Indonesian and Philippine Scripts and extensions

not yet encoded or proposed for encoding in Unicode as of version 6.0

A report for the Script Encoding Initiative

Christopher Miller

2011-03-11

Table of Contents

1.   Introduction[1]

A large number of lesser-known scripts of Indonesia and the Philippines are not as yet represented in Unicode. Many of these scripts are attested in older sources, but have not yet been properly documented in the available scholarly literature. This report attempts to synthesize the available historical literature and information based on original texts (where available) to provide a clearer picture of the potential encoding needs for indigenous scripts of Indonesia and the Philippines. Although there is a wealth of information for many scripts in the region, the existence of certain alleged historic scripts cannot (as yet) be verified on the basis of clearly authentic original documents. These will be presented on the basis of the available documentation with a critical evaluation of their likely authenticity.

Only a minority of the scripts described here are fully fledged independent scripts that stand apart from their neighbors. A large number of the regional scripts described in this report are varieties of scripts that have already been encoded: these are most appropriately dealt with as extensions to the already encoded scripts consisting either of previously unrecognised characters and variants attested in the literature or of additional characters developed in adaptations to languages with different sound systems.

For some script varieties described here, issues arise with respect to the arrangement of characters. In several South Sumatran scripts in particular, as in Batak script, vowel signs change position with respect to each other or to their host letter in certain contexts. This implies an eventual need for special attention to reordering strategies either in software or at the user interface.

I have tried as much as possible to include the following information in the entry for each script:

- Location and historical period of the script
- Name or names used in the literature
- Typological classification of the script
- Inventory according to available information, arranged by function, and/or estimated number of characters belonging to the script
- Individuals, where known, who may serve as expert contacts for the script in the event of a proposal being prepared

When referring to character values, as distinct from the names of encoded Unicode characters, I will use single guillemet (angled quote) ‹ › notation. In addition, breaking with the tradition in Indic script studies of referring to consonant letters with the alleged "inherent vowel" (generally /a/), I will only represent the invariant content of consonant characters, namely the consonant sound they represent with or without additional vowel signs or the virama that specifies the lack of a following vowel.

---

[1] I would like to thank Debbie Anderson for her valuable help with several difficult to obtain sources used in this report.

2.   The Philippines

The Indic scripts of the Philippines are among the earliest minor scripts to be encoded in Unicode. These include the Tagalog script that was widely used during the 1500s and 1600s before it was largely supplanted by Latin script, the Tagbanwa script of Palawan Island, and the two Mangyan script varieties of Mindoro Island: the northern *Buhid* and the southern *Hanunóo*. Tagalog script is also widely known as *Alibata*, a 1914 coinage derived from the first three letters of the Arabic script (Verzosa 1939), *Baybayin*, a historically earlier and now increasingly widely used name derived from a root with the meaning "arranged in order or sequence" (Pardo de Tavera 1884), and finally, at least in the Bikol region of southeastern Luzon, the script was known as *Basahán*, derived from the verb *basa* "read" (Mintz 2004).

With the notable exception of the Eskayan syllabary, the Philippine scripts are all Indic alphasyllabaries. The historically recorded scripts are all left to right in direction (apart from the mirror writing of left handed individuals attested for Hanunóo and Tagalog). There has been a long-running misconception that the scripts were historically written vertically from bottom to top, however this has been shown to be merely the mechanical consequence of incising letters into bamboo segments with a knife while holding the bamboo angled away from the midline of the body and making cutting strokes away from the body for safety's sake.

Over the past decade, the Tagalog script has been the subject of renewed interest as a part of Filipino cultural heritage, and this has led to two kinds of proposals for modifications of the script.

The first includes a number of individual proposals for expanding the script with newly created or derived letters capable of representing the sounds of English; these proposals have met with little acceptance given that the script is most widely used to write Tagalog and other Philippine languages either in online forums, emblematic inscriptions or, in another popular application, tattoos. These modern uses of the script have however led to a small number of common extensions to the script not provided for in the original encoding, in order to represent three sounds in modern Tagalog that have become phonemically contrastive in the language largely as a result of borrowing from Spanish and English.

The second modification is a fully fledged script for the Kapampangan language that is distinct in its basic structure from Tagalog and most linear Indic scripts. The character glyphs used are variants of Tagalog script characters recorded in sample alphabets by European observers for the Kapampangan region, but the structure and arrangement of the letters appear to be inspired in part by character arrangement in Han'gul syllable blocks.

These two sets of potential extensions to the Philippine scripts are described below.
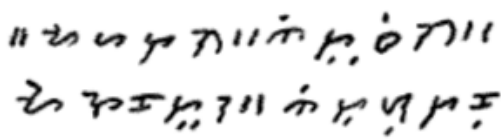
2.1.   Encoded script blocks

The currently encoded scripts occupy blocks 1700-171F (Tagalog), 1720-173F (Hanunóo), 1740-175F (Buhid), and 1760-177F (Tagbanwa).

2.1.1. Tagalog

The Tagalog script as currently encoded is limited to the set of distinct characters used to write Tagalog and other Philippine languages at the time of the first Spanish contact with the script. Over the past century, a number of proposals have been made to extend the Philippine script, both the Tagalog variety and the Tagbanwa variety, to adapt it to writing sounds not originally present in the languages for which it was used.[2] Most of the proposed additional characters are intended to represent Spanish and English sounds extrinsic to Philippine languages and have not found any acceptance among modern day users of the scripts. Two kinds of extensions have gained fairly widespread acceptance, especially among users of Tagalog script on Internet sites, and on this basis appear to be necessary at this point: these include the addition of an extra letter and two additional vowel signs. Apart from the addition of these three characters, previously overlooked data on spelling conventions in the script reveal the need for a way to encode the doubling of vowel signs on a vowel letter.

Vowel sign doubling

Tagalog script historically shares with Buginese a spelling convention that can be called "vowel sign doubling". This consists in marking a consonant letter with two vowel signs, either the same or two different ones, to represent two succeeding syllables beginning with the same consonant. Since /a/ is not marked but is supplied by default in the absence of any vowel marking, vowel sign doubling only involves the ‹-i› and ‹-u› signs. Although documents in the old Philippine Tagalog script are too few in number to furnish examples as numerous as those found in the Bugis and Makassarese Bird scripts in Sulawesi, there are nevertheless a handful of examples of this convention illustrated and noted in Santamaría (1938) and Villamor (1922), including two neighboring examples in the same document where the consonant letter ‹t› bears two ‹-u› dots underneath:

|| a ha la ga || ni *tuu* bi ga ||

a   pa ka *tuu* u || ni tu su la ku

Ang halaga ni[*tong tu*]bigan (…)
at pagka[*toto*]o nitong sulat ko (…)
'The price of this arable land (…)
and that this writing of mine is true (…)'

Figure 1. Two occurrences of vowel sign doubling in Tagalog script.
(Excerpted from a 1635 land deed, University of Santo Tomás Archives, reproduced in Villamor 1922)

---

[2] Among them Mendoza de Leon (1992) and Romuáldez (1914).

Although this does not involve the addition of any extra characters to the inventory of code points for Tagalog script, it does make it necessary to provide for the need to superimpose up to two identical or different vowel signs on a single consonant letter in the same way as in Buginese script.

A new ‹r› letter

When the Tagalog language was originally encountered by Spanish explorers, they heard the phonetic allophones of the Tagalog /d/ phoneme as two separate sounds, [d] and [r], which they transcribed separately in Latin script. However, since [r] was merely a non-contrastive variant of /d/ in intervocalic position, Tagalog script used the same letter TAGALOG LETTER DA for both phonetic variants. As a result of several centuries of language contact and borrowing from Spanish and English, [r] in Tagalog and other Philippine languages has expanded from being a positional variant of /d/ to all peripheral syllabic positions, thus becoming a distinct phoneme. As a result of this, there have been several proposals over the past century to add a new ‹r› letter to the script. The shapes of the proposed characters are widely divergent: some are borrowed from a pre-existing source and others are derivatives of the existing ‹d› letter shape. The four proposals known to me are illustrated below. The choice of glyph can be left to font-level variation; however the most widely used current variant is the fourth.

Glyph	*Origin*	*Source*
ẞ	Anonymous example alphabet for Zambales, 1601	Marcilla y Martín (1896)
⌒	From Buginese ⌃	Verzosa (1939)
⊏	Derived from ⊃	Mendoza de Leon (1992)
⊂	Derived from ⊃	De los Santos (2010)

Table 1. Four proposed glyphs for a new ‹r› letter.

Two new vowel signs

Similarly, the language originally only had three contrastive vowels: /i/, /u/ and /a/: the two high vowels had the mid-vowel positional variants [e] and [o], usually found in the final one or two syllables of a word. Although the Spanish transcribed these positional variants with ‹e› and ‹o› as they heard them, the Tagalog script represented them with the same two vowel marks (TAGALOG VOWEL SIGN I and TAGALOG VOWEL SIGN U) as their high counterparts [i] and [u]. As a result of widespread borrowing from Spanish and English, the two mid vowels have now become contrastive in Tagalog and other Philippine languages.

Value	Glyph	Status
-i	⊙	Traditional glyph
-e	⊙	Proposed innovation
-u	⊙	Traditional glyph
-o	⊙	Proposed innovation

Table 2. Two new vowel sign proposals.

Due to the addition of these three phonemes to the Tagalog inventory, many modern users of the script feel a need to distinguish them with distinct glyphs, and numerous modern Tagalog fonts have been created in the past decade with contrasting vowel marks for ‹e› versus ‹i› and ‹o› versus ‹u›, as well as a contrastive glyph to represent ‹r› as opposed to ‹d›. Although the form of the innovative glyphs varies, most particularly with respect to proposals for a new ‹r› letter, it is clear that there is general agreement among users that the script needs three new characters to accommodate these three new phonological contrasts in the language. It would therefore be useful to envision the addition of three new characters to the Tagalog block.

2.1.2. The modern *Súlat Kapampángan* script

The Kapampangan language, closely related to Tagalog, is spoken in Pampanga province north of the Greater Manila area. Historical documentation of the script variants used in this region is sparse. Several example abecedaries by early observers are extant, and agree to such an extent on particulars of the shapes of the glyphs that they may all go back to the same original (*cf.* examples reproduced in Marcilla y Martín 1895). In these samples, several letters (‹g›, ‹t›, ‹s›) differ notably in shape from their Tagalog equivalents.

Only three original documents are known to me to contain authentic specimens of Kapampangan handwriting variants. One is a case for beatification described in Santiago (2002) in which three witness signatures are recorded as being in Kapampangan handwriting: this document is currently unavailable to me. The other two are documents provided by Antoon Postma of the Mangyan Heritage Center in Mindoro, which contain signatures for five individuals, all five in the same hand in one document and a signature for one of the five in a different hand in the second document (Anonymous 1615). The character shapes agree approximately with the example abecedaries of observers, with the notable exception of ‹l›, which is a simple vertical line in both hands, unlike the abecedary version, which is much more similar to the Tagalog shape.

The most important difference from Tagalog script is an orthographic convention in the Anonymous (1615) document that represents a syllable final consonant at the end of each signature with the base consonant letter that in this case, is not pronounced with the default /a/ vowel. Elsewhere, syllable-final consonants are not represented as such but are resyllabified as the initial letter of a following syllable or by adding an ‹-i› vowel sign, presumably to reflect epenthesis in speech.

The admittedly sketchy evidence from available sources does not seem to justify encoding a historical Kapampangan script distinct from Tagalog script, since there is little reason to believe that the character shapes attested in available documents were anything other than variants cooccurring with more widely distributed shapes, and the spelling rule does not require any special encoding.

However, the past decade or so has seen spreading interest among segments of the Kapampangan community in a script known either as *Kulitan* (writing) or *Súlat Kapampángan* (Kapampangan script). The script has found popularity in marginal uses similar to the Tagalog script, including calligraphy, tattoos, and emblematic uses such as commercial and official logos. *Súlat Kapampángan* differs from the historically attested variety, and from Tagalog script, in several important ways described below, which may prove to be a complicating factor when it comes to encoding the script for computer entry. The character inventory and orthographic conventions are described in a series of papers by Michael Pangilinan *alias* Siuála ding Méangûbie of the Kapampangan Studies Center at Holy Angel University in Angeles City, Pampanga.

The standard letter glyphs appear to be derived fairly directly from those attested in the abecedaries reproduced in Marcilla y Martín (1895). The inventory has two letters fewer than Tagalog script: no distinct letters are given for the glides /y/ and /w/, although a shape very similar to Tagalog WA is used to represent the vowel /u/.

Figure 2. First historical example alphabet of letter shapes used in Kapampangan region, as reproduced in Marcilla y Martín (1892).

Figure 3. Example alphabet for Kapampangan region from de Mas (1843), reproduced in Marcilla y Martín (1892).

Unlike Tagalog and other Philippine script varieties, letters are not arranged in consistent linear order but rather in syllabic blocks analogous to those of Han'gul. The organization of letters within a block is itself linear but orthogonal to the flow of text within the line. The arrangement of letters inside a block is illustrated schematically in Figure 4.

C[v] = initial consonant + applicable vowel sign

V = initial vowel

(V2) = second vowel

(C) = coda consonant

Figure 4. Arrangement of letters in syllable blocks: (a) in horizontal text (b) in vertical text.

A minimal syllable block consists of either a bare vowel letter or a single consonant letter (with an ‹-i› or ‹-u› vowel sign or without, in which case it is read with /a/). A long vowel is represented by a base letter (with or without the relevant vowel mark as the case may be) followed by the independent letter for the same vowel. The sounds /e/ and /o/ are represented respectively by the independent letter ‹a› or a bare consonant letter as the case may be, followed by the independent ‹i› or ‹u› letter, respectively. A syllable-final consonant is written in final position and by virtue of its coda position is always read without a vowel.

Although left to right lines ordered downward on the page are acknowledged as a possibility, the recommended and most widely used writing direction is similar to that traditional in East Asian scripts: vertical lines in a right to left sequence.

If modern Kapampangan script differed from Tagalog script only in the shapes of some of its characters, it would be preferable to deal with it as a simple font variant under the same encoding, as indeed could well have been done for Buhid, Hanunóo and Tagbanwa[3]. However, the convention of arranging characters into complex syllable blocks perpendicular to the flow of the text, whether horizontal or vertical, requires a more complex encoding that would require its treatment as a fully distinct script in the event a proposal for its addition to Unicode is submitted.

Contact:

Michael Pangilinan, Center for Kapampangan Studies
Holy Angel University, Angeles City
The Philippines

siuala@yahoo.co.uk

---

[3] Although they appeared to be four distinct scripts with different if related letter shapes when the original proposals for the Philippine scripts were submitted and approved, examination of sources containing authentic 17th century Philippine handwriting makes it clear that the modern Tagbanwa, Buhid and Hanunóo scripts are modern stylistic variants that fall well within the range of 17th century handwriting variation, with very few substantive changes since that period.

2.1.3. The characters of the Calatagan pot inscription

The Calatagan pot is an inscribed artifact sold to the National Museum of the Philippines in 1961. Although the circumstances surrounding the pot's origin are unclear, it was claimed to originate from a site being excavated at the time by archeologists associated with the National Museum. Because of the the inscription around its neck, the pot has puzzled researchers and enthusiasts for the past five decades. The inscription is characterized by several puzzling and unusual features. It contains a set of characters some of which are recognizable as (mostly upside down) equivalents of the curvilinear Tagalog script variants found in the 1593 *Doctrina Christiana, en lengua española y tagala* (the oldest known document in Tagalog script). Combined with these are a second set of angular, rectilinear characters that bear strong resemblances to known Tagalog letter shapes, and a third set of characters not clearly related to any characters found in any attestations of Philippine script known to previous researchers. Finally, the inscription bears a set of anomalous features that give the appearance of left handed mirror writing.
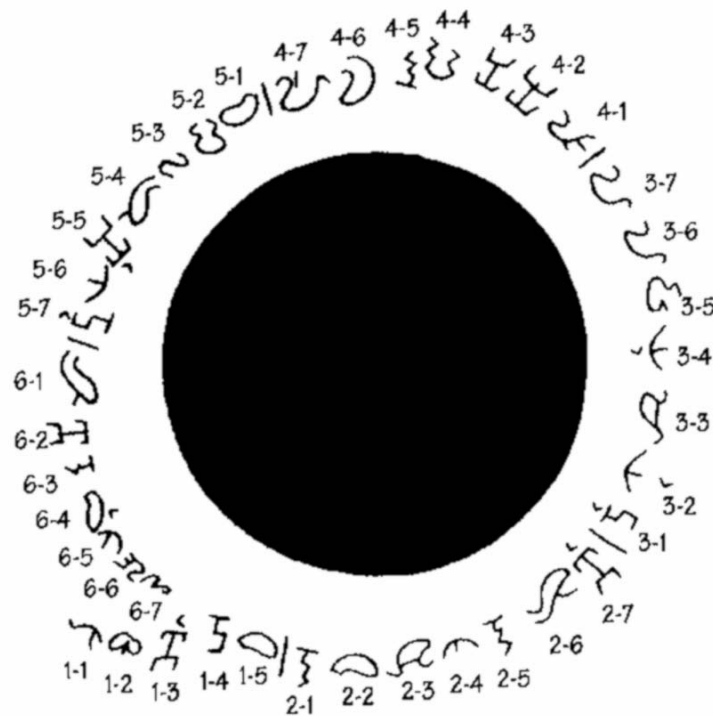
Figure 5. Drawing of characters on the Calatagan Pot (by permission, Santos 1996).

Because of the aberrant features of the inscription and the suspicious circumstances under which it was sold to the Museum, the possibility has long been considered that it may be a forgery. However, since the associated site yielded artifacts plausibly dated to the 14th or 15th centuries (Oropilla 2008), the artifact was declared a National Cultural Treasure of the Philippines on the basis that it appeared to be one of the earliest examples of pre-Hispanic writing yet known in the Philippines.

The inscription has been the subject of several attempts to determine the values of its constituent characters and additional attempts to reconstruct a reading. All such attempts, however, are subject to the same caveats that apply to attempts to decipher other unique artifacts of the sort. My own research (Miller, in prep.) has uncovered strong evidence that the inscription is modelled on a 1599 document from Magaldan in Pangasinan province, which is a record of the 1599 royal sovereignty referendum accessible only in the *Archivo de Índias* in Seville and only recently reproduced in publicly available form in Villarroel (2000) and Oropilla (2008). This document contains three Tagalog script signatures, two in left-handed mirror script whose letter shapes and orientation are as a result distorted enough to be unrecognizable without the realization that they are in mirror writing, and the third a normal left to right signature with some nonetheless unusual and rare early letter shape variants. It appears that almost all the aberrant and unexplainable features of the Calatagan Pot Inscription are modelled on the unusual features of the three signatures in the 1599 Magaldan document, supplemented with rotated glyphs from the 1593 *Doctrina Christiana*. Due to the relative public inaccessibility of these two documents in 1961, it seems plausible that the inscription was forged by someone at that time who was familiar with the two old documents. I must caution that the evidence I describe here has not yet been presented to an audience of specialists for evaluation.

Although the likelihood seems very high that the inscription is not an authentic representative of old Philippine writing, the possibility of an encoded message remains. Given the sustained though sporadic interest in finding a "decipherment", it seems worthwhile to consider setting aside a block of characters for the use of researchers interested in working on this inscription, as has been done for example in the case of the Phaistos disc.

The distinct characters are listed in Table 3, oriented as they appear when looking toward the neck of the pot, numbered by position according to the system used by Santos (1996), shown in Figure 5. Due to the shallowness of the inscription and the difficulty involved in reading some of the characters, there is slight disagreement among analysts as to the precise symbols. For this reason, some characters treated by individuals as the same, but not by others, are treated as distinct here. The characters include approximately eighteen apparent letters, one apparent textual division mark, and two vowel marks, one above and one below.

1-1	1-2	1-3	1-4	1-5	2-1	2-3	2-4	2-6	3-5	3-6	4-1	4-4	4-6	4-7	5-2	5-3	6-7	2-0	1-3	3-2
		2-7	3-1	2-2	2-5	3-3	3-2	5-4		3-7	6-6							3-0	2-7	5-7
		4-2	5-7	5-1	4-5		3-4	6-1										4-0	3-1	
		4-3		6-4	6-3		5-6											5-0	3-4	
		5-5																6-0	5-5	
		6-2																	6-4	

Table 3. Inventory of character glyphs for the Calatagan pot inscription.

2.1.4. The (non-Indic) Eskayan syllabary

The Eskayan script is used by a cultural minority on Bohol Island in the Visayas to write a language peculiar to the community. Although they speak Cebuano in daily life, the Eskayan language is taught in school and is apparently used in specific contexts as an emblem of community identity. The language itself appears to be related to Cebuano at the morphosyntactic and phonetic levels of structure, but with extensive relexification based on a variety of external sources, yielding a vocabulary whose phonotactics differ markedly from the usual (C)V(C) syllable structure of Cebuano and other Philippine languages (Kelly 2006). It appears likely that the language and the script were both consciously constructed, perhaps in the late nineteenth to early twentieth centuries (Kelly 2006).

Figure 6. Sample of characters from the Eskaya syllabary.
(Image from Wikimedia Commons: http://commons.wikimedia.org/wiki/File:Eskayasripttable.GIF).

The language is written in a complex syllabary that consists of perhaps a thousand or more distinct characters (Santos 1996), some representing syllable-internal segment sequences, and others particular affixes or sound sequences specific to the Cebuano language. A small subset of the characters in the script is illustrated above in Figure 6, in which it can be seen that connected Latin script handwriting appears to be an important source for the shapes of many characters.

Contact:

Piers Kelly
The Australian National University, Canberra
piers.kelly@anu.edu.au

2.1.5. Summary

In the event that proposals are considered for the scripts of the Philippines, the code spaces required would number three for extensions to Tagalog script, 18 for the Calatagan Pot Inscription glyphs, 17 (letters plus two vowel signs) for Kapampangan script, and in excess of 1000 for the Eskayan syllabary. In addition, special coding issues are raised by the use of vowel sign doubling in Tagalog script and by the distinct vertical directionality and Han'gul-like syllable blocks of the modern Kapampangan script.

3.   Sumatra

Scripts currently encoded for Sumatra include Batak script variants (1BC0–1BFF) and the South Sumatran Rejang script (A930–A95F). Both are Indic alphasyllabaries. The Batak encoding appears to be complete, as the encoding proposal was prepared with the advice and knowledge of Uli Kozok, the foremost contemporary specialist in these scripts. However, many South Sumatran varieties, as well as the Kawi-derived medieval "Malayu" script,[4] remain unencoded. Given the historical importance and variety of literature written in the scripts of the South Sumatran group, extensions or new character blocks would merit especially high consideration. Similarly, the Malayu script of medieval Sumatra is represented not only in a number of historically important inscriptions, but also in the Tanjung Tanah book of laws, the earliest known handwritten manuscript in a variety of the Malay language.

The South Sumatran group and the Malayu script group are dealt with in turn below.

3.1.   The South Sumatran script group

This group is currently represented in Unicode by one of its several varieties, the Rejang script that was previously in wide use in the Bengkulu area of southwestern Sumatra. However, this group contains several other varieties with distinct identities, distinct character-sound mappings, and in the case of two varieties, special encoding requirements analogous to those of Batak but not found in Rejang. Associated with Rejang specifically is a numeral notation system used into the 20th century. The issue of whether to encode particular varieties under an expanded Rejang-based block or as distinct scripts must take into account the fact that some have specific encoding requirements not needed for Rejang.

The scripts of this group are often referred to as *Rencong* (*Rentjong* in older orthography)[5] or *Surat Ulu* 'upriver writing'.

---

[4] I take the term "Malayu" from de Casparis (1975).

[5] See 3.1.5 for a discussion of possible etymologies for this term.

3.1.1. The Rejang Unicode block

This is the currently encoded representative of the group. The block contains the following specific character groups:

- 18 simple consonant onset letters
- one null consonant vowel bearer
- four letters for prenasalized voiced consonant initials
- eight dependent vowel signs
- four dependent signs for syllable-final consonants
- one virama
- one punctuation mark

Eleven positions in the Rejang block are empty (A954–A95E).
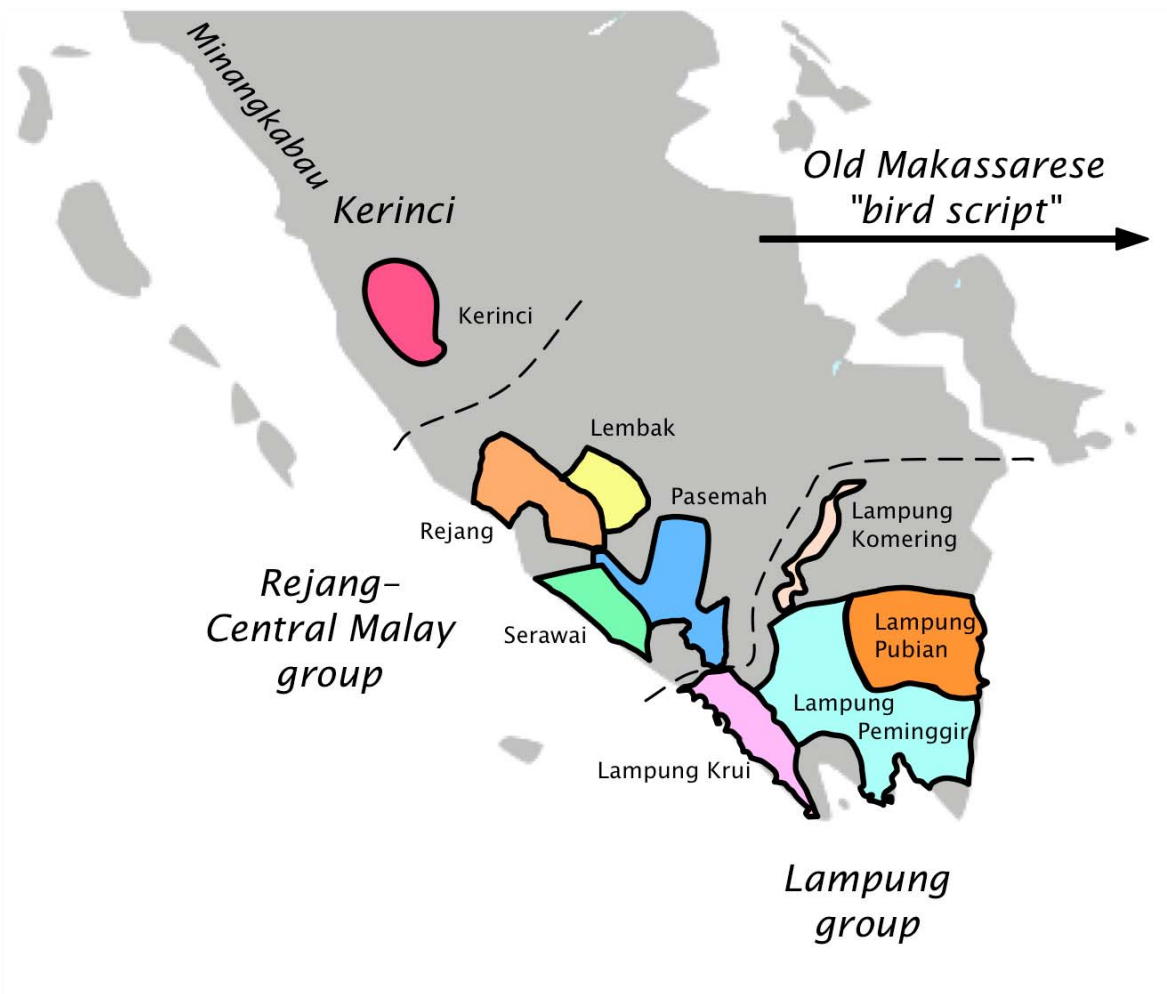
Figure 7. Locations of language/script groups in South Sumatra.

3.1.2. Central Malay extensions (Lembak, Pasemah, Serawai)

Rejang is only one regional representative of a set of closely related script varieties in Central South Sumatra, of which the others were used to write Central Malay dialects[6]. Attestations of varieties with characters representing sounds beyond those encoded for Rejang are found in de Sturler (1843) and Voorhoeve (1971). These include prenasalized voiceless consonant letters (absent in Rejang) and a velar/uvular rhotic "gra" given for several varieties by Voorhoeve.[7] Besides the same voiceless prenasalized consonant letters, extra characters given in de Sturler include a letter he names "erra" (which may be the same letter as Voorhoeve's "gra") and a subordinate consonant sign named *suku*[8] used in the Musi river valley east of the Kedukan Bukit mountain range. Unlike the other consonant signs, which represent coda consonants (including *kajunjung* ‹-r›), *suku* represents /r/ when it appears in second position in an onset, between an initial stop and the syllabic vowel. The additional characters, which number seven or eight, depending on the status of de Sturler's ‹rr›, are illustrated in Table 4.

Value	*Character variants*
ŋk	
ɲc	
nt	
mp	
ŋs	
rr	
ʁ	
r / C_V	

Table 4. Character extensions to the Rejang block needed for Central Malay varieties.[9]

---

[6] I use this equivalent for Voorhoeve's Dutch term *Midden Maleisch* to avoid ambiguity with "middle Malay", a stage in the historical development of the Malay language.

[7] Anderbeck (2008) notes that in the relevant Malay dialects, /r/ is phonetically a velar or uvular liquid. The Dutch digraph "gr" appears to be an attempt to represent this velarity/uvularity in contrast to the standard coronal trill. It may be safely assumed that there was a contrast, or a variable realization of /r/, before the 20th century.

[8] Spelled *soekoe* in the Dutch-based orthography used by Sturler.

[9] According to the comparative table in Voorhoeve (1971), the same ⱴ glyph is used both for ‹ŋs› in the Pasemah region and for ‹mp› in the Serawai and Lembak Bengkulu regions.

3.1.3. Tanjung Tanah manuscript extensions

The oldest known handwritten Malay text is found in a manuscript known as the Tanjung Tanah Code of Laws (Kozok 2004, 2006). The final two pages of this text are written in a South Sumatran script variety whose letter shapes diverge somewhat from Rejang and Central Malay shapes but show similarities to Lampung (see 3.1.4). These are compared with the Rejang range in Table 5.

Value	Tanjung Tanah	Rejang
k		
g		
ng		
c		
j		
ny		
t		
d		
n		
p		
b		
m		
n		
h		
l		
r		
y		
w		
ʔ		
ʔ		

Table 5. Letters in the Tanjung Tanah Book of Laws variety compared to Rejang.[10]

---

[10] The Tanjung Tanah character shapes are drawn after originals in photographs in Kozok (2004b), (no date).

Most letter shapes are a subset of the Rejang range. However, one letter, ‹ʔ›, is unrelated in shape to the functionally similar letter found in all other South Sumatran varieties. It appears, according to Kozok (2004a), to be derived from an uncommon ligature of ‹h› and the virama. Although this particular variant appears only once in any known Sumatran manuscript, it is clearly a token of a shape that at some time must have been in widespread enough use in Sumatra for it to become the evident antecedent of the cognate letter in Makassarese Bird script (see section 4.2):
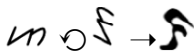
Figure 8. Shape of Tanjung Tanah ‹a› as antecedent to Makassarese Bird script ‹a› when rotated.

The subordinate vowel, coda consonant and punctuation signs found in the Tanjung Tanah document's two pages of South Sumatran script are for the most part a subset of the Rejang range; however two cases appear to justify new code points. The virama corresponds semantically to the Rejang code point; however, the shape of the Tanjung Tanah glyph is derived from the old Kawi/Javanese shape whereas the Rejang raised circle shape is a relatively recent borrowing of the Arabic *sukuun/jazma*. The Tanjung Tanah variety also uses a punctuation mark, analogous to the Latin comma, which occurs more frequently than the REJANG SECTION MARK (A95F) and within lines rather than at the beginning of a section of text. This has no clear counterpart in other attestations of South Sumatran scripts and seems to require a separate code point due to its distinct semantics.

Value	Glyph	Rejang code point
-i	◌̊	A947
-u	◌̞	A948
-ŋ	◌̄	A94F
-h	◌ ̎	A952
-ø	◌꜓	(A953)
Phrasal punctuation	꜀	

Table 6. Vowel, coda consonant and punctuation signs in Tanjung Tanah South Sumatran script.[11]

Although the majority of characters in the Tanjung Tanah variety of South Sumatran script can be subsumed under the Rejang block, it would be preferable, because of their distinct forms and histories, for the ‹ʔ› variant and virama to be assigned to code points distinct from their Rejang counterparts, and for the phrasal punctuation marker to be encoded separately from the existing REJANG SECTION MARK.

---

[11] Drawn after originals in photographs in Kozok (2004b), (no date).

Contact:

Uli Kozok

Indonesian Studies Program

University of Hawai'i at Manoa

http://www2.hawaii.edu/~kozok/index.html

http://ulikozok.com/

kozok@hawaii.edu

3.1.4. Lampung
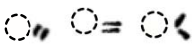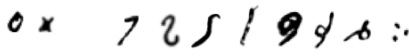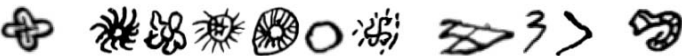
Lampung script is a set of variants ranging from relatively conservative shapes similar to Rejang-Central Malay and Tanjung Tanah variants to others that have gone through one or more stages of structural change. This script group was previously used in the Lampung speaking regions at the southern end of Sumatra. The most complete available description of Lampung script is van der Tuuk (1868), which illustrates ten different character sets from several manuscripts. The inventory of characters is a subset of the Rejang range, as shown in Table 7.

Value	Rejang	TT	Lampung samples from van der Tuuk									
			I	II	III	IV	V	VI	VII	VIII	IX	X
k												
g												
ng												
c												
j												
ny												
t												
d												
n												
p												
b												
m												
n												
h												
l												
r												
y												
w												
ʔ												
ʔ												
ʔ												

Table 7. Comparison of Rejang, Tanjung Tanah, and Lampung letter variants.

Despite the variety of forms illustrated in van der Tuuk, he gives no indication that they are distinct scripts. The most appropriate solution would be to represent diffferent co-occurring variants by different fonts. An alternative 𝟹ᵛ shape for ‹ʔ› found in van der Tuuk's Alphabet IV has no plausible relationship to the more common Lampung variant nor to its counterparts in the other South Sumatran scripts; however, considering its close resemblance to the early 17ᵗʰ century Javanese 𐀀𐀀 [12] or to modern Balinese 𐀀𐀀 [13] it is possible that it is a loan from the Javanese script. This being the case, the existence of this alternative and distinct character shape would seem to justify a distinct code point. Funke (1961) gives an extra ↜ "gra" letter in a modern variety of the script that appears to represent a velar or uvular liquid, as is similarly likely for the "gra" given in Voorhoeve.

Vowel marks, punctuation

Value	Lampung variants	Equivalent Rejang code point
-i		A947
-u		A948
-e		A949
-ə		
-ai		A94A
-au		A94C
-ng		A94F
-n		A950
-r		A951
-h		A952
-∅		A953
Section mark		A949

Table 8. Lampung vowel, coda consonant and punctuation signs.

---

[12] Handwritten shape from folio 325 in Ricklefs (1976), reproduced from Dutch *Nationaalarchief* inventory number 1070: 1.04.02, *Inventaris van het archief van de Verenigde Oost-Indische Compagnie (VOC)*, 1602-1795 (1811).

[13] The Balinese print glyph, as is generally the case, more closely resembles the early Javanese shape than does the modern Javanese print style.

The dependent vowel and consonant signs used in Lampung correspond for the most part to those encoded for Rejang. However, Lampung has a sign for a mid-central vowel /ə/ that appears not to exist in Rejang; close reading of van der Tuuk reveals this to have an alternative pronunciation [o]. It appears however that there is no independent /o/ vowel phoneme. Van der Tuuk also notes the existence of a sign for /e/ in three pages of one manuscript: he explicitly notes that Lampung contains no /e/ phoneme and points out elsewhere that these three pages are written in Malay and not Lampung (though the script variety is Lampung). Unlike Rejang, where the raised circle shape of the virama is clearly a borrowing of the Arabic *sukuun* or *jazma*, most shapes of virama in Lampung are derived — like most other vowel and coda consonant signs — from a Javanese family script.[14]

A peculiarity of stacked vowel and coda consonant signs on the same base letter should be taken into account in the event that Lampung extensions are considered for inclusion in Unicode. Three dependent coda consonant signs (‹-n›, ‹-ng› and ‹-r›) are written above the base letter, similar to the vowel signs ‹-i› and ‹-ə›. In many if not most of the manuscripts reproduced in van der Tuuk, when a letter is marked for one of these vowels together with one of the three consonant signs, the consonant sign is written closer to the base letter than the vowel sign, which is displaced either upward or rightward as illustrated in Figure 9. In examples (a) and (b) of Figure 9, the vowel sign is displaced to the right of the coda consonant sign, and in example (c) it is written above the consonant sign, which remains closer to the base letter.

(a) /pər/ = ‹p-r-ə›   (b) /sin/ = ‹s-n-i›   (c) /niŋ/ = ‹n-ŋ-i›

Figure 9. Displacement of ‹-i›, ‹-ə› with consonant sign written closer to base letter.

An orthographic vowel sign movement rule

Since the Lampung character inventory is a subset of the Rejang range, it might appear appropriate to treat this script as a set of variants included in the former. However, van der Tuuk records an orthographic rule nearly identical in its realization to one found in the Batak scripts, which in the context of a closed syllable marked by a virama on the coda consonant, displaces any vowel sign from the syllable-initial letter onto the coda consonant letter. The only detail of difference is that in Lampung the ‹-i› vowel sign is displaced not onto the second consonant letter but into a position between the initial and final letters. A similar issue was resolved with respect to the input model for Batak script by choosing a logical/phonetic input ordering method coupled with visual reordering in software (Everson and Kozok 2008). Since the Lampung scripts are no longer the primary script of Lampung speakers, who are literate in Indonesian and accustomed to the iconically linear sound to letter mapping of Indonesian Latin script orthography, it would be desirable to determine if the Lampung user community would prefer a similar solution to be implemented in the event this script is included in Unicode.

---

[14] It bears mentioning that Voorhoeve (1972) notes the presence of similar Javanese-type shapes only in Old Lebong script among the various Rejang-Central Malay varieties. In modern Lebong and the others, only the Arabic-derived shape is attested.

Possible contacts:

Uli Kozok

Indonesian Studies Program

University of Hawaiʻi at Manoa

http://www2.hawaii.edu/~kozok/index.html

http://ulikozok.com/

kozok@hawaii.edu

Titik Pudjiastuti

Program Studi Jawa, Fakultas Ilmu Pengetahuan Budaya/

Departemen Sastra

Universitas Indonesia

titikpuji@yahoo.com

3.1.5. Kerinci script

The northern area of South Sumatra is represented by the Kerinci script, a variety whose characters diverge slightly in shape and properties from those in other South Sumatran varieties. This variety of script is often referred to as *surat incung* "cursive writing" or *rèncong* "dagger writing" (Kozok 2004a). Kozok points out that the latter term appears to be based on the name of a ceremonial dagger from Aceh, which seems an unlikely connection given the geographical distance between the two regions. However, the two terms may plausibly be traced to folk etymologies or local cognates of the Malay/ Indonesian word *runcing* 'sharp, pointed'. Most known existing manuscripts are horn or bamboo inscriptions,[15] however one ink on bark paper manuscript is described, illustrated and transliterated in Voorhoeve (1970). Several horn manuscripts are reproduced, transliterated and translated in Westenenk (1922). Comparison of the script variants from these two sources reveals slight differences in the shape of certain characters due to the writing medium used. In comparison with Rejang and Central Malay script variants illustrated in Tables 4-6, Figure 10 reveals differences are generally not clearly significant enough to merit encoding in a separate block.
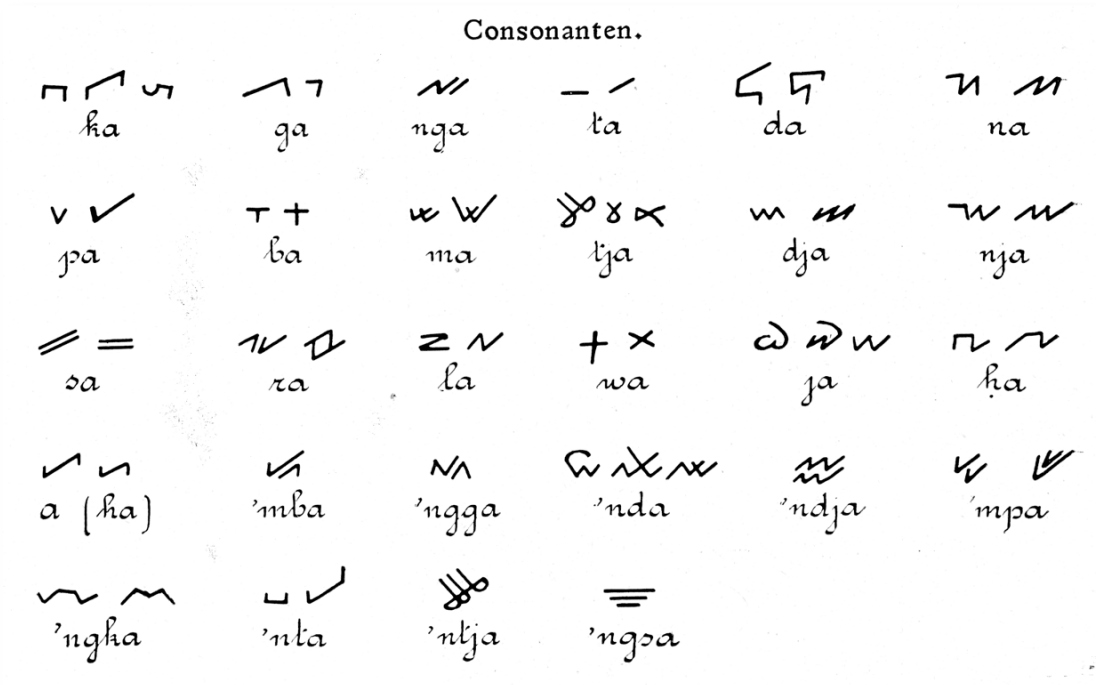
Figure 10. Consonant letters in Kerinci script (Westenenk 1922).

Vowel marks, punctuation

Unlike their cognates in Rejang-Central Malay and Lampung, the ‹-i› vowel character and ‹-h› coda consonant character both appear not as marks attached to their host consonant letter, but as distinct

---

[15] Examples are available at the *Pusaka Kerinci* web site linked to in the bibliography.

characters occupying a full letter space to the right of the host letter, as illustrated in the final two rows of Westenenk's Vowel table (*Klinker-teekens*) and in the examples (*Voorbeelden*) in Figure 11).

Figure 11. Kerinci vowel signs (Westenenk 1922).

Although this is not mentioned in Westenenk, Voorhoeve (1970) explicitly points out that the vowel /o/ is written as a digraph with the dependent ‹-u› vowel sign beneath the host letter followed by the spacing ‹-i› sign. He notes that this digraphic vowel notation is analogous to the compound vowel sign used in Rejang, which combines *kamica* ‹-e› and *kamitan* ‹-u› on the same host letter. Although Unicode reserves a distinct code point for each of these two signs (A948 for ‹-u› and A949 for ‹-e›), the Rejang digraph is given a separate code point (A94B). Adopting the same approach for Kerinci might lead to unnecessary complications, since the ‹-i› character is involved in an idiosyncratic orthographic metathesis rule, described immediately below, which also affects it when it is part of the digraph that represents /o/, and even across a line break.[16]

---

[16] I have observed this in the bark paper manuscript in Voorhoeve (1970).

A limited orthographic metathesis rule

As noted above, ‹-i› and ‹-h› appear not as vowel signs on their host letter but in separate letter spaces to the right of the host[17]. These two characters are also involved in an orthographic metathesis rule that reverses the logical phonetic order ‹-i-h› to ‹-h-i›. Figure 12 illustrates one version of the rule given in Westenenk (1922), in which the two characters switch position in a straightforward manner.

/tabih/   /sirih/

t[a]   b   h   i   s   i   r   h   i
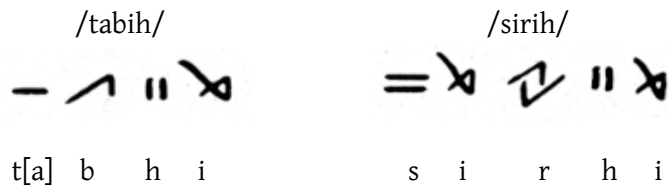
Figure 12. The Kerinci orthographic metathesis rule.

In the text illustrated in Voorhoeve (1970), the metathesis has the further effect of removing one of the two vertical bars in ‹-h› when it is reordered before ‹-i›: as a result this positional variant of the character is identical in shape to the virama in the same text (a single vertical bar). However, there is no possibility of confusion between the two because the virama itself can never appear to the left of a dependent vowel sign: the only characters that can follow to its right are a full consonant or a vowel bearer letter.

This rule lacks the general scope of the Batak and Lampung closed syllable vowel sign movement rule, and according to Westenenk and to my own observations, it only appears to affect these two characters in Kerinci script. It seems most clearly related to the Lampung convention that displaces any vowel sign written above the letter to a position further away if it co-occurs with a dependent coda consonant sign above a letter (*cf*. Figure 9 in 3.1.4). Considering the generally distinct letter shapes of Kerinci script and taking into account the distinct spacing properties of these two characters and the reordering that affects them when they co-occur, it appears that the script poses certain challenges to being encoded under the Rejang block, unless these two characters are assigned their own code points; the reordering rule could either be left to those who are knowledgeable about the script or dealt with in software.

Possible contact person:

Rina Syafitri Rusdi
Chair of *Yayasan Kajanglako Art Center*, Kerinci
Sumatra, Indonesia

---

[17] In this respect, these two signs are comparable to two vowel signs in the Batak scripts: ‹-i› and ‹-o› (in their Toba readings).

3.1.6. Alleged indigenous Minangkabau scripts

Images have been circulating on the World Wide Web of an alleged indigenous script used in the Minangkabau-speaking area of central Sumatra, between South Sumatra proper and the Batak lands to the north, attributed to Rais Kamardi (1987) and Sango (no date).

The letter shapes, illustrated in Table 9, appear to be plausibly related to those in other South Sumatran scripts, though the forms they take in the drawings are likely distorted. According to the reports reproduced on the web sites, there are fifteen base letters, four vowel signs and and one virama. The letter inventory corresponds to the basic Rejang inventory without prenasalized consonant letters, but also lacks a series of palatal consonant letters: this may be a reporting error since Minangkabau is closely related to Malay and would most likely have the same set of palatal consonants as Malay (namely /c/, /j/, /ɲ/, /y/), increasing the set to 19 letters.

According to a recent news article (Rinaldi 2010), Dr. Herwandi, Dean of the Faculty of Literature at Universitas Andalas in Padang, has found similarities in the Batulih Borobono stone inscription located in the region with illustrations of alleged Minangkabau letters published by Datuk Rajo Darwas Malano and Zuber Usman (no references available) and with other South Sumatran scripts. However, the lack of any known manuscripts in an indigenous Minangkabau script has made it difficult to determine the inscription's letters accurately.

There seem to be no original documents available in any earlier Minangkabau script according to the Rinaldi article, but from the available evidence, if it is accurate, it appears that this script could be treated as a font variant under the Rejang block.

A second script is reported to be found in a book from Sulit Air in Solok Regency (Rinaldi 2010, Rais Kamardi 1987). From the online reproduction of illustrations in Rais Kamardi, this appears to be an invented cipher script only superficially related in structure to Indic alphasyllabaries. Letter shapes in this script are completely unrelated to those of Indic scripts, as are the shapes and positions of vowel signs. Although these signs are placed above, below, to the left and to the right of consonant letters like dependent vowel signs in Indonesian scripts, these placements are unrelated to those of any Indic script; there is also a vowel sign for /a/ instead of the unmarked default reading in Indic scripts, and no virama: bare consonant letters are read without a vowel. The script also has a series of glyphs to replace Latin script punctuation marks such as the comma, question mark and others, as well as a series of clearly invented signs for arithmetical operations and relations. Since this script appears to be attested only from a single source and is clearly an invented alphabetic cipher that superficially mimics the appearance of an Indic script, it almost certainly does not merit further consideration.

Value	Other South Sumatran	Minangkabau
k	∧	⊤ʔ
g	∧	>
ng	N/ N N	⌄
c	⋏	
j	⋀	
ny	⋀	
t	⬧ ⬧ ∕	—
d	⋀	ℨ
n	M	⍭
p	√	∨
b	⋏	⊤
m	⋌ ⋎ ⋎	⅂'
s	⋀	⅂
h	⋁	⋁
l	⋀	⋀
r	⋀	⋀
y	W	
w	⋀ ⋀	⋋
ʔ	⋏	Y
-i		⊙
-u		⊙
-e		⊙
-o		⊙
-Ø		⊙

Table 9. A partial alleged Minangkabau script (from Datuk Rajo Darwas Malano and Zuber Usman) compared with other South Sumatran variants.

3.1.7. The *Angka bejagung* numeral system

An unusual quinary numeral notation system named *Angka bejagung* is attested for the Rejang speaking area in Jaspan (1967) and Sani (1988).[18] According to Jaspan, this numeral system may have been less common than the use of Rejang letters as numerals, an application reserved mostly for pages in manuscripts. Jaspan states that he saw samples of *Angka bejagung* used by various people, but the best sample was provided by one Adjai (= Haji) Sari who used it for his village shopkeeping accounts.

The basic characters for the system represent 1, 5, 10, 50, a base character for the 100s that combines with a score mark for each 100 up to 300, and base characters for 1000 and 10 000. These are combined in additive fashion in two-dimensional blocks to derive the other quantities. The illustrations in Jaspan and Sani agree but for slight differences in the positioning of characters in some cases. Jaspan, unlike Sani, illustrates an alternative glyph for the 5 character. Examples of individual and combined characters are reproduced in Table 10, based on the examples given in Jaspan's Figure 3 and on page 20 in Sani.

---

[18] *Angka* means number; I have not been able to find a translation for *bejagung.*

/	**//**	**ʲ/ʲ**	**ʲ/ʲ/**	
1	2	3	4	
T or **ᑫ**	**ᴛ**	**ᴛᴛ**	**ᴛᴛᴛ**	**ᴴᴛ**
5	6	7	8	9
d	**dʲ**	**dʲʲ**	**ʲdʲ**	**ʲdʲ**
10	11	12	13	14
d	**d**	**d**	**d**	**d**
15	16	17	18	19
dd	**ddʲ**	**ddʲʲ**	**dd**	**dd**
20	21	22	23	24
dᴛd			**djd**	**djdj**
25			30	40
⊕	**⊕d**	**d⊕d**	**d⊕d**	**d⊕d**
50	60	70	80	90
⤬d	**⤬dd**	**⤬⊕**		
100	110	150		
⤬Ọ⤬	**⤬Ọ⤬**	**⤬Ọ⤬Ọ**	**⤬Ọ⤬**	**⤬Ọ⤬**
200	300	400	500	600
/	**//**	**///**	**////**	**/////**
1000	2000	3000	4000	5000
//////	**///////**			**ᴎ**
6000	7000			10 000

Table 10. The *Angka bejagung* numeral system associated with Rejang script.[19]

---

[19] The **d** form of the right hand portion of **⤬d** '100', as illustrated both in Jaspan and Sani, appears to be a graphic variant of the circle plus diagonal straight line found in the other numerals for the hundreds (*cf.* 400).

3.1.8. Summary

The South Sumatran group of scripts require a number of extensions to the existing Rejang block. The extra letters needed total 9. Additional dependent vowel or consonant signs needed because they are not present in the Rejang inventory or have different spacing properties total 6. Taken together with an additional punctuation mark, these extensions number 16, five more than the current 11 unassigned code points. The most important extension to the block consists of the *Angka Bejagung* numerals: although these can be decomposed into eight base symbols, the complex manners in which they are combined may justify treating at least some, if not all of them, as single compound glyphs with their individual code points.

For the most part, variant South Sumatran character shapes that are formally cognate with those in Rejang could be treated as font variants rather than as code points in separate blocks. It is mainly in cases where characters with the same semantics demonstrably have a different origin that it seems justified to give them a distinct code point. This is mainly the case for the Tanjung Tanah ⟨ʔ⟩, which is derived from ⟨h⟩, and the various shapes of the Lampung virama, of which one set (cognate with the Tanjung Tanah variant) derives directly from the Old Javanese shape, contrasting with the raised circle that is a direct borrowing of the Arabic *sukuun/jazma*.[20]

Apart from the need for extensions to account for as yet uncoded distinct characters in the South Sumatran script group, it is also necessary to take into account the existence of an orthographic vowel sign movement rule for Lampung that is analogous to the rule found in Batak, a rule in Lampung that ensures that the coda ⟨-n⟩, ⟨-ng⟩ and ⟨-h⟩ signs are placed closer to the base letter when co-occurring with the ⟨-i⟩ and ⟨-ə⟩ vowel signs which are also placed above the letter, and a rule for Kerinci script analogous to both rule types that switches the order of ⟨-i⟩ and ⟨-h⟩ when they are written together to the right of the base letter.

---

[20] In Rejang-Central Malay, where the Arabic-derived glyph is the usual shape, there are nevertheless a group of Javanese-derived variants attested in older Lebong (Rejang) texts (Voorhoeve 1971).

3.2. Sumatran post-Pallava[21] or "Malayu" varieties

Stone inscriptions are found throughout highland Sumatra in a script that was in use around the 13th through 14th centuries in the Dharmasraya kingdom (Kern 1917a, b), and a similar variety is found on some 13th century tombstones in the Aceh region of northern Sumatra (Stutterheim 1936). The recently rediscovered 14th century Tanjung Tanah manuscript from Kerinci in South Sumatra (Kozok 2004, 2006) uses a third, very similar variety and with its 32 pages in this script, it is the most extensive representative of this group. Although de Casparis (1975) uses the term "Malayu" to refer to these varieties, it is not yet established if they should be considered a single script or closely related scripts: determining their precise relationships requires research that has yet to be done. Although they are clearly related to earlier Kawi (Old Javanese) script, the letter shapes of the Sumatran scripts clearly are later developments that have undergone important structural changes. Although many letters show strong similarities to shapes in contemporary Javanese inscriptions, ‹m› has a similar divergent shape both in the various inscriptions (as noted in de Casparis) and in the Tanjung Tanah manuscript.

Relation to historical Javanese script varieties

Certain characteristics of these varieties seem to place them (with the possible exception of the ‹m› shape) as representatives of an intermediate stage between Old Javanese or "Kawi" (*cf.* Holle 1882/1999, de Casparis 1975 and Postma 1992) and modern Javanese-Balinese. Some letter shapes, for example ‹n› and ‹w›, are for all intents and purposes embryonic versions of modern Javanese shapes that can be seen in early form in letters from the Sultan of Banten to the Dutch written in 1619. Others relate directly to modern Javanese shapes by reanalysis of stroke order and elaboration of existing on- and off-strokes.

Inventory

The character inventory, especially in the T(anjung) T(anah) manuscript, seems upon superficial overview to correspond generally to the Javanese-Balinese inventory. TT script contains autonomous letters that correspond fairly directly to the Javanese-Balinese range, comparable dependent vowel signs and conjunct letters, textual punctuation signs, and a restricted range of numerals.

Summary

In view of the above, although much work remains to be done in establishing the relationship of the Sumatran Malayu-type scripts to each other and to related scripts of Java, they do seem to share common characteristics that set them apart from both the Old Javanese/Kawi attested from Java and modern Javanese script. It seems worthwhile to envision the eventuality of a Sumatran Post-Pallava block equivalent in range and content to the existing Javanese or Balinese blocks.

---

[21] The term "Post-Pallava" is based on suggestions in Kozok (1996) and (2004a).

4.   Sulawesi, Sumbawa and Flores islands

The island of Sulawesi in eastern Indonesia is the historical home of one major script, Bugis (encoded under the name *Buginese* in Unicode) and two well-attested minor scripts: Makassarese *Jangang-jangang* or "Bird" script and the *Lontara' bilang-bilang* ("counting" or "number") script. These three are attested for the southwestern peninsula of the island, and a fourth script of unknown status has recently been described for the Minahasa region on the tip of the northeastern peninsula.

Earlier varieties of Buginese script that have come to light in recent research by Noorduyn (1993) and Tol (2008) use a set of variant glyphs that can cooccur in the same text as well as three alternative vowel glyphs that may, because of their interrelationships, require treatment as distinct code points. The Makassarese Bird script was noted in passing in an illustration in the Buginese final proposal (Everson 2003), but was the distinct national script of the Makassarese during the 17th century and perhaps earlier, used to write large portions of the entries in the Chronicles of the kingdoms of Gowa and Tallo'; the script is also encountered in the signatures of several Makassarese witnesses to the 1667 Treaty of Bungaya between the Makassarese and the Dutch, reproduced in Tol (1996). *Lontara' bilang-bilang* is a slightly different case: although used in a genre of poetry, it is a specialized script based on a common Arabic cipher, with sophisticated adaptations based on Jawi and Buginese scripts.

Buginese script was also used with extensions not present in the parent script to represent distinct properties of the phonologies of the Bima language in eastern Sumbawa and the Ende language of central Flores. Apart from these two extended versions of the script, there is a single attestation in Raffles (1817) of an entirely distinct older script used in Bima, for which no other supporting evidence has yet been uncovered.

4.1.   Buginese extensions

The Buginese script as currently encoded contains code positions for those letters, vowel signs and punctuation signs used in the script after its standardization in the mid-19th century. Prior to this time, variation in the script included stylistic variants in an old "palm leaf script" style that commonly cooccur with each other and standard variants in the same text, as well as a trio of dot notations for independent vowels that depart from the simple model of base letter plus dependent vowel sign. A separate set of characters were used in the northern kingdom of Luwu' for complex consonants including eight with no equivalent in other Buginese script varieties. In the islands of Sumbawa and Flores, which adopted Buginese script, extensions were made to the script to represent sounds in the respective languages that are absent in Bugis itself.

4.1.1. The Buginese Unicode block

The current Buginese block occupies positions 1A00-1A1F. Two positions in the block (1A1C, 1A1D) are empty.

4.1.2. Obsolete palm leaf script letter variants

An old set of variants predating standardization of the script were used in manuscripts written on palm leaf strips and are also found in various other manuscripts. Palm leaf script distinguishes itself from the later standard variety most notably by the greater verticality of its character glyphs. In particular, the ‹k› is formed by two vertical straight lines as opposed to the two angled lines of the standard variety, and the arches and cups that form most letters are much higher and deeper in this variety.

Beyond the global stylistic differences, palm leaf script contains a number of variants whose shapes differ in important ways from those in other varieties. These include a set of variant shapes for certain letters that co-occur freely with others in a single text, including with standard shapes, and three variant shapes for syllable-initial ‹a›, ‹i› and ‹u› whose unusual characteristics may justify their encoding as distinct letters.

Three letters with free shape variation in palm leaf style

Three letters illustrated in Table 11 have special variants in palm leaf style that can freely co-occur in a single text, both with each other and with the standard variants (Noorduyn 1993, Tol 2008). Of these, the 'long S' shape and its reversed counterpart can also be found in ordinary manuscripts that display a relatively older style.

Value	*Standard variant*	*Palm leaf variants*
‹j›	⌒	∴ ∴
‹s›	○	I Ո ʃ
‹l›	∾	∾

Table 11. Palm leaf variant glyphs for three Buginese letters.

Because of the free variation typical of these three letters in palm leaf style documents, treating the palm leaf variants as alternative font letters might entail more complications than treating them as extra variant letter shapes with their own code points.

Unusual initial vowel letter variants in palm leaf style

Standard Buginese script includes a single null (or glottal stop) syllable onset letter ‹ʔ› that has a dotted double arch shape similar to several other letters in the Buginese inventory. The bare letter represents a syllable with /a/, and the other vowels are represented straightforwardly by combining the requisite dependent vowel sign with the consonant in their canonical positions above, below, to the left or to the right.

Palm leaf style uses an alternative letter shape, a middle dot (•), whose form leads to a number of complications in the representation of vowels other than /a/. The ‹-e› and ‹-o› signs, rather than being

placed completely to the left and right of the palm leaf ‹ʔ› shape, respectively, are placed so that their stems are to the left or the right of the dot, with the arch extending from the top of the stem covering and enclosing the dot (Table 12). While the spacing behavior of these vowel signs is different from the way they behave with other letters, it can be dealt with most appropriately as a matter of kerning in a palm leaf style font.

Less straightforward is the way the palm leaf style of representing /i/ and /u/ affects the form of the resulting glyphs. Since ‹-i› and ‹-u› are respectively represented as a dot above or below the host letter, the expected result would be to place a dot above the palm leaf dot ‹ʔ› for ‹ʔi› and below for ‹ʔu›. However, this was apparently felt to be too ambiguous, and a different means was chosen for distinguishing the two (Table 12): ‹ʔi› was represented with two vertical dots, and ‹ʔu› with three. In this case then, the normally expected straightforward combination of vowel sign with base letter does not take place, and ‹ʔu› (at least) is represented irregularly by a suppletive glyph.[22] although this is not entirely clear from the handwriting in palm leaf manuscripts, it may also be the case that ‹ʔi› is represented by a glyph that is not a simple combination of the dot ‹ʔ› with a dependent ‹-i› dot placed above the letter space, but rather a vertical double dot spaced within the normal letter space. If this is the case, then ‹ʔi› would also be a suppletive character; in consequence, this possibility should be investigated in the event of a proposal for palm leaf style extensions to Buginese script.

Value	*Standard variant*	*Palm leaf variant*
ʔa	∿	•
ʔi	∿	⋮
ʔu	∿	⋮
ʔe	‹ ∿	ʕ˙
ʔo	∿˄	ʔ
ʔə	∿	˘

Table 12. Distinct representations of ‹ʔa›, ‹ʔi›, ‹ʔu›, ‹ʔe›, ‹ʔo› and ‹ʔə› in palm leaf script.

---

[22] Although there is no clear evidence to support this hypothesis, it is possible that the triple dot may in fact be a reflection of an earlier, otherwise disappeared, independent ‹u› letter related in shape to the three fragmented bars of the Batak ‹u› or the three connected horizontal strokes of the Tagalog/Hanunóo/Buhid letter.

4.1.3. Luwu' variants of Buginese script

Noorduyn (1993) discusses a set of special letter shapes found in a single manuscript from the kingdom of Luwu' in the northeastern end of the Bugis lands. The shapes he illustrates are reproduced from Noorduyn and Salim (1988) with clarifications and a correction. Six of the variants illustrated in the chart are variant shapes for existing letters that can be dealt with by means of alternative font glyphs[23], however there are eight glyphs for prenasalized or geminate consonants that have no standard Buginese equivalents (Table 13).

Value	Glyph	Base glyph	Base glyph value
mb	𑜒	𑜒	b
nt	𑜒	𑜒	ny
nd	𑜒	𑜒	d
nj	𑜒	𑜒	j
nn	𑜒	𑜒	n
ss	𑜒	𑜒	s
tt	𑜒	𑜒	t
mm	𑜒	𑜒	m

Table 13. Special Luwu' variants with no equivalents in standard Buginese script.[24]

Each of these extended letters is derived by adding a short descending diagonal stroke to an existing standard letter. Except for ‹nt›, which is derived by adding the dash to the end of the cup shape on the underside of ‹ny› (apparently because the base ‹t› shape was already used to derive ‹tt›), these are based on the letter for the corresponding plain consonant. If the base shape begins with an arch, the diagonal is attached to the left of the letter; if it is a cup shape, the diagonal is attached on the right. The three less straightforward cases are ‹nt› as mentioned above, where the cup on the underside is treated as the base for the additional stroke, ‹b›, where the stroke is added on the left side of the join between the lower arch and the upper adjunct hook, and ‹s›, where the upper arch was chosen as the base for the added stroke.

Although these extra glyphs appear to be attested only for a single manuscript, their existence is evidence that they must to some degree be representative of the script as used in the kingdom of Luwu' and thus merit consideration as extensions to the current Buginese script.

---

[23] Some of these variants, in particular for ‹s› and ‹h›, are in fact encountered in Bugis texts from other regions.

[24] These are after from the shapes illustrated in Noorduyn and Salim (1988) and Noorduyn (1993).

4.1.4. Ende script extensions

The Ende language is spoken in central Flores island (located in the Lesser Sunda/Nusa Tenggara archipelago east of Bali and Sumbawa, and south of Sulawesi) and small islands off the Flores coast. The language formerly used an extended variety of Buginese script which has been supplanted, like other indigenous Indonesian scripts, by Latin script. Ende script is derived from an older variety of Buginese script. The name of the script, *Lota Ende*, is derived from *Lontara',* the alternative Bugis name for Bugis and Makassarese script. Though still known by a minority of Endenese, it is apparently in decline (Oktora and Anwar 2010, Wahono 2010).

The original Buginese letter inventory was extended to represent sounds of the Ende language that are not present in Bugis (Table 14). These include two implosive voiced stops and three prenasalized voiced stops[25] that contrast with plain voiced egressive stops and nasals, and some fricatives and approximants not found in Bugis. Ende has strict CV syllable structure except for some rare loans from Indonesian and other sources, and the above-mentioned consonants all occur in syllable initial position. In addition to a glottal stop consonant represented by the same ᨖ glyph as in Bugis, Ende has a contrasting onset described by van Suchtelen as a voiced [ɦ] and by McDonnell (2009) as varying between breathiness in absolute initial position and between breathiness and epenthetic [w] and [y] glides word-internally, depending on the quality of the neighboring vowels[26]. I will represent this contrasting smooth glottal onset as ‹ɦ›.[27]

The letters that correspond directly to Bugis letter shapes reflect the more archaic shapes of palm leaf script (see 4.1.2). Two consonant letters (Ɡ ‹ɓ› and ꞯ ~Ɪ ‹ᵐb›) appear to have been borrowed from earlier variants of Bugis Ꞩ ‹c› and ꞩ ‹ɲc› with reassigned values. The /c/ phoneme is marginal in Ende and according to van Suchtelen, the Ɪ glyph used to represent this sound is of extraneous origin: it does show similarities in shape to an earlier variant of Bugis ‹ɲc› and may be a secondary borrowing. Although Roos (1887) and van Suchtelen (1921) disagree on which of ꞯ and Ɡ represent ‹ɦ› and ‹y›, respectively, it appears from the discussion in McDonnell that [y] is a positional phonetic variant of underlying ‹ɦ›. Banda (2005) notes only that these two glyphs are found, along with a third, more horizontally oriented variant of ꞯ.

Four letters (‹ɣ›, ‹ɗ›, ‹f› and ‹h›) appear plausibly to be derived from Arabic counterparts representing the same or similar sounds: ‹ɣ› ‎ ؅ < غ, ‹ɗ› ‎ ؆ < ض, ‹f› ‎ ؇ < ف, and ‹h› ‎ ؈ ~ ؉ < ه. The second ‹h› variant, attested by Roos, is also found in early Bugis texts; the first variant, given by van Suchtelen, is close in shape to another variant in early texts illustrated in Noorduyn (1993) and Tol (1996).

---

[25] At a syllable boundary, the nasal portion of a prenasalized stop is realized phonetically as a coda consonant in the preceding syllable.

[26] McDonnell (p. 210) notes that the epenthetic [w] is phonetically less turbulent than its phonologically underlying counterpart. He does not report a phonemic /y/ in the inventory of Ende.

[27] Due to the need for precise representation of the Ende consonantal inventory, I use IPA notation in this section, but retain 'y' for the palatal glide and 'c' and 'j' for the alveo-palatal voiceless and voiced affricates.

Value	van Suchtelen	Roos	Bugis	Code point
k	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A00
g	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A01
ŋg	⟨glyph⟩	⟨glyph⟩		
ɣ	⟨glyph⟩	⟨glyph⟩		
ŋ	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A02
ŋk			⟨glyph⟩	1A03
t	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A08
d	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A09
ɗ	⟨glyph⟩	⟨glyph⟩		
ⁿd	⟨glyph⟩	⟨glyph⟩		
n	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A0A
c	⟨glyph⟩		⟨glyph⟩	1A0C
j	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A0D
ɲ	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A0E
ɲc			⟨glyph⟩	1A0F
p	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A04
b	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A05
ɓ	⟨glyph⟩	⟨glyph⟩		
ᵐb	⟨glyph⟩	⟨glyph⟩		
f	⟨glyph⟩	⟨glyph⟩		
m	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A06
mp			⟨glyph⟩	1A07
h	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A16
s	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A14
r	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A11
nr			⟨glyph⟩	1A0B
ɹ	⟨glyph⟩	⟨glyph⟩		
l	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A12
y	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A10
ɦ	⟨glyph⟩	⟨glyph⟩		
ʔ	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A15
w	⟨glyph⟩	⟨glyph⟩	⟨glyph⟩	1A13

Table 14. Ende *Lota* script letters compared with encoded Buginese letters.

Ende script uses the same four vowel signs as Makassarese (Table 15). Although /a/ and /ə/ contrast phonemically, it does not use the *ɔccə'* ‹-ə› vowel sign of Bugis (BUGINESE VOWEL SIGN AE – 1A1B).[28] However, van Suchtelen (p. 225) reports the use of a virama ("*tanda mata*")[29] to indicate vowelless consonants in loanwords where these occur.

Value	van Suchtelen	Roos	Bugis	Code point
i	⠈	⠈	⠈	1A17
u	⠠	⠠	⠠	1A18
e	⌐	Γ	‹	1A19
o	⌐	⌐	⌐	1A1A
ə			⸴	1A1B
-∅	⌃			

Table 15. Ende *Lota* script vowel and virama signs compared with encoded Buginese signs.

Possible contact persons:

Maria Matildis Banda
Author of *Aksara Lota Ende,* lecturer at the Faculty of Letters, Udayana University, Bali

Mustafa Saleh Nggae
Lota preservation activist, District of North Ende

Prof Stephanus Djawanai
Professor of Linguistics, Universitas Gadjah Mada, Ende

---

[28] The default vowel read with bare consonant letters is /ə/, and /a/ is expressed by following the unmarked consonant letter with an unmarked ‹ɦ› (van Suchtelen 1921:224).

[29] It is unclear if this is the actual name used locally rather than a typographic or lexical error. The term seems clearly to be Malay/Indonesian; however, the second word would normally be *mati,* which has the meaning 'die'. The *mata* in van Suchtelen's *tanda mata* normally means 'eye' in Malay, but the sign's shape has no obvious visual resemblance to an eye.

4.1.5. Bimanese variants

The Bimanese language is spoken in the eastern half of Sumbawa island in the Lesser Sunda/Nusa Tenggara islands. Before the Latin alphabet, it was written in adaptations of Arabic and Buginese scripts and there is a single example alphabet illustrating syllable onset letters (but no vowel signs) for a script alleged to have been used at some earlier time in Bima.

According to Jonker (1896), Bimanese in Buginese script (which he called "Makassaarsch schrift") could either be written without any specific additions to the original script, new letters derived by the addition of extra strokes added to the original Buginese forms indicated prenasalization and other distinctions necessary to the Bimanese sound system. The only sources available to me with illustrations of the specific letter shapes used for Bimanese are the second-hand reproductions in Holle (1882)[30] and an original document in Bimanese-Buginese script illustrated in Chambert-Loir (1996).

It can be concluded from Jonker that the Bimanese inventory contains the following syllable onset consonants: k g ŋg p f b ᵐb m t d ⁿt ⁿd c ᶮc j ᶮj l r y w h s. He states that when plain Buginese script style is used, prenasalization can be indicated by placing *anca'* ◌ (1A1B, BUGINESE VOWEL SIGN AE) over the relevant consonant, adding that it also serves as a virama when writing Bimanese. However, Jonker says that prenasalization is rarely indicated this way in writing. Otherwise, prenasalized consonants can also be indicated with an extra stroke added to the base letter. A separate letter ‹f› is also a modification of the basic ‹p› letter according to his discussion. As well, the original Buginese ‹y› is used to represent /h/ with an ✗ shaped character used to represent /y/ instead.[31]

From the illustrations in Holle and Chambert-Loir, it is difficult to determine the actual values of the various characters encountered; however, there do appear to be a number of characters in use that contrast visually with others and probably, in line with Jonker's discussion, represent prenasalized consonants and possibly also /f/. The inventory certainly seems — at least on the basis of these two documents — to be smaller than the Ende script inventory, and also seems to require fewer extensions than Ende script. For lack of any clearer information at the time of writing, it seems prudent not to venture further than to point out the eventual likelihood of a need for a number of extensions to take into account the way Buginese script was used to write Bimanese, for the benefit of scholars in the area.

4.1.6. "An alphabet formerly adopted in Bima but not now used"

Raffles (1817) reproduces a set of characters (Figure 13) allegedly formerly used in Bima. However, this is the only attestation of such a script and no other documentary evidence of any similar letters has as yet been found.

---

[30] Also recently republished, with a translation into English of Holle's preliminary essay, as Holle (1999).

[31] This recalls the confusion between character glyphs for zero onset or /y/ onset syllables in Ende.

Figure 13. Alleged older Bimanese letters from Raffles (1817).

4.2.   Makassarese *jangang-jangang* (bird) script

Although Makassarese has generally been written in the same script as the Bugis language, with a few minor differences in the way specific characters were used, it was written in its own script, distinct from Buginese script, during the 17th century and to a somewhat lesser degree in the 18th century. The script used before the move to Buginese script was commonly referred to as "bird letters" (*ukiri' jangang-jangang* in Makassarese (Jukes 2006) or *uki' manu'-manu'* in Bugis (Tol 1996). This name refers to the shapes of many of the letters, which resemble birds in various postures as can be seen in Table 16.

The relationship of the Makassarese Bird Script to Buginese script and their ultimate origins have long been mysterious. I present evidence in Miller (in press; in preparation) that Bird Script derives mainly from South Sumatran script varieties, including the unique token of ‹ʔ› in Tanjung Tanah script.[32] Where letters show no direct relationship to South Sumatran counterparts, a demonstrable relationship exists with early modern Javanese counterparts.[33] Although the structural relationships are systematic, Bird Script differs from these cognate scripts by the prevalence of curving arches in its lettershapes, a feature it shares with Buginese script. In addition, its dependent vowel signs are identical in composition and shape to those of Buginese script (with the exception of ‹-ə›, which is not used because there is no /ə/ vowel in Makassarese) and are distinct overall from those attested for South Sumatran scripts. These facts taken together lead to the hypothesis that the script was intentionally devised for the Makassarese language by adapting Sumatran and Javanese scripts to the pre-existing model of Buginese script with its arching strokes, and retaining the already extant vowel marking system.

That this script was apparently conceived to replace the Buginese script likely already known to the person(s) who devised it does not however justify treating this script as a cipher of Buginese. It was used for at least a century as the national script of the Makassarese kingdoms (Noorduyn 1991), is the script of the signatories to the 1667 Treaty of Bungaya and other documents, and stands in a relation to

---

[32] The letter shapes used as a basis for comparison, reproduced in Table 16, are drawn from the earliest dated entries (from the early decades of the 1600s) in the Chronicles of Gowa and Tallo' (Tropenmuseum 668-216).

[33] The ‹d› letter has a clear relationship both to South Sumatran and Javanese cognates and may have been partial inspiration for the borrowing of Javanese letters elsewhere in the script.

Buginese script similar to that between Cyrillic and Latin scripts as used for Serbian, compared to the monoscriptal Croatian standard.

Value	Bird script	South Sumatran	Javanese Early 17th century	Javanese Modern. Balinese	Javanese Modern Javanese	Buginese
k						
g						
ng						
c						
j						
ny						
t						
d						
n						
p						
b						
m						
s						
l						
r						
y						
w						
ʔ						

Table 16. Makassarese Bird script letters compared with South Sumatran, Javanese and Buginese counterparts.[34, 35]

---

[34] Early 17th century Javanese variants from Ricklefs (1976), reproduced from Dutch *Nationaalarchief* inventory number 1070: 1.04.02, *Inventaris van het archief van de Verenigde Oost-Indische Compagnie (VOC)*, 1602-1795 (1811).

[35] Makassarese Bird script characters in Tables 16 and 17 and Figures 14-16 from KITLV Tropenmuseum, Amsterdam 668-216.

The script's inventory consists of 18 consonant letters, one common character that appears to be a variant of *angka* (the Arabic numeral ٢ (2) occasionally used in Buginese script as a reduplication marker), the Buginese ‹-i›, ‹-u›, ‹-e› and ‹-o› vowel signs, and three text-level punctuation signs.

Since the vowel signs are identical in shape and usage to Buginese 1A17-1A1A illustrated in Table 15 for Ende *Lota*, they will not be reproduced here. In common with the Philippine Tagalog script and Buginese script, Makassarese Bird script uses the vowel sign doubling abbreviation. This is illustrated in Figure 14: the first line of transcription shows the base value of each orthographic character and the second illustrates the base reading (without syllable final consonants) of the written characters.

| nu | p | l*ii* | k | | m | p | l | p *oo* | | r | o | | n | k | e | t | e | n | | y | d*uu* | m | k | o |
| nu | pa | *lili* | ka | | ma | pa | la | *popo* | | ro | | | na | ka | te | | ne | | | ya | *dudu* | ma | | ko |
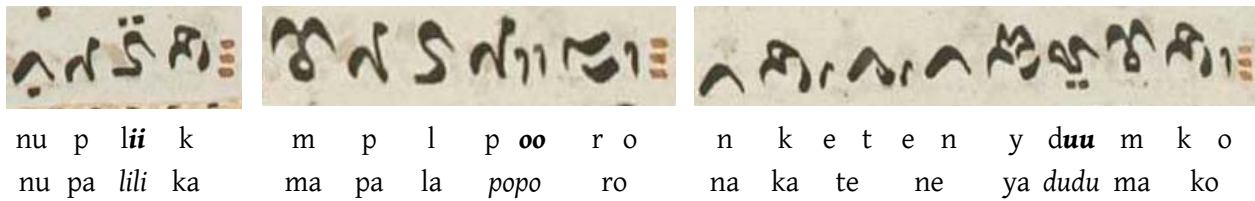
Figure 14. Vowel sign doubling in Bird script.[35]

Six examples of the apparent *angka* reduplication sign ⸲ from the Tropenmuseum 668-216 manuscript are illustrated in Figure 15. In the sixth example, the character bears an ‹-i› vowel sign (a dot above).

Figure 15. Occurrences of apparent *angka* (reduplication sign) in Bird script.[35]

Four textual punctuation signs are used in the Makassarese texts I have seen. Three of these are illustrated in Table 17. The fourth is an image of two palm trees found to my knowledge only in an 18th century manuscript illustrated in Jukes (2006), which appears to function as a variant glyph for a section end marker.

Value	*Character*	*Buginese equivalent*	*Buginese code point*
Divider (*passimbang*)	⁞	•∴ (*pallawa*)	1A1E
Section end marker	⁞⁞	◌̂	(1A1F)
End of text marker	قَمّت	(قَمّت)	

Table 17. Text markers in Makassarese Bird script.[35]

The divider or *passimbang* is directly equivalent to the Buginese *pallawa* and serves in principle, but variably, to separate words or phrasal segments. The section end marker is related in form but consists of six dots following the final letter in the section, unlike the Buginese character whose two segments

bracket the final letter from above and below. The third marker, used consistently in the texts in Tropenmuseum 668-216, is also found in Buginese texts illustrated in Gallop (1991) and Noorduyn (1993). It takes the form of the Arabic word تمّت *tammat* 'it is completed' but need not cooccur with any other Arabic text. In the Noorduyn text and the Makassarese examples, it has a decorative, elongated shape out of proportion to the rest of the text, and in the Chronicles of Gowa and Tallo' (Tropenmuseum 668-216), it is often illuminated in contrast with the rest of the text (Figure 16).

Figure 16. The *tammat* (تمّت) end of text marker in Makassarese Bird script.[35]

Due to its semantics as a stereotypical end of text marker, unlike elsewhere, where Arabic script is simply used to write segments of Arabic or Malay text, as well as the decorative aspect it takes on in order to set it apart from the text as a whole, this marker should be encoded as a distinct glyph as a part of Makassarese Bird script as well as of Buginese script.[36]

4.3.   The *Lontara' bilang-bilang* cipher script

A curious script, used in Buginese poetry, is described in Matthes (1883) and Tol (1992). This script is derived from an Arabic "abjad" cipher attested from British India (now Pakistan and Afghanistan) in Leitner (1882). The abjad cipher takes as its basis the numeric values of the Arabic letters in their old order similar to Phoenician, Aramaic and Hebrew.[37] Each letter in the cipher is replaced by the corresponding Hindu-Arabic numeral, which stops above the baseline for 1-9, touches the baseline from 10 to 99, crosses the baseline for 100 to 999, and crosses with an added curl below for 1000 and up.

The Sulawesi version, whose Buginese name means "counting" or "number" script, adds extensions for Bugis sounds absent in Arabic while discarding letters for Arabic sounds unused in Bugis. The first level of extensions uses derived letters added in Jawi Arabic script for specifically Malay sounds, and reassigns numeric values to certain ones of these in place of the dropped letters for Arabic sounds. Where a specific Jawi letter is derived from the corresponding Arabic one by addition of dots, the same number of dots are added to the respective base Hindu-Arabic numeral to form the corresponding character. In a second level of extensions, letters are derived for the prenasalized Buginese letters by adding a crossbar through the stem of the corresponding digit. The base characters of this script are illustrated in Table 18 together with their Buginese, Jawi and Hindu-Arabic equivalents.

---

[36] An analogous case is found in Arabic script, where the Arabic word سنة *sanah* is encoded as a distinct character ﺳﻨﺔ (0601 ARABIC SIGN SANAH), in the context where it is written underneath the digits for a year.

[37] Hence "abjad", derived from the first four letters: ا ب ج د ه و ز ح ط ى ك ل م ن س ع ف ص ق ر ش ت ث خ ذ ض ظ غ.

Value	Bugis letter	Lontara' bilang-bilang	Jawi	(Arabic base)	Numeric value
'			ا	ا	١
y			ي	ي	١٠
ny			ث	(ن) , ي	١٠
b			ب	ب	٢
k			ك	ك	٢٠
ngk					٢٠
g			ݢ	ك	٢٠
r			ر	ر	٢٠٠
nr					٢٠٠
j			ج	ج	٣
c			چ	ج	٣
nyc					٣
l			ل	ل	٣٠
d			د	د	٤
m			م	م	٤٠
t			ت	ت	٤٠٠
n			ن	ن	٥
w			و	و	٦
s			س	س	٦٠
ng			غ	ع	٧
p			ڤ	ف	٨
mp					٨

Table 18. *Lontara' bilang-bilang* cipher script.

Lontara' bilang-bilang has no special vowel signs and uses the same signs, in the same positions relative to the base letters, as in Buginese script (*cf.* Table 15). Text with vowel marks is illustrated in Figure 16.
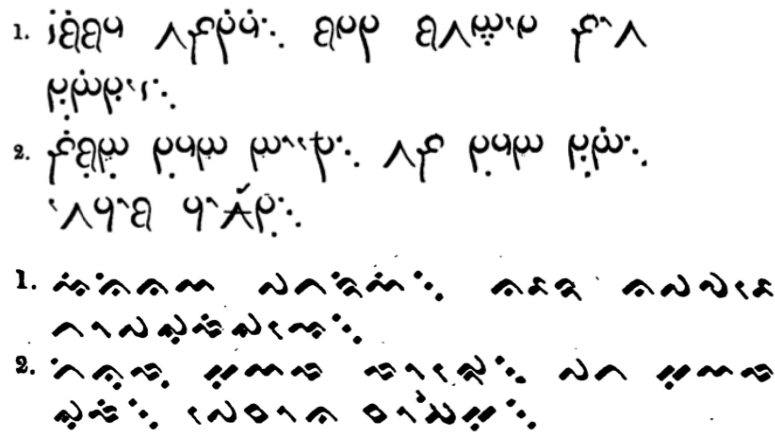
1. 'i ni na wa   pa ma ri wi   \   na ba ra   na pa ca _be   to pa gu li ga _'e   \
2. ti nu lu   ku wa la   lo _nre   \   ga ta   ku wa la gu li   \   _pe so na   so mpə ku   \

Figure 16. *Lontara' bilang-bilang* with vowel marks, standard Buginese script equivalent, and transcription of orthographic characters (Matthes 1883). The left hand position of the ‹-e› vowel sign is indicated by an underscore; the *pallawa* separator is indicated with a backslash.

Although this script is a cipher for the Buginese script, it had a specialized function for writing poetry, as described in Matthes and Tol, which may be an argument in favor of encoding it.

4.4.   Old Minahasa script

In general, the literature dealing with writing in Sulawesi has only spoken of indigenous phonographic scripts in the south of the island. For the rest of Sulawesi, including the Minahasa region of the northern peninsula, there is no clear evidence for the existence of writing. For example, Lapian (1987:104) speaks of an oral culture in Minahasa territories with no known written manuscripts, where the only known record of what may be writing consists of undeciphered petroglyphs carved into the famous landmark, the *Watu Pinawetengan* (Stone of Division). However, a recent posting on a blog about the Minahasa region in the eastern end of Sulawesi's northern peninsula (Talumewo 2009) has brought to light an alleged indigenous script described in a book by a local historian (Taulu 1980).

The book by Taulu is the only attestation of this script that I know of or have seen referred to anywhere. The book, whose title translates to "History of the founding of the Stone of Division and its charters", contains a short discussion of the alleged Minahasa script and numerous examples of texts in this script in a Minahasan language or languages, accompanied by Indonesian translations, whose content generally refers to pledges of unity between regional chiefs.

c↓ \ v→	-Ø	-a	-i	-u	-e	-o
Ø-						
k						
g						
ng						
t						
d						
n						
p						
b, w						
m						
s						
h						
r						
l						
y						
c						

Table 19. The alleged old Minahasa syllabic script.

The characters of this script are reproduced in Table 19 in drawings based on those in the table in Taulu, supplemented with other characters that were omitted from his table but found elsewhere in his example texts. Unfortunately, the available microfiche of the book was of very poor quality due to the already bad quality of the original printed typescript, so very little of the text was legible enough to extract much useful information.[38]

In the discussion that follows, it must be kept in mind that all characters appear to have been reproductions in Taulu's own hand and — on the assumption that they were drawn from authentic original documents — may not be fully accurate renditions of any original letter shapes. Apparently on the basis of theories claiming that the ancestors of the Minahasa migrated from Japan by way of the Philippines, Taulu compares the shapes of the Minahasa characters with their orthographic counterparts in Japanese kana and Tagalog script; however, his drawings of the Japanese and Tagalog

---

[38] My thanks go to Debbie Anderson for providing printouts of the microfiche text.

characters are very distorted and it seems justified to assume this may be the case for the Minahasa characters as well (if it is indeed the case that they are reproduced from authentic originals). To compound the difficulty, my redrawings of Taulu's drawings may be missing important details due to the poor quality of the microfiche.

There is enough information in the character shapes given by Taulu to make certain generalizations about the structure of the script. These concern the representation of vowels on consonant-based letters, the derivation of the ‹e› and ‹o› vowel letters from ‹i› and ‹u› bases, and the plausible origin of the letter shapes themselves.

First and most importantly, the character inventory Taulu gives departs from most Indic scripts by not including any distinct dependent vowel signs. Instead, apart from the expected series of independent (null onset) vowel letters, vowels are indicated by a variety of modifications to the shapes of the base consonant characters. This in itself is quite unusual: the only remotely similar Indic case I know of is the way Marathi Modi script and the Buhid and Hanunóo script varieties of Mindanao combine certain dependent vowel signs with their host letters by means of suppletive ligatures. Unlike these, there appears to be no clear, regular pattern to the way the base consonant glyphs are modified to indicate distinct vowels. This might however be due in part to inaccurate copying of the originals if such exist.

Second, the script contains independent vowel letters for ‹e› and ‹o›, which is relatively unusual for Indic scripts in general and especially for scripts of Indonesia and the Philippines, which normally only have independent letters for ‹a›, ‹i› and ‹u›. The ‹e› and ‹o› letter shapes are clearly mirror images of ‹i› and ‹u›, respectively.

Third, the glyphs are unlike those of the other Sulawesi scripts, but for the most part show enough convincing and regular structural relationships to 16th century Philippine handwriting variants for it to be relatively plausible that they are derived from Philippine models. It is worth noting here that the Philippine letter shapes Taulu used for comparison in his table are awkward copies of shapes specific to the print font used in the 1620 Ilocano *Doctrina Christiana*, which have become the most widely known exemplar of the old Tagalog script and are mistakenly believed to be typical of the script. Some correspondences between the typographic Philippine letter shapes and the Minahasa ones are apparent, but others become more apparent when the Minahasa shapes are compared to authentic Philippine handwriting letter shapes. This is significant in that Taulu appeared not to have access to Philippine script handwriting samples, which is not surprising considering that even at the beginning of the second decade of the 21st century, the sources in which these are available are still not widely known.

It is unclear at this point if this is an authentic script, especially given the lack of information about original texts. However, the structural evidence alone indicates quite strongly that it may be derived from a Philippine script variety. Despite the lack of a systematic explanation for the way the letter shapes change to represent different vowel readings, the letter shapes themselves show enough systematic structural similarities to older Philippine script varieties that there is good reason to suspect the script is not an outright invention. Pending further research, the question must remain open, but it seems prudent to leave open the possibility that the script may need to be encoded.

4.4.1. Summary

The Buginese script as currently encoded represents the core of characters that were retained for use in the standard printing fonts developed in the mid 19th century for the Bugis and Makassarese languages. The choice of characters made at that time may have represented what were currently the most widespread character variants in handwriting, but numerous other variants with historical importance were left out, not only those specific to palm leaf script style but also earlier variants found in numerous paper manuscripts. Most prominent among these are variants used in Luwu', but apart from these the end of text *tammat* sign found also in Makassarese Bird script merits consideration for inclusion, as well as several alternative earlier letter shapes not discussed as such above. These extensions number between 14 and 17 new code points depending on the decisions made.

The most important set of extensions to Buginese script are those needed for the *Lota* variant of the script used for the Ende language of Flores. These number between 7 and 14 or more, depending on eventual decisions on the status of particular characters, and include six letters completely distinct from the Buginese inventory and one virama sign not provided for in Buginese. Other possible extensions include characters that correspond to pre-standardization Buginese variants or to letters with reassigned values in Ende *Lota*.

Bimanese variant letters needed to represent prenasalized consonants in that language could in principle be subsumed under the same code points as Ende *Lota*, but their shapes are different enough to justify separate encoding. However, it is not yet clear what all the shapes are and further research needs to be done to determine the precise inventory of the Bimanese variety of Buginese script.

Among the Sulawesi region scripts, the Makassarese *Jangang-jangang* "Bird" script most clearly merits a new block of its own. Although it can be mapped directly onto the Buginese range as used to write the Makassarese language, it constituted a distinct national script for at least a century and possibly as many as three. It uses the same vowel signs as Buginese, but this is analogous to the identical shapes of Latin and Cyrillic ‹a›, ‹e›, ‹o›, and ‹y›, as well as ‹j› in the case of Serbian Cyrillic. This block requires 18 distinct letters, one reduplication sign, four vowel signs, and three textual punctuation signs.

Lontara' bilang-bilang is a cipher, but there is a possible justification for assigning it to a new block or extensions due to its specialized use in literature. (It would for the most part subsume the basic abjad cipher used elsewhere.)

Although it is not clear whether the alleged Old Minahasa letters published in Taulu (1980) are authentic representatives of a historical script, an eventual character block might need as many as 80 distinct code points for syllabograms representing the range of possible consonant plus vowel combinations. If further research succeeds in identifying authentic historical manuscripts in this script and in determining if the modifications to base letter shapes follow a systematic, regular pattern, the number of code points in an eventual proposal might be considerably smaller and more along the lines of those included in Buginese or the Philippine scripts.

Since no further evidence has surfaced of an ancient Bimanese script illustrated in Raffles (1817), it seems that no extra provision needs to be made for this script unless such evidence eventually surfaces.

5.   Cipher scripts

Apart from *Lontara' bilang-bilang*, other devised/cipher scripts are attested for the region: *Gangga Malayu* (Kern 1908)[39] and Javanese ciphers (Behrend 1996). Behrend describes the existence of several cipher scripts popular in Java up until the early 20th century, for which keys or "textual gateways" were published, but gives no details on their character inventories, their structure, or the domains in which they were used. The Javanese ciphers are undescribed as yet, but *Gangga Malayu*, *Lontara' bilang-bilang* and Makassarese Bird script all share the peculiar feature of combining letter shapes of exotic origin with a dependent vowel sign set from an existing locally used script.

6.   Related Indian scripts

The earliest attested script of Indonesia and the Philippines is the Pallava script that is found in 7th century inscriptions on sacrificial *yupa* posts in Kutai, eastern Borneo (Kern 1882, de Casparis 1975). Pallava is attested from inscriptions in various locations in island and mainland Southeast Asia, and although it is clearly related to Old Javanese/Kawi script and to the mainland Mon, Khmer and Cham scripts, I am not acquainted with any descriptions in the literature of the precise relationships. De Casparis describes Pallava as a monumental script compared to Kawi, whose forms are those of ordinary handwriting. However, this distinction is one of style or register and cannot plausibly be applied to whole scripts. It is clear that Pallava script originated in southeastern India and is closely related to the Kadamba and Cālukya scripts from which modern Telugu and Kannada scripts developed (Burnell 1878). Like Kawi and modern Javanese-Balinese scripts, it appears that Pallava likely contains a full complement of Brahmic letters, dependent conjuncts, vowel signs and numerals.

Both for their importance for the history of writing in southeast Asia and for the development of the Telugu-Kannada script group in India, Pallava and the closely related scripts deserve consideration for encoding as distinct scripts or as members of a single script block, subject to further research on their precise relationships.

More or less contemporaneous with Kadamba and Cālukya is the Gupta script of northern India, from which the modern North Indian scripts, including the Eastern and Western Nagari, Tibetan and Sharada groups, developed. As the first stage of development in the North of a script clearly distinct from Brahmi, Gupta merits consideration for distinct encoding in its own block.

---

[39] *Gangga Malayu* is the subject of a preliminary proposal (Pandey 2010).

7.   An extended Arabic-Indic numeral shape used in the Malay archipelago

A minor point concerns the shape of the numeral ‹5› in Jawi Arabic script in the Malay Archipelago region. Jawi numeral shapes are essentially similar to Eastern Arabic-Indic digits (06F0-06F9) except for ‹5›, whose preferred Jawi shape resembles a reversed 'B' or 'β', *cf. Lontara' bilang-bilang* (4.3). Another example of this, out of many, is illustrated in Figure 17.
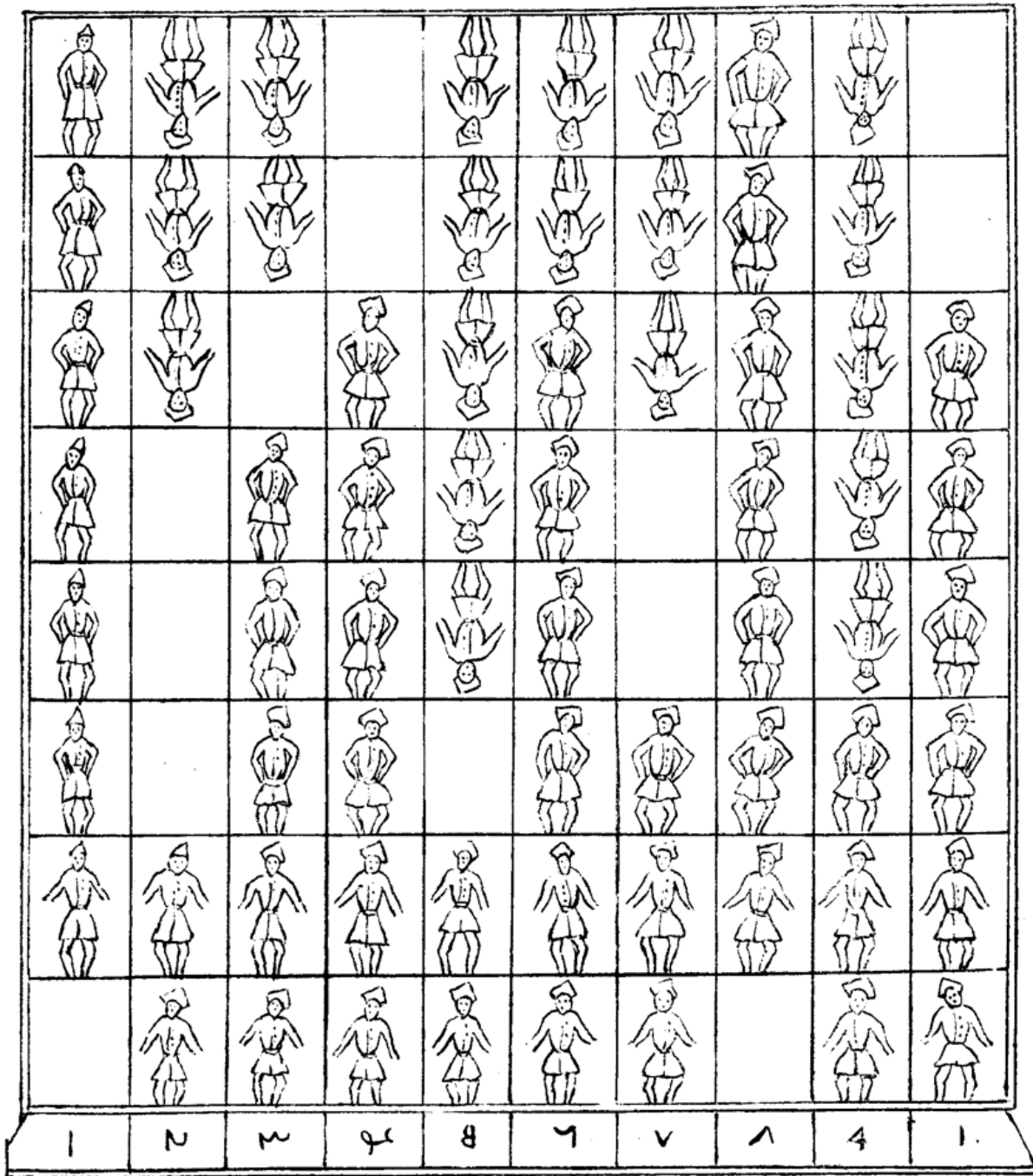

Figure 17. An example of Jawi Arabic numeral shapes, from Matthes (1868).

8.   Final summary

In most cases reviewed in this report, previously unrepresented regional script varieties share the majority of their characters with already encoded blocks and any unencoded characters can be treated as extensions to the existing Unicode blocks. In a very few cases, some extensions are required for new letters that have come into use in modern applications of the scripts or have proved to be needed to correctly represent the range of characters used in earlier versions of some scripts. These extensions would only add three new characters to the Tagalog script but would entail major additions of code points to the Rejang and Buginese blocks to support characters specific to related script varieties in their respective regions. Potentially the largest single set of extensions to be envisioned is the addition of the *Angka bejagung* numeral characters to the Rejang block.

Five sets of characters would require entirely new blocks as distinct scripts. These are the Eskayan syllabary, Makassarese Bird script, the Minahasa script, the Buginese *Lontara' bilang-bilang* cipher, and the Sumatran "Malayu" or "Post-Pallava" script group intermediate between Old Javanese/Kawi and modern Javanese-Balinese scripts. Of these, the clearest candidate for encoding in the short term is Makassarese Bird script, the national script of Makassar during the 17th century and at least for several decades before and after that time. The major effort required for an Eskayan script proposal involves cataloguing the characters of the syllabary, which apparently number in the region of 1000 and more. The Minahasa and Malayu scripts require further research to determine their precise status, and although it was specialized for a specific use in Buginese literature, the *Lontara' bilang-bilang* risks rejection on the basis that it is a cipher of the existing Buginese script.

To these can be added Pallava and the related Kadamba-Cālukya and Gupta scripts of India, which are historically important to the development of writing in India and its spread to continental and insular Southeast Asia. As is the case with the Malayu script group, further research is needed to determine their precise relationships and character inventories.

Two scripts of the Philippines are plausible candidates for distinct script status on the basis of unusual properties: the character set of the Calatagan Pot Inscription, which is partially related to Tagalog script but whose ultimate origin is unclear enough that the values of its characters are uncertain; and the modern Kapampangan script, which while using a subset of character variants that fall within the Tagalog range, uses a set of principles for combining characters and a directionality that are comparable to traditional Han'gul and utterly different from any other Philippine or Indonesian script.

Finally, although these issues are far less wide-reaching than the Kapampangan case, both the Kerinci and Lampung varieties of the South Sumatran script group clustered around Rejang make use of reordering rules that affect the placement of dependent vowel signs in specific contexts. These need to be taken into account in evaluating whether Kerinci and Lampung should be treated as subsets of/ extensions to the Rejang block or as distincts scripts in their own right.

9.   References

Anderbeck, Karl Ronald (2008). *Malay Dialects of the Batanghari River Basin (Jambi, Sumatra).* SIL International: SIL e-Books.

Anonymous (1615) Martínez de Rojo. (Two handwritten legal documents in Kapampangan obtained by Antoon Postma, Panaytayan, Mindoro Oriental, Philippines).

Banda, Maria Matildis (2005). *Aksara Ende: Flores Nusa Tenggara Timur.* Jakarta: Djambatan.

Behrend, Timothy E. (1996). Textual Gateways: The Javanese Manuscript Tradition. In Kumar and McGlynn (eds):161-200.

Burnell, Arthur Coke (1878). *Elements of South-Indian Palaeography from the Fourth to the Seventeenth Century A.D.* London: Trübner & Co.

Casparis, Johannes de (1975). *Indonesian Palaeography: A History of Writing in Indonesia from the Beginnings to C. A.D. 1500.* Leiden/Köln: E. J. Brill.

Chambert-Loir, Henri (1996). Bima on the Edge of Tradition. In Kumar and McGlynn (eds):74-75.

Cummings, William (2002). *Making Blood White: Historical Transformations in Early Modern Makassar.* Honolulu: University of Hawai'i Press.

De los Santos, Norman. (2010). Ten for the 10s. In *Baybayin Modern Fonts (Anak Bathala Project).* http://nordenx.blogspot.com/2010/01/ten-for-10s.html (accessed 2011-02-21).

Everson, Michael (2003). Revised final proposal for encoding the Lontara (Buginese) script in the UCS. Unicode Working group document ISO/IEC JTC1/SC2/WG2 N2633R, 2003-10-05. http://std.dkuug.dk/jtc1/sc2/wg2/docs/n2633r.pdf.

Everson, Michael and Uli Kozok (2008). On the inputting model for Batak. Unicode Working group document JTC1/SC2/WG2 N3489, 2008-08-04. http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3489.pdf.

Funke, Friedrich W. (1961). *Orang Abung: Volkstum Süd-Sumatras im Wandel.* Leiden: E. J. Brill.

Gallop, Annabel Teh (1991). *Golden Letters: Writing Traditions of Indonesia/Surat Emas: Budaya Tulis di Indonesia.* London: The British Library and Jakarta: Yayasan Lontar.

Guillermo, Ramon G. (2008). Ina Bisa Kata: An Experimental Decipherment of the Calatagan Pot Inscription. Manuscript, University of the Philippines, Diliman.

Guillermo, Ramon G. and Myfel Joseph D. Paluga (2011). *Barang King Banga:* A Visayan Language Reading of the Calatagan Pot Inscription (CPI). *Journal of Southeast Asian Studies* 42:121-159.

Holle, K. F. (1882). *Tabel van Oud- en Nieuw-Indische Alphabetten: Bijdrage tot de Palaeographie van Nederlandsch-Indië.* Batavia (Jakarta): W. Bruining & Co. and The Hague: M. Nijhoff.

Holle, K.F. (1999). Table of Old and New Indic Alphabets: Contribution to the Paleography of the Dutch Indies. *Written Language & Literacy*, Volume 2, Number 2:167-245(79).

Jaspan, Mervyn A. (1964). *Folk Literature of South Sumatra. Rejang Ka-Ga-Nga Texts.* Canberra: The Australian National University.

Jaspan, Mervyn A. (1967). Symbols at work: Aspects of kinetic and mnemonic representation in Redjang ritual. *Bijdragen tot de Taal-, Land- en Volkenkunde* 123 (1967), Number 4:476-516.

Jonker, J. C. G. (1896). *Bimaneesche Spraakkunst.* In *Verhandelingen van het Bataviaasch Genootschap van Kunsten en Wetenschappen, Deel XLVIII.* Batavia (Jakarta): Landsdrukkerij and The Hague: M. Nijhoff.

Jukes, Anthony (2006). *Makassarese (*Basa Mangkasara'*): A Description of an Austronesian Language of South Sulawesi.* PhD thesis, The University of Melbourne.

Kelly, Piers (2006). The classification of the Eskayan language of Bohol. A research report submitted to the National Commission on Indigenous Peoples, Bohol, The Philippines. http://www.scribd.com/doc/8137028/The-Classification-of-the-Eskayan-Language-of-Bohol (accessed 2011-02-09).

Kelly, Piers (2009). Towards a history of the Eskayan auxiliary language and script of Bohol, Philippines. Paper presented at the Eleventh International Conference on Austronesian Linguistics, 22–26 June, 2009, Aussois, France. http://www.vjf.cnrs.fr/11ical/data/11ical_Abstract-book_AlexF_rev.pdf#page=58

Kern, Hendrik (1882). Over de Opschriften uit Koetei in Verband met de Geschiedenis van het Schrift in den Indischen Archipel, *Verslagen en Mededeelingen der Koninklinke Akademie van Wetenschappen,* Afdeeling Letterkunde 2:182-203.

Kern, Hendrik (1917a [2009]). The Inscription Commemorating the Consecration of the Amoghapāśa Statue of Padang Candi (Central Sumatra); 1269 Śaka. Annotated translation by Uli Kozok and Eric van Reijn of Kern (1917), "De wij-inscriptie op het Amoghapāça-beeld van Padang Candi (Midden Sumatra); 1269 Çaka", *Verspreide Geschriften* Volume 7.

Kern, Hendrik (1917b [2009]). The Sanskrit-Inscription of King Ādityawarman at Kubu Rajo (Minangkabau; ±1300 Śaka). Annotated translation by Uli Kozok and Eric van Reijn of Kern (1917), Het Sanskrit-inschrift op den grafsteen van Vorst Ādityawarman te Kubur Raja (Menangkabau; 1300 Çaka), in *Verspreide Geschriften* Volume 7. (Previously published as Kern (1913), Grafsteenopschrift van Koeboer Radja, *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlands-Indië* Volume 67:401-404).

Kern, R. A. (1908). A Malay Cipher Alphabet. *Journal of the Royal Anthropological Institute of Great Britain and Ireland,* Volume 38:207-211.

Kozok, Uli (1996). Bark, Bones, and Bamboo: Batak Traditions of Sumatra. In Kumar and McGlynn (eds): 231-246.

Kozok, Uli. (2004a). A 14th Century Malay Manuscript from Kerinci. *Archipel* 67:37-55.

Kozok, Uli. (2004b). *The Tanjung Tanah Code of Law. The Oldest Extant Malay Manuscript.* Cambridge, England: CUP.

Kozok, Uli. (2006). *Kitab Undang-Undang Tanjung Tanah. Naskah Melayu yang Tertua.* Jakarta: Yayasan Naskah Nusantara and Yayasan Obor Indonesia.

Kozok, Uli. (no date). TT-ms Photographs. http://ulikozok.com/images/album/index.html. Web page linked to in *Naskah Tanjung Tanah. The TT Code of Law – Kitab Undang-Undang TT.* http://ulikozok.com/tanjung-tanah.html (accessed 2011-02-21).

Kumar, Ann and John McGlynn (editors) (1996). *Illuminations: The Writing Traditions of Indonesia.* Jakarta: The Lontar Foundation and New York/Tokyo: Weatherhill, Inc.

Lapian A.B. (1987). Point de vue sur l'historiographie de Minahasa. *Archipel* 34:103-118.

Leitner, G. W. (1882). *Linguistic Fragments Discovered in 1870, 1872 and 1879 Relating to the Dialect of the Magadds and Other Wandering Tribes, the Argots of Thieves, the Secret Trade-Dialects and Systems of Native Cryptography in Kabul, Kashmir and the Punjab.* Selection from the Records of the Punjab Government. Section I. Lahore: Punjab Government Civil Secretariat Press.

Mas, Sinibaldo (Sinibald) de (1843). *Informe sobre el Estado de las Islas Filipinas en 1842.* Madrid (no publisher).

Matthes, B. F. (1868). *De Makassaarsche en Boegineesche Kotika's.* Makassar: Brill.

Matthes, B. F. (1883). *Eenige Proeven van Boegineesche en Makassaarsche Poëzie.* The Hague: Martinus Nijhoff.

McDonnell, Bradley (2009). A preliminary description of Ende phonology. *Journal of the Southeast Asian Linguistics Society* 2:195-226.

Marcilla y Martín, Cipriano (1895). *Estudio de los Antiguos Alfabetos Filipinos.* Malabón: Tipo-litografía del Asilo de Huérfanos.

Mendoza de Leon, Bayani (1992). *Baybayin: Ancient Script of the Philippines: A Concise Manual.* Paramus, New Jersey: Bycynthium Treasures.

Miller, Christopher (in press). A Gujarati origin for Scripts of Sumatra, Sulawesi and the Philippines. To appear in the Proceedings of Berkeley Linguistics Society 36.

Miller, Christopher (in preparation). Post-Nagari scripts of Indonesia and the Philippines. Ms.

Mintz, Malcolm (2004). *Bikol Dictionary. Diksionáriong Bíkol.* Perth, Western Australia: Indonesian/Malay Texts and Resources.

Noorduyn, Jacobus (1991). The Manuscripts of the Makasarese Chronicle of Goa and Talloq: An evaluation. *Bijdragen tot de Taal-, Land- en Volkenkunde, Manuscripts of Indonesia* 147, Number 4:454-484.

Noorduyn, Jacobus (1993). Variation in the Bugis/Makasarese script. *Bijdragen tot de Taal-, Land- en Volkenkunde, Manuscripts of Indonesia* 144, Number 2/3:351-352.

Noorduyn, Jacobus and M. Salim (1988). Special Buginese characters in Luwu' manuscripts. *Bijdragen tot de Taal-, Land- en Volkenkunde, Manuscripts of Indonesia* 149, Number 3:533-570.

Oktora, Samuel and Khaerul Anwar. (2010). Ahli Waris Aksara Lota Ende. *Kompas.com* 2010-12-09, http://oase.kompas.com/read/2010/12/09/14575920/Ahli.Waris.Aksara.Lota.Ende (accessed 2011-02-18).

Oropilla y Fortich, Quintin (2008). *Deciphered Secrets: The Calatagan Pot Ancient Inscriptions.* Quezon City (no publisher).

Pandey, Anshuman (2010). Preliminary Proposal to Encode the Gangga Malayu Script. ISO/IEC JTC1/SC2/ WG2 N3798, 2010-05-06.

Pangilinan, Michael (2009). Assessing the current status of the "pre-Hispanic" Kapampangan script. Paper presented at the Eleventh International Conference on Austronesian Linguistics, 22–26 June, 2009, Aussois, France. http://www.vjf.cnrs.fr/11ical/data/ 11ical_Pangilinan_Assessing_the_current_status_of_the_Kapampangan.pdf

Pangilinan, Michael (n.d.) *Sulat Kapampangan.* Web site accessed 9 February 2011. http://www.siuala.com/read_blog.php?id=15

Pardo de Tavera, Trinidad Hermenegildo (1884). *Contribución para el Estudio de los Antiguos Alfabetos Filipinos.* Losana: Imprenta de Jaunin Hermanos.

Postma, Antoon (1992). The Laguna Copper-Plate Inscription: Text and Commentary. *Philippine Studies* Volume 40, Number 2:183-203.

Pusaka Kerinci web site. http://pusaka-kerinci.com/en/index.php?id=news (accessed 2011-02-10).

Raffles, T. S. (1817). *The History of Java.* London: Black, Parbury and Allen.

Rais Kamardi Datuk Panjang Simulie, Datuk [Distinction between given and family names and titles unclear] (1987). Ditemukan di Pariangan dan Sulit Air. Dua Macam Huruf Asli (Aksara) Minangkabau. *Limbago: Majalah Adat dan Kebudayaan Minangkabau* Number 5, 1987, reproduced online at http://sasdaminangkabau.wordpress.com/2010/05/12/aksara%C2%A0minangkabau/ (accessed 2011-02-10).

Ricklefs, M. C. (1976). Bantĕn and the Dutch in 1619: Six Early 'Pasar Malay' Letters. *Bulletin of the School of Oriental and African Studies, University of London,* Volume 39, Number 1:128-136.

Rinaldi, Ingki (2010). Huruf Asli Minangkabau Raib. *Kompas*, 5 September 2010 (accessed 2011-02-10). http://regional.kompas.com/read/2010/09/05/1951487/Huruf.Asli.Minangkabau.Raib-4.

Romuáldez, Norberto. (1914). *Tagbanwa Alphabet with some Reforms. Alfabeto Tagbanwa con Algunas Reformas.* Manila: Cultura Filipina.

Roos, S. (1877). Iets over Endeh. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 24:481-580.

Sango, Datoek Batuah (*alias* Hadji Datoek Batoeah) (no date). *Tambo Alam Minangkabau.* Payakumbuh, Indonesia: Limbago. Images reproduced online at http://aswilblog.wordpress.com/2010/03/29/ huruf-asli-minangkabau/ (accessed 2011-02-10).

Sani, Abd(ullah?) (1988). *Belajar menulis dan membaca tulisan Rejang.* (No publishing information).

Santamaría, Alberto (1938). El "Baybayin" en el Archivo de Santo Tomás. *Unitas* Volume XVI, Number 8:441-480.

Santiago, Luciano P. R. (2002). *Laying the Foundations: Kapampangan Pioneers in the Philippine Church, 1592-2001.* Angeles City, Philippines: Juan D. Nepomuceno Center for Kapampangan Studies, Holy Angel University.

Santos, Hector (1996). The Calatagan Pot. In *A Philippine Leaf* web site. http://www.bibingka.com/dahon/mystery/pot.htm (accessed 2011-02-23).

Santos, Hector (1996). The Eskaya Script. In *A Philippine Leaf* web site. http://www.bibingka.com/dahon/mystery/eskaya.htm (accessed 2011-02-09).

Sato, Takayuki K. (2000). Philippino [sic] characters (status report). ISO/IEC JTC/1 SC/2 WG/2 N2194 (2000-02-22).

Sturler, W. L. de (1843). *Proeve eener beschrijving van het gebied van Palembang, (Zuid-oostelijk gedeelte van Sumatra).* Leiden: E. J. Brill.

Stutterheim, W. F. (1936). A Malay *sha'ir* in Old-Sumatran Characters of 1380 A. D. *Acta Orientalia ediderunt Societates Orientales Batava Danica Norvegica* Volume XIV:268-279.

Talumewo, Bodewyn Grey (2009). Alfabeth Minahasa Purba. *Bangsa Minahasa* blog, 16 February 2009. http://bode-talumewo.blogspot.com/2009/02/galeri-foto-minahasa-alfabeth-minahasa.html (accessed 2011-02-26).

Taulu, Hersevien M. (1980). *Sejarah terciptanya Watu Pinawetengan dan piagam-piagamnya.* Manado: Yayasan Membangun.

Tol, Roger (1992). Fish food on a tree branch: Hidden meanings in Bugis poetry. *Bijdragen tot de Taal-, Land- en Volkenkunde* 148, Number 1:82-102.

Tol, Roger (2006). A separate Empire: Writings of South Sulawesi. In Kumar and McGlynn (eds):213-230.

Tol, Roger (2008). Rolled up Bugis Stories: A parakeet's song of an old marriage calendar. Paper presented at the 17th Biennial Conference of the Asian Studies Association of Australia, Melbourne, 1-3 July 2008.

Van Hasselt, A. L. (1881). *De Talen en Letterkunde van Midden-Sumatra.* In Veth, P. J. (1881). *Midden-Sumatra. Reizen en onderzoekingen der Sumatra-Expeditie, uitgerust door het Aardrijkskundig Genootschap, 1877–1879.* Part 3. *Volksbeschrijving en Taal.* Leiden: E. J. Brill.

Van der Tuuk, Hermannus Neubronner. (1868). *Les manuscrits lampongs, en possession de M. le Baron Sloet van de Beele (Ancien gouverneur-général des Indes Néerlandaises).* Leiden: Hooiberg et fils.

van Suchtelen, B.C.C.M.M. (1921). *Endeh (Flores). Mededeelingen van het Bureau voor de Bestuurszaken der Buitengewensten, bewerkt door het Encyclopædisch Bureau. Aflevering XXVI.* Weltevreden, Java: Papyrus.

Verzosa, Paul Rodriguez (1939). *Pangbansang Titik nang Pilipinas (Philippine National Writing).* Manila (no publisher).

Villamor, Ignacio (1922). *La Antigua Escritura Filipina. The Ancient Philippine Writing.* Manila: Tip. Pontificia del Colegio de Santo Tomás.

Villarroel, Fidel (2000). The Church and the Philippine Referendum of 1599. *Philippiniana Sacra* Volume 35, Number 3:89-128.

Voorhoeve, Petrus (1970). Kerintji documents. *Bijdragen tot de Taal-, Land- en Volkenkunde* 126 (1970), Number 4:369-399.

Voorhoeve, Petrus (1971). *Südsumatranische Handschriften.* Wiesbaden: Franz Steiner Verlag.

Wahono, Tri (editor) (2010). Lota Ende Terasing di Negeri Sendiri. *Kompas.com* 2010-12-09, http://oase.kompas.com/read/2010/12/09/13495362/Lota.Ende.Terasing.di.Negeri.Sendiri (accessed 2011-02-18).

Westenenk, Louis Constant (1922). Rèntjong-schrift. II. Beschreven hoorns in het landschap Krintji. *Tijdschrift voor Indische Taal-, Land-, en Volkenkunde* Deel 61.

Wikipedia article. Eskaya. http://en.wikipedia.org/wiki/Eskaya (accessed 2011-02-09).

Yudono, Jodhi (2011). Patenkan Segera Aksara "Incoung"!. *Kompas.com* 2011-02-04, http://oase.kompas.com/read/2011/02/04/18061925/Patenkan.Segera.Aksara.Incoung. (Accessed 2011-02-18).