

Representing Bima in Unicode

Anshuman Pandey
pandey@umich.edu

May 3, 2016

1 Introduction

This document offers an approach for representing the ‘Bima’ script in Unicode. This script is used for writing Bima (ISO 639: bhp). This Malayo-Polynesian language, which is known indigenously as ‘Ngahi Mbojo’, is spoken on the Indonesian island of Sumbawa. The script is an extension of Buginese with language-specific letters and signs.

2 Script Details

The following letters are used for writing Bima:

<hr/>	<hr/>	<hr/>			
∕	/k/	∩	/n/	◊	/s/
≈	/g/	≈	/tʃ/	≈	/ʔ/, /a/, 0
∟	/ŋ/	↪	/dʒ/	≈	/h/
≈	/p/	≈	/ɲ/	≈	/f/
↪	/b/	×	/j/	∩	/mp/
∩	/m/	=	/r/	∩	/nt/
∩	/t/	≈	/l/	∩	/ntʃ/
∩	/d/	≈	/w/		
<hr/>		<hr/>		<hr/>	

The script contains the following vowel signs:

◌̇ /i/	/◌̊ /e/
◌̈́ /u/	◌̊\ /o/

and two modifier signs:

◌̂ vowel silencer
◌̣ gemination mark

The structure of Bima is similar to that of Buginese. Each consonant letter possesses the inherent vowel /a/. This vowel is changed by applying dependent vowel signs, which attach to the left, right, above, and below the base consonant. A bare consonant is indicated by use of a sign, which indicates the silencing of the inherent vowel. Some prenasalized consonants are represented using distinctive letters. Doubled consonants may be indicated by a specific mark.

The letter ㄨ is a vowel carrier and represents the independent form of the vowel /a/. Independent forms of vowels are represented by attaching vowel signs to ㄨ , as shown below.

/a/	ㄨ	ㄨ
/i/	ㄨ̇	$\text{ㄨ} + \text{◌̇}$
/i/	ㄨ̈́	$\text{ㄨ} + \text{◌̈́}$
/e/	/ㄨ̊	$\text{ㄨ} + \text{/◌̊}$
/o/	ㄨ\	$\text{ㄨ} + \text{◌̊\}$

The vowel-silencing sign is used as follows:

/ka/	ㄨ	ㄨ
/k/	ㄨ̂	$\text{ㄨ} + \text{◌̂}$
/ga/	ㄨ	ㄨ
/g/	ㄨ̂	$\text{ㄨ} + \text{◌̂}$

The gemination mark is used as follows:

/ka/	ㄨ	ㄨ
/kka/	ㄨ̣	$\text{ㄨ} + \text{◌̣}$
/ga/	ㄨ	ㄨ
/gga/	ㄨ̣	$\text{ㄨ} + \text{◌̣}$

3 Comparison of Bima and Buginese repertoires

Several letters are shared between Bima and Buginese, but there are differences in the forms and values of letters, as well as letters used in Bima for sounds that are not represented in the standard Buginese script. A comparison is shown below, using Buginese as the basis for comparison as it is already encoded in Unicode.

3.1 Current repertoire for Buginese

3.1.1 Consonants

The Buginese block contains 23 letters. Of these, 12 are identical or nearly similar in Bima:

	Buginese	Bima		Buginese	Bima
KA	∕∕	∕∕	JA	↷	↷
BA	↷	↷	NYA	↷	↷
MA	∨	∨	RA	↷	=
TA	∧	∧	VA	↷	↷
DA	∨	∨	SA	◇	◇
NA	∧	∧	A	↷	↷

Nine consonants have different forms:

	Buginese	Bima		Buginese	Bima
GA	↷	↷	NYCA	↷	↷
NGA	↷	↷	YA	↷	×
PA	↷	↷	LA	↷	↷
MPA	↷	↷	HA	∞	↷
CA	↷	↷			

The block contains 2 letters that are not used in Bima. These represent prenasalized consonants of the Bugis language:

	Buginese	Bima
NGKA	↷	—
NRA	↷	—

3.1.2 Vowel signs

Buginese contains 5 vowel signs, of which 2 are identical in Bima:

	Buginese	Bima
VOWEL SIGN I	◌̇	◌̇
VOWEL SIGN U	◌̈	◌̈

while 2 have different forms:

	Buginese	Bima
VOWEL SIGN E	◌̈̈̈	◌̈̈̈
VOWEL SIGN O	◌̈̈̈̈	◌̈̈̈̈

The following sign is not used in Bima:

	Buginese	Bima
VOWEL SIGN AE	◌̈̈̈̈̈	—

3.1.3 Punctuation

Both Buginese marks of punctuation are used in Bima:

	Bima	Buginese
PALLAWA	⋯	⋯
END OF SECTION	⋄	⋄

The Υ U+A9CF JAVANESE PANGRANGKEP is used for marking repetition of syllables.

3.2 Missing characters

The ‘Buginese’ block does not have characters for the following Bima letters:

ƒ	/f/
nta	/nta/

The block also lacks characters corresponding to the following Bima signs:

◌̂	vowel silencer
◌̃	gemination mark

4 Approach for encoding

The ‘Buginese’ block contains 30 characters: 23 consonant letters, 5 vowel signs, and 2 punctuation signs. Representing Bima in Unicode requires 31 characters: 25 letters, 4 combining vowel signs, 1 VIRAMA, and 1 gemination mark. Of these letters, 10 are distinctive, while 15 can be represented using existing Buginese characters. Of the vowel signs, 2 are identical, 2 may be considered to be alternate forms, and 1 does not occur in Bima. The following actions are required:

1. As the ‘Buginese’ block in the BMP has only two spaces remaining, and as there is no free space in the BMP, a new block should be created in the SMP with the name ‘Buginese Extensions’. The block should encompass at least 5 columns to accommodate characters from other Buginese orthographies.
2. Encode the following 10 letters in ‘Buginese Extensions’. As some letters may also be used in Buginese orthographies for other languages, character names should be generic and not specific to Bima:

◌̃	<i>ga</i>
◌̂	<i>nga</i>
◌̃	<i>pa</i>
◌̂	<i>mpa</i>
◌̂	<i>nta</i>
◌̃	<i>ca</i>
◌̂	<i>nyca</i>
×	<i>ya</i>
◌̃	<i>la</i>
◌̃	<i>fa</i>

3. Add the following annotation to the names list for the ‘Buginese’ block:

1A10 ◌̂ ◌̂ BUGINESE LETTER YA
 • used in Bima for /h/

4. Encode the vowel-silencing sign in the existing ‘Buginese’ block (see L2/16-075):

◌̂ BUGINESE SIGN VIRAMA-2

5. Encode the gemination mark in the new ‘Buginese Extensions’ block:

◌̣ *gemination mark*

6. Determine whether the following Bima vowel signs are distinctive characters or glyphic variants:

/◌◊ *vowel sign e*

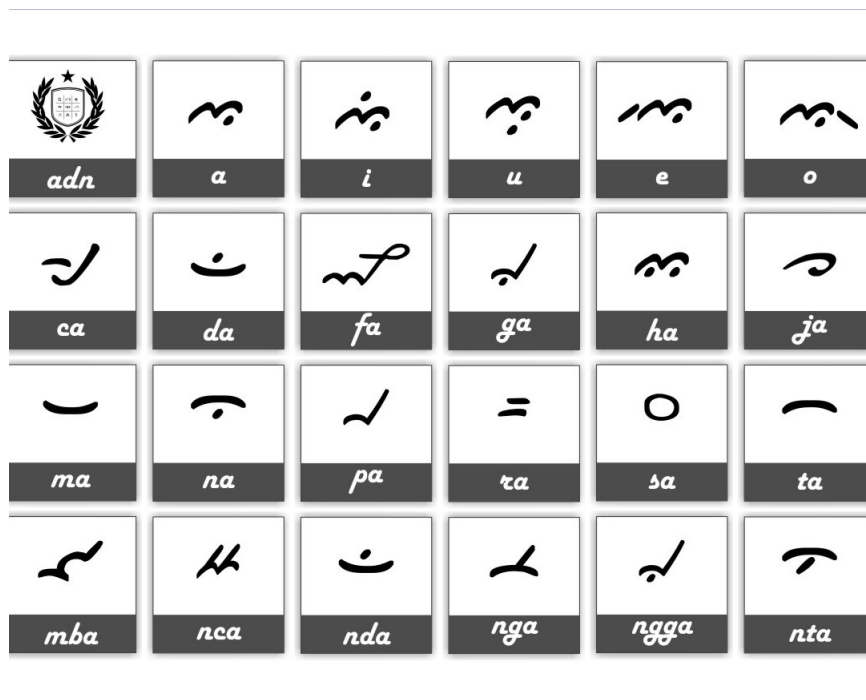
◌◊ *vowel sign o*

A formal proposal for encoding letters of Bima and other Buginese-based scripts is forthcoming.

5 References

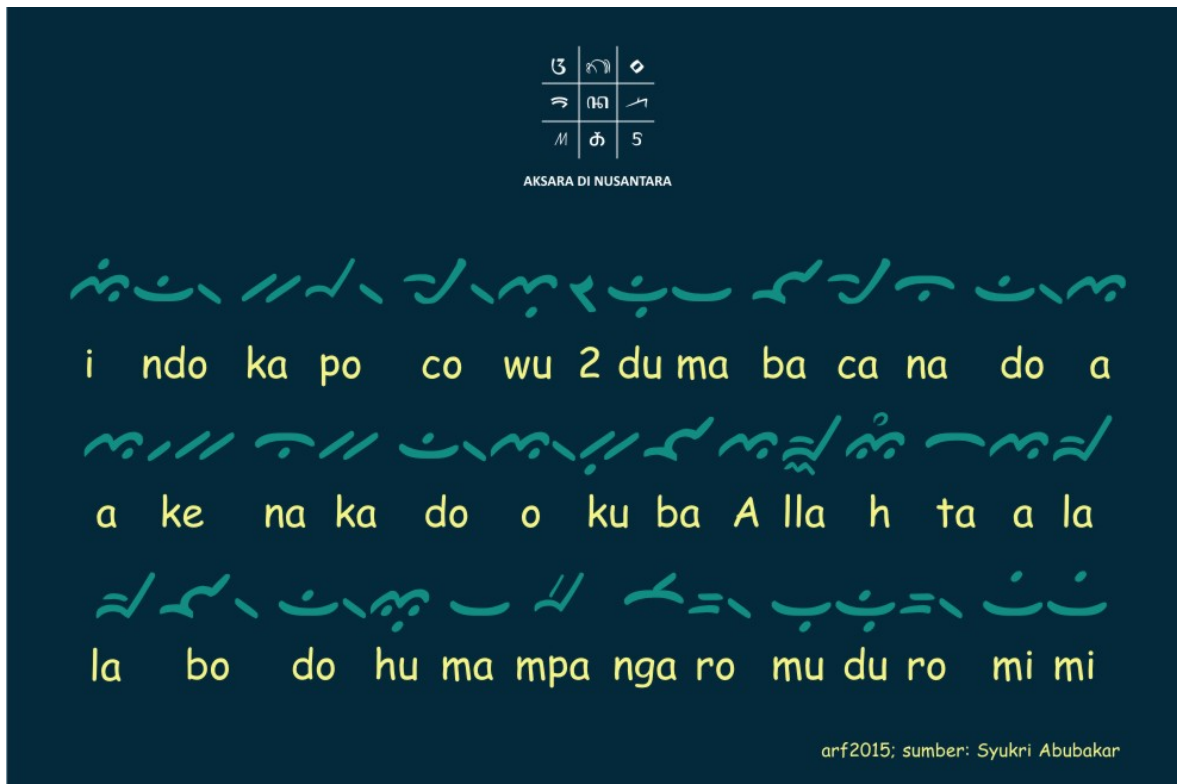
Pandey, Anshuman. 2016. “Proposal to encode VIRAMA signs for Buginese”. L2/16-075.

<http://www.unicode.org/L2/L2016/16075-buginese-virama-signs>



Source: <http://aksaradinusantara.com/2015/09/bentuk-aksara-bima/>

Figure 1: Glyphs of a Bima font.



Source: <http://aksaradinusantara.com/2015/08/contoh-penulisan-mbojo/>

Figure 2: Bima text typeset using a digitized typeface.