Annotation Requests for Sundanese Characters

Febri Muhammad Nasrullah <niomi13@pm.me>
2025-05-09

Introduction

The Sundanese script is a traditional writing system from West Java, Indonesia. It was in use from the 14th to the 18th century before being revived as a standardized modern script in the late 20th century. While Unicode accommodates both historical and contemporary forms, several innovations introduced in modern standardization—such as the addition of new letters and modifications to the vowel system—lack historical precedent. As a result, supplementary annotations are necessary to distinguish between authentic historical orthographic features and modern adaptations, ensuring a more accurate representation of the script.

Notations of the vowel -o and other uses of VOWEL SIGN PANOLONG

One key aspect that requires annotation is the function of the SUNDANESE VOWEL SIGN PANOLONG. In the modern standard Sundanese script, PANOLONG is introduced solely to represent the vowel /o/. However, historical evidence from pre-Islamic Sundanese manuscripts indicates that PANOLONG had a broader range of functions. Specifically, PANOLONG was not restricted to denoting /o/ but was also used to indicate long vowels, such as /aː/.

A recent study identifies four different notations used to represent the vowel /o/ in Sundanese manuscripts. Among these, two prominent notations exist: (1) a lone PANOLONG and (2) a combination of PANAELAENG and PANOLONG placed around the base consonant.[1] The choice of notation varies depending on the manuscript material.

Older manuscripts and inscriptions, such as the Batutulis (Figure 5) and Kebantenan (Figure 6) inscriptions, as well as the Siksa Kandaṅ Karəsian (SKK) *gebang* manuscript (Figure 3), predominantly employ the second notation. This practice appears to be a continuation from the Kawi script. In contrast, lontar manuscripts more frequently use the first notation, with some exceptions, such as the Pabyantaraan manuscript, which retains the second notation.[2]

Beyond its function in transcribing the vowel /o/, PANOLONG also serves as a marker of vowel length.[3] This usage is analogous to the KAWI VOWEL SIGN AA and JAVANESE VOWEL SIGN TARUNG,

---

[1] Gunawan (2019)
[2] Gunawan and Griffiths (2021:141)
[3] Gunawan and Griffiths (2021:142)

which indicates vowel elongation ā /aː/. However, unlike its Javanese counterpart, PANOLONG can be combined with other vowel signs to denote long vowels ī /iː/ and ū /uː/. This feature is particularly evident in loanwords, especially in *gebang* manuscripts, though it is also occasionally found in lontar manuscripts. For instance, in SSK, PANOLONG appears in words of Sanskrit origin, such as *bhayū* (from Skt. *vāyu*) (Figure 1), as well as in Old Javanese loanwords, such as *rāma* (Figure 2) and *ajī* (Figure 3). This function is also documented in two other sources: the Batutulis inscription (Figure 5) and the Pabyantaraan text (Figure 4).

<u>New invented letters</u>

One key aspect requiring further analysis is the introduction of new letters in the standardized Sundanese script. These letters—EU, VOWEL SIGN PANEULEUNG, QA, ZA, FA, VA, and XA—were incorporated as part of the early 21st century standardization process to accommodate the Sundanese language's continuous absorption of foreign vocabulary and accuracy of modern phonology. According to the team responsible for this modification, these letters were based on variants of Old Sundanese script that had previously seen limited use.[4]

However, this modification raises several critical concerns. First, there is a lack of transparency regarding the identities and expertise of the involved bureaucrats and scholars, making it unclear whether philologists or linguists specializing in Sundanese script were part of the decision-making process. Furthermore, there is insufficient evidence to support the claim that these letters are derived from pre-existing variants of the Old Sundanese script. Second, ambiguity remains regarding whether these letters were introduced to represent phonemes in the Sundanese language or merely to facilitate transliteration from the Latin script.

A significant challenge arises from Sundanese phonology which tends to adapt foreign phonemes by substituting them with more familiar sounds. The phonemes /f/, /q/, /v/, /x/, and /z/ are absent in Sundanese and are systematically replaced in speech. For example, the phoneme /f/ in the word *fitnah* ('defamation') is commonly substituted by /p/, resulting in *pitnah*. This pattern reflects a broader linguistic adaptation in which Sundanese speakers simplify non-native sounds to align with their phonological system.[5]

A particularly complex case involves the letter X, whose pronunciation varies across languages. In Sundanese and Indonesian, it is typically realized as /s/ at the beginning of words, as in *xenon* [sɛnɔn], and as the consonant cluster /ks/ in loanwords, such as *taxi* [taksi]. However,

---

[4] Baidillah et al. (2008:65–66)
[5] Robins (1953)

the phoneme /x/ itself does not exist in Sundanese phonology and is not consistently represented in speech. For instance, the name of Alexandra Asmasoebrata, a female Indonesian national racer of Sundanese descent, would be written in Aksara Sunda as ᮃᮜᮦᮊ᭄ᮞᮔ᭄ᮓᮢ (Alé**xa**ndra), even though it is pronounced as [alɛ**ksa**ndra].

Given these linguistic constraints, it is more accurate to state in the annotations that these newly introduced letters function as a transliteration system of Sundanese Latin alphabet rather than as representations of distinct Sundanese phonemes.

Deprecated letters

When the Sundanese script was first encoded into Unicode in January 2006, it included the newly invented letters KHA and SYA.[6] However, based on the results of the Sundanese Script Standardization Seminar held at Padjadjaran University on May 28, 2008, these two letters were removed from the standard repertoire.[7] According to Unicode® Character Encoding Stability Policies, once characters have been encoded, they cannot be moved or deleted. Consequently, annotations indicating that these letters are deprecated and should no longer be used must be added.

Annotations

The proposed annotations are as follows:

1B89   SUNDANESE LETTER EU
- only used in 21st century document

1B8B   SUNDANESE LETTER QA
- only used in 21st century document to transliterating 0051 and 0071

1B90   SUNDANESE LETTER ZA
- only used in 21st century document to transliterating 005A and 007A

1B96   SUNDANESE LETTER FA
- only used in 21st century document to transliterating 0046 and 0066

1B97   SUNDANESE LETTER VA
- only used in 21st century document to transliterating 0056 and 0076

1B9F   SUNDANESE LETTER XA
- only used in 21st century document to transliterating 0058 and 0078

1BA9   SUNDANESE VOWEL SIGN PANEULEUNG
    = eu
- only used in 21st century document

1BA4   SUNDANESE VOWEL SIGN PANGHULU
    = i
- combines with AA to form II

---

[6] Everson (2006:1)
[7] Baidillah et al. (2008:90)

1BA5   SUNDANESE VOWEL SIGN PANYUKU
      = u
- combines with AA to form UU

1BA6   SUNDANESE VOWEL SIGN PANAELAENG
      = ae
- combines with AA to form O

1BA7   SUNDANESE VOWEL SIGN PANOLONG
      = o
- also used for AA

1BAE   SUNDANESE LETTER KHA
- deprecated; don't use it.

1BAF   SUNDANESE LETTER SYA
- deprecated; don't use it.

1BBD   SUNDANESE LETTER BHA
      ※ SUNDANESE LETTER ARCHAIC I
- character was originally misidentified, so its name is a misnomer
- for actual use, use 1B84

However, the suggestions from SAH and other parties are welcomed to handle the better annotations.

Attestations

Figure 1. SKK *gebang* manuscript. Red box: ᮘᮚᮅ [ᮘᮚᮥ] *bhayū*.

Figure 2. SKK *gebang* manuscript. Red box: ᮛᮙ [ᮛᮙ] *rāma*.

Figure 3. SKK *gebang* manuscript. Red box: ᮃᮏᮤ [ᮃᮏᮤ] *Ajī*. Blue box: �ñᮠᮧ [ᮑᮠᮧ] *ñaho*.

Figure 4. Pabyantaraan lontar text. Blue box: ꦏꦠꦢꦠꦺꦴꦤ꧀ *katadaton*. Red box: ꦢꦺꦤ꧀ꦠ *dentā*.

Figure 5. Batutulis inscription. Red box: ꦥꦸꦫꦤ *purāna*. Blue box: ꦩꦺꦴꦏ꧀ꦠ *mokta*. [8]

Figure 6. Kebantenan II inscription. Red box: ꦕꦶꦲꦺꦴꦗꦺ *cihoje*. [9]

References

Baidillah, Idin, Undang A. Darsa, Oman Abdurahman, Tedi Permadi, Gugun Gunardi, Agus Suherman, Taufik Ampera, Harja Santana Purba, Dian Tresna Nugraha, and Dadan Sutisna. 2008. *Direktori Aksara Sunda untuk Unicode*. Bandung: Pemerintah Provinsi Jawa Barat.

Everson, Michael. 2006. *Proposal for Encoding the Sundanese Script in the BMP of the UCS*.

Gunawan, Aditia. 2019. *Bhīma Svarga: Teks Jawa Kuna Abad Ke-15 Dan Penurunan Naskahnya*. Jakarta: Perpustakaan Nasional Republik Indonesia.

Gunawan, Aditia, and Arlo Griffiths. 2021. "Old Sundanese Inscriptions: Renewing the Philological Approach*." *Archipel* 101:131–208. doi: 10.4000/archipel.2365.

Noviana, Eka. 2020. "The Sundanese Script: Visual Analysis of Its Development into a Native Austronesian Script." Hochschule fuer Bildende Künste Braunschweig, Braunschweig.

Robins, R. H. 1953. "The Phonology of the Nasalized Verbal Forms in Sundanese." *Bulletin of the School of Oriental and African Studies* 15(1):138–45. doi: 10.1017/S0041977X00087309.

---

[8] Noviana (2020:193)
[9] Noviana (2020:195)