

Basic rules for producing a *honkoku*

A *honkoku* 翻刻 intends to: 原本を忠実に伝えること. From this point of view, it is not dissimilar to the idea of ‘diplomatic transcription’ as defined in Western palaeography.

In our summer school we apply the following rules. You must treat these as our *hanrei* 凡例. The following rules are for texts written in *wabun* (vernacular Japanese) and to a certain extent *sōrobun*. They do not apply for text in *kanbun*.

1. *Hiragana* variants are to be transcribed using the *hiragana* shapes used in contemporary Japanese.

It is important to keep in mind the following points:

- 1.1. *Katakana*-looking ハ・ミ・ニ are to be transcribed as *hiragana* は・み・に (deriving from the *jibo* 字母: ハ・三・ニ). You need to be mindful that this rule is not shared by everyone, particularly in historical fields, so in other transcriptions these same *hiragana* can be transcribed as *katakana*. Needless to say, when they are clearly used as *katakana* (i.e., in onomatopoeia, loan words, etc.) they should be transcribed as *katakana*.
- 1.2. *Hiragana* ゑ・お and *katakana* エ・ヰ need to be kept as such.
- 1.3. The や・ゆ・よ in compounds such as しや・しゆ・しよ must be kept as full-size hiragana. For example: いしや should remain like this despite the fact that the word in modern script would be 医者 = いしゃ.
- 1.4. つ must be kept full size at all times. For example: ひつそり (with the meaning of ひっそり).

2. Transcribing *kanji* presents a number of challenges and rules to bear in mind:

- 2.1. You must transform *kyūji* 旧字 (old variants) into *shinji* 新字 (modern variants). For example: 佛 → 仏
醫 → 医 賣 → 売
- 2.2. Transform *itaiji* 異体字 (alternative variants) into mainstream shapes. For example: 桒 → 松

Making decisions around *itaiji* is complicated as different dictionaries label the same shape in different ways (*itaiji*, *zokuji*, etc.). There are no hard rules about how to treat these variants, so you will need to make decisions on a case by case basis. For example, the following variant shapes are retained in transcriptions: 躰 (normally classified as an *itaiji* of 体), 脉 (normally classified as an *itaiji* of 脈) 迺 (normally classified as an *itaiji* of 迄)

There are also shapes that might appear as *itaiji* but are not. They often appear in dictionaries as *dōji* 同字. They must be retained.. For example, 嶋 as compared to 島.

You can use a number of online materials in your decision-making process:

<https://jigen.net/>

8画 / 教育漢字 4年生 / 常用漢字

松

常用音訓 □ シヨウ □ まつ

成り立ち □ 形声#1

漢字構成 □ 口木公

発音 □ sōng

表示 □ U+677E □ 松

異体字 柏 鬆 糜 𣎵 窠 窠 忝 髻 髻 髻 髻

繁体字 鬆

声符「公」 公 佶 舩 翁 空 訟 蚣 頤

声符「松」 𣎵 蠅 淞 鬆 菘

<http://en.glyphwiki.org/wiki/GlyphWiki:MainPage>

Please be careful in using this site as it is not clear what dictionaries it uses as main references.

	松	(otakusei_hkrm-03086623) (=hdc_hkrm-03086623)
	松	(otakusei_hkrm-03086730) (=hdc_hkrm-03086730)
Variants: 𣎵 (u37e3) [入管正字]	𣎵	(u37e3)
	𣎵	(hkcs_m37e3)
窠 (u3ba4) [異体字 (戸籍統一文字)] [異體字 (漢語大字典)]	窠	(u3ba4) (=hdc_hkrm-03086741) (=u3ba4-ue0101) (=u3ba4-ue0100) (=koseki-172160) (=juki-3ba4) (=jmj-001772) (=dkw-15227) (=aj1-17834)
	窠	(toki-01045380) (=u3ba4-ue0102)
	窠	(hkcs_m3ba4)
糜 (u5eba) [入管正字] [異体字 (戸籍統一文字)] [異體字 (民國教育部)] [異體字 (漢語大字典)]	糜	(u5eba) (=dkw-09383) (=koseki-106080) (=jmj-011021)
	糜	(hkcs_m5eba)
𣎵 (u6780) [入管正字] [異体字 (戸籍統一文字)] [異體字 (民國教育部)] [異體字 (漢語大字典)]	𣎵	(u6780) (=u6780-ue0100) (=koseki-163510) (=juki-6780) (=jmj-013670) (=gt-18477) (=dkw-14517) (=aj1-19445)
	𣎵	(zlhai-011414)
	𣎵	(hkcs_m6780)

You can also use any 漢和辞典 that you are familiar with.

Be mindful that the above rules regarding the transcription of *kanji* are not universally agreed upon and may differ depending on the field you work in or the series you are transcribing for.

3. *Okurigana* remain unaltered. For example, if the text we have contains 云ける (where clearly the *okurigana* ひ is missing) we must retain what is in the original document. Needless to say, when reading the text aloud we must reconstruct and vocalize the *okurigana*: we read as if the *okurigana* were there but we do not transcribe them.
4. *Dakuten* 濁点 are to be displayed as they are in the original text. For example: わか身. Even if the context tells us that the meaning is 我が身, we cannot transcribe it as わが見. We need to retain the か without *dakuten*. Of course when we read the text aloud we need to vocalize this as *wagami*. Another common example is べからず - we know that it is supposed to be べからず but if the final *hiragana* does not have a *dakuten* in the text, it doesn't have one in our transcription. If the *dakuten* appears on the wrong syllable, please retain where it is. If you want to signal that it is a mistake, please see point 8 below.
5. Diacritics remain unaltered. Please refer to the following chart:

ゝ	Repetition of one sound (= one <i>hiragana</i>).
ゞ	Repetition of one sound (= one <i>hiragana</i>) that has a <i>nigori</i> .
ゞ	Repetition of one sound (= one <i>katakana</i>).
ゞ	Repetition of one sound (= one <i>katakana</i>) that has a <i>nigori</i> .
〈	Repetition of two or more sounds (=two or more <i>kana</i>).
〈ゞ	Repetition of two or more sounds (=two or more <i>kana</i>), the first of which has a <i>nigori</i> .
々	Repetition for one character (not that in the original it appears as 々 and that there are certain <i>honkoku</i> retain this shape).
ゝ	Sign used from time to time to record a poem, a song or speech.

ㇿ for より should be transcribed as ㇿ.

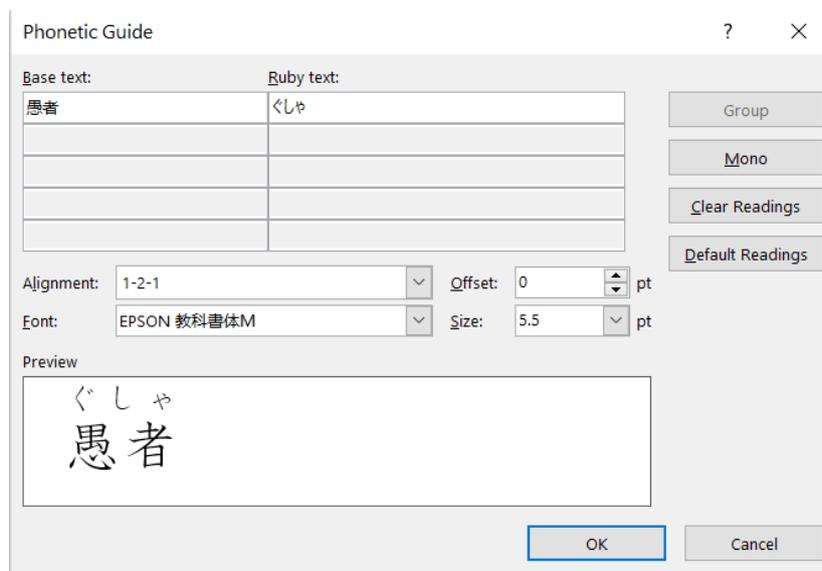
When こと or さま appears in highly abbreviated forms that look like one character, they need to be transcribed as こと and さま.

6. When transcribing *furigana* you must follow the rules explained for converting *hiragana* variants. In addition, remember the following.

6.1. If you are using Word, you would normally click the following button:

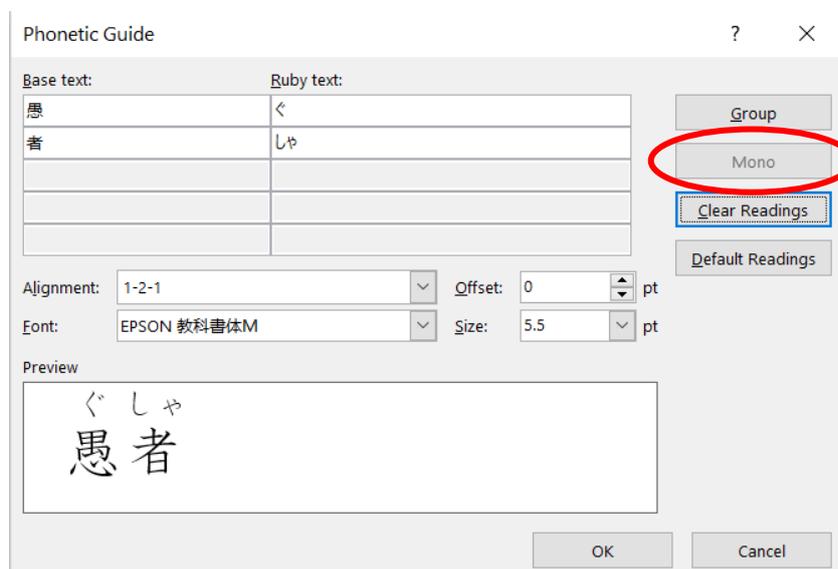


When inputting *furigana* in Word, the danger is that you go with the suggestions given by the software. That can lead to all sorts of mistakes. For example, let's say that in the original you have 患者^{ぐしや}. When you input this compound in Word, the software will give you the following:

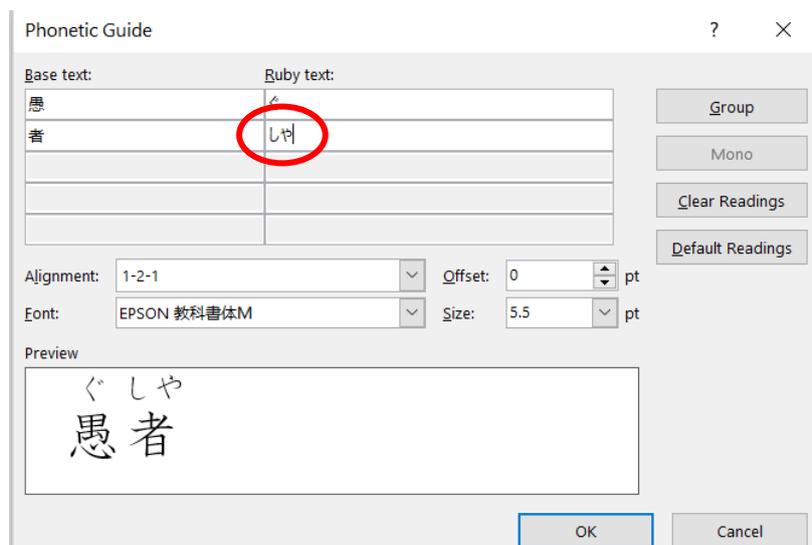


This is incorrect on two accounts: 1) しや should be transcribed as しや (see above point 1.3.); 2) the *furigana* gets distributed unevenly on the two characters: 患者^{ぐしや}. To avoid these problems, follow these two steps.

Step 1: click on "Mono".



Step 2: amend the *hiragana* as required.



You can also decide to click on “clear readings” in Step 2 and insert the *furigana* anew. Please note, publishers are **very** picky about *furigana* placement, so you must get it right first time or you will be spending hours redoing it all.

6.2. For compounds that work as *ateji* 当て字, the *furigana* needs to be spread on both characters. For example:

いなか くすし いしゃ
田舎、医者 (but 医者). This applies also when unconventional readings are given to *kanji* compounds.

6.3. Be careful about the combination of *okurigana* and *furigana*. For example you could have: ^{いひ}云ける or ^{いひ}云ひける. Any of this, albeit incorrect from the point of view of contemporary Japanese, needs to be retained. The *okurigana* can be part of the *furigana* only or repeated both in the *furigana* and in the *okurigana*.

6.4. When there is *furigana* on both sides of characters, they need to be retained as in the original.

7. Punctuation (*kutōten* 句読点) is left as it appears in the original. ○ remain as such, even if they indicate a comma and not a period. When you have ● you must transcribe them as ○. No research suggests that these black ● had a specific meaning. They were simply not cut inside by the woodcutter. On some occasions you may find something akin to this shape: ◌ Unfortunately there is no way to retain this in Word. There are options on *Minna de honkoku*.

8. Errors are not amended and are indicated by (ママ) (positioned in the *furigana*). For example 云 ^(ママ)はける.

If you are not sure that it is a mistake, it is advisable to just transcribe what is in the original without adding (ママ) .

9. You must start a new line whenever the original text starts a new one. Where possible try to align the text in your transcription with the text in the original, for example if there is a space at the start of a line then try to retain it.
10. At the end of each half folio indicate that the text ends for that half folio by putting the number of the folio in Japanese numbers and the indication whether it is the *recto* or the *verso* (e.g. 九ウ). *Recto* must be rendered as オ and *verso* as ウ.
11. When a passage cannot be read because physical damage, please use □. If multiple *kanji* or *hiragana* are missing, please add multiple □. If you are able to infer what is missing, please add this in furigana. For example: ^(読むカ) □ □ .
12. If there are *kuuten* marks, please make sure to insert them.

For ways of inputting your *bonkoku* in the platform *Minna de bonkoku* please refer to the “Input rules” on the system.