Syllabic category of Balinese Surang, Javanese Layar, and Sundanese Panglayar

Aditya Bayu Perdana / ᬨᬟᬶᬢ᭄ᬬᬵᬩᬬᬸᬧᬾᬃᬤᬦ

Ilham Nurwansah

Norbert Lindenberg

2020-06-05

Proposal

We propose:

- To change the Indic Syllabic Category of the characters U+1B03 BALINESE SIGN SURANG (◌ᬃ), U+A982 JAVANESE SIGN LAYAR (◌ꦂ), and U+1B81 SUNDANESE SIGN PANGLAYAR (◌ᮁ) to Consonant_Final.
- To change the annotations for these three characters in the code charts for Balinese, Javanese, and Sundanese from "= repha" to "= final r", and for BALINESE SIGN SURANG and JAVANESE SIGN LAYAR to add "• also used for repha in transliteration of Kawi".
- To update the text in the paragraph "Behavior of ra" as well as Figure 17.2 in [section 17.3 of The Unicode Standard](#) to use the code sequence 1B25, 1B28, 1B03 for the Kawi word *dha-rma*, the same code sequence as for the Balinese word *dha-mar*, and explain that the same code sequence is used because of the identical rendering, despite the different pronunciation.
- To mention in the same section that positioning of *surang* as repha to the left of other above-base marks can be implemented as a font feature.

Error reports

This proposal takes up two [error reports](#) that Richard Wordingham and R.S. Wihananto sent to the Unicode Consortium.

Richard Wordingham on 2015-03-31:

> While the current candidates for category Consonant_Succeeding_Repha may descend from repha, only one of them, U+17CC KHMER SIGN ROBAT is clearly still a repha.

> Reading the script descriptions in TUS makes it abundantly clear that U+1B03 BALINESE SIGN SURANG and U+A982 JAVANESE SIGN LAYAR are actually final consonants. The TUS also states that U+1B81 SUNDANESE SIGN PANGLAYAR is a final consonant, but without going into any details.

R.S. Wihananto on 2015-04-09:

> Indic Syllabic Category of U+1B03, U+1B81, and U+A982
>
> I agree with Mr. Richard Wordingham's feedback. U+1B03 BALINESE SIGN SURANG, U+1B81 SUNDANESE SIGN PANGLAYAR, and U+A982 SIGN LAYAR all were historically repha; but in modern writings, these characters are final -r consonant sign.
>
> Because Balinese, Sundanese, and Javanese characters are encoded with logical order, I think categorized these characters for visually ordered repha is wrong. For consistency with other Indic scripts that only use repha in older texts (such as Telugu), old repha in Balinese, Sundanese, and Javanese should also be encoded with RA + VIRAMA + ZWJ.

Contemporary use of *surang*, *layar*, and *panglayar*

In contemporary use, *surang*, *layar*, and *panglayar* are always used as final -r consonants within their orthographic syllables:

- For Balinese *surang*, this is already documented in the paragraph "Behavior of ra" in section 17.3 of The Unicode Standard.
- For Javanese *layar*, this is already documented in the paragraph "2.5. Syllable-final -r" in the Proposal for encoding the Javanese script in the UCS.
- For Sundanese *panglayar*, Direktori Aksara Sunda untuk Unicode is a key reference for the modern use of the Sundanese script. It describes *panglayar* (◌̃) on page 66: "panglayar berfungsi menambah konsonan /+r/ pada akhir aksara dasar. Contoh: ꦏꦏ = ka menjadi ꦏ̃ꦏ = kar." (*Panglayar* adds the consonant /+r/ at the end of the base character. Example: ꦏꦏ = ka becomes ꦏ̃ꦏ = kar.)

Because of this contemporary use, the three characters should be categorized as Consonant_Final.

## Past use of *surang*, *layar*, and *panglayar* as repha

As the paragraph "Behavior of ra" in section 17.3 of The Unicode Standard and the paragraph "2.5. Syllable-final -r" in the Proposal for encoding the Javanese script in the UCS note, *surang* and *layar* were used as *repha* in transliterations of Kawi, i.e., as the initial consonant of an orthographic syllable rather than the final one. For *panglayar*, such use is not attested.

The glyphs used for *surang* and *layar* did not change when used as repha vs. a final consonant. For *surang* (◌̃), however, the position of *surang* as repha in some older documents is left-most among above-base marks (e.g., ◌̃) rather than the right-most position of *surang* as final consonant in more recent documents (◌̃). The images below show usage in three versions of Nagarakrtagama, a Kawi poem in Balinese script: ① in a lontar manuscript held at Perpustakaan Nasional Indonesia (leaf 4B), ② in Brandes's edition

(NÂGARAKRĔTÂGAMA: Lofdicht van Prapanjtja op koning Radjasanagara, Hajam Wuruk, van Madjapahit, 1904, page 54), both of which use *surang* as repha, and ③ in a [lontar manuscript](#) from about 1995, held at Pusat Documentasi – Dinas Kebudayaan Provinsi Bali (leaves 4A and 6A), which uses *surang* as final consonant on the previous syllable.
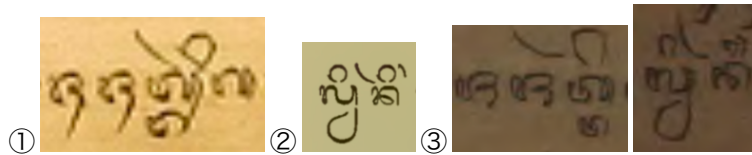

We know of no such divergent positioning for *layar*. Javanese versions of similar Kawi literature kept the *layar* on the right hand side of vowels such as U+A9B6 ᴊᴀᴠᴀɴᴇsᴇ ᴠᴏᴡᴇʟ sɪɢɴ ᴡᴜʟᴜ and final consonants such as U+A981 ᴊᴀᴠᴀɴᴇsᴇ sɪɢɴ ᴄᴇᴄᴀᴋ, as here seen in the word *anjur ning* in [Bharatayuddha: Oudjavaansch Heldendicht (1903, page ꦩꦎꦴ)](#).
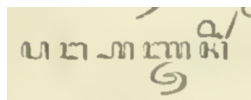
As in most cases there's no visual difference between *surang* and *layar* used as final consonant or used as repha, we think it would be inappropriate to offer different encodings for the two uses. It is not unusual that the same character sequence is pronounced differently in different languages. Converting a sequence of U+1B2D ʙᴀʟɪɴᴇsᴇ ʟᴇᴛᴛᴇʀ ʀᴀ and U+1B44 ʙᴀʟɪɴᴇsᴇ ᴀᴅᴇɢ ᴀᴅᴇɢ into *surang* as repha, as proposed in section 17.3 of The Unicode Standard, seems particularly bad, as such sequences can occur, but must be processed differently, in normal Balinese text, e.g. in the word ᬳᬃᬓᬭᬯᬬᬓ. We know of no font for Balinese, Javanese, or Sundanese that does this conversion, except for one font that uses Javanese code points to encode the Kawi script. Using the sequence U+1B2D ʙᴀʟɪɴᴇsᴇ ʟᴇᴛᴛᴇʀ ʀᴀ U+1B44 ʙᴀʟɪɴᴇsᴇ ᴀᴅᴇɢ ᴀᴅᴇɢ U+200D ᴢᴇʀᴏ ᴡɪᴅᴛʜ ᴊᴏɪɴᴇʀ, as Wihananto proposed, would not be harmful to Balinese language text and would enable the representation of prepositioned *surang*, as shown in Nagarakrtagama. However, as there's no contrasting use of different *surang* positioning in the document, the specific positioning can also be implemented via an OpenType feature in the font to produce ᬓᬭ.

We therefore propose to remove the inappropriate encoding of *surang* as repha from the standard, and to recommend the same encoding sequence for *surang* as repha as for *surang* as final consonant.

Panglayar in Old Sundanese was written in a variety of positions relative to other above-base marks: to the left, above, above-right, or to the right. The images below show ① *gĕgĕr* in the manuscript Carita Parahyangan, held at Perpustakaan Nasional Indonesia, and ② *ciluwĕr* in the manuscript [Bujangga Manik](#) held at the Bodleian Library (leaf 2R). Text analysis however has shown that in all cases it is used as a final consonant, not as repha. Fonts designed for Old Sundanese could support different positions as stylistic variations.

① ②
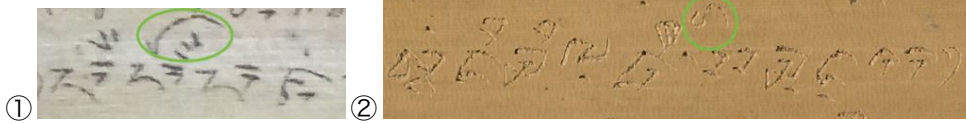
Impact on Universal Shaping Engine

The Universal Shaping Engine relies on Unicode data including the Indic syllabic category to render a large number of scripts, including Balinese, Javanese, and Sundanese. The changes proposed here have no immediate impact on the USE, as it currently treats the Indic syllabic categories Consonant_Succeeding_Repha and Consonant_Final the same way. (In an additional quirk, the USE recategorized *layar* as Tone_Mark until a change made in March 2020.) The treatment has been appropriate for *surang*, *layar*, and *panglayar*. The treatment would not have been correct for U+17CC KHMER SIGN ROBAT, which is a real repha and is conventionally encoded between the base consonant and any vowels; however, this hasn't caused problems yet because Khmer has been implemented by a separate Khmer shaping engine. Separating the Indonesian final consonants from Robat enables the USE to handle Robat correctly, if it were to implement Khmer, or any other consonant-succeeding repha that might be encoded in the future. ꦒꦂ