

On Unencoded Characters for Old Sundanese

[illegible]

Febri Muhammad Nasrullah <niomi13@pm.me>

$\frac{7}{\sqrt{7}} = \frac{7}{\sqrt{7}} = \frac{7}{\sqrt{7}} = \frac{7}{\sqrt{7}}$

2025-05-09

ကုမ္ပဏီမှန်မှန်

Introduction

The Sundanese script is an indigenous writing system from western Java, historically used for writing Old Sundanese and related languages, including Old Javanese. Emerging during a period of strong Hindu-Buddhist influence, Old Sundanese incorporated a significant amount of Sanskrit vocabulary, necessitating adaptations in the script.¹ Derived from the Kawi script, the Sundanese script retained several characters originally designed for foreign phonemes, resulting in a relatively extensive character set with a near one-to-one correspondence to the Javanese and Balinese scripts.

Over time, the Sundanese script was gradually replaced, first by the Javanese and Pegon scripts, and later by the Latin alphabet. The last known manuscript written in the script, *Carita Waruga Guru*, dates to the early 18th century and was inscribed on European paper.² In an effort to revive the script, the late 20th and early 21st centuries saw the development of the modern Sundanese script (*Aksara Sunda Baku*).³ However, this revival lacked a full repertoire of the original script, omitting several characters used to represent phonemes borrowed from Old Javanese and Sanskrit.

In 2009, Michael Everson proposed encoding several Old Sundanese characters into Unicode, but his submission lacked references to scholarly research and contained inaccuracies.⁴ Some of proposed characters was later revised in 2021 by Ilham Nurwansah, incorporating more comprehensive academic sources.⁵ Despite these efforts, many Old Sundanese characters remain unencoded in the current Unicode standard, underscoring the need for further research and updates to digital representation.

This document serves as a preliminary proposal to introduce the missing Old Sundanese characters into Unicode. However, as ongoing research continues to uncover new insights, the proposal remains subject to revision. A comprehensive understanding of the Old Sundanese

¹ Gunawan and Griffiths (2021)

² Baidillah et al. (2008: 50)

³ Baidillah et al. (2008: 61–65)
















⁴ See L2/09-251R.

⁵ See [L2/21-221](#).

script requires continued scholarly investigation to ensure its accurate representation and preservation.

Glyphs

Based on available sources, glyphs for Old Sundanese that may be relevant for encoding are as follows:




No.	Glyph	Possible Name	Correspondence		
			Kawi	Javanese	Balinese
1		SUNDANESE LETTER GHA	𑌕	𑌕𑌃	𑌕𑌃
2		SUNDANESE LETTER TTA	𑌔	𑌔𑌃	𑌔𑌃
3		SUNDANESE LETTER DDA	𑌖	𑌖𑌃	𑌖𑌃
4		SUNDANESE LETTER NNA	𑌗	𑌗𑌃	𑌗𑌃
5		SUNDANESE LETTER PHA	𑌙	𑌙𑌃	𑌙𑌃
6		SUNDANESE LETTER SHA	𑌛	𑌛𑌃	𑌛𑌃
7		SUNDANESE LETTER SSA	𑌝	𑌝𑌃	𑌝𑌃
8		SUNDANESE LETTER JNYA	𑌟	𑌟𑌃	𑌟𑌃
9		SUNDANESE LETTER FINAL N	N/A	N/A	N/A
10		SUNDANESE VOWEL SIGN AI	𑌱	𑌱𑌃	𑌱𑌃
11		SUNDANESE VOWEL SIGN VOCALIC R	𑌳	𑌳𑌃	𑌳𑌃
12		SUNDANESE PUNCTUATION DOT	𑌵	𑌵	𑌵
13		SUNDANESE PUNCTUATION DOUBLE DOT	𑌷	𑌷	𑌷
14		SUNDANESE PUNCTUATION SPACE FILLER	3	N/A	3
15		SUNDANESE LETTER RO	𑌹	N/A	N/A

The explanation of each glyph is as follows:

1. Sundanese Letter GHA

This letter represents the phoneme /g^ha/ in Sanskrit.⁶ This letter is attested in the Kebantenan No. 1 copper plate inscription, specifically in the opening invocation *avighnamastu* (Figure 1).⁷

2. Sundanese Letter TTA

This letter theoretically corresponds to the phoneme /ṭa/ in Sanskrit and Old Javanese. However, over time, this letter also came to also represent consonant cluster /tra/ depending on the language of the text. The /ṭa/ value is usually preserved in texts using Old Javanese language. For instance, Sanskrit *bhaṭāra* appears in Old Javanese language texts as *baṭara* , whereas in Old Sundanese language texts it appears as *batara*  due to assimilation of /ṭ/ into /t/.⁸ In Old Sundanese texts, TTA is usually used for /tra/ sound, for example Sanskrit *śāstra* may appear as *sastra* .⁹ This letter frequently occurs in sources such as the Kabuyutan Ciburuy manuscripts and the Wirit Nur Muhammad text (Figure 11). According to the DHARMA transliteration guide, the strict transliteration of this character is *ṭa*, while a more flexible transliteration renders it as *tra*.¹⁰

3. Sundanese Letter DDA

This letter corresponds to the phoneme /ḍa/ in Sanskrit and Old Javanese.¹¹ For example, the Sanskrit word *samiḍa* 'firewood'; was absorbed into Sundanese as *samiḍa*.¹² The letter frequently occurs in sources such as the Batutulis, Kebantenan, and Kawali inscriptions (Figure 2).¹³

4. Sundanese Letter NNA

This letter corresponds to the phoneme /ṇa/ in Sanskrit. It occurs in the Pabyantaraan text (Figure 3).¹⁴

5. Sundanese Letter PHA

This letter corresponds to the phoneme /p^ha/ in Sanskrit. It occurs in a conjunct form in the Pabyantaraan text (Figure 4).¹⁵

6. Sundanese Letter SHA

This letter corresponds to the phoneme /śa/ in Sanskrit. It occurs in the Pabyantaraan text (Figure 6) and Kebantenan inscriptions where it used to represent the honorific prefix *śrī* (Figure 2).¹⁶

⁶ Gunawan and Griffiths (2021:159)

⁷ Gunawan and Griffiths (2021:160)

⁸ Gunawan and Griffiths (2021:190; Nurwansah 2020:11)

⁹ Nurwansah (2021:11; Sopian 2020:135)

¹⁰ Balogh and Griffiths (2020:21; Noviana 2020:161)

¹¹ Gunawan and Griffiths (2021:147–48)

¹² Gunawan and Griffiths (2021:154)

¹³ Gunawan and Griffiths (2021:195)

¹⁴ Mamat Ruhimat, Aditia Gunawan, and Tien Wartini (2014:9)

¹⁵ Mamat Ruhimat et al. (2014:9)

¹⁶ Gunawan and Griffiths (2021:159)

7. Sundanese Letter SSA

This letter corresponds to the phoneme /ṣa/ in Sanskrit. It occurs in the Pabyantaraan text (Figure 5).¹⁷

8. Sundanese Letter JNYA

This letter corresponds to the consonant cluster /d͡ʒna/ in Sanskrit. In the Kabuyutan Ciburuy Manuscripts and Pabyantaraan text, it appears in the name of a narrative figure, *Ajñana* (Figure 10).¹⁸ Additionally, the use of a conjunct stacking form to represent the same phoneme is also attested within the same manuscript (Figure 5).¹⁹

9. Sundanese Letter FINAL N

This letter is unique as it has no equivalent in its parent script, Kawi, nor in its sister scripts, Javanese and Balinese. The LETTER FINAL N functions as a substitute for the LETTER NA when its inherent vowel has been removed. It is found in Kabuyutan Ciburuy manuscripts (Figure 9).²⁰

10. Sundanese Vowel Sign AI

This vowel sign corresponds to the phoneme /ai/ in Sanskrit. It occurs in the Pabyantaraan text (Figure 4).²¹

11. Sundanese Vowel Sign Vocalic R

This vowel sign corresponds to the phoneme /r̥/ in Sanskrit which is often realized as /rə/ in Indonesian context. It occurs in the Pabyantaraan text (Figure 4).²²

12. Sundanese Punctuation DOT

The punctuation serves the same function as the DEVANAGARI DANDA, namely as a marker for commas, pauses, or line breaks (Figure 4). Additionally, it is used to enclose Sundanese numerals, preventing potential confusion, as Sundanese numerals share visual similarities with Sundanese script characters (Figure 7).²³

13. Sundanese Punctuation DOUBLE DOT

Along with PUNCTUATION DOT, this punctuation used to enclose BINDU punctuations (Figure 6). It occurs in Pabyantaraan text.²⁴

14. Sundanese SPACE FILLER

It's derived from the KAWI PUNCTUATION SPACE FILLER and used to justify texts or to fill gaps that are too small to fit another letter at the middle or end of a line (Figure 6).²⁵

¹⁷ Mamat Ruhimat et al. (2014:9)

¹⁸ Mamat Ruhimat et al. (2014:9)

¹⁹ Sopian (2020:138)

²⁰ Sopian (2020:136)

²¹ Mamat Ruhimat et al. (2014:9)

²² Mamat Ruhimat et al. (2014:9)

²³ Nurwansah (2020:18, 133)

²⁴ Mamat Ruhimat et al. (2014:9)

²⁵ Balogh and Griffiths (2020:21; Nurwansah 2020:133; Sopian 2020:132)

15. Sundanese Letter RO

Like KAWI DIGIT TWO found in West Javanese *Gebang* manuscripts, SUNDANESE DIGIT TWO also possesses a unique property, as it can function as a letter representing the syllable *ro*, effectively replacing the combination of LETTER RA + VOWEL SIGN O (Figure 8). When used in this capacity, dependent signs may be attached to the numeral.²⁶ According to the DHARMA transliteration guide, the strict transliteration of this character is 2 (number two), while a more flexible transliteration renders it as *ro*.²⁷

Figures

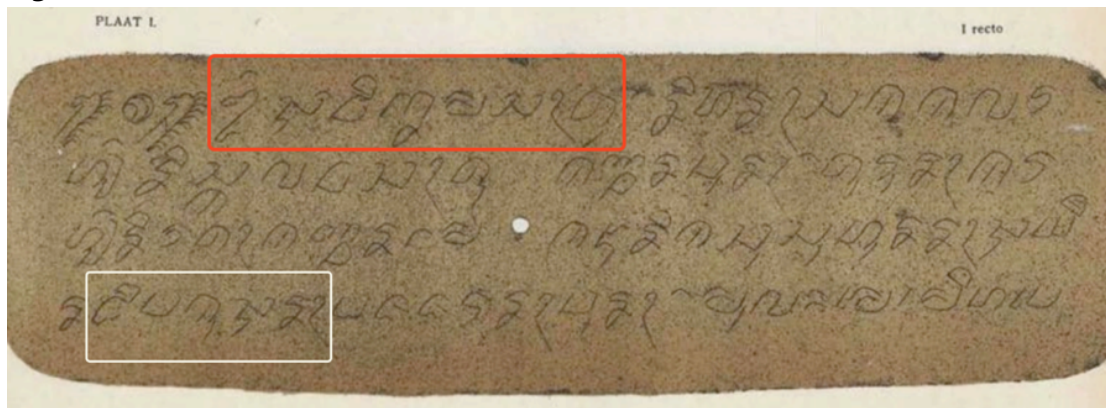


Figure 1. Kebantenan no. 1 inscription. Red box: *Oṃ Avighnamastu* ᨧᩢ᩠ᨦᩈᩢᩣᨶᩥᩬ᩵ᨦᩈᩢᩣᨶᩥᩬ᩵ᨦᩈᩢᩣᨶᩥᩬ᩵. White box: *ḍi pakuAn* ᨧᩢ᩠ᨦᩈᩢᩣᨶᩥᩬ᩵.²⁸

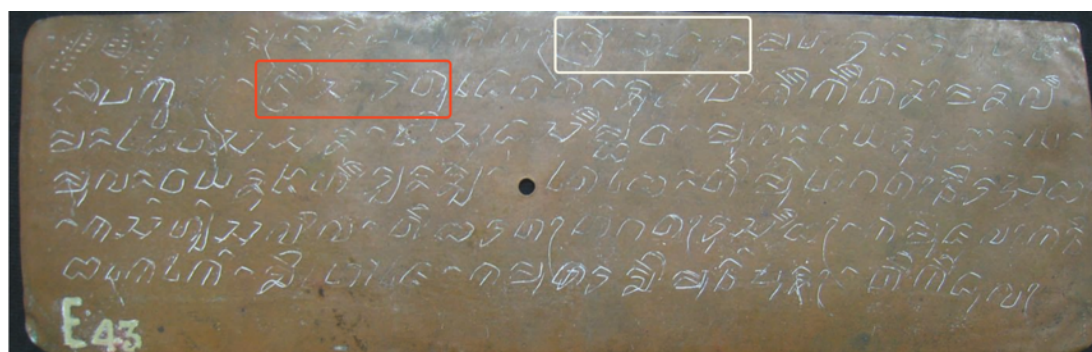


Figure 2. Kebantenan no. 2 inscription. White box: *śrī buduga* (𑀲𑀺𑀢𑀸𑀭𑀺𑀓). Red box: *śrī sam ratu* (𑀲𑀺𑀓𑀢𑀺𑀭𑀺𑀓).²⁹

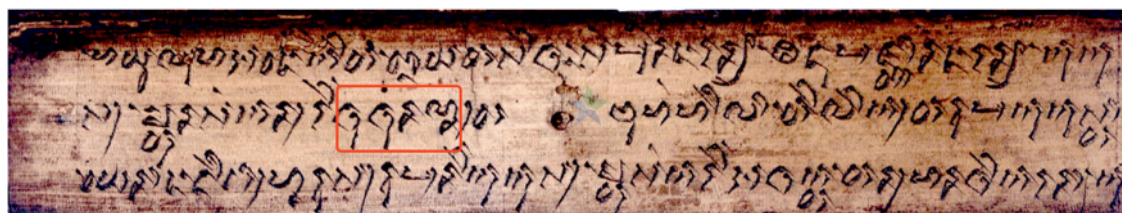



Figure 3. Pabyantaraan text. Red box: *vuvunaṇṇa* .

²⁶ Sopian (2020:128)

²⁷ Balogh and Griffiths (2020:21)

²⁸ Noviana (2020:196)

²⁹ Noviana (2020:195)

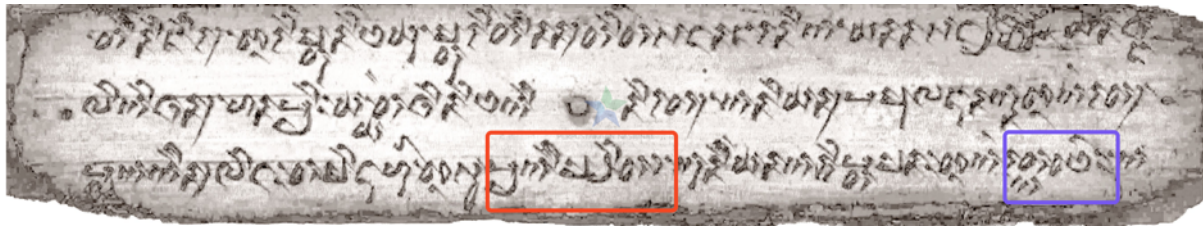


Figure 4. Pabyantaraan text. Red box: *prkəmphitā*, ပုဒ်ဒ်ပုဒ်ဒ်. Blue box: *rat kabaih*, ရာတုကပါး.

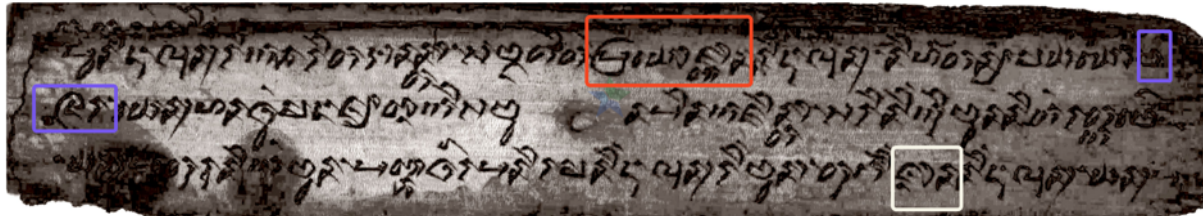


Figure 5. Pabyantaraan text. Red box: *prayojñāna* (ပုဒ်ဒ်ပုဒ်ဒ်). Blue box: *ṣadrara* (ဗဒ္ဒရာ). White box: *jñāna* (ညီနာ).

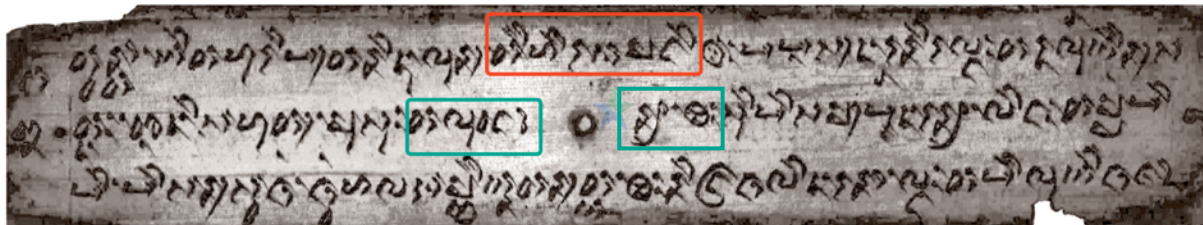


Figure 6. Pabyantaraan text. Red box: *tahāraśamadi* (တရားရံဆမာဒိ). Green box: *talū de* (SPACE FILLER) *nya*, (BINDU) ,, *ta* (ပုဒ်ဒ်ပုဒ်ဒ်).

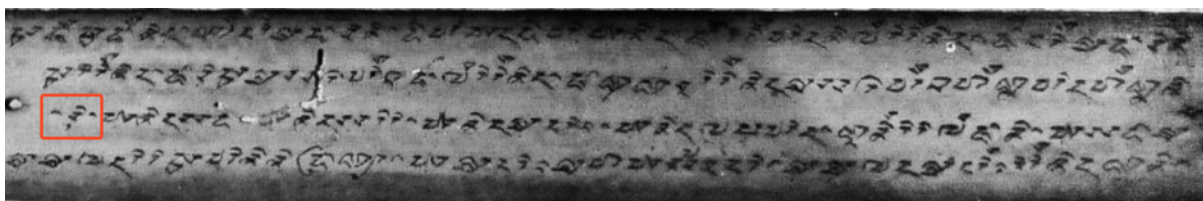


Figure 7. Siksa Kandang Karesian text. Red box: ,4, ဒ်.³⁰

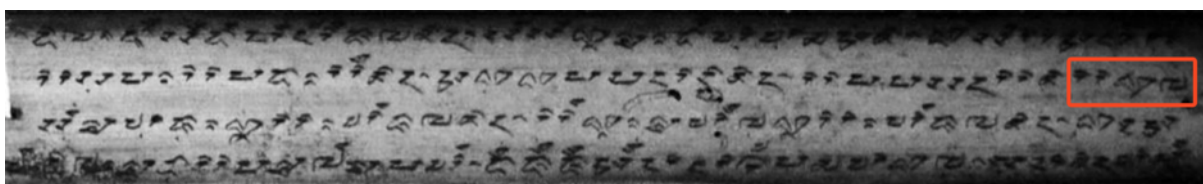


Figure 8. Siksa Kandang Karesian text. Red box: *karola* ကာလာ.³¹

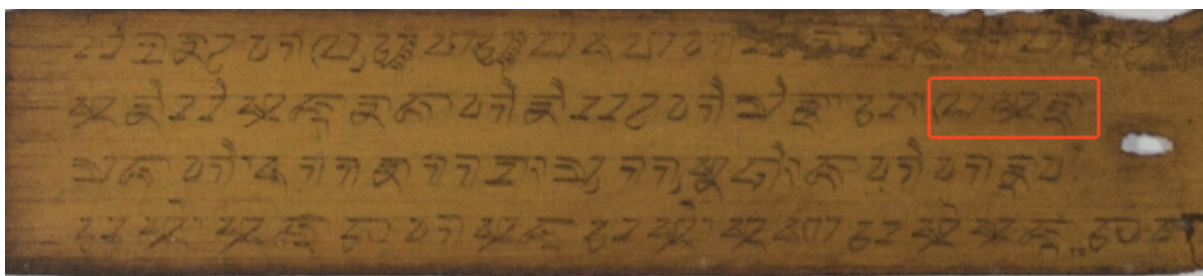


Figure 9. Kabuyutan Ciburuy manuscripts. Red box: *laln* (လှလှ).

³⁰ Nurwansah (2020:50)

³¹ Nurwansah (2020:19)

manuscript, call number L420, which is currently housed at the National Library of the Republic of Indonesia. I express my heartfelt gratitude to him for his typeface. Furthermore, I appreciate his technical advice and attempts to examine and correct the substance of this early report. I would also want to thank Ilham Nurwansah ([ᮘᮞ᮪ᮒᮥᮔ᮪ᮓᮦᮕᮥᮑᮧᮒ](#)) for notifying me about the existence of these unencoded Sundanese characters. [ᮘᮞ᮪ᮒᮥᮔ᮪ᮓᮦᮕᮥᮑᮧᮒ](#).

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.
- Balogh, Dániel, and Arlo Griffiths. 2020. *DHARMA Transliteration Guide*.
- Gunawan, Aditia, and Arlo Griffiths. 2021. "Old Sundanese Inscriptions: Renewing the Philological Approach*." *Archipel* 101:131–208. doi: 10.4000/archipel.2365.
- Mamat Ruhimat, Aditia Gunawan, and Tien Wartini. 2014. *Kawih Pangeuyeukan: Tenun Dalam Puisi Sunda Kuna Dan Teks-Teks Lainnya*. Jakarta: Perpustakaan Nasional Republik Indonesia & Pusat Studi Sunda.
- Noviana, Eka. 2020. "The Sundanese Script: Visual Analysis of Its Development into a Native Austronesian Script." *Hochschule fuer Bildende Künste Braunschweig, Braunschweig*.
- Nurwansah, Ilham. 2020. *Siksa Kandang Karesian: Teks Dan Terjemahan*. Jakarta: Perpustnas Press.
- Nurwansah, Ilham. 2021. *Wrong Identities of Three Historical Sundanese Character*.
- Sopian, Rahmat. 2020. "Old Sundanese Script in Kabuyutan Ciburuy Manuscripts." *言語・地域文化研究* 26:117–48. doi: info:doi/10.15026/94310.