Title: Preliminary Proposal to Encode the Makassarese Bird Script in Unicode
Author: Anshuman Pandey (pandey@umich.edu)
Date: 2015-06-24

1 Introduction

This is a preliminary proposal to encode the 'Makassarese Bird Script' in the Unicode standard. The script was described by Christopher Miller in Unicode Technical Note #35 "Indonesian and Philippine Scripts and Extensions" and recommended for encoding (2011: 43–46, 51). In an attempt to implement that recommendation, the present document provides a brief description of the script, a tentative character repertoire, character properties, and specimens of usage. Issues requiring resolution are enumerated in section 5. The proposal author seeks feedback from scholars regarding the information presented here. A formal proposal is forthcoming.

2 Background

The 'Makassarese Bird Script' is a historical writing system that was used in South Sulawesi, Indonesia for writing *basa mangkasara'* or Makassar (ISO 639-3: mak), a Malayo-Polynesian language currently spoken by 2.1 million people. The script was maintained for official purposes in the kingdoms of Makassar in the 17th century. It was used for writing a number of historical accounts, such as the "Chronicles of Gowa and Tallo'". Metal types were developed for the script in the 19th century. The script is known indigenously in Makassar as *ukiri' jangang-jangang* "bird letters" and in Bugis as *uki' manu'-manu'*.

The Makassarese Bird Script is one of two Indic scripts used for representing the Makassar language. The other is *lontara' beru* "new writing", which is known popularly as 'Bugis' or 'Buginese' (see figure 9). The Buginese script is also referred to as the Bugis-Makassar script on account of its usage for writing both the Bugis and Makassar languages. The character repertoire of the Makassarese Bird Script is similar to that of the Buginese script; however, it lacks letters for the pre-nasalized clusters /ŋka/, /ɲca/, /mpa/, /nra/ and the consonant /h/ that are found in Buginese. Similar to Buginese, the Makassarese Bird Script does not mark syllable codas. A comparison of the two script is given in figures 6–8. A folio showing usage of the two scripts in a single source is given in figure 2.

3 Script Details

3.1 Structure

The Makassarese Bird Script is an alphasyllabary that is written from left to right. It is based upon the Brahmi model and is related to various scripts of Indonesia and Philippines. The only independent vowel letter is MAKASSARESE LETTER A, which has the default value /a/, but also functions as a vowel carrier. Vowels are represented using dependent combining signs. These signs are written with the vowel carrier for expressing independent forms of vowels. Each consonant possesses the inherent vowel /a/. The inherent vowel is changed by applying a vowel sign to a consonant. There is no VIRAMA-like sign for silencing the inherent vowel. Vowel signs may occur to the left, right, above, and below a consonant letter. Two vowel signs may occur with a base letter.

The script has a system for abbreviating syllables and reduplicating onset consonants. Abbreviation of syllables is performed by doubling the vowel sign of a base consonant (see section 3.4). Reduplication of an onset consonant is marked using a placeholder, which also functions as a vowel carrier (see 3.3.3).

The structures of orthographic syllables in the Makassarese Bird Script are:

Vowel	Consonant
$\mathbf{V_{carrier}}$ [V_{sign}] [V_{sign}]	**C** [V_{sign}] [V_{sign}]
	$\mathbf{C_{placeholder}}$ [V_{sign}]

Various forms of punctuation are used (see section 3.3.4). Words are generally separated using spaces. Sentences are delimited using three vertical dots, text sections are marked using a triangle consisting of six dots, and end of text may be marked using a stylistic rendering of the Arabic word *tammat* 'it is complete'.

3.2 Encoding model

The chief complexity of the Makassarese Bird Script is the visual ordering of the vowel sign ◌ MAKASSARESE VOWEL SIGN E. Although the vowel represented by this sign is pronounced after a consonant, the sign is written before the consonant. The prepending behavior of MAKASSARESE VOWEL SIGN E is identical to that of the corresponding character in Buginese, ◌ U+1A19 BUGINESE VOWEL SIGN E. There are two possible models for managing such behavior:

- *Logical order* This approach follows the current model for Buginese. The VOWEL SIGN E would be encoded as a combining sign and it would be placed in its logical position after a base consonant in an encoded sequence, but it would be prepended to the base consonant in the visual output:

  <KA, ◌ VOWEL SIGN E> → ◌KA

  <KA, ◌ VOWEL SIGN E, KA, ◌ VOWEL SIGN E> → ◌KA◌KA

  Placing the vowel sign manually before the consonant would result in incorrect rendering:

  <◌ VOWEL SIGN E, KA> → ◌○KA

  <◌ VOWEL SIGN E, KA, ◌ VOWEL SIGN E, KA> → ◌○KA◌○KA

  This model depends upon proper support from a rendering engine to re-position the sign before the consonant in the visual output.

- *Visual order* This approach would require manual placement of the VOWEL SIGN E before the consonant in the encoded sequence. The sign would be encoded as a 'letter', not as a 'combining sign' because such signs must be placed after the base letter to which they attach. According to this model the so-called 'VOWEL MARK E' would be used in an encoded sequence as follows:

  < VOWEL MARK E, KA> → ◌KA

  < VOWEL MARK E, KA, ◌ VOWEL MARK E, KA> → ◌KA◌KA

Placing this vowel mark after the consonant letter would result in incorrect rendering:

<𑻞 KA, 𑻳 VOWEL MARK E> → 𑻞𑻳

<𑻞 KA, 𑻳 VOWEL MARK E, 𑻞 KA, 𑻳 VOWEL MARK E> → 𑻞𑻳𑻞𑻳

This model does not require support from a rendering engine.

Of the above, the logical model is considered the more advantageous and is adopted here. It enables the VOWEL SIGN E to be treated properly as a combining sign like the other vowel signs in the script, instead of as a letter. This model also provides for easier identification of syllables, searching, and collation. Additionally, the encoding for Buginese in Unicode is based upon the logical model. Given the relationship between the two scripts and the potential overlap of their user communities, it is practical that the model for Makassarese Bird Script be the same as that for Buginese.

3.3 Tentative repertoire

The script block is tentatively named 'Makassarese Bird Script', but the descriptor 'Makassarese' is used in character names. Character names are patterned upon names used for Buginese characters in Unicode. The ordering of letters also follows that of the Buginese block. The character repertore consists of 18 consonant letters, 4 combining vowel signs, 1 consonant reduplication sign, and 3 punctuation marks. Digits used in manuscripts resemble Latin and Arabic-Indic forms, but do not appear to be entirely distinctive. Representative glyphs for the proposed characters are based upon forms used in manuscripts.

3.3.1 Consonants

Eighteen consonant letters are proposed for encoding:

	Character name	Phonetic value
𑻞	MAKASSARESE LETTER KA	/k/
𑻠	MAKASSARESE LETTER GA	/g/
𑻡	MAKASSARESE LETTER NGA	/ŋ/
𑻢	MAKASSARESE LETTER PA	/p/
𑻣	MAKASSARESE LETTER BA	/b/
𑻤	MAKASSARESE LETTER MA	/m/
𑻥	MAKASSARESE LETTER TA	/t/
𑻦	MAKASSARESE LETTER DA	/d/
𑻧	MAKASSARESE LETTER NA	/n/
𑻨	MAKASSARESE LETTER CA	/t͡ʃ/
𑻩	MAKASSARESE LETTER JA	/d͡ʒ/

𑻠	MAKASSARESE LETTER NYA	/n/
𑻡	MAKASSARESE LETTER YA	/j/
𑻢	MAKASSARESE LETTER RA	/r/
𑻣	MAKASSARESE LETTER LA	/l/
𑻤	MAKASSARESE LETTER VA	/v/
𑻥	MAKASSARESE LETTER SA	/s/
𑻦	MAKASSARESE LETTER A	/a/, 0

Several glyphic variant forms of consonants are attested. Some are shown below:

	Regular	Variant
DA	𑻣	𑻧
RA	𑻢	𑻨

3.3.2 Vowel signs

Four combining vowel signs are proposed for encoding:

	Character name	Phonetic value
◌𑻩	MAKASSARESE VOWEL SIGN I	/i/
◌𑻪	MAKASSARESE VOWEL SIGN U	/u/
𑻮◌	MAKASSARESE VOWEL SIGN E	/e/
◌𑻯	MAKASSARESE VOWEL SIGN O	/o/

These signs are applied to consonants as follows:

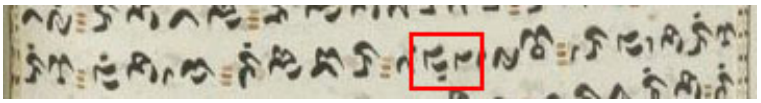
ka	𑻞	<𑻞 KA>
ki	𑻞𑻩	<𑻞 KA, ◌𑻩 VOWEL SIGN I>
ku	𑻞𑻪	<𑻞 KA, ◌𑻪 VOWEL SIGN U>
ke	𑻮𑻞	<𑻞 KA, 𑻮◌ VOWEL SIGN E>
ko	𑻞𑻯	<𑻞 KA, ◌𑻯 VOWEL SIGN O>

The 𑻳◌ VOWEL SIGN E is rendered before the consonant in the visual sequence, but it is placed after the base consonant in the encoded sequence, as shown above. The glyph reordering will be performed by the rendering engine.

3.3.3 Consonant reduplicator

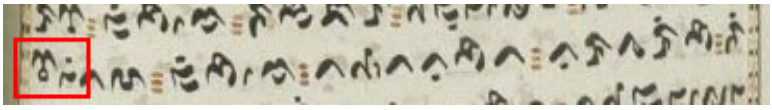
The 𑻴 MAKASSARESE ANGKA is used for reduplicating the onset consonant of the previous syllable. Its usage is based upon a convention opposite that of doubling vowel signs for syllable abbreviation (see section 3.4). As there is no sign or other means for marking the inherent vowel of a consonant, it is not possible to abbreviate two contiguous syllables consisting of identical consonants by doubling their vowel signs. Instead, the consonant following the onset is replaced with the ANGKA.

The usage of ANGKA is illustrated in the following examples. The boxed text in the excerpt below is 𑻵𑻴, which is the syllable 𑻵 <𑻲 RA, ◌𑻵 VOWEL SIGN U> followed by 𑻴 ANGKA:[1]



This text is to be read as 𑻵𑻲 *rura*. As shown, the ANGKA reduplicates the onset consonant RA of the previous syllable *ra*, but does not carry the acccompanying vowel *u*; it retains the inherent vowel *a*.

The ANGKA may also serve as a vowel carrier, as shown below. The boxed text shows 𑻰𑻴, which is the syllable 𑻰 MA followed by an 𑻴 ANGKA carrying the ◌𑻶 VOWEL SIGN I.



This text is to be read as 𑻰𑻰 *mami*. In this case, the two syllables have identical consonants, but only the second has a vowel sign.

The usage of ANGKA is based upon the notion of using the digit '2' for indicating repetition. The form of ANGKA is derived from ꧒ U+A9CF JAVANESE PANGRANGKEP, which is itself based upon ٢ U+0662 ARABIC-INDIC DIGIT TWO. A similar system of syllable reduplication is used in Buginese. However, there is no distinct ANGKA-type character for Buginese, which instead uses the Javanese PANGRANGKEP for this purpose.

As pairs of base letters and combining vowel signs belonging to different script blocks may complicate rendering, syllable identification, collation, and other processing, it may not be practical to use Javanese PANGRANKEP as a base letter in Makassarese contexts. For this reason, the ANGKA is proposed for encoding as a separate character for the Makassarese Bird Script.

3.3.4 Punctuation

Three punctuation signs are proposed for encoding:

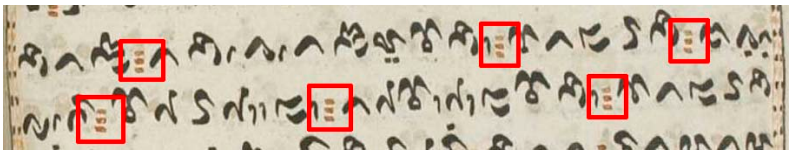
[1] Unless otherwise stated, all excerpts are from KIT 668-216 (see figure 1).

MAKASSARESE PASSIMBANG

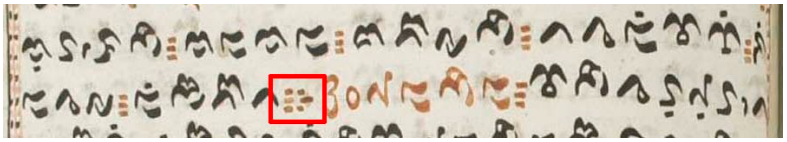
MAKASSARESE END OF SECTION

MAKASSARESE END OF TEXT

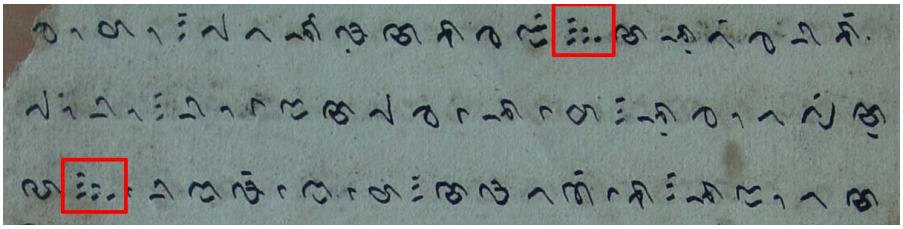
The Makassarese PASSIMBANG consists of three dots oriented in a vertical column. It is similar to U+1A1E BUGINESE PALLAWA.



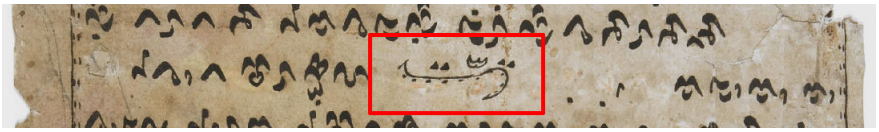
The Makassarese END OF SECTION consists of six dots oriented in the shape of a right-pointing triangle:



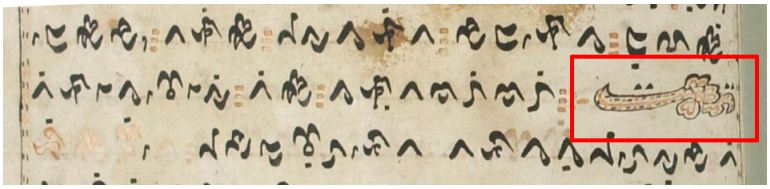
The dots in the END OF SECTION mark are also oriented in the form of a right triangle (TM Or545.232, reproduced in Jukes 2014):



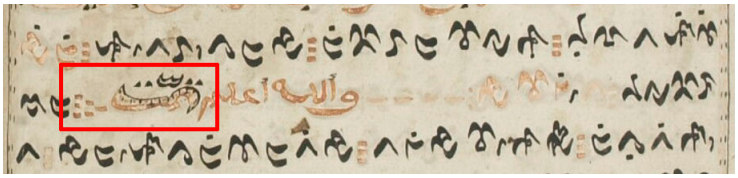
The Makassarese END OF TEXT is a stylized representation of the Arabic word تمّت *tammat* 'it is complete':



It is also written with decoration, as shown below:

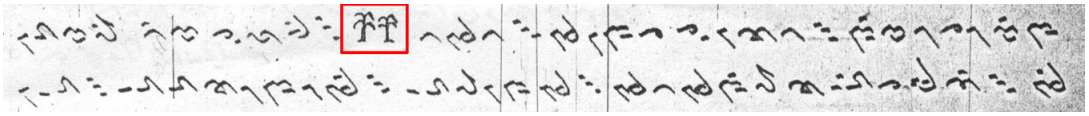


Here it follows the END OF SECTION mark:



Although the end-of-text marking word تمّت could be represented as a sequence of Arabic letters — ت U+062A ARABIC LETTER TA, م U+0645 ARABIC LETTER MEEM, ّ U+0651 ARABIC SHADDA, ت U+062A ARABIC LETTER TA — it is practical to treat it as an atomic character. Encoding it as a character will preserve its function as a mark of punctuation with appropriate character properties, which cannot be easily captured with a sequence of letters. This approach will also facilitate input of the character within the left-to-right environment of Makassarese and will avoid the need for switching to an Arabic script context.

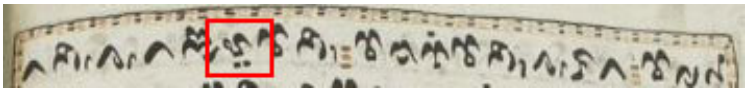
Another end-of-text marker is attested in a manuscript (microfilm at Australian National University) from the 1834–1858 that is written is a variant form of the Makassarese Bird Script (Jukes 2014: 5). It uses motifs resembling palm trees for marking sections:

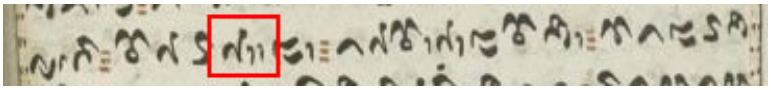


Additional research is required for determining the extent to which such motifs are used in the manuscript and in other sources.

3.4 Syllable abbreviation

Two contiguous and identical graphical syllables may be abbreviated by deleting the consonant of the second syllable and grouping its vowel sign with the first syllable, resulting in two vowel signs attached to a single base consonant. For example:





du·u	*dudu*
po·o	*popo*
li·i	*lili*

The abbreviated syllables shown above would be represented in encoded text as follows:

du·u <DA, VOWEL SIGN U, VOWEL SIGN U>

po·o <PA, VOWEL SIGN O, VOWEL SIGN O>

li·i <LA, VOWEL SIGN I, VOWEL SIGN I>

3.5 Multiple vowel signs

In order to accommodate this system of abbreviation, rendering engines should consider the contiguous occurrence of two of the same vowel sign as valid input. Moreover, the engine should provide appropriate spacing for sequences of a left-side vowel sign:

	Visual order	Logical encoded sequence
kake	𑻠𑻵𑻠	<𑻠 KA, 𑻵◌ VOWEL SIGN E, 𑻠 KA>
kake·e	𑻠𑻵𑻵𑻠	<𑻠 KA, 𑻵◌ VOWEL SIGN E, 𑻵◌ VOWEL SIGN E, 𑻠 KA>

If more than two vowel signs occur contiguously in an encoded sequence, then the additional signs should be displayed using a dotted circle:

<𑻠 KA, ◌̇ VOWEL SIGN I, ◌̇ VOWEL SIGN I, ◌̇ VOWEL SIGN I> → 𑻠◌̈◌̇

<𑻠 KA, 𑻵◌ VOWEL SIGN E, 𑻵◌ VOWEL SIGN E, 𑻵◌ VOWEL SIGN E> → 𑻵𑻵𑻠𑻵◌

Although the available sources do not show evidence of syllable abbreviation occuring with dissimilar vowel signs, sequences of such signs should be considered valid:

ku·i 𑻠◌̣◌̇ <𑻠 KA, ◌̣ VOWEL SIGN U, ◌̇ VOWEL SIGN I>

ko·e 𑻵𑻠𑻶 <𑻠 KA, ◌𑻶 VOWEL SIGN O, 𑻵◌ VOWEL SIGN E>

3.6 Digits

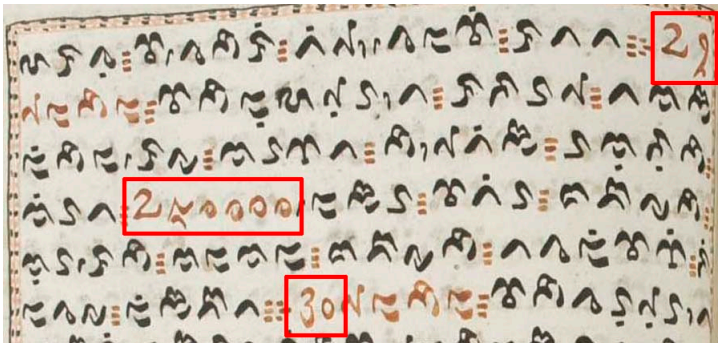
Digits resembling Latin and Arabic-Indic forms are attested in manuscripts. These are shown below:

	Latin-like	Arabic-like
zero		
one		
two		
three		
four		
five		
six		
seven		

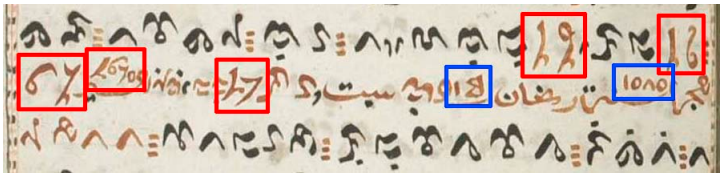
eight	
nine	

The shapes of ‘one’ and ‘nine’ differ from the corresponding Latin forms by the addition of a hook to the bottom right of the stem. This hook resembles that found in 𑴳 U+1B53 BALINESE DIGIT THREE. The first form of ‘five’ could be a modified version of ‘5’ in which the bottom curve is truncated, while the second form could be related to ৪ U+A9D4 JAVANESE DIGIT FOUR or ৫ U+A9D4 JAVANESE DIGIT FIVE, or a rotated form of ۵ U+06F5 EXTENDED ARABIC-INDIC DIGIT FIVE. The first form of ‘seven’ contains the same bottom hook as ‘one’ and ‘nine’, while the second form is nearly identical to ‘7’.

Numbers occur quite frequently in manuscripts. The excerpt below shows the numbers 29, 250000, 30 written using Latin-like digits:



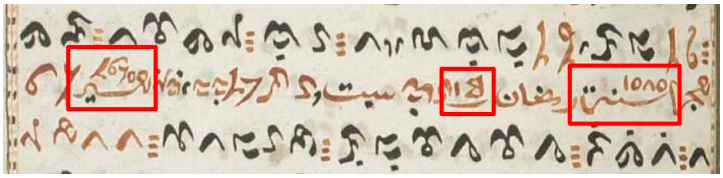
The following excerpt shows Latin-like digits in the numbers 19, 16, 67, 1670, and 17 (boxed in red), and Arabic-like digits in 15 and 1080 (boxed in blue):



: no we bu ru 19 ’e ra [30] 16
67 1670 hijîr [sic] pi bi re ru 17 ’a lo sabt bu la 15 ramaDân sanah 1080 hijr
pa ka na na :[2]

The numbers 1670, 15, and 1080 deserve further notice. They are written above what appear to be date and number signs:

[2] Transliteration courtesy of Christopher Miller.



The number 1670 represents the Gregorian year 1670 and is written above the Arabic text ہیر *hīr* (meaning is not known at present). The number 15 is written above a line that might be the ـۙـ U+0600 ARABIC NUMBER SIGN; however, the '5' does not resemble either ٥ U+0665 ARABIC-INDIC DIGIT FIVE or ۵ U+06F5 EXTENDED ARABIC-INDIC DIGIT FIVE. The number 1080 is written above the Arabic word سنة *sanah* (or a dotted form of ـۜـ U+0601 ARABIC SIGN SANAH and represents the Hijri year 1080. This number may be interpreted as Arabic-Indic ١٥٨٥, but what appears to be ٥ U+0665 ARABIC-INDIC DIGIT FIVE is actually the Latin-like '0'. The cases for 15 and 1080 suggest that numerical sequences might be written using digits from different script blocks.

Further research is needed for determining how to treat digits found in Makassarese manuscripts. Some forms, such as the Latin-like 'one' and 'nine', are distinctive enough to warrant separate encoding, but on the whole these forms could be unified with Latin digits 0..9. The Arabic-like forms could be unified with Arabic-Indic digits ٠..٩. These sets of digits should be specified as script extensions for the Makassarese Bird Script. The potential usage of non-Arabic-Indic digits with ـۙـ U+0600 ARABIC NUMBER SIGN and ـۜـ U+0601 ARABIC SIGN SANAH also needs to be better understood.

3.7 Linebreaking

Linebreaking generally occurs after an orthographic syllable; however there is the potential that syllables containing ◌ᨗ VOWEL SIGN E may be split across lines, such that the vowel sign remains the last character on the line and the consonant is written at the beginning of the next line. It is not clear at this time whether such occurrences should be considered normative or idiosyncratic. Hyphens or other marks indicating continuance are not used.

3.8 Collation

Collation for the Makassarese Bird Script follows the sort order for Buginese:

ᨀ KA < ᨁ GA < ᨂ NGA < ᨃ PA < ᨄ BA < ᨅ MA < ᨆ TA < ᨇ DA < ᨈ NA <
ᨉ CA < ᨊ JA < ᨋ NYA < ᨌ YA < ᨍ RA < ᨎ LA < ᨏ VA < ᨐ SA < ᨑ A <
◌ᨘ VOWEL SIGN I < ◌ᨙ VOWEL SIGN U < ◌ᨗ VOWEL SIGN E < ◌ᨚ VOWEL SIGN O

The sort order for ᨕ ANGKA needs to be determined. If possible, the ANGKA should be sorted using the same weight as for the consonant letter of the preceding syllable. In cases where two identical consonants occur alongside a sequence of the same consonant and ANGKA, then the sequence containing the ANGKA should be sorted after the sequence containing the two identical consonants:

ᨀᨀ *kaka*, ᨀᨕ *kaka*, ᨀᨗ *kaki*, ᨀᨕᨗ *kaki*, ᨀᨘ *kika*, ᨀᨗ *kiki*, ᨀᨕᨗ *kiki*, ᨀᨙ *kuka*, ᨀᨙ *kuku*,
ᨀᨕᨙ *kuku*, ᨁᨀᨁᨀ *keke*, ᨁᨀᨁᨕ *keke*, ᨀᨚᨀᨚ *koko*, ᨀᨚᨕᨚ *koko*

4 Tentative Character Data

4.1 Character Properties

Properties in the format of `UnicodeData.txt`:

```
11880;MAKASSARESE LETTER KA;Lo;0;L;;;;N;;;;;
11881;MAKASSARESE LETTER GA;Lo;0;L;;;;N;;;;;
11882;MAKASSARESE LETTER NGA;Lo;0;L;;;;N;;;;;
11883;MAKASSARESE LETTER PA;Lo;0;L;;;;N;;;;;
11884;MAKASSARESE LETTER BA;Lo;0;L;;;;N;;;;;
11885;MAKASSARESE LETTER MA;Lo;0;L;;;;N;;;;;
11886;MAKASSARESE LETTER TA;Lo;0;L;;;;N;;;;;
11887;MAKASSARESE LETTER DA;Lo;0;L;;;;N;;;;;
11888;MAKASSARESE LETTER NA;Lo;0;L;;;;N;;;;;
11889;MAKASSARESE LETTER CA;Lo;0;L;;;;N;;;;;
1188A;MAKASSARESE LETTER JA;Lo;0;L;;;;N;;;;;
1188B;MAKASSARESE LETTER NYA;Lo;0;L;;;;N;;;;;
1188C;MAKASSARESE LETTER YA;Lo;0;L;;;;N;;;;;
1188D;MAKASSARESE LETTER RA;Lo;0;L;;;;N;;;;;
1188E;MAKASSARESE LETTER LA;Lo;0;L;;;;N;;;;;
1188F;MAKASSARESE LETTER VA;Lo;0;L;;;;N;;;;;
11890;MAKASSARESE LETTER SA;Lo;0;L;;;;N;;;;;
11891;MAKASSARESE LETTER A;Lo;0;L;;;;N;;;;;
11892;MAKASSARESE VOWEL SIGN I;Mn;230;NSM;;;;N;;;;;
11893;MAKASSARESE VOWEL SIGN U;Mn;220;NSM;;;;N;;;;;
11894;MAKASSARESE VOWEL SIGN E;Mc;0;L;;;;N;;;;;
11895;MAKASSARESE VOWEL SIGN O;Mc;0;L;;;;N;;;;;
11896;MAKASSARESE ANGKA;Lo;0;L;;;;N;;;;;
11897;MAKASSARESE PASSIMBANG;Po;0;L;;;;N;;;;;
11898;MAKASSARESE END OF SECTION;Po;0;L;;;;N;;;;;
11899;MAKASSARESE END OF TEXT;Po;0;L;;;;N;;;;;
```

4.2 Linebreaking

Linebreaking properties in the format of `LineBreak.txt`:

```
11880..11891;AL   # Lo  [18] MAKASSARESE LETTER KA..MAKASSARESE LETTER A
11892..11895;CM   # Mn   [4] MAKASSARESE VOWEL SIGN I..MAKASSARESE VOWEL SIGN O
11896;AL          # Lo       MAKASSARESE ANGKA
11897..11899;AL   # Po   [3] MAKASSARESE PASSIMBANG..MAKASSARESE END OF TEXT
```

4.3 Syllabic Categories

Syllabic categories given in the format of `IndicSyllabicCategory.txt`:

```
# Indic_Syllabic_Category=Vowel_Dependent
11892..11893    ; Vowel_Dependent       # Mn  [2] MAKASSARESE VOWEL SIGN I..VOWEL SIGN U
11894..11895    ; Vowel_Dependent       # Mc  [2] MAKASSARESE VOWEL SIGN E..VOWEL SIGN O

# Indic_Syllabic_Category=Consonant
11880..11890    ; Consonant             # Lo [17] MAKASSARESE LETTER KA..LETTER SA

# Indic_Syllabic_Category=Vowel_Independent
11891           ; Vowel_Independent     # Lo       MAKASSARESE LETTER A

# Indic_Syllabic_Category=Consonant_Placeholder
11896           ; Consonant_Placeholder  # Lo       MAKASSARESE ANGKA
```

4.4 Positional Categories

Positioning data for combining signs in the format of `IndicPositionalCategory.txt`:

```
# Indic_Positional_Category=Right
11895          ; Right                  # Mc       MAKASSARESE VOWEL SIGN O

# Indic_Matra_Category=Left
11894          ; Left                   # Mc       MAKASSARESE VOWEL SIGN E

# Indic_Matra_Category=Top
11892          ; Top                    # Mn       MAKASSARESE VOWEL SIGN I

# Indic_Matra_Category=Bottom
11893          ; Bottom                 # Mn       MAKASSARESE VOWEL SIGN U
```

4.5 Script Extensions

The following characters should be extended for usage with the present script:

```
0660..0669     ;   # Nd  [10] ARABIC-INDIC DIGIT ZERO..ARABIC-INDIC DIGIT NINE
```

4.6 Confusables

```
11884 MAKASSARESE LETTER BA     ;  1A0E BUGINESE LETTER NYA
11888 MAKASSARESE LETTER NA     ;  1A08 BUGINESE LETTER TA
11892 MAKASSARESE VOWEL SIGN I  ;  1A17 BUGINESE VOWEL SIGN I
11893 MAKASSARESE VOWEL SIGN U  ;  1A18 BUGINESE VOWEL SIGN U
11894 MAKASSARESE VOWEL SIGN E  ;  1A19 BUGINESE VOWEL SIGN E
11895 MAKASSARESE VOWEL SIGN O  ;  1A1A BUGINESE VOWEL SIGN O
11896 MAKASSARESE ANGKA         ;  A9CF JAVANESE PANGRANGKEP
11897 MAKASSARESE PASSIMBANG    ;  1A1E BUGINESE PALLAWA
```

5 Outstanding Issues

The following issues must be resolved before a formal proposal can be produced:

- *Name* The script block is tentatively named ‘Makassarese Bird Script’ and the descriptor in character names is ‘Makassarese’. Is this acceptable? Is there a formal English designation for the script? Another possibility is ‘Makassarese’, which may be preferrable because the word ‘script’ is not used in Unicode identifiers (therefore ‘Makassarese Bird Script’ may be referred to as ‘Makassarese Bird’ in the standard, which is not entirely suitable, in my opinion). Other possibilities include ‘Makassar’ or ‘Mangakasar’. The indigenous name ‘Makassar Jangang-jangang’ is another option. In any case, the identifer for the script should be a name used in English and one that signifies its linguistic, cultural, or regional affiliation.

- *Encoding model* The logical model with vowel reordering is recommended here. Is there a preference instead for the visual model?

- *Digits* What is the most appropriate method to handle the digits found in manuscripts? Is usage of regular Latin digits suitable, despite the stylistic differences?

- *Punctuation* Are the palm tree motifs used throughout the manuscript in which they occur or only in the single instance shown here?

- *Linebreaking* Do examples exist of linebreaking where syllables containing VOWEL SIGN E are split across lines?
- *Collation* The collation for ANGKA needs to be specified.

6 References

Faulmann, Carl. 1880. *Das Buch der Schrift: Enthaltend die Schriftzeichen und Alphabete aller Zeiten und aller Völker der Erdkreises*. Zweite Vermehrte und verbesserte Auflage. Wein: Der Kaiserlich-Königlichen Hof- und Staatsdruckerei.

Holle, K. F. 1882. *Tabel van Oud- en Nieuw- Indische Alphabetten*. Bijdrage tot de palaeographie van Nederlandsch-Indië. Batavia: W. Bruining & Co.; 's Hage: M. Nijhoff.

Jukes, Anthony. 2014. "Writing and reading Makassarese". Presented at the International Workshop on Endangered Scripts of Island Southeast Asia, Tokyo University of Foreign Studies, February – March 2014.
`http://lingdy.aacore.jp/doc/endangered-scripts-issea/anthony_jukes_paper.pdf`

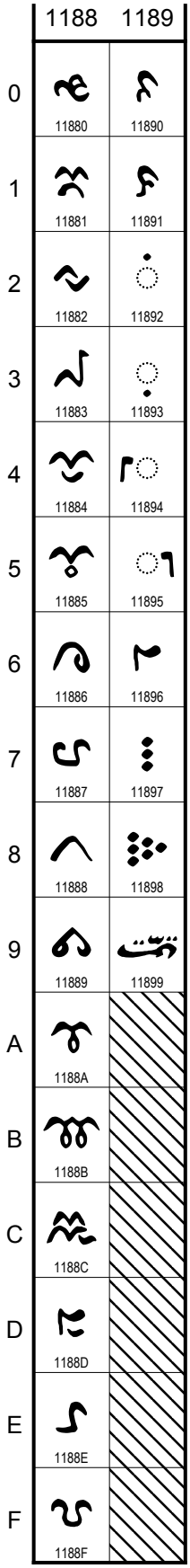
Miller, Christopher. 2011. "Unicode Technical Note #35: Indonesian and Philippine Scripts and Extensions". `http://www.unicode.org/notes/tn35/`

Raffles, Thomas S. 1817. "The History of Java", vol. 2. London: Black, Parbury, and Allen.

Rahman, Nurhayati. 2014. "Sejarah dan dinamika perkembangan huruf lontaraq di Sulawesi selatan". Presented at the International Workshop on Endangered Scripts of Island Southeast Asia, Tokyo University of Foreign Studies, February – March 2014.
`http://lingdy.aacore.jp/doc/endangered-scripts-issea/nurhayati_rahma_paper.pdf`

7 Acknowledgments

This proposal would not be possible without Christopher Miller, who graciously shared both his knowledge of the Makassarese Bird Script and source materials, and responded to my numerous questions with insight and patience.



This script is also known as Ukiri' Jangang-jangang.

Consonants

11880 MAKASSARESE LETTER KA
11881 MAKASSARESE LETTER GA
11882 MAKASSARESE LETTER NGA
11883 MAKASSARESE LETTER PA
11884 MAKASSARESE LETTER BA
11885 MAKASSARESE LETTER MA
11886 MAKASSARESE LETTER TA
11887 MAKASSARESE LETTER DA
11888 MAKASSARESE LETTER NA
11889 MAKASSARESE LETTER CA
1188A MAKASSARESE LETTER JA
1188B MAKASSARESE LETTER NYA
1188C MAKASSARESE LETTER YA
1188D MAKASSARESE LETTER RA
1188E MAKASSARESE LETTER LA
1188F MAKASSARESE LETTER VA
11890 MAKASSARESE LETTER SA
11891 MAKASSARESE LETTER A

Vowel signs

11892 MAKASSARESE VOWEL SIGN I
11893 MAKASSARESE VOWEL SIGN U
11894 MAKASSARESE VOWEL SIGN E
11895 MAKASSARESE VOWEL SIGN O

Syllable reduplicator

11896 MAKASSARESE ANGKA

Punctuation

11897 MAKASSARESE PASSIMBANG
11898 MAKASSARESE END OF SECTION
11899 MAKASSARESE END OF TEXT
= tammat

	Makassarese Bird Script	Buginese
KA		
GA		
NGA		
NGKA	—	
PA		
BA		
MA		
MPA	—	
TA		
DA		
NA		
NRA	—	
CA		
JA		
NYA		
NYCA	—	
YA		
RA		
LA		
VA		
SA		
A		
HA	—	

Table 6: Comparison of Makassarese Bird Script and Buginese consonants.

	Makassarese Bird Script	Buginese
VOWEL SIGN I	◌	◌
VOWEL SIGN U	◌	◌
VOWEL SIGN E	◌	◌
VOWEL SIGN O	◌	◌
VOWEL SIGN AE	—	◌

Table 7: Comparison of Makassarese Bird Script and Buginese vowel signs.

Makassarese Bird Script	Buginese
PASSIMBANG	PALLAWA
END OF SECTION	END OF SECTION
ANGKA	(U+A9CF JAVANESE PANGRANGKEP)

Table 8: Comparison of Makassarese Bird Script and Buginese punctuation and other characters.

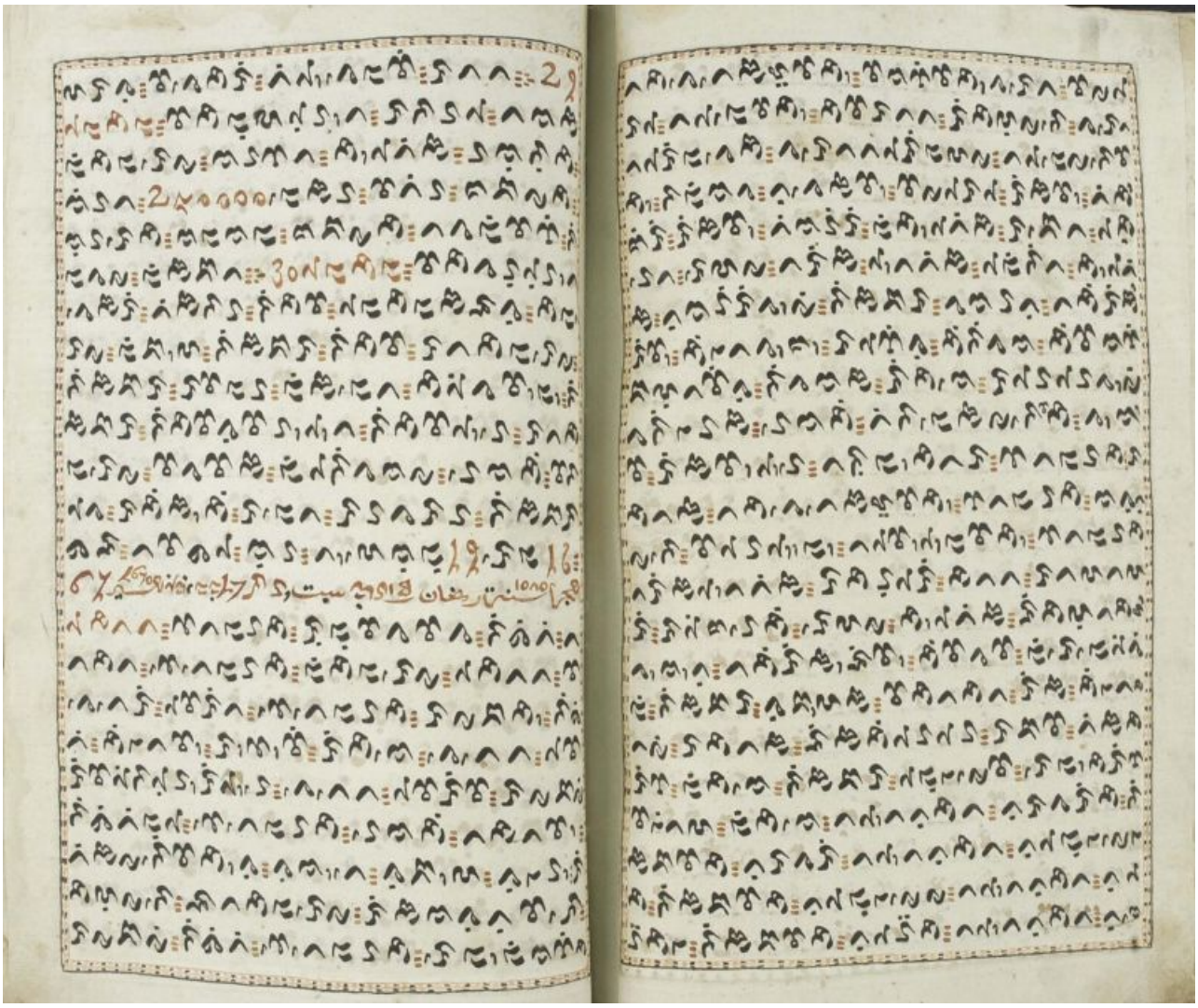


Figure 1: Excerpt from hand-written book in the Makassarese Bird Script (KIT 668-216). Image from WikiMedia Commons, provided by the Tropenmuseum of the Royal Tropical Institute (KIT). Source: `http://commons.wikimedia.org/wiki/File:COLLECTIE_TROPENMUSEUM_Gedeelte_van_het_dagboek_van_de_Vorsten_van_Gowa_in_oud_Makassaarschrift_TMnr_668-216.jpg`.

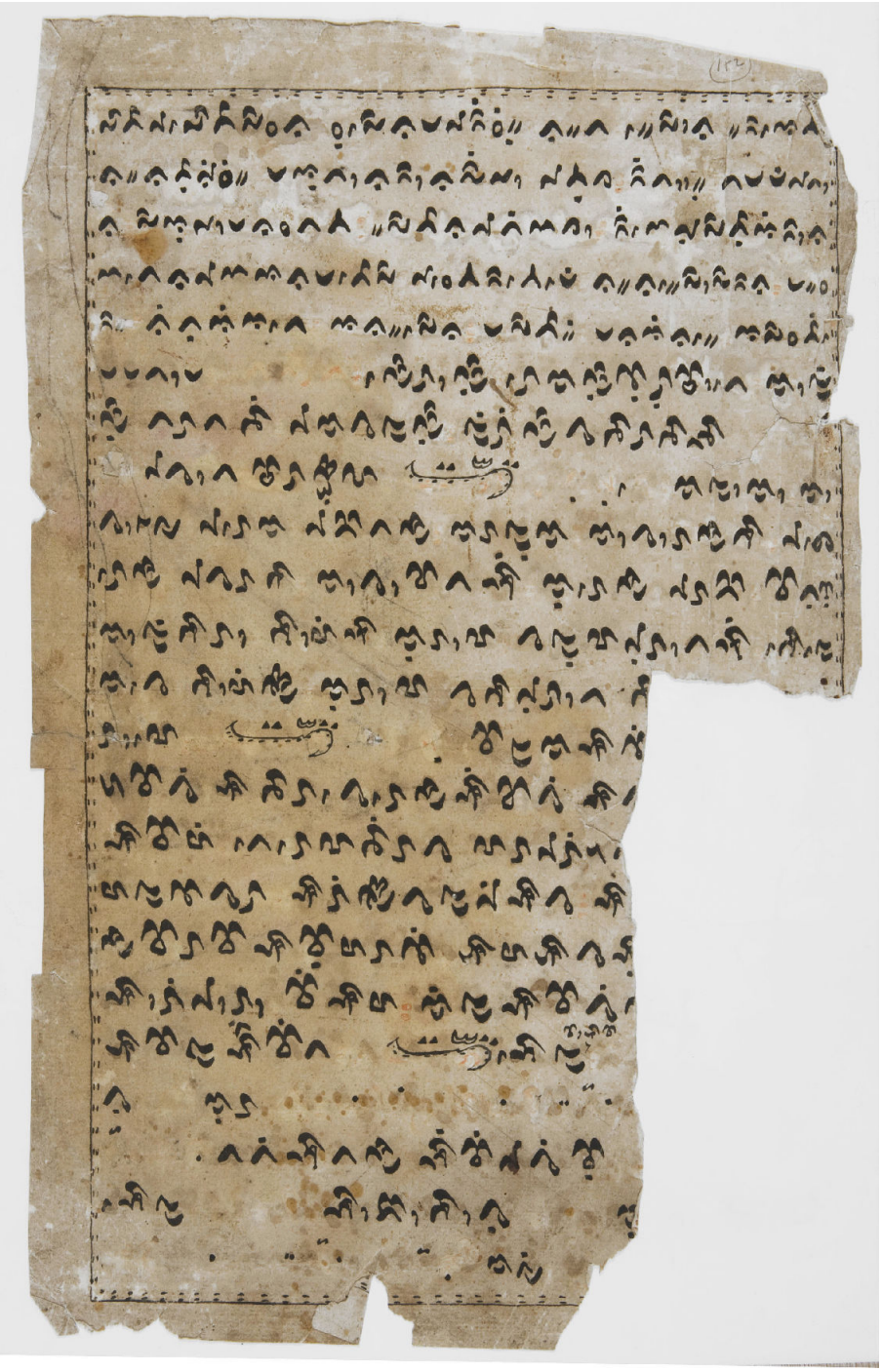


Figure 2: A folio containing text written in both the Buginese (first five lines and beginning of line six) and Makassarese Bird Script (Tropenmuseum 668-216 no. 119). Image courtesy of Christopher Miller.

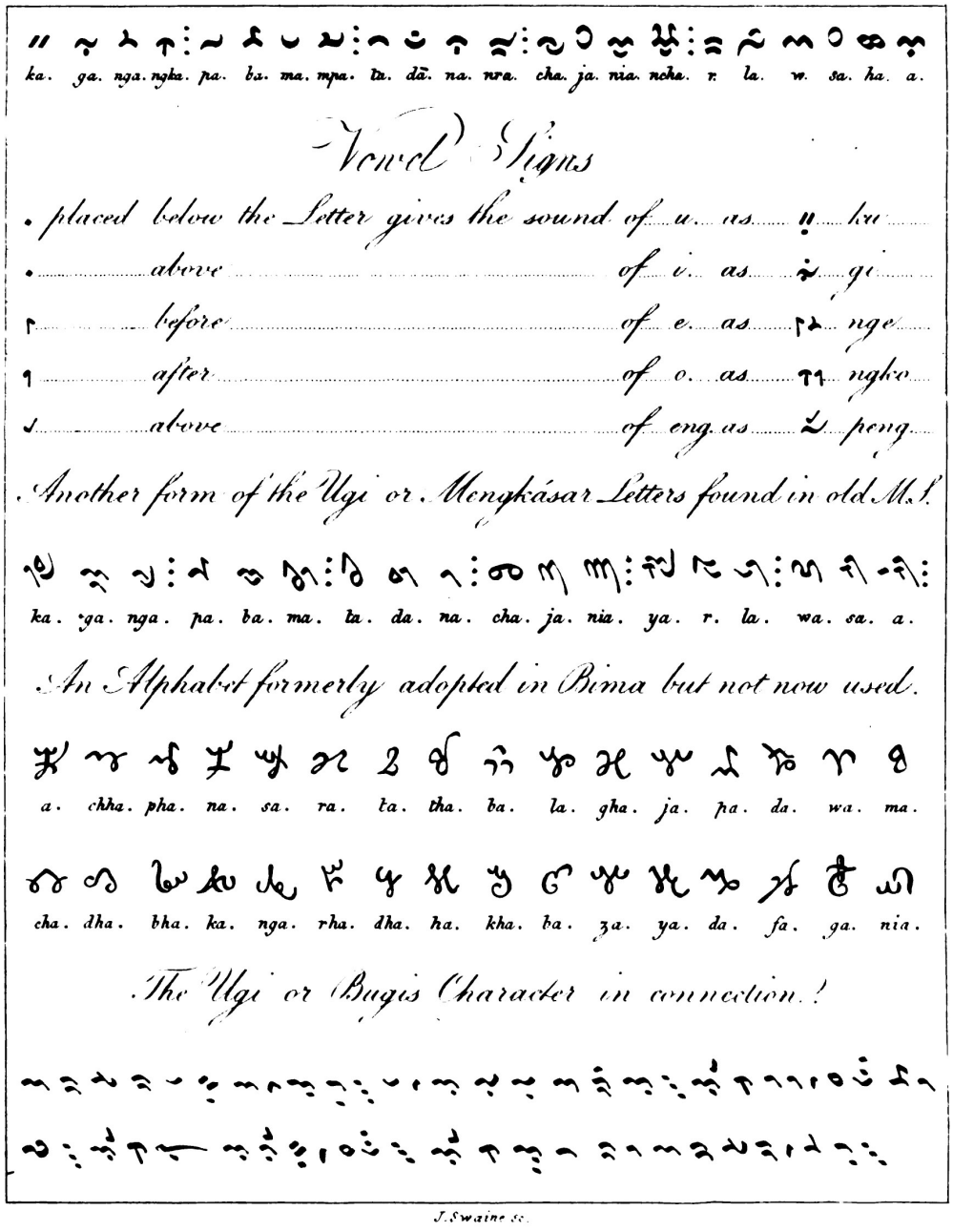
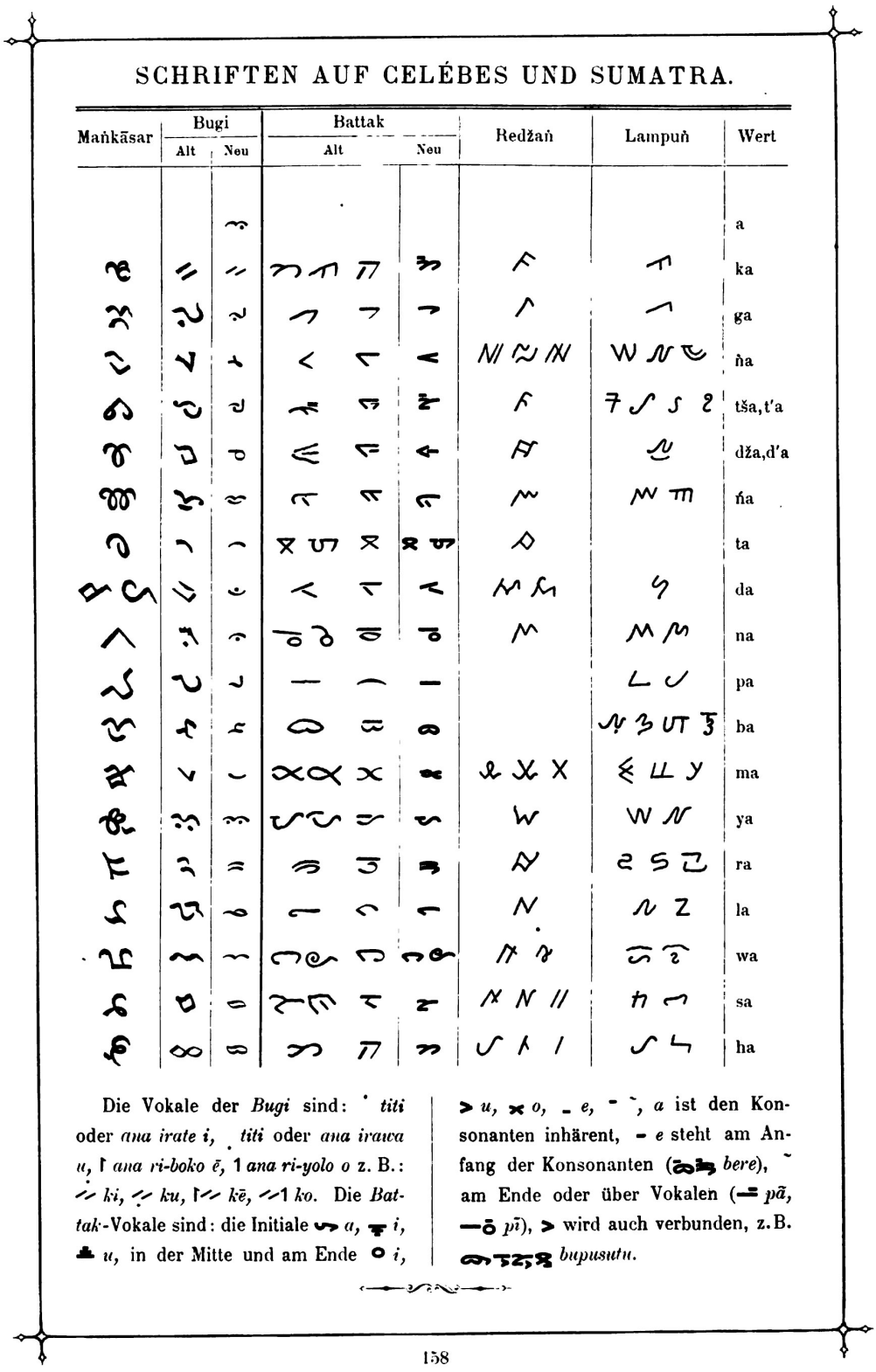


Figure 3: Chart showing Makassarese scripts (from Raffles 1817, plate after p. clxxxviii) The Makassarese Bird Script is shown under the heading "Another form of the Ugi or Mengkásar Letters found in old M. S.". The character repertoire shown here is identical to the proposed repertoire. Some glyph appear to be different, but the underlying graphical structure is evident.



SCHRIFTEN AUF CELÉBES UND SUMATRA.

Maṅkāsar	Bugi Alt	Bugi Neu	Battak Alt	Battak Neu	Redžaṅ	Lampuṅ	Wert
							a
							ka
							ga
							ña
							tša,t'a
							dža,d'a
							ńa
							ta
							da
							na
							pa
							ba
							ma
							ya
							ra
							la
							wa
							sa
							ha

Die Vokale der *Bugi* sind: ◌ *titi* oder *ana irate i*, ◌ *titi* oder *ana irawa u*, ◌ *ana ri-boko ē*, ◌1 *ana ri-yolo o* z. B.: ◌ *ki*, ◌ *ku*, ◌ *kē*, ◌1 *ko*. Die *Battak*-Vokale sind: die Initiale ◌ *a*, ◌ *i*, ◌ *u*, in der Mitte und am Ende ◌ *i*, ◌ *u*, ◌ *o*, ◌ *e*, ◌ ◌, *a* ist den Konsonanten inhärent, ◌ *e* steht am Anfang der Konsonanten (◌ *bere*), ◌ am Ende oder über Vokalen (◌ *pā*, ◌ *pī*), ◌ wird auch verbunden, z. B. ◌ *bupusutu*.

158

Figure 4: Chart showing Makassarese Bird Script consonants (from Faulmann 1880: 179). Faulmann erroneously equates 𑻠 MAKASSARESE LETTER A with ᜀ U+1A16 BUGINESE LETTER HA.

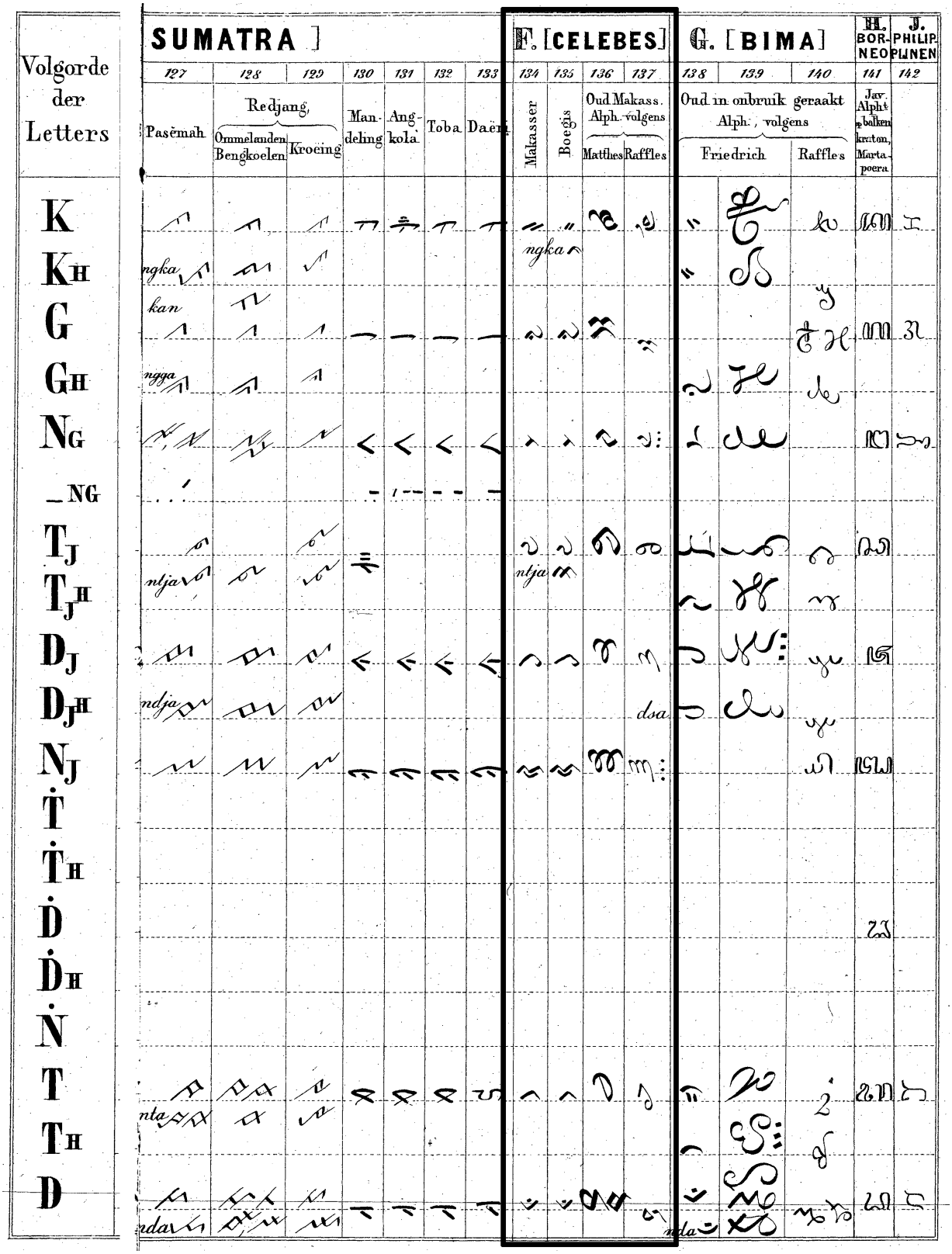


Figure 5: Chart showing scripts from "Celebes" or Sulawesi (from Holle 1882: 11) Columns 136 and 137 show the Makassarese Bird Script. The column showing transliteration ("Volgorde der Letters") has been stitched from the previous page in Holle.

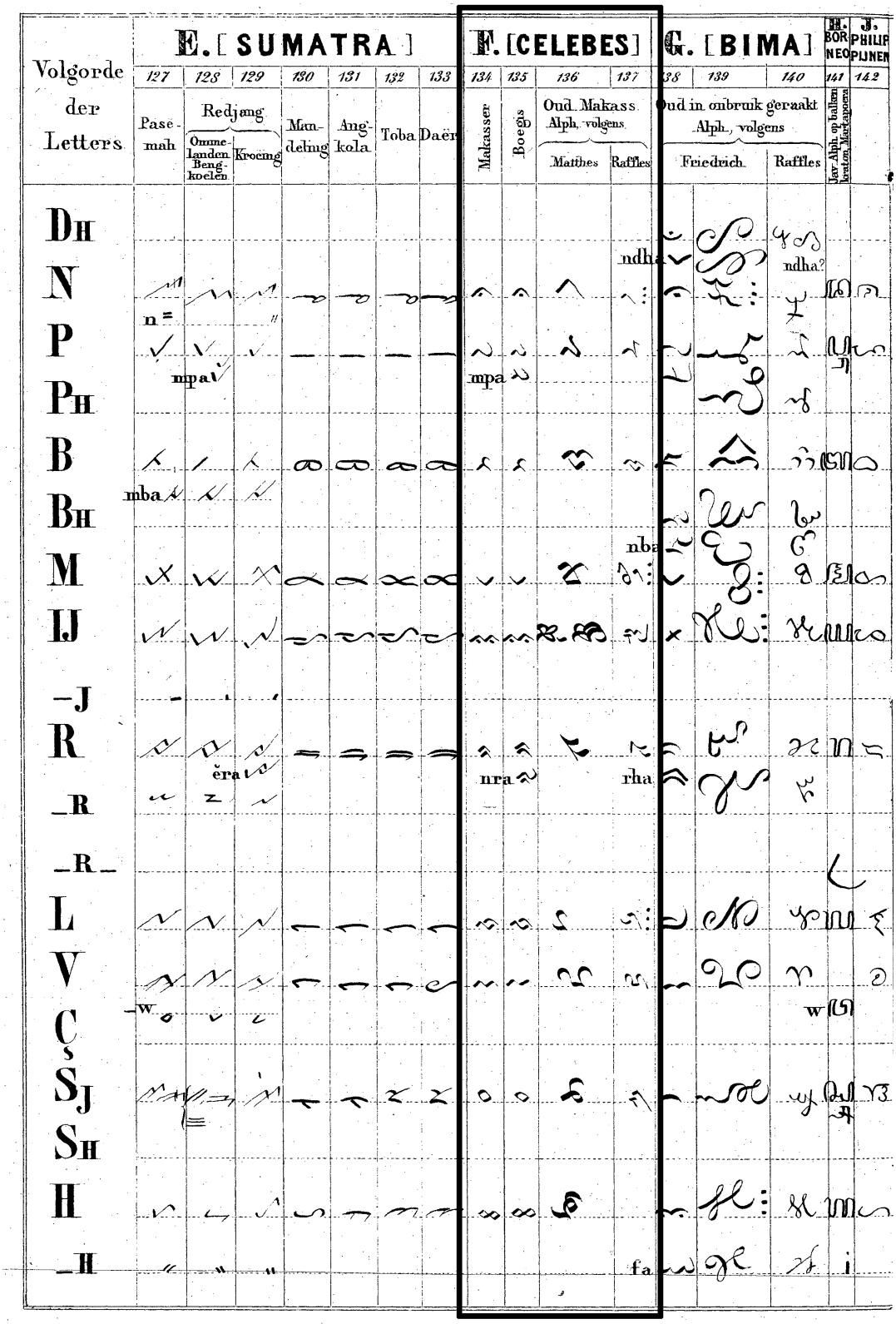


Figure 6: Chart showing scripts from "Celebes" or Sulawesi (from Holle 1882: 20). Columns 136 and 137 show the Makassarese Bird Script.

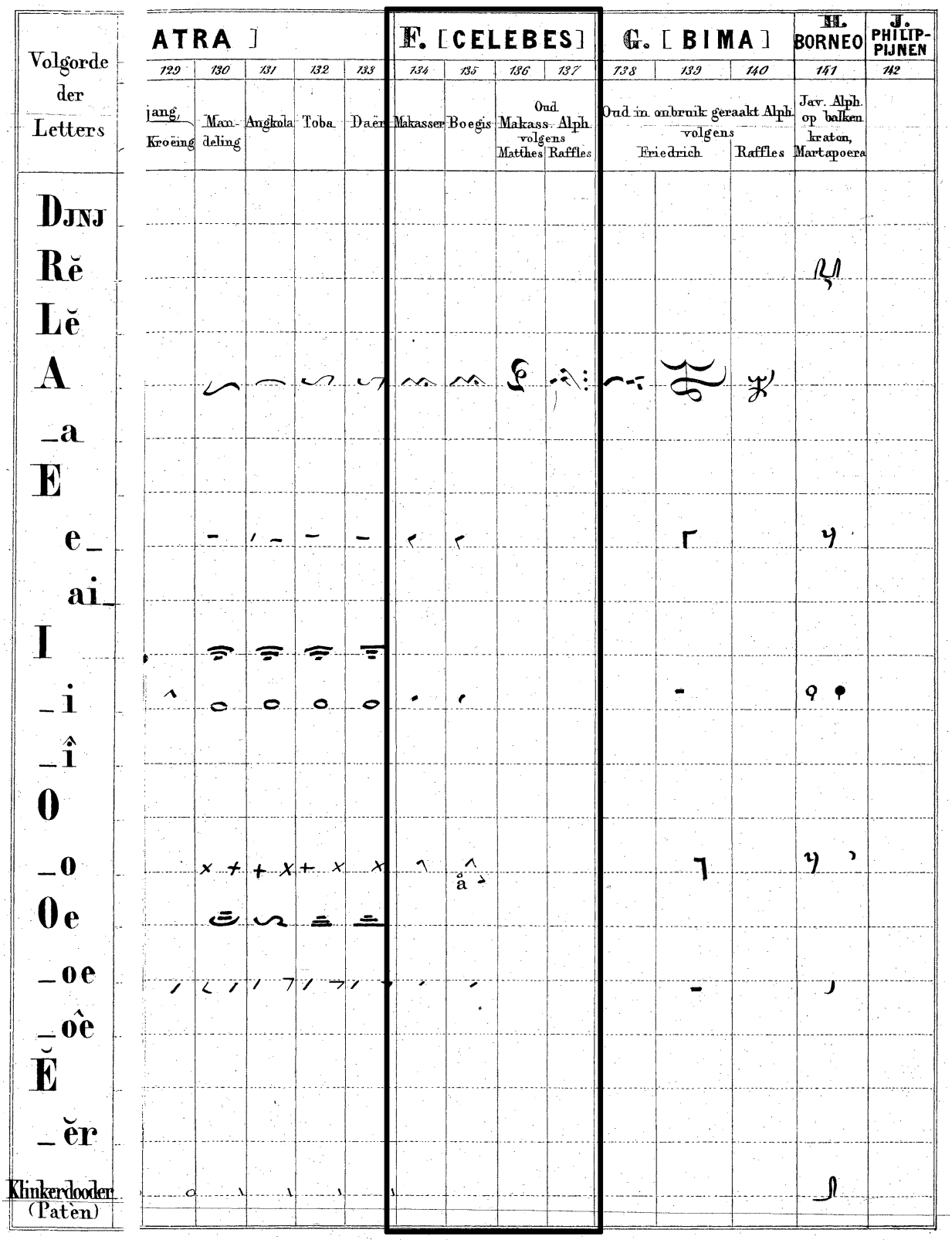


Figure 7: Chart showing scripts from "Celebes" or Sulawesi (from Holle 1882: 29). Columns 136 and 137 show the Makassarese Bird Script. The column showing transliteration ("Volgorde der Letters") has been stitched from the previous page in Holle.

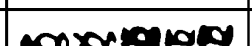
Value	*Bird script*	*South Sumatran*	*Javanese*			*Buginese*
			Early 17th century	*Modern. Balinese*	*Modern Javanese*	
k						
g						
ng						
c						
j						
ny						
t						
d						
n						
p						
b						
m						
s						
l						
r						
y						
w						
ʔ						
						

Figure 8: Chart showing Makassarese Bird Script and related scripts (from Miller 2011: 44).

Figure 9: Charts showing "Aksara Lontara Toa jangang-jangang" = Makassarese Bird Script (left), "Aksara Lontara Baru" = Buginese (center), and "Aksara Lontara Bilang-bilang" = Counting Script (right). From a display at Balla Lompoa Museum, Sungguminasa, Gowa. Image from WikiMedia Commons, provided by Sandjaja Kosasih (User:Sanko). Source: `http://commons.wikimedia.org/wiki/File:Lontara_script.jpg`.