

## **UC Berkeley**

### **Proceedings of the Annual Meeting of the Berkeley Linguistics Society**

#### **Title**

A Gujarati Origin for Scripts of Sumatra, Sulawesi and the Philippines

#### **Permalink**

<https://escholarship.org/uc/item/4q6906nz>

#### **Journal**

Proceedings of the Annual Meeting of the Berkeley Linguistics Society, 36(36)

#### **ISSN**

2377-1666

#### **Author**

Miller, Christopher

#### **Publication Date**

2016

Peer reviewed

PROCEEDINGS OF THE THIRTY SIXTH ANNUAL MEETING OF THE  
**BERKELEY LINGUISTICS SOCIETY**

February 6-7, 2010

**General Session**

**Special Session**

Language Isolates and Orphans

**Parasession**

Writing Systems and Orthography

**Editors**

Nicholas Rolle  
Jeremy Steffman  
John Sylak-Glassman

Berkeley Linguistics Society  
Berkeley, CA, USA

Berkeley Linguistics Society  
University of California, Berkeley  
Department of Linguistics  
1203 Dwinelle Hall  
Berkeley, CA 94720-2650  
USA

All papers copyright © 2016 by the Berkeley Linguistics Society, Inc.

All rights reserved.

ISSN: 0363-2946

LCCN: 76-640143

# Contents

<b>Acknowledgments</b>	<b>v</b>
<b>Foreword</b>	<b>vii</b>
<i>Basque Genitive Case and Multiple Checking</i> Xabier Artiagoitia .....	1
<i>Language Isolates and Their History, or, What's Weird, Anyway?</i> Lyle Campbell .....	16
<i>Putting and Taking Events in Mandarin Chinese</i> Jidong Chen .....	32
<i>Orthography Shapes Semantic and Phonological Activation in Reading</i> Hui-Wen Cheng and Catherine L. Caldwell-Harris .....	46
<i>Writing in the World and Linguistics</i> Peter T. Daniels .....	61
<i>When is Orthography Not Just Orthography? The Case of the Novgorod Birchbark Letters</i> Andrew Dombrowski .....	91
<i>Gesture-to-Speech Mismatch in the Construction of Problem Solving Insight</i> J.T.E. Elms .....	101
<i>Semantically-Oriented Vowel Reduction in an Amazonian Language</i> Caleb Everett .....	116
<i>Universals in the Visual-Kinesthetic Modality: Politeness Marking Features in Japanese Sign Language (JSL)</i> Johnny George .....	129
<i>Equative and Predicational Copulas in Thai</i> Nancy Hedberg and David Potter .....	144

<i>On the Reflexive-Antipassive Polysemy: Typological Convergence from Unrelated Languages</i> Katarzyna Janic .....	158
<i>Position and Height Asymmetries in Hiatus Resolution: A Case Study of Korean VV Sequences</i> Hijo Kang .....	174
<i>Negative Concord in Western Armenian</i> Hrayr Khanjian .....	188
<i>Emergent Hidden Grammar: Stochastic Patterning in Korean Accentuation of Novel Words</i> Hyun-Ju Kim .....	203
<i>Evidentiality in Korean Conditional Constructions</i> Iksoo Kwon .....	218
<i>The Source-Goal Asymmetry in SLA</i> Wojciech Lewandowski .....	233
<i>Subject Relatives and Expletives in Early New High German</i> Caitlin Light .....	247
<i>An Embodied Account of Argument Structure Development</i> Josita Maouene, Nitya Sethuraman, Mounir Maouene, and Linda B. Smith .....	261
<i>A Gujarati Origin for Scripts of Sumatra, Sulawesi and the Philippines</i> Christopher Miller .....	276
<i>A Quantitative Analysis of Nominative/Genitive Alternation in Japanese</i> Satoshi Nambu .....	292
<i>A Sibling Precedence Approach to the Linearization of Multiple Dominance Structures</i> David Potter .....	307
<i>Surface Faithfulness Phenomena and the Consonantal Root in the Modern Hebrew Verb System</i> Tom Recht .....	322
<i>Partial Wh-Movement and Wh-Copying in Dutch: Evidence for an Indirect Dependency Approach</i> Ankelien Schippers .....	338
<i>Computational Cognitive Morphosemantics: Modeling Morphological Compositionality in Hebrew Verbs with Embodied Construction Grammar</i> Nathan Schneider .....	353

<i>Some Hypotheses About Possible Isolates within the Atlantic Branch of the Niger-Congo Phylum</i> Guillaume Segerer .....	368
<i>From Relativization to Nominalization and More: Evidence from the History of Okinawan</i> Reijirou Shibasaki .....	382
<i>A Cross-linguistic Study of Sound Symbolism: The Images of Size</i> Kazuko Shinohara and Shigeto Kawahara .....	396
<i>Testing for Frequency and Structural Effects in an English Stress Shift</i> Morgan Sonderegger .....	411
<i>Neighborhood Density in Phonological Alternations</i> Sverre Stausland Johnsen .....	426
<i>Person Indexicals in Uyghur Indexical Shifting</i> Yasutada Sudo .....	441
<i>Metathesis and Reanalysis in Ket</i> Edward Vajda .....	457
<i>An Empirical Investigation of Typicality and Uniqueness Effects on Article Choice in Attributive-Possession NPs</i> Gregory Ward, Christopher Ahern, and Tom Hayden .....	472
<i>Perception of Illegal Contrasts: Japanese Adaptations of Korean Coda Obstruents</i> James D. Y. Whang .....	488
<i>Diglossia versus Register: Discursive Classifications of Two Sinhala Varieties</i> Cala Zubair .....	499



# Acknowledgments

The editors of the 36th Annual Meeting of the Berkeley Linguistics Society are grateful to conference participants, our volunteers, session chairs, and the faculty, all of whom made the event an intellectually stimulating and enriching event. We would like to extend our sincere gratitude to the contributors of this volume for their professionalism, responsiveness, attention to detail, and patience in the editorial process, without which this would not have been possible.

Within our department, special thanks go to Paula Floro and Belén Flores for all their support with BLS 36 and this annual conference in general. The editors wish to also thank the executive committee of BLS 36 who organized and ran the conference, and Zachary O'Hagan for expertise in the final compilation using LaTeX.

Finally, we would like to thank the following organizations for their generous financial support:

Department of Linguistics  
Graduate Assembly  
Social Sciences Division  
Student Opportunity Fund  
International Computer Science Institute (ICSI)





# Foreword

This monograph contains 34 of the 51 talks given at the 36th Annual Meeting of the Berkeley Linguistics Society (BLS 36), held in Berkeley, California, February 6-7, 2010. The conference included a General Session, one Special Session entitled *Language Isolates and Orphans*, and one Parasession entitled *Writing Systems and Orthography*. It was planned and run by the second-year graduate students in the Department of Linguistics at the University of California, Berkeley. The members of this executive committee were Jessica Cleary-Kemp, Clara Cohen, Stephanie Farmer, Melinda Fricke, Laura Kassner, and John Sylak-Glassman.

The papers contained herein were edited principally for style by the three editors Nicholas Rolle, Jeremy Steffman, and John Sylak-Glassman, and then given back to contributors to make changes. Nicholas Rolle took upon primary editorial responsibilities, Jeremy Steffman was an undergraduate editorial assistant, and John Sylak-Glassman helped to edit papers. Upon the final resubmission, the final versions of these papers were incorporated by Zachary O'Hagan and Nicholas Rolle into the monograph found here. Our goal has been the speedy publication of these proceedings, and as such, certain aspects – e.g., the complete unification of formatting – have been sacrificed. It is our belief that this does not detract from the final publication in any way.

Nicholas Rolle  
Jeremy Steffman  
John Sylak-Glassman

January 2016



# **A Gujarati Origin for Scripts of Sumatra, Sulawesi and the Philippines**

CHRISTOPHER MILLER

*Independent Scholar*

## **Introduction**

Indigenous scripts of Sumatra, Sulawesi, and the Philippines (SSP) have proved difficult to trace to a specific origin. The first serious investigations begin with Holle (1882), whose detailed alphabet charts clearly established all the indigenous scripts of the archipelago as members of the Indic family, and Kern (1882) who, while noting numerous individual points of similarity with old Javanese Kawi or north Indian Nagari scripts, concluded that all Indic scripts of the archipelago descended from Kawi. Diringer (1948) concurred on a Kawi source for Sumatran scripts, but placed the source of the Sulawesi and Philippine scripts in northeast India on the basis of partial affinities with early Nagari script from Assam. Based on letter shapes or evidence from its spelling system, the origin of Philippine scripts has been placed in northeast India (Gardner 1943), south Sumatra (Francisco 1975), Cham script (Wade 1993), or south Sulawesi (Scott 1984).

While each proposal has brought forth new partial correspondences, none has provided a cogent case for the putative origins. Current specialist consensus is that absent further discoveries, the similarities between the modern scripts and putative relatives are still too unclear for any convincing demonstration of their origins (de Casparis 1975, Noorduyt 1993, Cummings 2002, Kozok 1996).

This paper presents evidence for a new theory that SSP scripts are descendants of an early variety of Gujarati script introduced into the archipelago, thus placing them in the Nagari group. This at first seems implausible, but the historical record attests to the major role of Gujaratis in the archipelago. Tomé Pires (Cortêsão 1944) reports a thousand Gujaratis in Malacca prior to 1512, Barnes (2004) shows that Gujaratis sold printed cloths manufactured to the tastes of customers in Sulawesi and the Moluccas around this time, and Gujaratis are known to have played a major role in introducing Islam to the archipelago. Though no records of Gujarati script are known in the region from that era, it would not be surprising if

similarities to Nagari scripts could be explained as developments from a locally adopted variety of Gujarati script.

## **1 Theoretical and Methodological Considerations**

This paper compares features of the internal structure of letters rather than only individual letters as is typically the case. Correspondences between structural features provide regular and systematic evidence for intra- and inter-script relationships. Seemingly irregular changes in form often appear, serving to maintain contrast between similar letters. Scripts may also develop characteristic stroke types and sequences leading to reanalysis of earlier stroke sequences. Both factors can obscure otherwise regular relationships and must be taken into account in evaluating correspondences.

The core of the arguments presented here is a hypothetical proto-script (henceforth PS) reconstructed initially to illustrate regular changes between old Gujarati and early Philippine letter forms. Augmented with a restricted range of attested variants, it provides the basis for an empirical demonstration of the common origin of the SSP scripts. Letters and other graphic elements whose form cannot be explained on the basis of a given script's internal evolution are attributed to contact with other scripts, and borrowing itself provides interesting evidence for the locus of their genesis and their subsequent development.

## **2 The Philippines**

The earliest known and dated document in an SSP script is the 1593 *Doctrina Christiana* (henceforth DC) from Manila, one of the most extensive known records of old Philippine (Ph) script. Though it is a woodcut print, the range of variants found in the book are likely modelled on handwriting from the same period or earlier. The only authentic samples of Ph handwriting that have survived are land deeds from 1613 and 1615 and several dozen individual signatures from miscellaneous documents. Contemporary versions of the script are in use among the Mangyan of Mindoro and, until recently, the Tagbanuwa of Palawan.

The last four rows of letters illustrated in (1) bring out their overall similarity of form. Samples are from the DC, images of Tagbanuwa (Tb) from Kroeber (1919) and drawings based on Francisco (1973), plus images of early 17<sup>th</sup> century handwriting (Hw) from Villamor (1922) and Potet (1987). Mangyan (Mn) is represented here with the publicly available Buhid and Hanunoo typefaces which represent fairly typical shapes. Tagbanuwa seems overall more conservative and similar to DC than Mangyan scripts, while the latter appear to have developed from Hw variants and are in practice more curvilinear than the typeface variants.

The wide range of variants shown and the general structure of the letters provide a basis for comparison with (likely early variants of) Gujarati script and

*Gujarati and Sumatra-Sulawesi-Philippine-scripts*

reveal relationships of a systematic nature observed with no other plausible antecedent. Hypothetical early Gujarati (Gj) forms (most attested, some reconstructed from correspondences with Kaithi and Devanagari) are shown in the first row. Reconstructed proto-script (PS) forms are shown in the second row.

	a	p	y	m	b	w	ng	g	t	l	n	s	d	k	i	u	h
Gj																	
PS																	
DC																	
Tb																	n/a
Hw																	
Mn																	

(1)

Strikingly, all the Gujarati letters with a cup plus a right-hand stem correspond to a Ph variant with a similar cup plus short downward tail. An old Gj <a><sup>1</sup> variant has a prefixed reversed “E” to the left of the cup, also found in DC and likely the origin of similar elaborations in other Ph variants. Gujarati <y> is distinguished from <p> by a left-concave arc at the top left of its cup; in Ph, the left hand side is identical, but the tail of <y> curls into a loop whereas the tail of <p> is distinguished by the addition of a diagonal tick on the right. It seems plausible that the initial loop of Gj <y> could have moved rightward to result in Ph <y>.

<sup>1</sup> Departing from the general practice in literature on Indic scripts, I represent letters by only their invariant content and without the default vowel, which is predictable by rule.

Gujarati ⟨m⟩ is basically cup+stem, with a loop forming a small closed counter whose Ph counterpart is formed by a cross-stroke closing the top of the cup. A plausible intermediate variant would have a loop larger than in Gujarati. Again basically cup-shaped, Gj ⟨b⟩ and ⟨w⟩ begin with a curve from the upper right; in Ph these end in a rising leftward curl instead of a tail (but note Tb ⟨w⟩).

Ph ⟨ng⟩ doesn't correspond directly to Gj S ⟨ng⟩, but does have a tail for the stem on the right of Gj 𑀓 ⟨ny⟩ (shown in (1) as the Gujarati model for Ph ⟨ng⟩). Gujarati ⟨ny⟩ and ⟨ng⟩ being rarely used letters, they could plausibly have been easily confused; in addition S ⟨ng⟩ is nearly identical to S ⟨ḍ⟩, thus favoring the choice of 𑀓 ⟨ny⟩. Gj 𑀔 ⟨g⟩ is similar but for the cross-stroke, and has a vertical stem attached at the top instead of the usual short tail. While clearly related to their Gj counterparts, an added squiggle on the tail of ⟨ng⟩ and elaborations to the left half of ⟨g⟩ help to contrast the two Ph letters.

Other Ph letters relate less directly to their Gujarati equivalents, but systematic correspondences can still be detected. Gujarati ⟨t⟩, ⟨l⟩ and ⟨n⟩ all begin with a stroke curving from the lower left to join the stem. Rotating ⟨t⟩ and ⟨l⟩ leftward and elaborating the rotated stem gives forms similar to DC. The typical Ph squiggle in some Hw ⟨l⟩ variants is especially similar to the (early) Gj double arch and may be an elaboration of the Gj structure. Ph ⟨n⟩ has a squiggle similar to that of ⟨l⟩ instead of the Gujarati curl; the arch may have developed for contrast.

Ph ⟨i⟩, ⟨k⟩, and ⟨h⟩ also relate to their Gj counterparts, if rotated to the left. ⟨i⟩ is segmented into two separate stroke sets (avoiding a large leftward return stroke), ⟨k⟩ has an additional horizontal (losing the vertical in Mn), and the bottom stroke of ⟨h⟩ is absent. Keeping the regular rotation effect in mind, it appears plausible that the tail on the cup-shaped (⟨p⟩ class) letters is what remains of a cursive rendition of the Gujarati stem when these letters are rotated to the left.

Assuming that Ph letters descend from a rotated rendition of early Gj helps explain the form of Ph ⟨s⟩. Rotated leftward and with a short tail for the stem, Gj ⟨s⟩ takes on a shape similar to the several Ph variants. However, instead of a tail *in situ*, there is a short squiggle attached to the the outside of the right arm. A plausible explanation would be that at some point the tail moved rightward (cf the initial curl of ⟨y⟩), then was elaborated into the typical squiggle.

These systematic structural correspondences are strong evidence that Ph variants likely derive fairly directly from an early version of Gujarati script. We have seen that the systematic correspondences in form between the two scripts are mediated by two principal systematic differences: Ph letters appear overall to be rotated leftward, and the Gujarati stem is reduced to a short tail in most Ph letters. Discounting unsystematic differences as likely later developments, we can derive plausible intermediate proto-script (PS) forms in the second row of (1), through which Gujarati letters would likely have evolved into the Philippine forms.

### **3 Sulawesi**


One important piece of evidence makes it unlikely that Gujarati script was introduced directly to the Philippines. Though like Gujarati, Philippine languages have coda consonants of all types, Ph scripts do not represent coda Cs despite the wide range of means available in Gujarati script (conjunct consonant shapes, virama, and full consonant letters read “bare,” without the default vowel).

The same case is found in the old script used to write the Bugis and Makassarese languages of south Sulawesi. While this is an anomaly for a script used to write Philippine languages, it is less so for Bugis and Makassarese. Coda consonants in these languages are restricted to a nasal final (word final [ŋ] or a nasal homorganic to a following onset) and a “stopped” final (the coda half of a geminate except voiced stops, otherwise [ʔ]). In Bugis-Makassarese (BM) as in Ph script, there is no means to represent coda Cs. The resulting frequency of homographs has in fact been exploited as part of a genre of riddle literature (Tol 1992). Given this anomaly and their geographic proximity, Scott (1984) proposes that the immediate origin of Ph script was likely Sulawesi.

Further evidence for a relationship comes from vowel doubling. Instead of writing a repeated consonant twice, the consonant could be written only once, but with a separate vowel mark for each syllable. This is fairly common in both BM and old Makassarese script, and to my knowledge is not attested elsewhere apart from the Philippines. It occurs twice in a 1615 land deed (Villamor 1922:97) and also in a signature given in Scott (1968:54). That this phenomenon exists only in the Philippines and Sulawesi is further evidence for a close relationship.

Kern (1882) and others have already noted similarities for certain letters, but the many differences between BM and Ph letter shapes lead to doubts about their relatedness (Postma 1972, cited in de Casparis 1975). However, if Ph script were adopted from Sulawesi a century or so before contact with the Spanish, it stands to reason that letter variants attested in the Philippines from that period would be similar to early BM forms, from which modern BM forms would logically descend. With the range of authentic Philippine variants illustrated in (1), a better basis for comparison is found in a large range of BM variants that disappeared after standardization of the script in the mid-1800s. These variants provide clues to older stages of the script that correlate with Ph and Gujarati forms.

Central to the argument for a relationship between BM and Ph/Gujarati variants are the placement of dots in BM letters and their relation to structural elements in corresponding Ph letters, and final rising swashes in certain letters.

Noorduyn (1993) discusses an old variant of ⟨a⟩ with a single dot under the left arch rather than under the right arch as in standard , which he attributed to individual scribal variation. However, several manuscripts illustrated in Tol (1996) show evidence for a left-dotted ⟨a⟩ and there is evidence from at least one manuscript (National Library of Indonesia No. VT 129, figure 231 in Tol 1996)



that at some point a left-dotted ⟨a⟩ may have contrasted with a ⟨y⟩ bearing a single dot under its right arch (distinct from the standard ʌ). Although this is the only clear example of a right-dotted ⟨y⟩ I have seen, this may simply indicate early disappearance of this variant, partly under the pressure for maximum contrast with ⟨a⟩. Just such pressure may explain why in manuscripts with double-dotted ⟨y⟩, it contrasts with any one of three variants of ⟨a⟩: left-dotted, right-dotted, or a single midline dot in palm leaf script and some old Makassarese documents. If the original form was right-dotted, variability in the dot position of ⟨a⟩ may have triggered the development of a double dotted ⟨y⟩ to avoid ambiguity.

A characteristic of modern BM script is an upward-leftward flourish at the end of ʌ ⟨g⟩, ʌ ⟨c⟩, ʌ ⟨p⟩ and the derivative ʌ ⟨mp⟩ and ʌ ⟨nr⟩. No contrast with other letters hinges on the flourish; in fact, it is a final position variant of a plain up-angled stroke that follows an up-down stroke sequence. It is absent in some variants shown in Noorduynd (1993), in palm leaf script, and some early documents, and in the modern script never occurs on letters with a final upstroke preceded only by a downstroke. We may assume it is a later development and the earlier forms more closely approximate the original letter shapes.

	(a)	y	g	n	d	l	p	m	b	ng	s	t	k	c	j	r	w	h
Gj	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	n/a
PS	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	n/a
OP	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	
Im	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	Ar.
BM	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ	ʌ

(2)

## *Gujarati and Sumatra-Sulawesi-Philippine-scripts*

BM, Ph, proto-script and Gujarati shapes are compared in (2) along with likely intermediate (Im) forms.<sup>2</sup> The first systematic correspondence involves a dot inside an enclosing cup or arch in BM script, which corresponds regularly to a straight stroke or a tick *in a similar position* in Ph. We can conclude that the BM dot is a simplification of more complex strokes in an earlier BM variety similar to Ph. This conclusion is reinforced by the dot on the palm-leaf script ⟨l⟩ (Noorduyn 1993, Tol 2008) that appears instead of the usual arch. The curling stroke on the underside of the Ph and proto-script ⟨t⟩ has disappeared rather than turn into a dot in BM, perhaps to maintain contrast with ⟨n⟩, whose PS and Ph shapes have more complex strokes under the arch. Similarly for ⟨m⟩, the cross-stroke above the Ph cup has no dot in BM, plausibly a deletion to ensure contrast with √ ⟨d⟩.

Other letters correspond to Ph or earlier forms in more individual ways. BM ⟨b⟩ is reminiscent of the Gujarati and proto-script shapes, but unlike Ph ⟨b⟩ it abbreviates the left side of the bottom up-down sequence so that the tail projects from the left-hand loop itself. The similar ⟨ng⟩ resembles the proto-script and Ph shapes but with a short tail. Standard ◊ ⟨s⟩ has several palm-leaf style variants, including the “J” shape. Variation between these shapes is consistent with the hypothesis that Ph ⟨s⟩ developed by moving the tail to the end of the letter: modern standard ◊ ⟨s⟩ corresponds to a form with the tail attached *in situ* to the “V” body of the letter, whereas the palm leaf ribbon variants, as in Ph, result from moving the tail. Variants of ⟨k⟩ include a cross shape (Noorduyn 1993) and the usual double dash, both with close counterparts in Mangyan scripts.

Since indigenous Philippine phonologies lack the corresponding phonemes, ↷ ⟨c⟩, ↶ ⟨j⟩, and ↷ ⟨r⟩ have no Ph counterparts. Their structure *is* strikingly similar to that of the corresponding Gujarati letters (ઞ, જ, ર) (the Gujarati stem disappearing in BM ⟨c⟩). Of all Indic scripts, only Gujarati has a set of forms that correspond so closely to the BM shapes, and જ is basically unique to Gujarati. Unlike the predicted proto-script forms, the stylus is lifted between the first and second strokes in ⟨c⟩ and ⟨r⟩, fragmenting the letter, and ⟨j⟩ is simplified. These letters are the first clear evidence from outside the Philippines of a Gujarati origin.

Apart from letters for prenasalized consonants, likely derivative or borrowed from another script, only ⟨ny⟩, ⟨w⟩ and ⟨h⟩ show no obvious relation to Ph or Gujarati letters. These also are likely borrowed from other scripts, as mentioned below; ⟨h⟩ clearly developed from variants of the Arabic letter.

In general, evidence from the early history of BM script connects it directly to older forms similar to Ph and/or reconstructed proto-script variants (including ⟨c⟩, ⟨j⟩ and ⟨r⟩ proto-forms), and thence ultimately to an early Gujarati variety. The following two sections illustrate the predictive power of the proto-script: although reconstructed on the basis of regular Gujarati-Philippine correspondences and augmented to account for the extra BM letters, it provides a clear basis for

---

<sup>2</sup> Photoreproductions of the topmost BM ⟨g⟩, ⟨l⟩, ⟨ng⟩ variants are from Tropenmuseum, Amsterdam 668-216; other BM photoreproductions are from Noorduyn (1993).

explaining the origins of scripts of Sumatra and clarifying their otherwise unclear connections with Ph and BM scripts.

#### 4 Batak (North Sumatra)

The north Sumatran Batak scripts are a family of closely related varieties that differ slightly from north to south. In what follows, I will refer only to the more conservative varieties, thus excluding Simalungun forms apart from the reversed tilde-like ⟨p⟩.<sup>3</sup> Several Batak letters have long been recognised for their similarity to Ph or BM counterparts, but resemblances are difficult to find for most others. Starting with the proto-script and interpolating changes for the most part already observed in BM and Ph, we can follow the likely evolution of Batak variants (3).

	h/a	k/h	t	y	m	w	ng	d	j	s	h	g	l	i	u	ny	n	p	b	r
PS	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	n/a	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	n/a	ᵇ	ᵇ	ᵇ	ᵇ
Im	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ
Bk	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ	ᵇ

(3)

Batak letters show two kinds of off-strokes that reflect the structure of the corresponding proto-script letter. Proto-script letters with minimal structure preceding the off-stroke correspond to Batak letters with an off-stroke curling down and back toward the body of the letter as in northern ᵇ ⟨t⟩, ᵇ ⟨h⟩, and ⟨k⟩ (both Karo ᵇ and the more widespread ᵇ). Proto-script letters with more structure preceding the off-stroke have Batak counterparts with either a short downward curl or a simple straight stroke sloping down to the right: ⟨y⟩, ⟨m⟩, ⟨w⟩ (where the tail is on the left branch of the last two variants), ⟨ng⟩ and ⟨d⟩. The off-stroke on ⟨s⟩ doesn't even turn downward.

The middle row shows that few cases require intermediate forms not already attested or likely in eastern scripts: this suggests that variants similar to 15<sup>th</sup> to early 17<sup>th</sup> century BM and Ph forms were used in Sumatra during that period.

<sup>3</sup> Most Simalungun forms are fragmented counterparts of variants with connected strokes in other varieties.

One striking deviation from the proto-script is the absence of the extra stroke on ⟨h⟩, ⟨g⟩, and ⟨l⟩. In Philippine script, only ⟨h⟩ shares this trait. Since Ph ⟨h⟩ could not have descended from any old BM form, this may be evidence that it was borrowed at a later period through contact with users of a Sumatran script.

The two older ⟨t⟩ variants plausibly developed from a shape little different from the proto-script form (in the intermediate row). The ⟨j⟩ letter appears to be derived from ⟨d⟩ by the addition of a stroke. The letter is unlike the BM (𑄎) and Gujarati (જ) shapes: Batak languages have no phonemic /c/ and (except for Karo 𑄏) no completely distinct ⟨c⟩ letterform. The evidence in van der Tuuk (1971) is not entirely clear, but it may be that originally there was no phonemic /dʒ~j/ either to motivate the adoption of the corresponding proto-script letters.

In general, the strong to direct shape correspondences for most letters, and the fact that distinctions in the shape of Batak off-strokes relate systematically to the structure of their proto-script counterparts, together form strong evidence for the hypothesis that the Batak scripts, like the Sulawesi-Philippine scripts, derive from the neo-Gujarati proto-script.

## **5 Surat Ulu (South Sumatra)**

South Sumatra is home to four closely related varieties known as *S(urat) U(lu)* ‘upriver writing’ among other names. Known from the early colonial period are Rejang (Marsden 1811, Jaspan 1964) and Kerinci (Westenank 1922) in the west, and a range of Lampung varieties in the southeast (van der Tuuk 1868). Kozok (2002a, 2004a,b) has published images of another variety, from Kerinci, plausibly from before the 16<sup>th</sup> century. The old Makassarese “bird” script of south Sulawesi turns out to be directly related to these four and only indirectly to BM script.

The basic (non-prenasalized) SU letters are compared in (4) together with conservative old Makassarese (Mk) and proto-script variants, and intermediate (Im) forms. Mk letters in general are directly related to SU equivalents rotated leftward or flipped over, with a slight reordering of strokes: whereas most (conservative) SU letters are made up of the base shapes found in ⟨g⟩ and ⟨p⟩, Mk letters are formed by drawing as much as possible of the letter in a single sequence from the lower left to the upper or lower right. Mk ⟨k⟩, ⟨g⟩, and ⟨y⟩ don’t correspond directly to SU but contain forms that *do* correspond quite closely. The intermediate forms of ⟨k⟩ and ⟨g⟩, similar to ⟨s⟩ and ⟨n⟩, seem to have developed more complex shapes influenced in part by Javanese counterparts in the bottom row (in a Balinese typeface similar to a late 16<sup>th</sup> century style). To start with, ⟨w⟩ seems to be a direct loan from the Javanese letter.<sup>4</sup> The first two arches of ⟨k⟩ equate to the first two arching strokes of Javanese ⟨k⟩, whose final two arches appear to have been added to the Mk base, the final downstroke being fragmented

---

<sup>4</sup> And likely the source for BM ⟨w⟩.

from the rest. The first arch of MK ⟨g⟩ corresponds to the first arch of the Javanese letter and the final two have been added on top. The plausible intermediate ⟨y⟩ is now essentially identical to this shape, and appears to have added extra strokes on the model of ⟨k⟩.

	a	k	g	ng	c	j	ny	t	d	n	p	b	m	s	y	r	l	w	h	h/a
PS							n/a													
Im																				
SU																				
Im																				
Mk																				
Jv																				

(4)<sup>5</sup>

Having shown that Mk script is essentially derived from a pre-17<sup>th</sup> century variety of SU, we turn to the relation between the original SU letterforms and Gujarati and/or the proto-script, one that in general is fairly direct.

⟨k⟩ shortens the cross-stroke, moving it below the main body of the letter, and ⟨a⟩ rotates leftward and distorts slightly to give the SU form. Just as in the Batak scripts, the extra stroke of ⟨g⟩, ⟨l⟩, and ⟨h⟩ is deleted. The ⟨b⟩ shape retained in Sulawesi is basically flattened out, though the rare Lampung variant (shared with Mk) points to early coexistence of a shape similar to the Batak/Ph form. As in Ph and BM palm-leaf script, the tail of ⟨s⟩ moves to the end of the letter. ⟨m⟩ has

<sup>5</sup> SU letters are reproduced from Marsden (1811), Westenek (1922), Kozok (2004b) and van der Tuuk (1868). Photo reproductions of OM letters are all from Tropenmuseum, Amsterdam 668-216.

flipped over: the loop is still present in one Rejang variant. The open counter followed by a tail of ⟨w⟩ relates to the proto-script form, but the bow has been flipped in the opposite direction in *Surat Ulu*. With this change, the letter can be written with two of the basic stroke types; not flipped, the resulting letter would likely have resembled ⟨d⟩. A third letter involving complex changes is ⟨ng⟩ in which the left-facing bow of the proto-script letter is reflected in the structure of the “underlined Z” shape, which plausibly then developed into the other “Z” forms shown. Incising the proto-script shape into bamboo would have been easiest by making a simple “Z” zigzag, then returning to add a line for the bottom segment of the bow.

⟨c⟩, ⟨j⟩ and ⟨r⟩ all share loops in the proto-script and enclosed counters in SU. Otherwise, all three are flattened out, similar to ⟨b⟩ when compared to BM. Proto-script ⟨j⟩ and ⟨r⟩ are very close in shape: the ⟨j⟩ has a larger loop though. ⟨c⟩ has a small loop like ⟨r⟩ plus a perpendicular tail. In SU, ⟨j⟩ also bears a tail that distinguishes it from the otherwise similar ⟨r⟩. It seems plausible that this tail might be a reflection of the phonological relationship between ⟨c⟩ and ⟨j⟩.

In ⟨n⟩, ⟨y⟩ and ⟨ny⟩, we see three closely related shapes. The shape of ⟨n⟩ shares a descending middle section reminiscent of Ph and BM variants, and the shape of ⟨y⟩ might possibly have derived from the proto-script form by flipping the letter vertically, preserving contrast with ⟨n⟩. Comparing these two with ⟨ny⟩, which has no obvious proto-script counterpart, it seems it might have been derived by merging their two lettershapes into one. The resulting shape is ⟨n⟩ (or alternatively ⟨y⟩) with an extra stroke, which might have been the inspiration for deriving prenasalized letters by adding an extra stroke, not only in *Surat Ulu* but also in Sulawesi, where the same means of deriving prenasalized consonant letters was used — though (as Noorduyn points out) at a relatively late stage in the evolution of the script. A SU source for the use of an extra stroke for BM derived nasal letters is all the more plausible considering that an early variant of BM ⟨nc⟩ has a shape very close to Kerinci ⟨c⟩ and may therefore possibly be borrowed from some SU variety.

## **6 Inter-script borrowing and the vowel marking system**

The partial Javanese script borrowings in Mk suggest an alternative explanation for the shapes of ⟨n⟩, ⟨y⟩ and ⟨ny⟩. In many cases, the Gujarati and proto-script letters are strikingly similar to one or more Kawi counterparts due to a common Brahmi origin: 5<sup>th</sup> century North Indian Gupta and ca. 10<sup>th</sup> century *kuṭila* letters closely resemble 9th century Kawi (cf. Faulmann 1880).

This suggests possible adstratal influence: during the introduction of Gujarati script, many users already familiar with a Kawi-based script, perceiving the similarities, may have modified the new script in ways bringing some letters closer to their Kawi-type equivalents. Hence the Sumatran loss of the extra line in

⟨g⟩, ⟨l⟩ and ⟨h⟩ and shortening of the cross-bar in ⟨k⟩, and possible borrowing of ⟨y⟩ and ⟨ny⟩ from Kawi models. However, none of this weakens the case for the systematic and regular correspondences between the early Ph/BM scripts and Gujarati, nor the way the reconstructed proto-script reveals otherwise opaque correspondences between scripts. These extensive regularities remain obscure when we compare OP, BM, Batak, OM and SU scripts with Kawi.

Until now we have dealt entirely with autonomous letters, ignoring the vowel marks. Following the structure of Indic scripts in general, postconsonantal vowels are represented by smaller marks above, below, or on either side of the base letter (5). Most vowel marks are similar across the SSP scripts, Gujarati, and Kawi-based Javanese and Sundanese, but where there are significant differences, the SSP scripts all agree with Kawi-type marks and not Gujarati. Although several Batak marks appear to be written in positions that correspond to Gujarati rather than Kawi, this may be due to vertical incision on bamboo. Batak vowel marks share their form with conservative Lampung vowel marks (van der Tuuk 1868), which may indicate direct common origin: Batak shares with Lampung a spelling rule observed by van der Tuuk that moves vowel marks rightward from their logical host to a following consonant marked for a zero vowel. These facts point to an early Kawi-derived vowel marking system combined with Gujarati-derived base letters, simplifying and changing as it was adapted by users of each script.

	i (ī)	u	ě	e	o	-m/-ng	-h	-r	-∅
Guj	𑀩𑀭	𑀭		𑀭	𑀭	𑀭	𑀭:	𑀭	𑀭
Kawi	𑀩	𑀭	𑀭	𑀭	𑀭	𑀭	𑀭:	𑀭	𑀭
Jav-Bal	𑀩	𑀭	𑀭	𑀭	𑀭	𑀭	𑀭	𑀭	𑀭
Sunda	𑀩	𑀭	𑀭	𑀭	𑀭	𑀭	𑀭	𑀭	𑀭
SU-Lpg	𑀩	𑀭	𑀭	𑀭	(𑀭)	𑀭	𑀭	𑀭	𑀭
Batak	𑀩	𑀭	𑀭	𑀭	𑀭	𑀭	𑀭		𑀭
Bugis	𑀩	𑀭	𑀭	𑀭	𑀭				
OP	𑀩	𑀭							

(5)

Given the strong evidence for a Gujarati origin for the consonant letters, the use of a Kawi vowel system is unexpected. One explanation might lie in the nature of North Indian mercantile scripts, which often did not mark word-internal

short vowels (Diringer 1948) Gujarati traders conceivably used such a style, and when their script was adopted by their trading partners, Kawi vowel marks would naturally have been called upon to represent otherwise unwritten vowels.

## **7 Conclusion**

Evidence seems strong for a mainly Gujarati origin for the SSP scripts. The key piece of evidence comes from resemblances between old Philippine and Gujarati scripts supplemented with reconstructions of early variants: these, with additional evidence from Sulawesi, form the basis for a reconstructed intermediate proto-script that clarifies previously unsuspected relationships between scripts of the SSP group. The evidence put forth in this paper strongly supports the view that (apart from the Kawi-type vowel markings) SSP scripts descend from an early, unvowelled mercantile variety of Gujarati script and are therefore closely related to modern Gujarati and Kaithi scripts within the Nagari script group.

## **References**

- (Anonymous) 1593 (2005). *Doctrina Christiana. The First Book Printed in the Philippines. Manila, 1593*. A Facsimile of the Copy in the Lessing J. Rosenwald Collection. Library of Congress, Washington. With an Introductory Essay By Edwin Wolf 2<sup>nd</sup>.
- (Anonymous) No date. *The Chronicles of Goa and Talloq. 668-216* KIT Royal Tropical Institute Tropenmuseum, Amsterdam.
- Barnes, Ruth. 2004. Indian Textiles for Island Taste: Gujarati Cloth in Eastern Indonesia. *Ars Orientalis* 34, *Communities and Commodities: Western India and the Indian Ocean, Eleventh-Fifteenth Centuries*, pp. 134-149.
- de Casparis, J. G. 1975. *Indonesian Palaeography. A History of Writing in Indonesia from the Beginnings to c. A.D. 1500*. Leiden: Brill.
- Cortese, Armando. 1944. *The Suma Oriental of Tomè [sic] Pires. An account of the East, from the Red Sea to Japan, written in Malacca and India in 1512-1515*. London: The Hakluyt Society.
- Cummings, William. 2002. *Making Blood White. Historical Transformations in Early Modern Makassar*. Honolulu: University of Hawai'i Press.
- Diringer, David. 1948. *The Alphabet. A Key to the History of Mankind*. New York: Philosophical Library.



Christopher Miller

- Francisco, Juan. 1973. *Philippine Palaeography*. *Philippine Journal of Linguistics* Special Monograph, Issue Number 3.
- Faulmann, Karl. 1880. *Illustrierte [sic] Geschichte der Schrift*. Vienna/Pest/Leipzig: Hartleben's Verlag.
- Gardner, Fletcher. 1943. *Philippine Indic Studies*. San Antonio, Texas: Witte Memorial Museum.
- Holle, K.F. 1882. *Tabel van Oud- en Nieuw-Indische Alfabetten. Bijdrage tot de palaeographie van Nederlandsch-Indië*. The Hague: M. Nijhoff.
- Jaspan, M. A. 1964. *Redjang Ka-Ga-Nga Texts*. Canberra: Australian National University.
- Kern, Hendrik. 1882. Over de Opschriften uit Koetei in Verband met de Geschiedenis van het Schrift in den Indischen Archipel. *Verlagen en Mededeelingen der Koninklijke Akademie van Wetenschappen*, pp. 182-203.
- Kroeber, Alfred Louis. 1919. *Peoples of the Philippines*. American Museum of Natural History Handbook Series No. 8. New York: American Museum of Natural History.
- Kozok, Uli. 1996. Bark, Bones, and Bamboo: Batak Traditions of Sumatra. In A. Kumar and J. H. McGlynn, eds.) pp. 231-246.
- Kozok, Uli. 2002. Surat Batak. <http://www.hawaii.edu/indolang/surat>. (26 May, 2010).
- Kozok, Uli. 2004a. A 14<sup>th</sup> Century Malay Manuscript from Kerinci. *Archipel* 67(1):37-55.
- Kozok, Uli. 2004b. *The Tanjung Tanah Code of Law. The Oldest Extant Malay Manuscript*. Cambridge: Cambridge University Press.
- Kumar, Ann and John H. McGlynn, eds. *Illuminations. The Writing Traditions of Indonesia: Featuring Manuscripts from The National Library of Indonesia*. Jakarta/New York: The Lontar Foundation/Weatherhill, 1996.
- Marsden, William. 1811. *The History of Sumatra* (Third edition). London: Longman, Hurst, Rees, Orme, and Brown.

*Gujarati and Sumatra-Sulawesi-Philippine-scripts*

- Noorduyn, J. 1993. Variation in the Bugis/Makasarese script. *Bijdragen tot de Taal-, Land- en Volkenkunde*, 149(3):533-570.
- Postma, Antoon. 1972. Contemporary Mangyan Scripts. *The Philippine Journal of Linguistics* 2(1):1-12.
- Potet, Jean-Paul. 1987. La pétition tagale : Caming manga alipin (1665). *Cahiers de linguistique - Asie orientale* 16(1):109-157.
- Scott, William Henry. 1984. *Prehispanic Source Materials for the Study of Philippine History*. Manila: New Day Publishers.
- Tol, Roger. 1992. Fish Food on a Tree Branch; Hidden Meanings in Bugis Poetry. *Bijdragen tot de Taal-, Land- en Volkenkunde* 148(1):82-102.
- Tol, Roger. 1996. A Separate Empire. Writings of South Sulawesi, in Ann Kumar and John H. McGlynn, eds., 213-230.
- Tol, Roger. 2008. Rolled up Bugis Stories: A Parakeet's Song of an Old Marriage Calendar. Paper read at the 17<sup>th</sup> Biennial Conference of the Asian Studies Association of Australia, 1-3 July 2008.
- Van der Tuuk, H. N. 1868. *Les manuscrits lampongs, en possession de M. Le Baron Sloet van de Beele (ancien Gouverneur-Général des Indes Néerlandaises)*. Leiden: T. Hooiberg et fils.
- Van der Tuuk, H. N. 1971. *A Grammar of Toba Batak*. Translation of H. N. van der Tuuk. 1864, 1867. *Tobasche Spraakkunst*. Koninklijk Instituut voor Taal, Land- en Volkenkunde Translation Series 13. The Hague: Martinus Nijhoff.
- Villamor, Ignacio. 1922. *The Ancient Filipino Writing Drawn by Dn. Ignacio Villamor from Belarmino and Other Ancient Documents*. Manila: Colegio de Santo Tomás.
- Wade, Geoff. 1993. On the Possible Cham Origin of the Philippine Scripts. *Journal of Southeast Asian Studies*, 24(1):44-87.

Christopher Miller

Christopher Miller  
43 chemin du Grand Moulin  
Deux-Montagnes QC J7R 3C3  
Canada

[christophermiller@mac.com](mailto:christophermiller@mac.com)