools of thought on how rōmaji should look for these extended vowels:
  - Both schools of thought agree that the ones with differing vowels should be written out explicitly, e.g. えいが should **always** be *eiga*, and never *ēga*
  - However, for the ones with the **same**-vowel duplications, there is a split philosophy:
    - Some people like to be consistent and write it out, e.g. おねえさん should always be *oneesan*, never *onēsan*
    - Others like to use the macron to suggest "long vowel": おじいさん should always be *ojīsan*, never *ojiisan*
  - What we will use: a hybrid philosophy
    - For some words we will use the macron, for others we will not—just whatever will convey the meaning better
    - We will also ignore the part that both schools of thought agree on: We spell it as *rōmaji*, not *roumaji*
    - ...but we will say *obaasan*, *ojiisan*, *sūji*, *sensei*, *eiga*, *oneesan*, *arigatou*, *hōritsu*, *too*
- Official rules for **our** romanization techniques:
  - *aa*: Written out as *aa*.
  - *ii*: Written out as *ii*.
  - *uu*: Written out as *uu* if at the end of some form of a verb, otherwise is indicated by a macron: *ū*.
  - *ee*: Written out as *ee*.
  - *oo*: Written out as *oo*.
  - *ou*: Written out as *ou* if at the end of some form of a verb, otherwise is indicated by a macron: *ō*.
  - *ei*: Written out as *ei*.
  - All other combinations: Written out explicitly, without macrons.
- These rules form a variation on what we know as the *traditional Hepburn romanization* of Japanese

- Some other basic problems related to pronunciation include how ん works:
  - Full syllable in terms of length, but pronunciation varies with phoneme that follows it
  - If followed by a vowel or if it **ends** a phrase, then ん indicates that the precending vowel is nasalized (like the word "bon" in French), i.e. [ɴ], [ũ], or [ĩ]:

    れんあい (*rẽai*, or romance) ほん (*hõ*, or book)
  - When followed by *n*-*, *t*-*, *d*-*, *s*-*, or *z*-* kana, ん is pronounced [n]:

    おんな (*onna*, or woman)
  - When followed by *m*-*, *b*-*, or *p*-* kana, ん is pronounced [m]:

    さんほ (*sampo*, or stroll)
  - When followed by *k*-* or *g*-* kana, ん is pronounced [ŋ]:

    まんが (*manga*, or comics)
  - That is, **ん is a *nasalizer*, while っ is a *palatalizer***
- We consider the fact that sometimes we would like to duplicate entire morae, not just consonants: e.g. *jiji*
  - We can use what is called an iteration mark (*odoriji*) to signify this
  - For example, instead of writing じじ, we could write じ ゞ, where the **voiced** iteration mark ゞ tells us to duplicate the previous mora and voice it (notice the dakuten on the iteration mark)
  - There is also an **unvoiced** iteration mark: ゝ
  - We can take the previous mora and unvoice it with the second mark: じ ゝ would be *jishi*
  - Makes writing certain longer words out a bit more convenient
- Other marks that are of importance include punctuation marks in Japanese:

  Period: 。 Comma: 、 Single quotations: 「 」 Double quotations: 『 』 Wave dash: ～ Question mark: ？
  Exclamation mark: ！ Interpunct: ・ Asterisk: ※ Groupings: （） , 〔 〕 , ［］ , ｛｝ , 〈〉 , 《》 , 【】 , etc.
- These marks (*sokuon*, ん, long vowels, *odoriji*, punctuation, etc.) are called **functional marks** of the language
- Some exceptions in pronunciation: when certain kana are used as *particles*
  - *Particle*: Part of speech that cannot be inflected, declined, or conjugated (i.e. adverbs, prepositions, conjunctions, and interjections are all particles)
  - The kana は (*ha*), を (*wo*), and へ (*he*) are pronounced as *wa*, *o*, and *e*, respectively, when used as particles
  - Other than these "standalone word" particles, Japanese is phonemically orthographic: one-to-one correspondence between kana and respective sounds
  - Pitch accent is a different story, of course... requires more advanced study
  - More on this later—it is a grammatical construct that we will learn in the future
- Lastly, some interesting history about hiragana's origins:
  - When first developed as a simplication, hiragana was not universally accepted
  - The elite and educated/scholarly few preferred to use the kanji system to write scriptures
  - However, women were not allowed this level of education at the time, and thus preferred hiragana for easier use
  - Thus, an alternative name for hiragana is often おんなで, or *onnade* (women's writing)
  - Over time, male authors began adopting hiragana for literature and unofficial writing, like personal letters
  - Thus, for all syllables, there was more than one possible hiragana, due to complex kanji interactions
  - In 1900 BCE, linguists simplified the system so that each syllable[4] had exactly one associated kana
  - The pre-1900 deprecated hiragana that was closer to kanji is now called *hentaigana*
  - We will later learn about a pangram poem called *Iroha-uta*, which uses each hiragana character exactly once

[4] Technically, each mora, not each syllable.

a	あ	一	𠂇	あ
i	い	丨	い	
u	う	丶	う	
e	え	一	え	
o	お	一	お	お
ka	か	丿	カ	か
ki	き	一	二	き
ku	く	く		
ke	け	丨	𠂆	け
ko	こ	一	こ	
sa	さ	一	さ	
shi	し	し		
su	す	一	す	
se	せ	一	廿	せ
so	そ	そ		
ta	た	丿	𠂇	𠂉	た

chi	ち	一	𠂉	ち
tsu	つ	つ		
te	て	て		
to	と	丶	と	
na	な	丿	𠂇	𠂉	な
ni	に	丨	𠂆	𠂉	に
nu	ぬ	丿	ぬ	
ne	ね	丨	ね	
no	の	の		
ha	は	丨	𠂆	𠂉	は
hi	ひ	ひ		
fu	ふ	丶	う	𠂉	ふ
he	へ	へ		
ho	ほ	丨	𠂆	𠂉	ほ
ma	ま	一	二	ま	
mi	み	み	み	

mu	む	こ	む	む
me	め	し	め	
mo	も	し	も	も
ya	や	つ	う	や
yu	ゆ	ゆ	ゆ	
yo	よ	ー	よ	
ra	ら	、	ら	
ri	り	い	り	
ru	る	る		
re	れ	|	れ	
ro	ろ	ろ		
wa	わ	|	わ	
wi	ゐ	ゐ		
we	ゑ	ゑ		
wo	を	ー	ナ	を
n'	ん	ん		