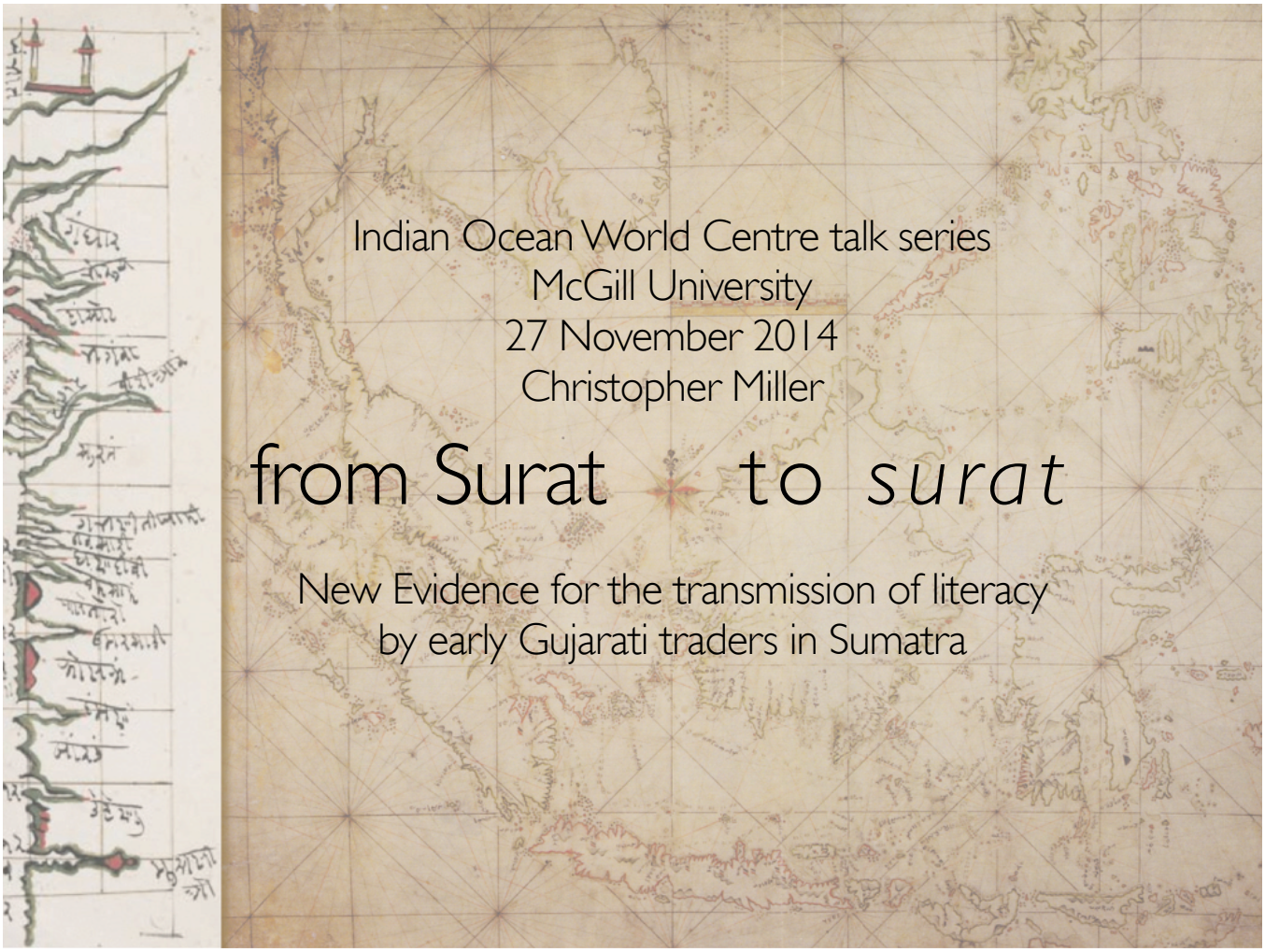
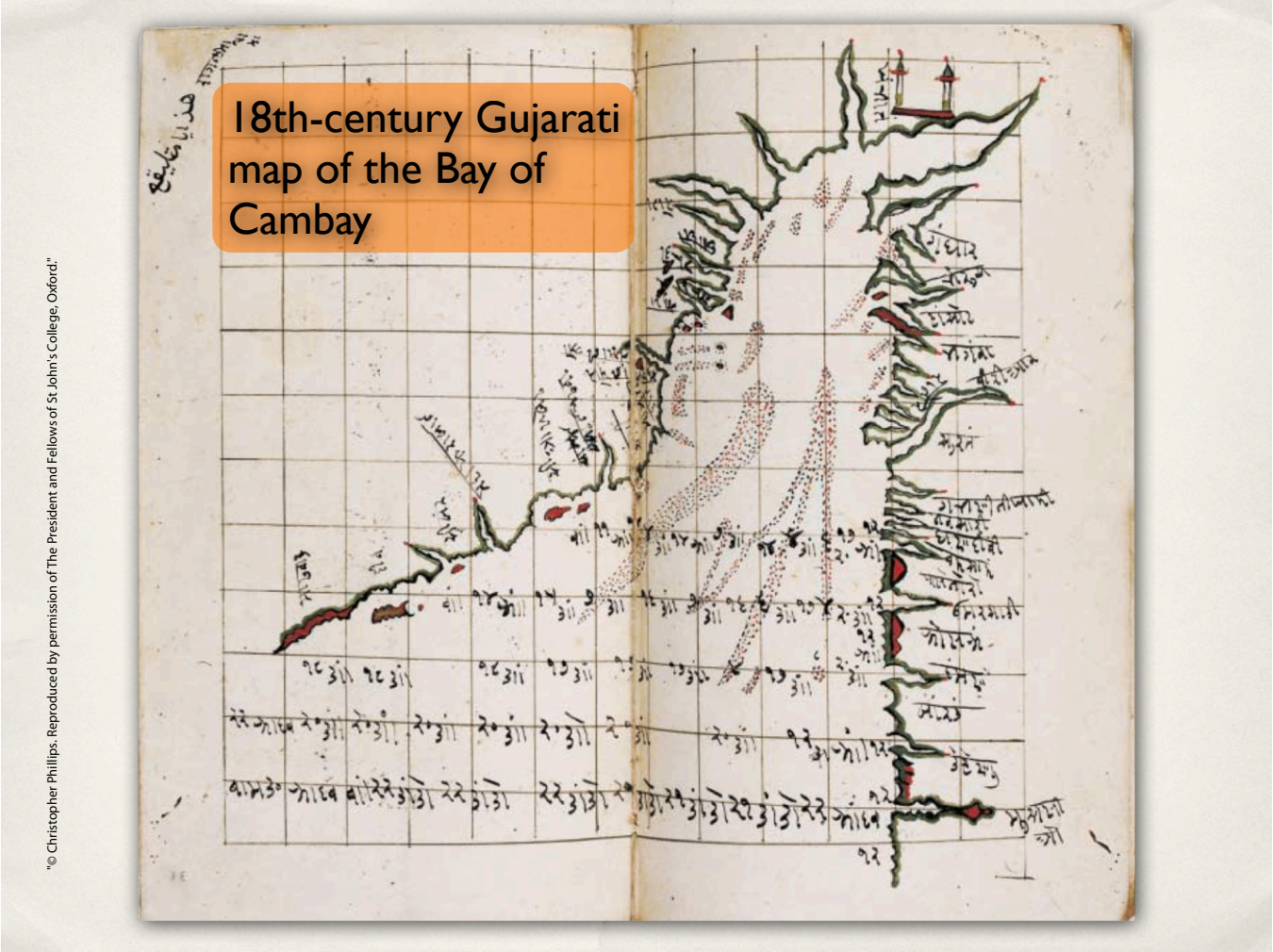
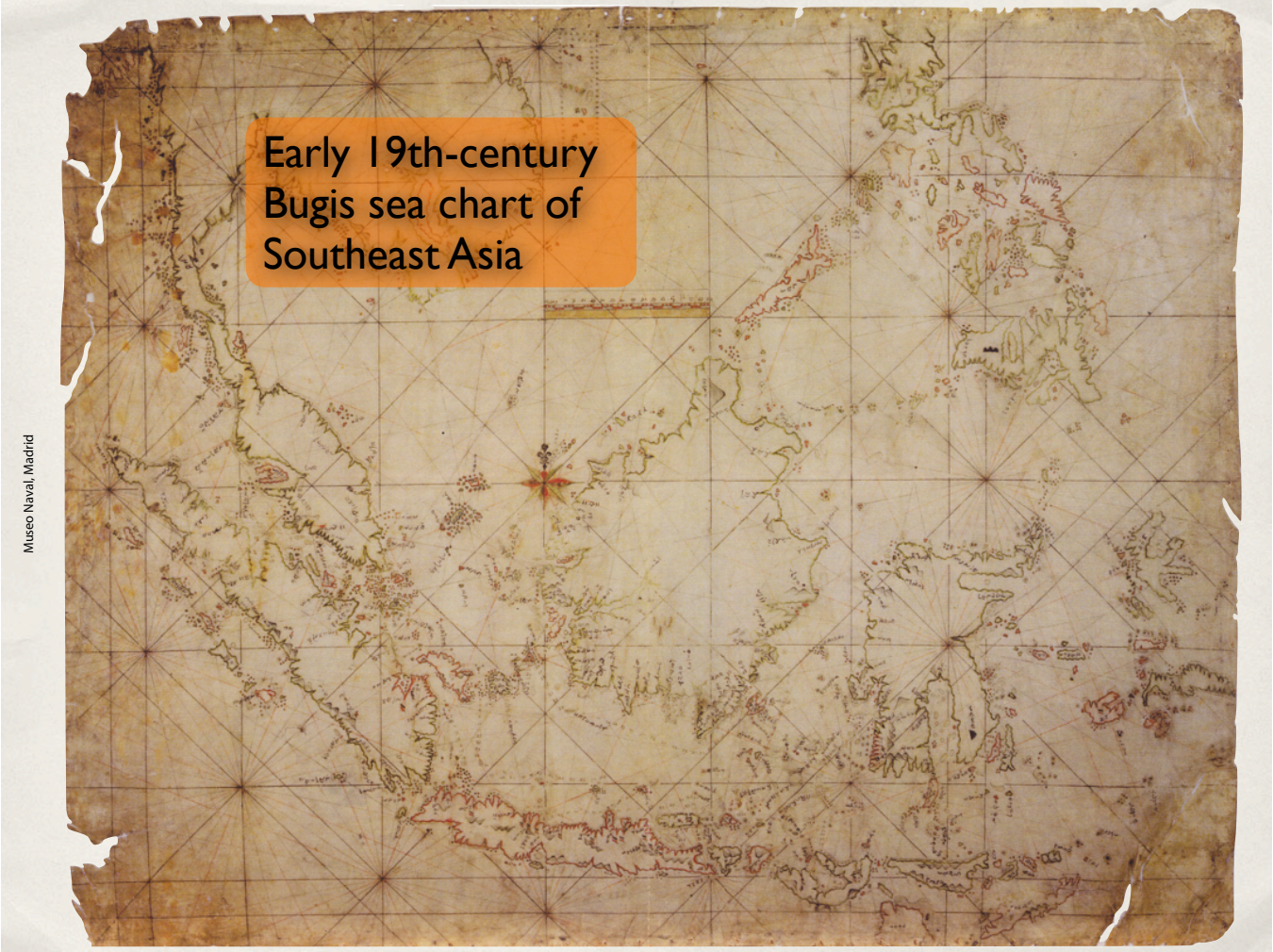


This talk is the most recent version of a series of presentations and papers over the past five years on the properties, relationships and history of several less well-known scripts of the Philippines and Indonesia. The core thesis of this research is that these scripts are derived, through a set of regular processes of chirographic (handwriting) change, from an informal Devanagari-derived shorthand introduced to the southeast Sumatran Malay homeland by Gujarati merchants, probably somewhere around the turn of the 14th century.

Hence the title "from **Surat**" (the then important Gujarati seaport Surat/સુરત, at mid-left, next to the title, on the section of an old Gujarati map of the Bay of Cambay) "to ***surat***" ('writing' in Malay and several other Malayo-Polynesian languages; *súlat* in Tagalog, *suwat* in Bisayan, *surĕ'* in Bugis, *soratra* in Malagasy).

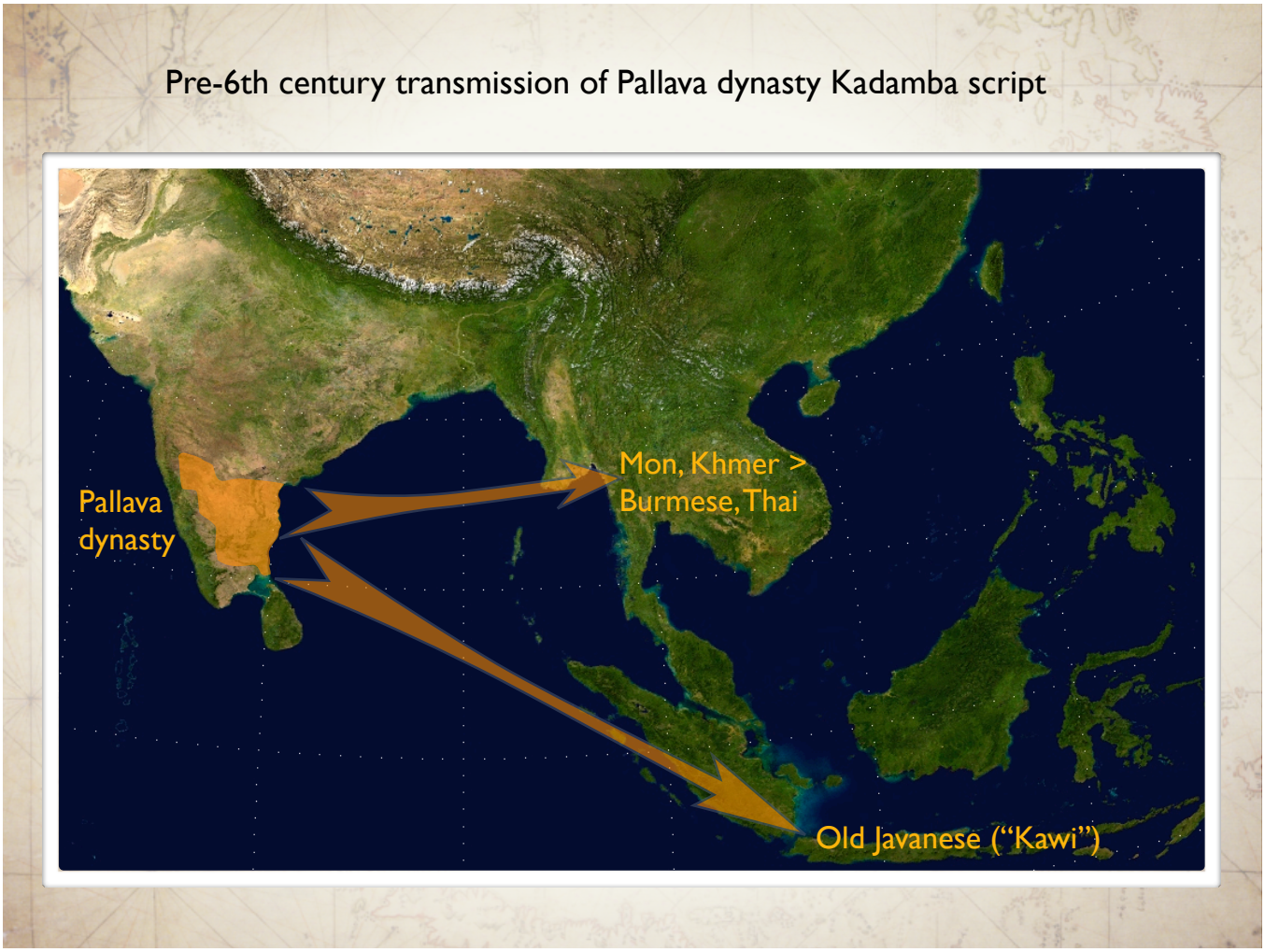


St. John's College, Oxford, acquired in 1754. Features attest to apparent borrowing of European mapmaking and navigation techniques. (Sheikh (2009) "A Gujarati Map and Pilot Book of the Indian Ocean", .)
The script variety itself (also used in the accompanying pilot book) has several interesting features different from modern Gujarati and other early Gujarati script varieties.



3

This 72 × 90 cm cowhide sea chart (portolan) in the Naval Museum in Madrid is one of three similar known maps (the other known copies are in Utrecht and Batavia/Jakarta) that date to before 1830 or earlier. Toponyms are Bugis, written in a variety of Bugis script. Few early documents in Bugis-Makasarese script are securely datable, so this gives a glimpse at the features of the script around the turn of the 19th century.
(Cf. Le Roux (1935), 'Boegineze Zeekaarten van den Indischen Archipel' and Schwartzberg (1994) 'Southeast Asian Nautical Maps'.)

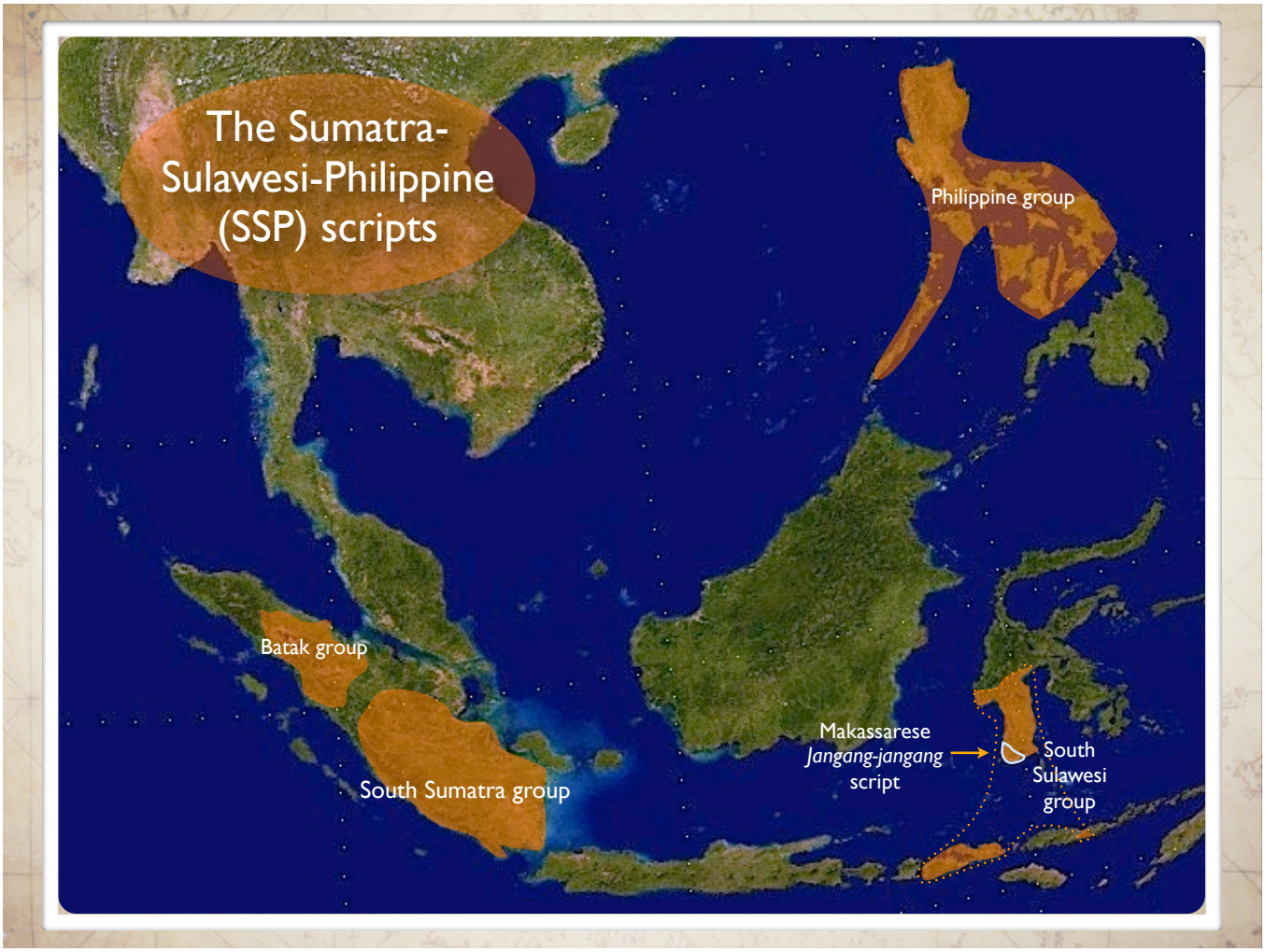


The earliest known inscriptions in southeast Asia, both on the mainland and in the archipelago, are in a script commonly known as "Pallava" after the Pallava dynasty of the 4-6th centuries but also referred to as Kadamba — the direct ancestor of the closely related Kannada and Telugu scripts of mid-south India. It was succeeded in SE Asia after a couple of centuries by other scripts assumed to descend from Pallava, though precise details of the Old Javanese script (commonly called "Kawi" after specialised poetic language it was used to write) correspond much more closely to Tibetan script than to Pallava, which appears to argue for a common origin for the two.

The derivation of the colonial-era Javanese-Balinese and Sundanese scripts can be followed fairly directly by comparing with standard Kawi of the 10th century (as written in the Laguna Copperplate Inscription from the Philippines) and an intermediate stage found in a 14th century Malay-language book of laws kept as an heirloom in Tanjung Tanah, Kerinci, mid-south Sumatra.
(Postma, Kozok references)

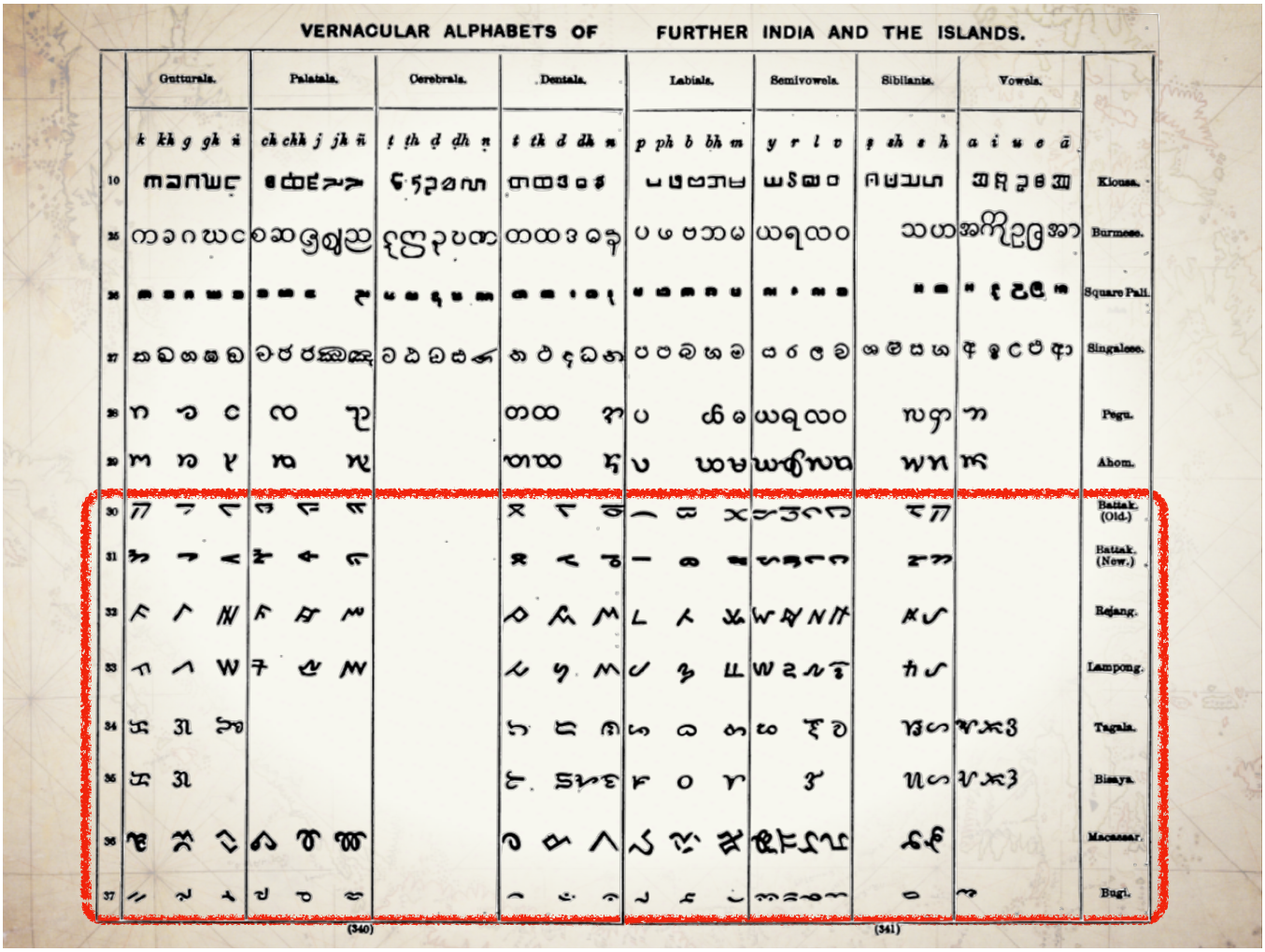
This clear continuity disappears when we compare indigenous scripts of Sumatra, Sulawesi and the Philippines. There are a number of sporadic points of resemblance between one or both of the Sumatran scripts and old Javanese shapes, but no obvious systematic correspondences. Resemblances are even sparser further east, for the Philippines and Sulawesi. Although Kawi (and before it Pallava) are the only clearly documented Indic precedents with any widespread use in the region, there is no clear case to be made (on the basis of any observable structural correspondences) that the letters of the Sumatran, Sulawesi and Philippine scripts are plausibly derived from a Pallava or Kawi antecedent.

At the same time — although the character sets illustrated here are stereotypical representatives of each script (in fact not fully representative of the range of variation and even of normal shapes in each) — there are enough sporadic similarities between one or more of them to raise the possibility that they might share some common historical antecedent.

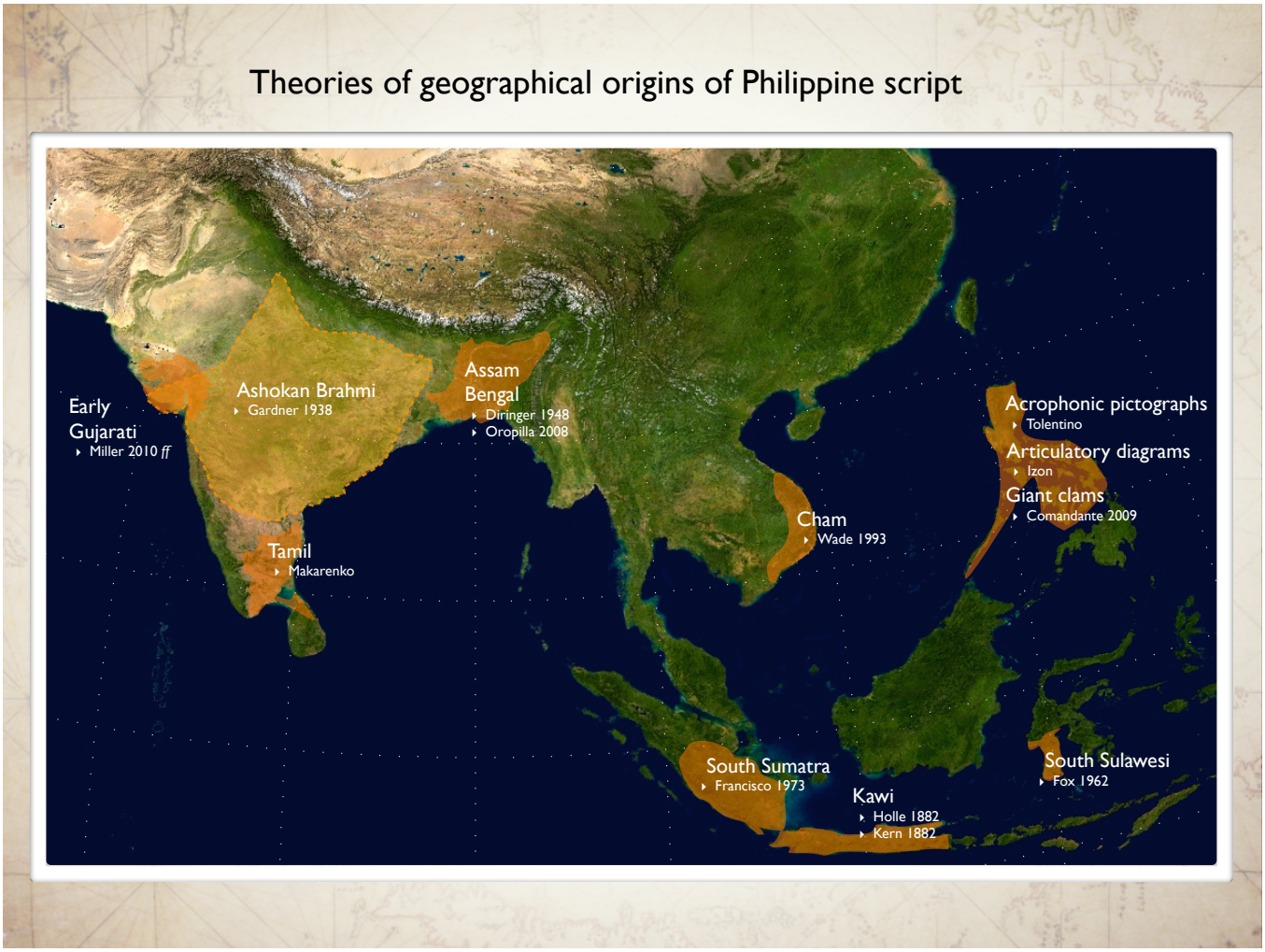


6

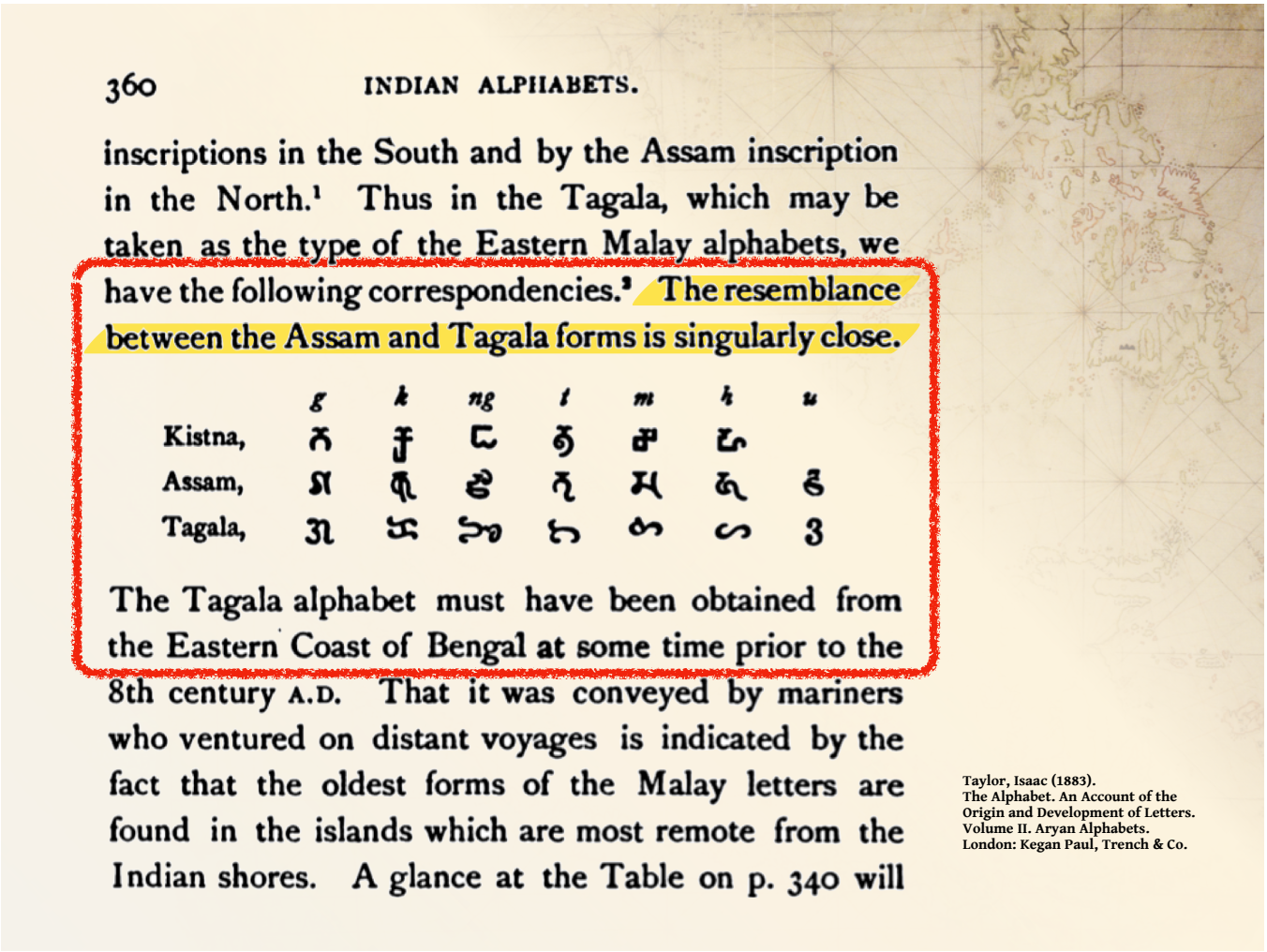
This map shows the locations of the four main groups of Indic scripts in Sumatra, Sulawesi and the Philippines.



Historical comparisons between scripts have generally relied on second- or third-hand reproductions rather than original data. These reproductions very often introduce distortions and other misrepresentations, and the character sets chosen (as mentioned above) are either not necessarily representative of the variety of different shapes to be found in each region or (in the Philippines) create a misleading impression of distinct regional scripts.



Apart from the Kawi hypothesis, the most widely-accepted as plausible, various other possibilities have been raised over the past three centuries for the origins of the various Sumatra-Sulawesi-Philippine scripts. The claims for the Philippines are particularly legion, as shown here. Some are clearly phantastical, whereas others, such as Fox's discussion of spelling peculiarities shared between the Philippines and South Sulawesi, or Wade's 1993 comparison of inscriptional Cham with reproductions of original archival letter shapes from the Philippine script (the first such attempt to use first-hand data in this way) show a soundness of methodology and argumentation that is lacking in many other proposals. Even Wade's laudable advance in methodology yielded only a tentative conclusion of a general similarity between the scripts, but one that could not be formulated in any precise terms.



360 INDIAN ALPHABETS.

inscriptions in the South and by the Assam inscription in the North.[1] Thus in the Tagala, which may be taken as the type of the Eastern Malay alphabets, we have the following correspondencies.[2] The resemblance between the Assam and Tagala forms is singularly close.

	g	*k*	*ng*	*t*	*m*	*h*	*u*
Kistna,							
Assam,							
Tagala,							

The Tagala alphabet must have been obtained from the Eastern Coast of Bengal at some time prior to the 8th century A.D. That it was conveyed by mariners who ventured on distant voyages is indicated by the fact that the oldest forms of the Malay letters are found in the islands which are most remote from the Indian shores. A glance at the Table on p. 340 will

Taylor, Isaac (1883).
The Alphabet. An Account of the Origin and Development of Letters.
Volume II. Aryan Alphabets.
London: Kegan Paul, Trench & Co.

Any illustration of the difficulties involved in past comparisons between scripts is found above in an excerpt from Taylor (1883). Although Taylor claims that there is a "singularly close" resemblance between Assamese and Tagalog ("Tagala") scripts, there are two methodological problems that cast doubt on his claim. First, the comparison is only partial: although the Assamese (early Bengali) script contains the full complement of over 40 distinct independent letters, even the Philippine (Tagalog) script has no less than 17 such independent letters. There is no way to know whether the other corresponding letters in each script can be described as similar, even impressionistically. Second, no independent metric of any sort is given to evaluate the degree of similarity or difference between paired letters: there is no way to decide between conflicting impressions on the part of different observers.

(Taylor, Isaac (1883). *The Alphabet. An Account of the Origin and Development of Letters. Volume II. Aryan Alphabets*. London: Kegan Paul, Trench & Co.)

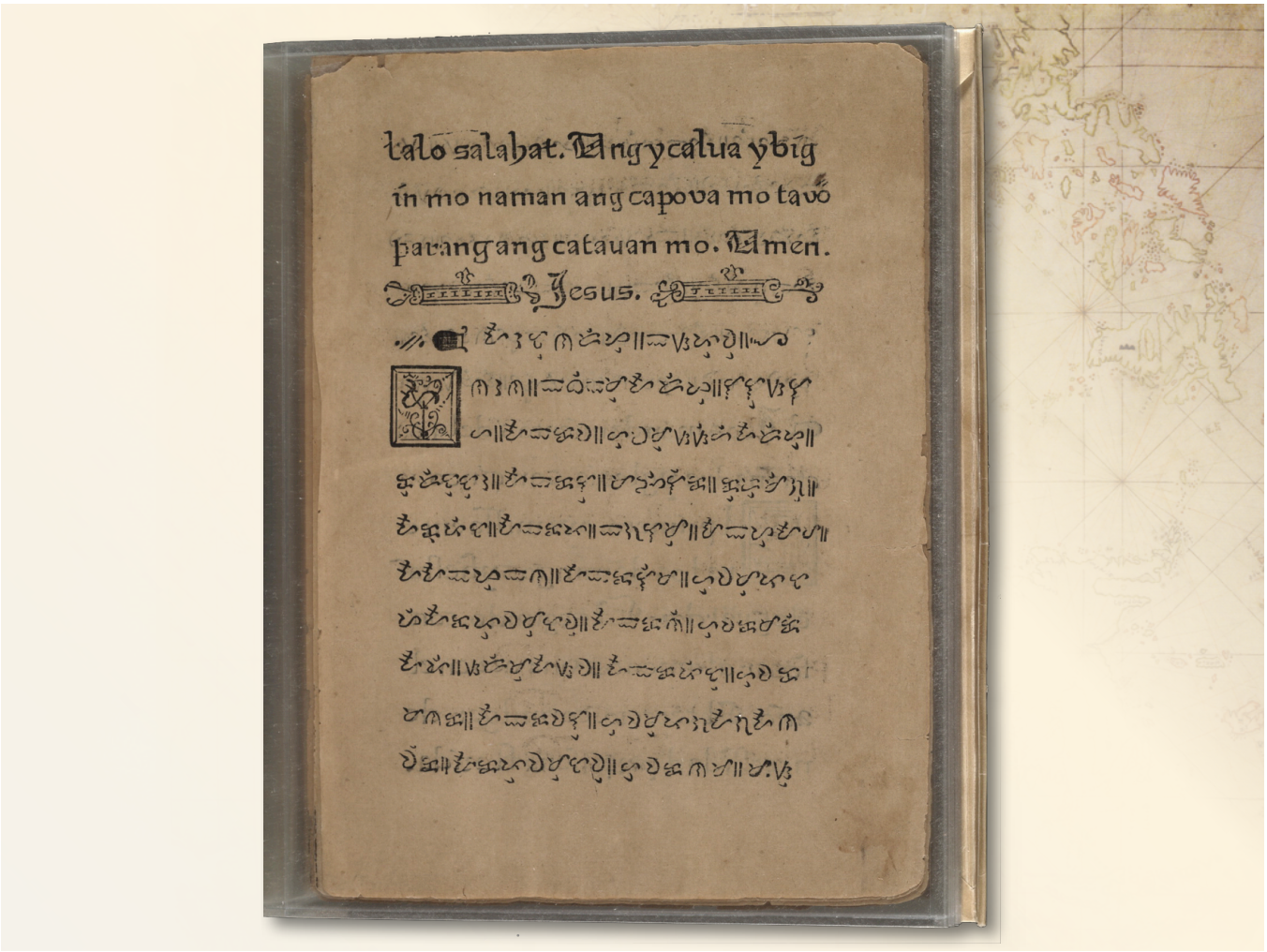
Returning to the (superficial) comparison of the Javanese group scripts (the first four columns) and the S(umatra)-S(ulawesi-)Philippine scripts (the last four columns), it is striking, even without relying on independent criteria to evaluate their similarity, how closely the first four relate to each other in form right through the (subset of) forms illustrated (about half the actual inventory of these scripts), while any such continuity of form is difficult to detect in the final four scripts.

Comparing vowel and final ('coda') consonant signs across the same two groups (with the addition of Kerinci and Lampung scripts for South Sumatra) reveals a much different picture. Although the inventory of signs in Bugis-Makassarese script and Philippine script in particular is much reduced, it is clear that there is a strong continuity across all the scripts — historical and later — in either the form of a given sign, its position, or both. Furthermore, there are clear relationships in the names of the various signs across scripts, even if less so in B(ugis-)M(akassarese) and Philippine scripts. (Even though it does not use counterparts of the Sumatran names for its different vowel signs, BM script shares with the North Sumatran Batak scripts the practice of describing independent letters as "mothers of writing" and the bound subordinate signs illustrated here as "children of writing".)

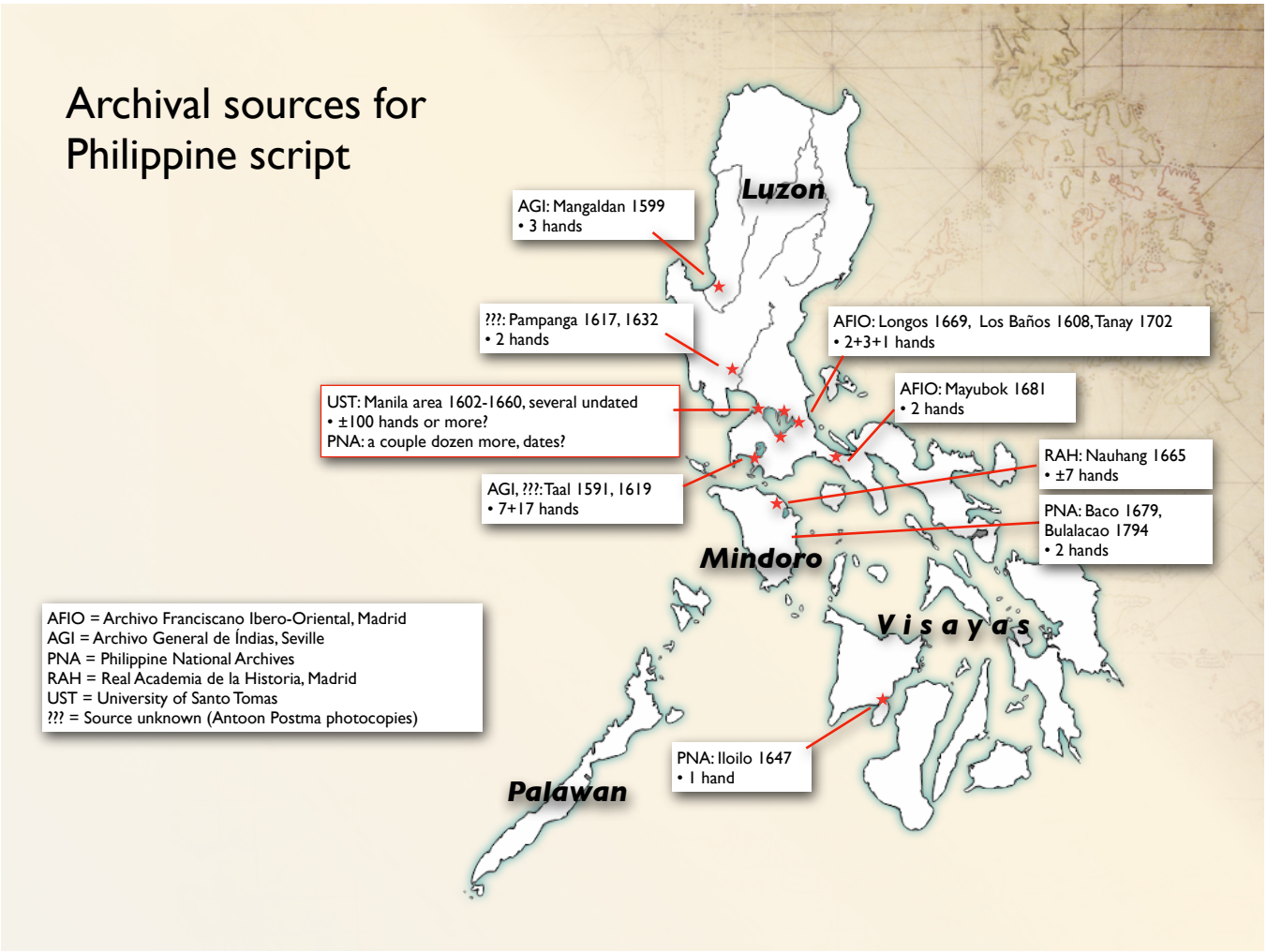
The clear continuity of form across all these scripts for the bound vowel and coda signs contrasts sharply with the lack of any such visible continuity for the SSP scripts compared to the Javanese group. This paradox calls out for an explanation.

A major difficulty in studying the history of the SSP scripts is that in general, securely datable texts are few and far between — especially for early dates. The earliest datable manuscript for Sumatra is an Islamic *Hikayat Nur Muhammad* legend written in Malay in a Lampung script variety, acquired by the Oxford Bodleian library in 1630; for Sulawesi, the earliest datable document is also datable to the early 1600s. Otherwise, few documents in an SSP script from either region give any overt clues to their date of composition.

The Philippines stand out in two ways. First, dated archival documents with the indigenous Philippine script are numerous from the late 16th century through the mid- to late 17th century, when its widespread use began to die out except for survivals on the islands of Palawan and Mindoro. Second, nearly all the dated samples we have are individual signatures rather than texts of any length. Although three documents with indigenous script (widely known as *Baybayin*) exist in the *Archivo General de Índias* in Seville (two from 1591 and one from 1599), the earliest known full text is one of the first two books printed in the Philippines and in fact one of the first printed books in any Indic script: the 1593 *Doctrina Christiana*, printed in Spanish...



...and in Tagalog, in both Latin script and *Baybayin*.



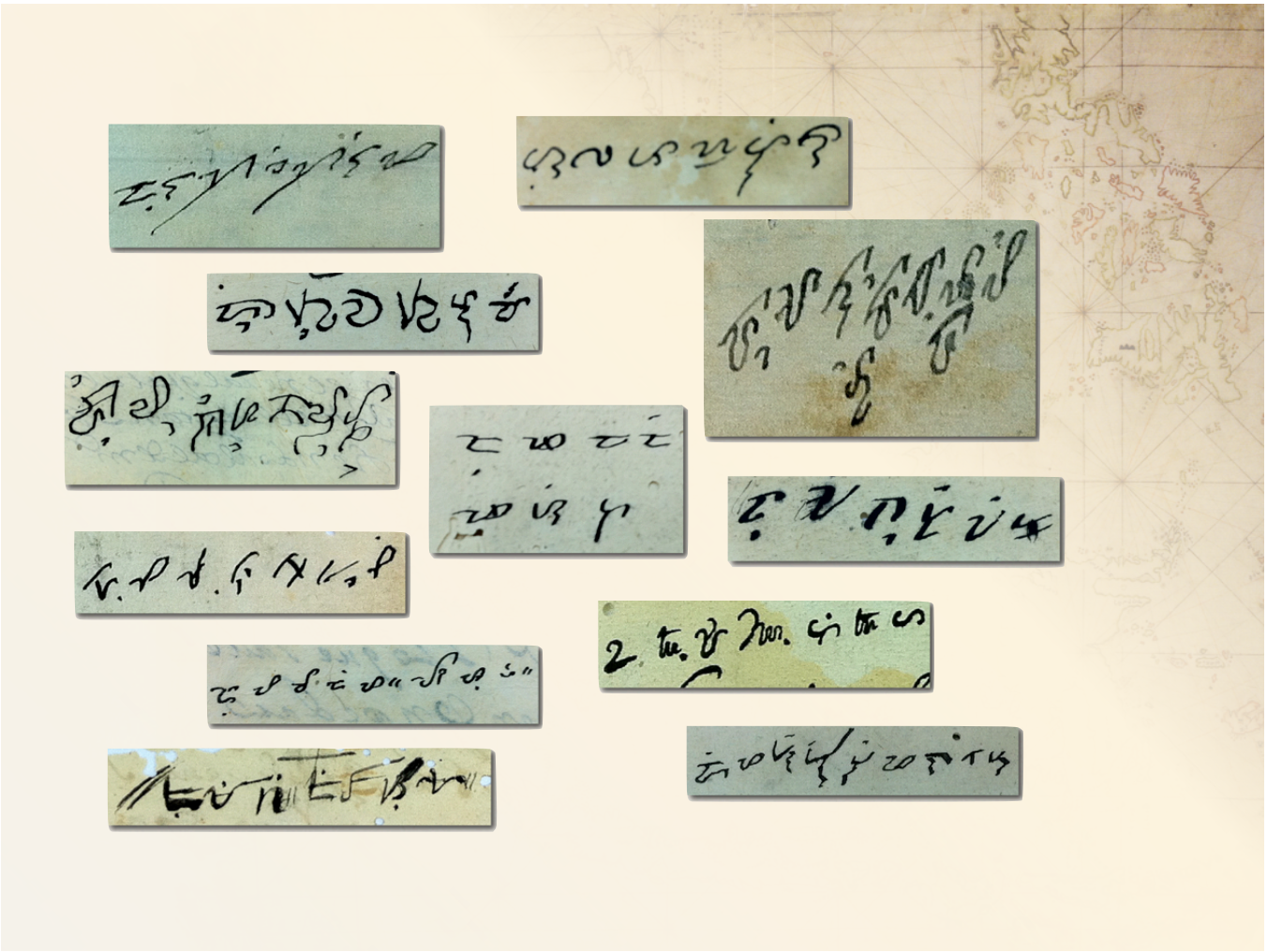
Although there are a number of other probable sources yet to be explored, this map illustrates the places of origin and archival locations for early document with Philippine script known to me in 2012.

Though the three oldest known documents are found in Seville, the most copious source, by far, is the Archives of the University of Santo Tomas in Manila, where a lengthy series of land deeds was preserved to demonstrate the University's land title, together with several other legal documents including a few petitions. Several dozen of these, though written in Tagalog or Spanish in Latin script, bear signatures in *Baybayin* script. Apart from these over 100 distinct signatures and a number of short annotations, there are two full land deeds written in *Baybayin*, that were declared National Historic Treasures in August of 2014.

Several examples of archival *Baybayin*. Except where otherwise noted, these originate in what is now metropolitan Manila.

Left: the December 4 1625 land deed, written by Doña María Silang. The earlier land deed in the script dates to 1613.

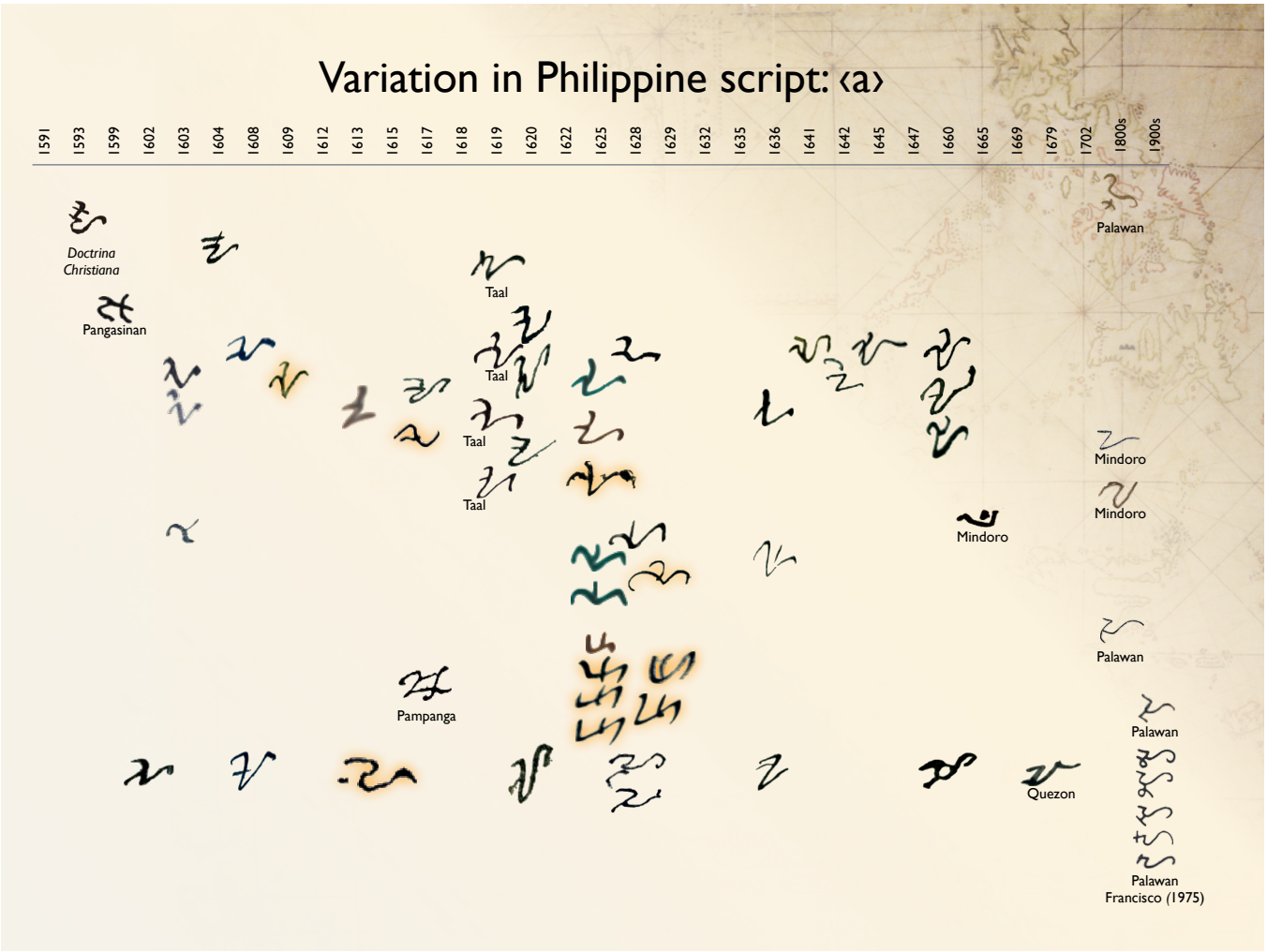
Right:
1) A summary note written on the back of a folded document naming the purchaser (Don Augustín Caso) and price (265 *salapî/tostones*) of a piece of land.
2) A second summary note by the same writer naming the seller (Don Xuan de la Sara) and purchaser (Don Ambrosio Hubag).
3) A 1665 petition from Nawhang (now Naujan), Mindoro. Although one or two hands are unusually ornate, and the top right *Baybayin* signature incorporates a European-inspired final flourish of curlicues, the overall style is fairly typical of handwriting found in Manila during the same period. One or two letters show peculiarities not found elsewhere, though.
4) A third summary note from Manila naming the seller (Don Domingo Alas) and purchaser (Don Augustín Carajas).
5) A 1599 report by Manila bishop Benavides from Mangaldan, Pangasinan describing the local proceedings in King Philip's referendum of submission to his rule. It contains three signatures: the two on the left are in left-handed mirror writing.



Signatures show a wide variety of styles from precise and utilitarian to flamboyant and ebullient to rough and plain.
The ones illustrated here are from the Santo Tomas archives alone.

Variability in the script is not restricted to individual styles or style types. The archives provide a rich enough source of precisely dated samples that it has been possible to map representative tokens of individual hands by year and structural features. The following slides illustrate, for some of the most interesting letters, how they varied between individuals across regions and across time. It becomes clear that the greatest variation is by time.

(The order of the slides follows the indigenous recital order given in the 1593 *Doctrina Christiana*:
a u i h p k s l t n b m g d y ng w.)

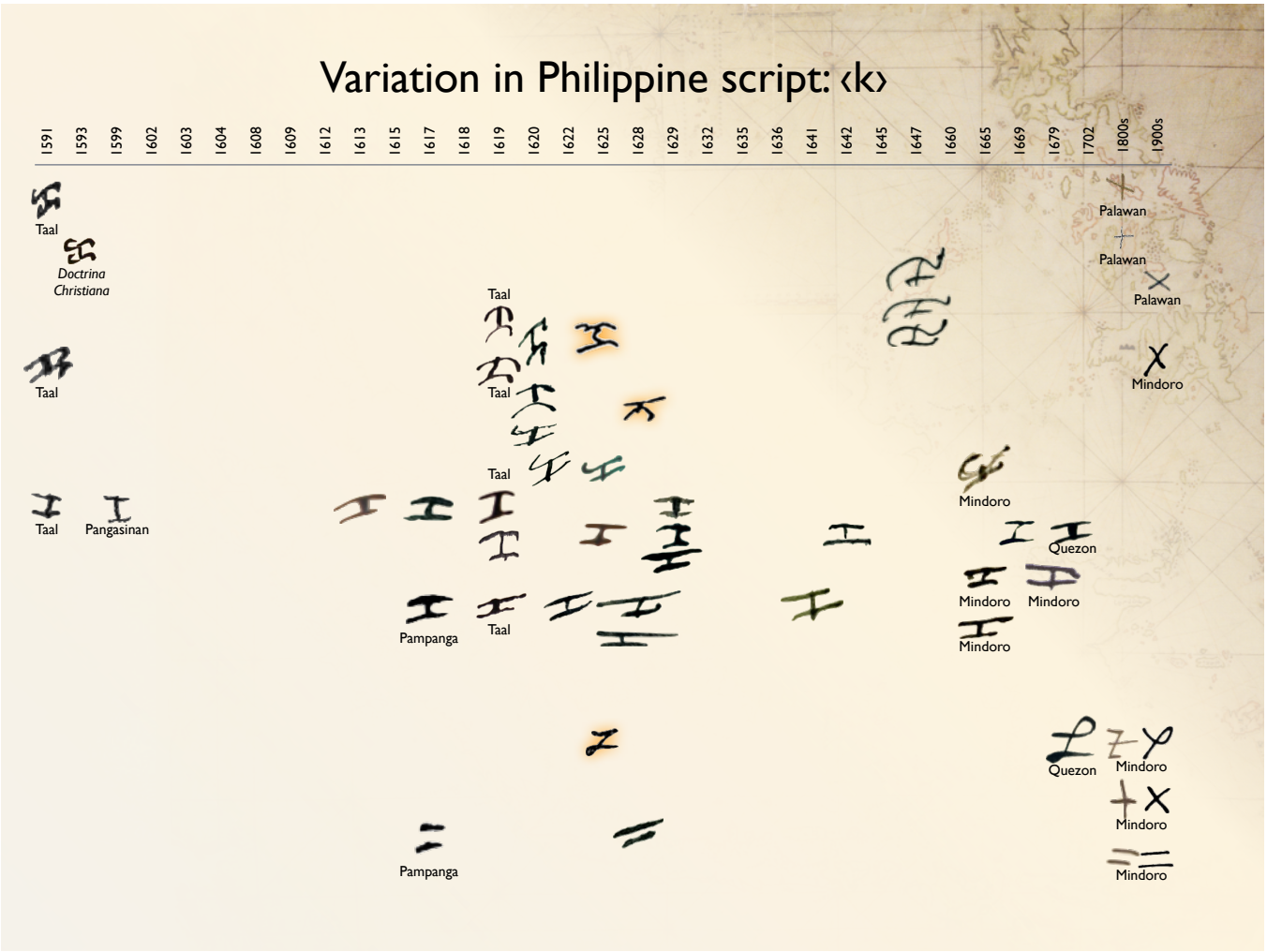


In common with most other Indic scripts, *Baybayin* uses independent vowel letters only to represent syllable-initial vowels. Otherwise, the vowel /a/ is pronounced by default on a consonant letter, and the /i~e/ and /u~o/ vowels are represented by separate dependent marks added above or below a consonant letter, respectively.

This time map shows two main variant shape classes for the letter. At the bottom, the body of the letter has an added "adjunct" dash on its left-hand side. The more widespread variants have an adjunct on the upper side of the letter; the complex curl+crossbar on the top left of the letter quickly simplifies to a plain curl beginning in the early 1600s, and this curl gradually drifts rightward inside the upper body of the letter over the remainder of the century, and disappears entirely in Mindoro.

The Palawan island Tagbanuwa variant given from the 1800s (top right) — with its 30th century developments at lower right — probably represents the original form that both variants discussed above derived from. The upper variant simply has the crossed adjunct drifting up to the top left of the letter's body, whereas the lower variant is an early simplification *in situ* from the crossed shape to a simple dash.

The letter for syllable-initial /i/ shows two kinds of change over time. On the one hand, the main body of the letter changes early mid-century from an earlier cursive '𑀇'-like shape to the stereotypical body shape shared by ‹a› and several other letters. On the other, the "adjunct" horizontal stroke written above the body "drifts" to below the letter in many hands, again starting mid-century, and some writers actually join it to the '𑀇' body with a transitional stroke. The two changes are combined in Mindoro, where only the change to a more stereotypical body change is found in Mindoro, which has various interesting archaic features as with ‹a›.



The most iconic *Baybayin* letter is ‹k›, with its two double curls joined by a vertical crossbar. This shape is taken from the printed version of the script and ultimately based on the shape found in the 1593 *Doctrina Christiana*. It is an extremely rare variant in archival handwriting samples, though, where it is only found in a single 1591 signature.

Much more common on Luzon are three related variants: 1) the most common version — with two straight horizontal strokes resembling a rotated 'H' — which is found right from the earliest manuscripts; 2) a version with an upper "Nike swash" joined to a lower straight stroke by the crossbar; and 3) a version with two arches — often resembling reversed "Nike swashes" — joined by the crossbar.

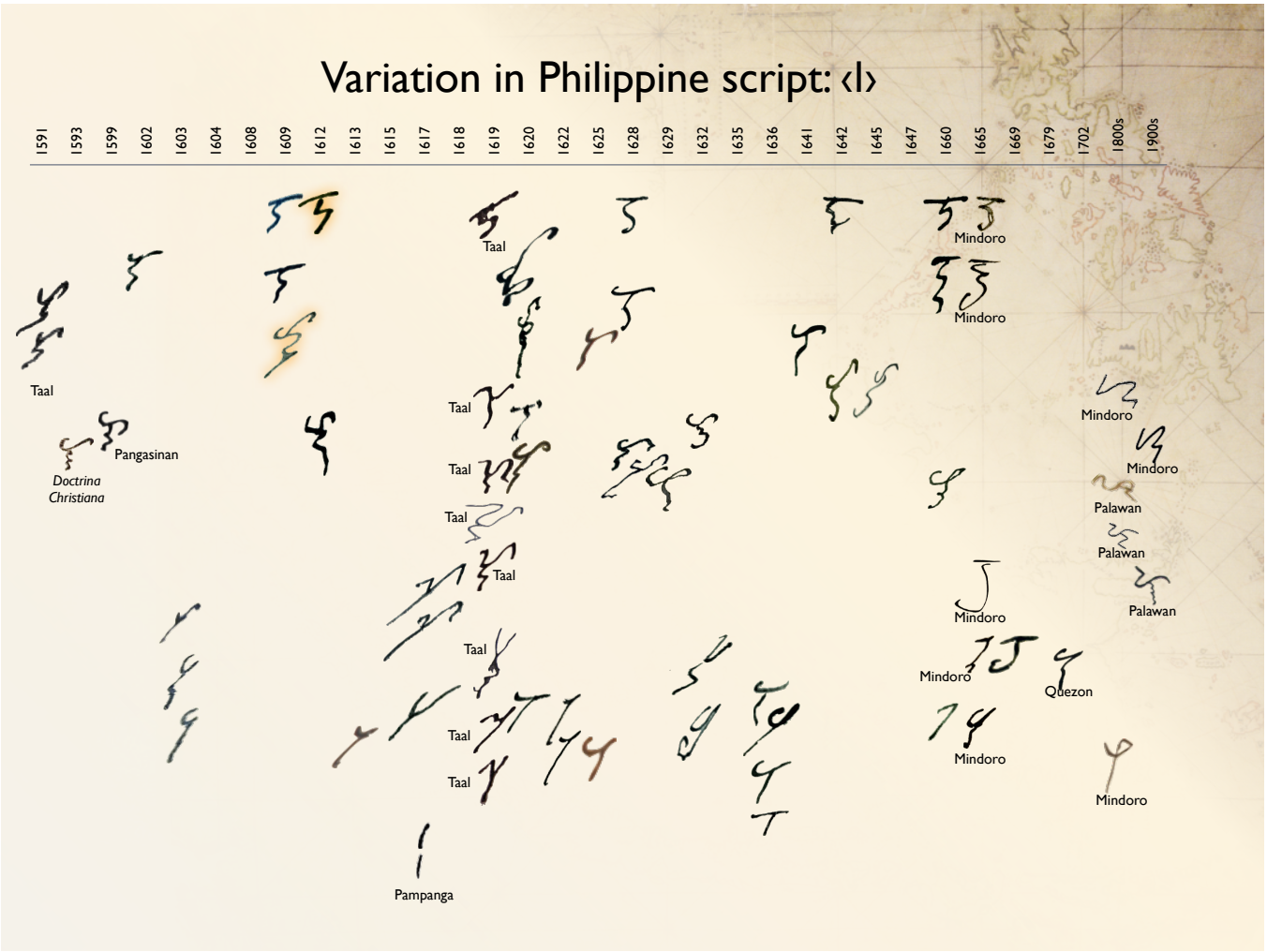
Mid-century and later, the transition to the loop and to the lower horizontal is made by some people by adding "excrescent" loops at top right and lower left, and it is plausible that the current most common Mindoro variant, with a top-right loop and now lower horizontal, derives from this. Early on in Pampanga, and later on in Mindoro, the crossbar disappeared from the first variant, leaving only two parallel horizontal strokes.

Palawan has only an 'X' shape — yet another indication of an early divergence of Luzon and Palawan varieties before the mid-15th century arrival of the Spanish. It is worth noting that the 'V' plus 'Λ' forming the upper and lower halves of the 'X', respectively, can be seen as angular counterparts of the "Nike swash" and arch, respectively, of the second and third Luzon variants. A plausible antecedent for both islands could be reconstructed as a more curvilinear combination of these last two, which would have subsequently changed in different ways on the two islands.

Like ‹a›, ‹s› (in most of its variants) is made up of a main body with an adjunct — except that in this case, the adjunct is always on the right-hand side of the body.

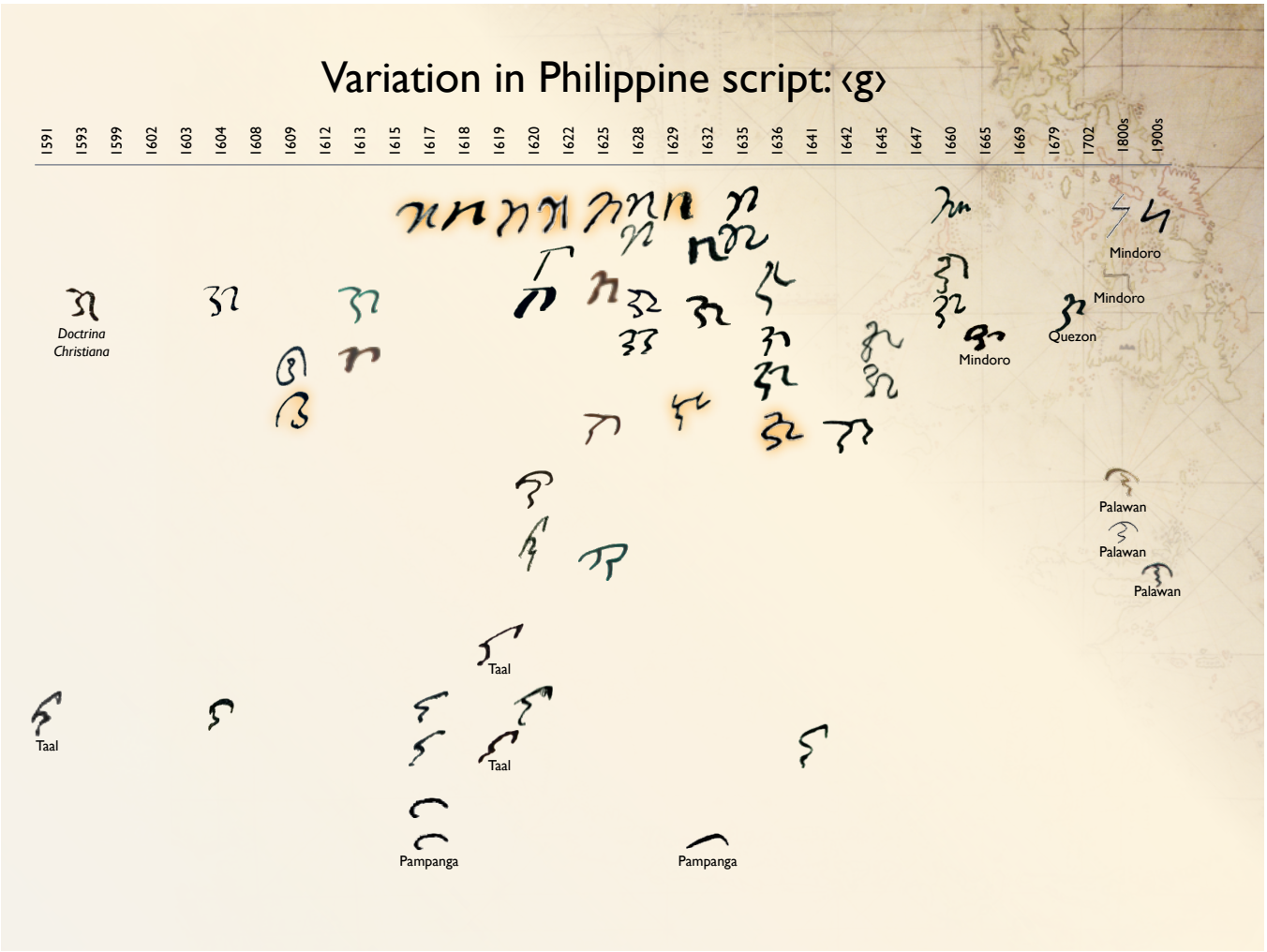
Over the century, the form of the letter varies in two main ways. Beginning with the earliest tokens available, the body can either be an angular ‘V’ shape or a more curvy ‘U’. Fairly early, this adds a short initial horizontal tick, conforming to the stereotypical body shape of many other letters in the script. The other important variation sees the intial backward ‘Ƨ’ or ‘ƨ’~‘Ƨ’ adjunct adding complexity by changing into a ‘3’-like shape.

These changes are followed by a couple of more minor changes: 1) writers in Pampanga delete the main body curl, leaving only the ‘3’-like shape; 2) in a 1665 document from Mindoro, different writers convert the transition between the top of the ‘U’ and the following ‘3’ to a loop similar to the one at the top of a Latin script handwritten ‹s›: ‘ᦰᦱ‘ — though it is not clear that this is anything but an independent development; 3) the ‘3’ changes to a vertical squiggle or zigzag — a change found analogously in other letters as well.



Like ‹s›, the next letter in order, ‹l›, consists of a main body and an adjunct. In this case, the main body is most commonly a double curl shared with many other letters in the script, and an adjunct written underneath that varies between a shape similar to an ‘S’ or the bottom part of ‘ʒ’ or ‘ƽ’ and a vertical squiggle or zigzag, similar to the change already described for ‹s›. In Palawan, and in rare variants in Luzon, an initial horizontal tick is added to the main body, changing it to the stereotypical shape of many letters on Luzon and most letters in the Palawan variety of the script.

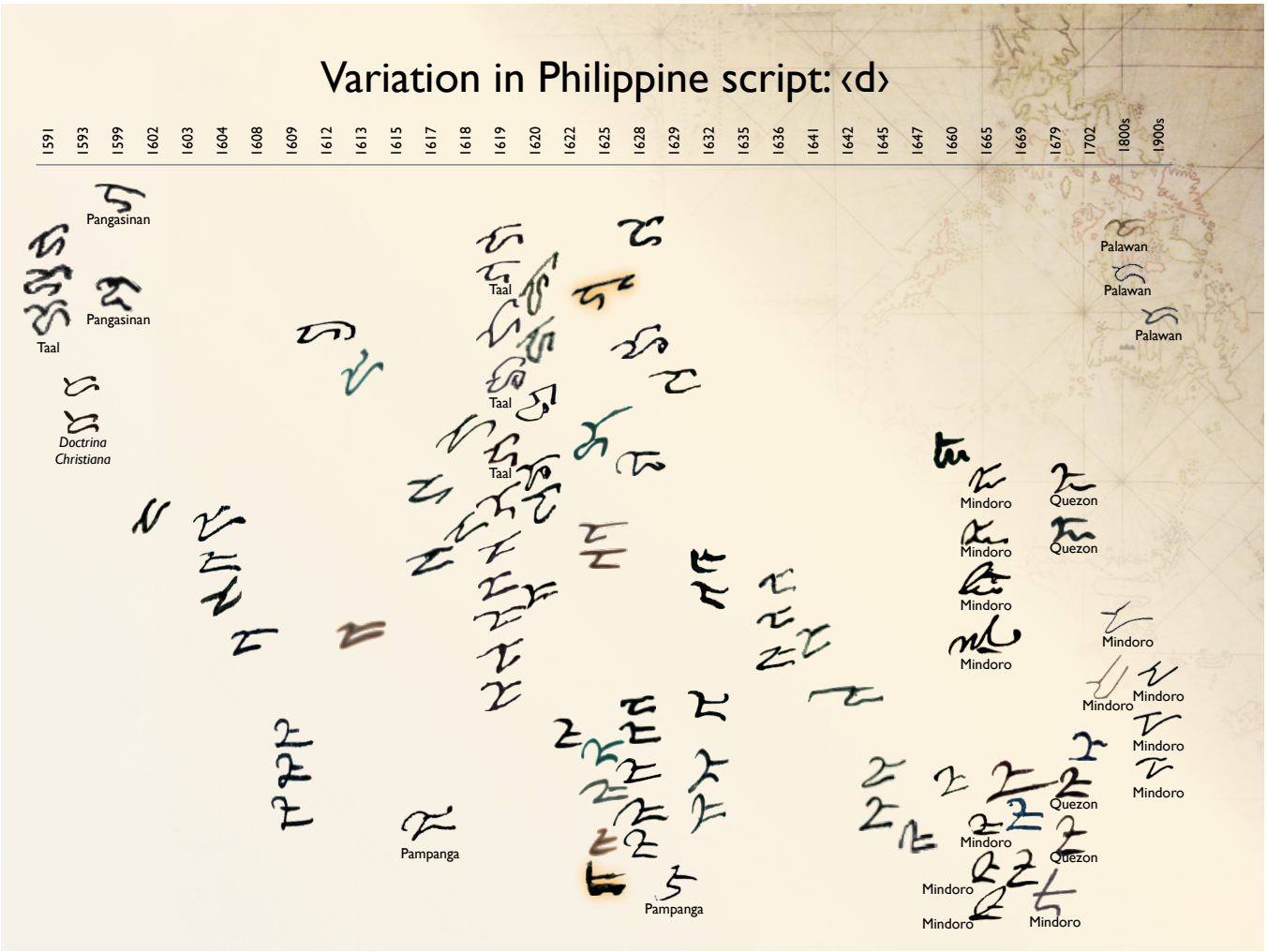
Over the century, the vertical squiggle commonly changes to a simple straight line, and the double curl simplifies in many hands to a ‘U’. In Pampanga, as with ‹s›, the ‘U’ is deleted, leaving only the straight vertical.



This letter has a rather unusual bipartite structure: a shape usually open to the left, with second shape either to its right or attached above it. The left-hand shape varies between something similar to a ‘ʒ’ and a simple curve open to the left. The second shape varies between a curve open to the left or a backward ‘ƨ’ or ‘ƨ’, and an arch above the letter often resembling a reversed “Nike swash”.

The simple curve plus right-hand curve or ‘ƨ’, most prominent around mid-century, is the most plausible antecedent of the zigzag shape found in 19th-21st century Mindoro scripts. Pampanga appears, starting with the arch variant, to have eliminated the initial portion of the letter, analogous to its treatment of ‹s› and ‹l›, leaving only the arch.

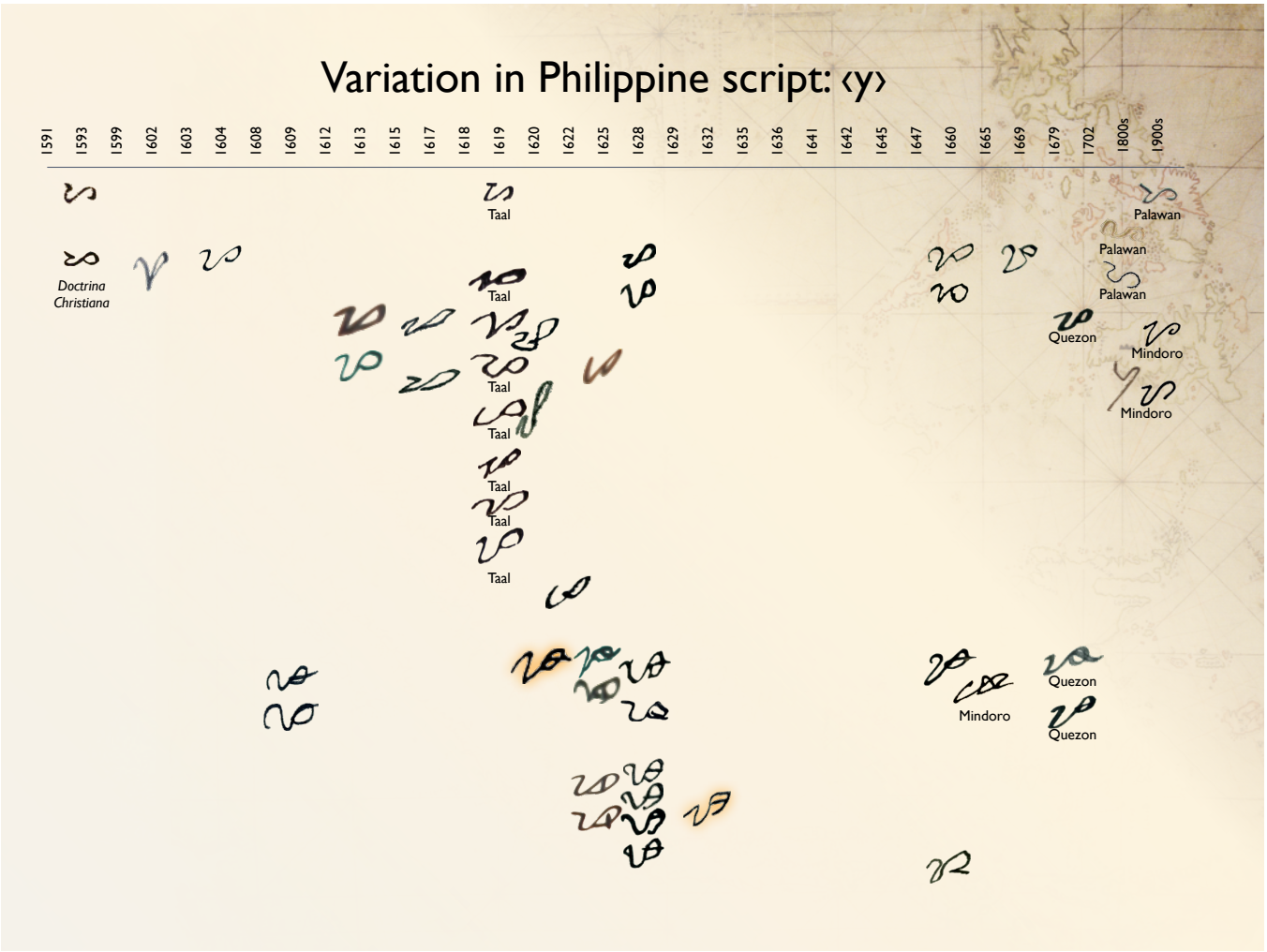
Palawan shows an arch with either a ’ʒ’ or a vertical squiggle or zigzag below it, closely related to the variants in the lower two-thirds of the time map.



By far the most common letter in the archival manuscripts is ‹d›. This is simply the fortuitous result of the fact that the vast majority of the signatures in archival documents begin with the Spanish honorific “Don”~”Doña”. This letter has a straightforward two-part structure: A main body followed with a more-or-less horizontal adjunct stroke extending rightward from the top left portion of the body.

In the earliest hands, the body has a backward ‘S’ or ‘ƨ’ that extends into a downward curve at its right end. Fairly early, the curl at the bottom left is replaced on Luzon by a sharp angle, yielding a ’2’-like body shape. Similarly, the downward curving or angled end of the adjunct simplifies almost universally to a straign and often quite short horizontal stroke.

As with many innovations taking place on Luzon, these changes are absent in Palawan, which conserves the earliest shape more or less intact, instead converting the top left curve to a short horizontal tick transitioning to a main body shape shared with most other letters in the script.

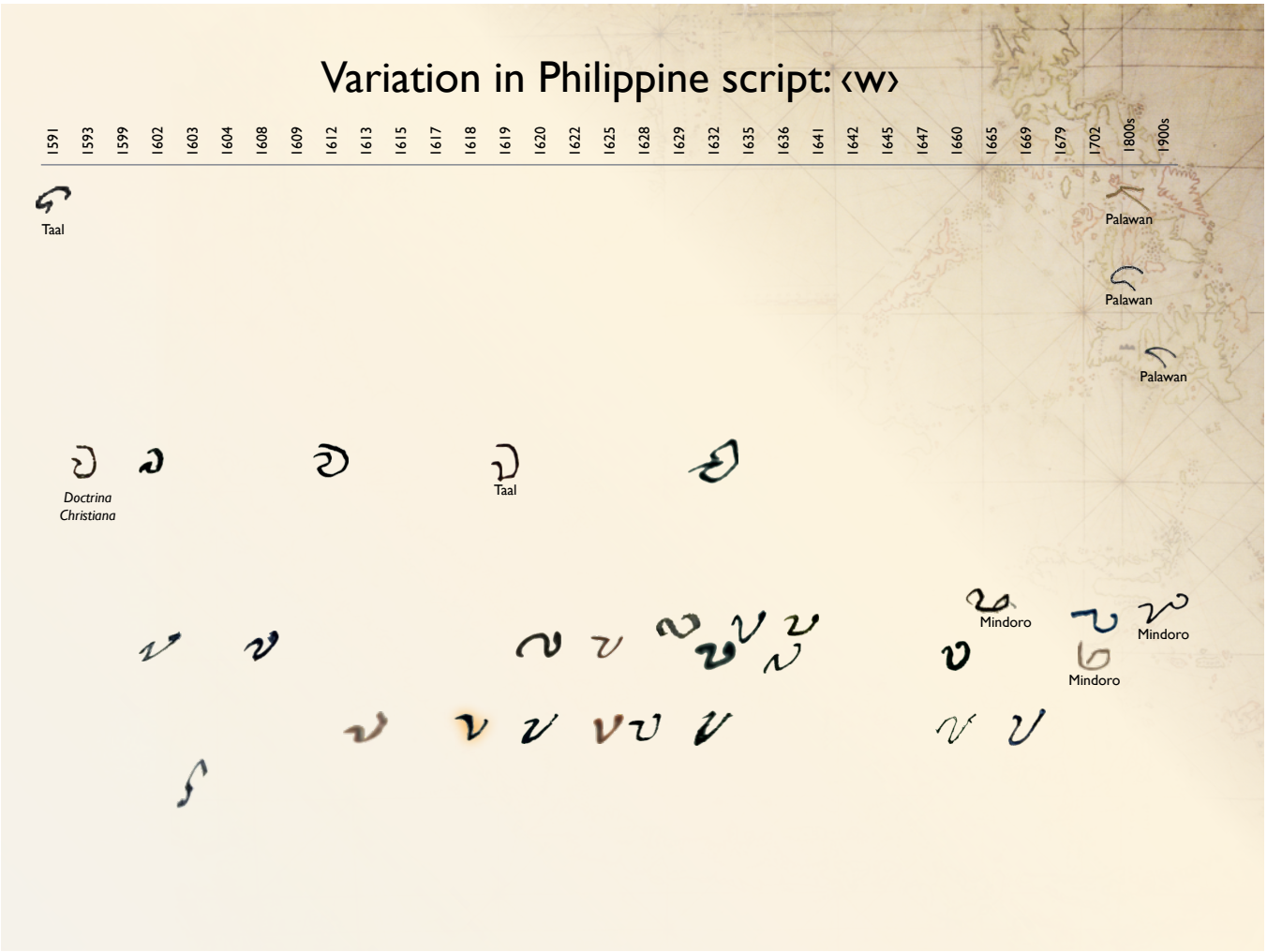


The second-to-last of *Baybayin*'s 17 letters, ‹y› shares the basic shape of ‹a› and several other letters, with the difference that the final tail curls around on the right to contact the right-hand side of tbe body (at least in carefully-formed writing), forming an enclosed white space ("counter" in typographic terminology).

In many variants, at least from the early 1600s, this extends to a further stroke that cuts through the counter and the right-hand side of the curl; alternatively, this added stroke segment is rendered as a separate short crossbar that cuts through the counter horizontally, diagonally or vertically, or even extends rightward from the bottom of the curl just before it contacts the body of the letter.

The 19th-21st century Mindoro shape opens up the enclosed curl, changing it into an extended downward stroke (as it does with the originally 'ᜓ'-shaped ‹b›, turning it into a '7' shape). The Palawan shape, on the other hand, preserves the shape used in Luzon, but without the extra extension.

Although the Palawan shape does not use the final extension found fairly commonly on Luzon, it seems a plausibly hypothesis that the reason ‹y› is the only letter whose tail is extended to curve around and backward toward the body of the letter — with or without contact — is that this was originally a transition to the beginning of the final extended stroke that generally cuts through the extended tail.



The final letter in the *Baybayin* recital order, ‹w› has two different shapes. The one nearly universally used on Luzon begins with a short horizontal tick and curves downward to the right like many other letters, but rather than ending in a tail curving down to the right, it continues its counterclockwise curve upward and to the left. Over the 17th century, this simplifies to a shape similar to a cursive or italic Latin '*v*'.

On Palawan, and in a single hand from a 1591 document in Luzon, it instead takes the form of a stroke curving counterclockwise to the left, then down to the right and ending in a short downward tail. This had already changed to two separate curving strokes connected at an angle on the left in 18th- and 19th-century Palawan, and this has changed at the end of the 20th century to two separated '᜔'-like strokes with sharp downturns, one above the other.

Similarities to Gujarati and Devanagari scripts

- There is little evidence that a relationship with Devanagari script, let alone with its descendant Gujarati, has ever been seriously entertained in the literature.
- Striking correspondences appear by comparing the range of old Philippine lettershapes with shapes and related structural changes observed in early 17th-19th century Devanagari and Gujarati (Meherji Rana Library and Bombay University Library) and 21st century handwriting samples (Indian Type Foundry).
- This discovery is unexpected, but it is known that Gujarati traders were active in Sumatra and Melaka in the 14th-16th centuries, and traded cloth as far east as Sulawesi and the Moluccas, and acted as *shahbandars* (harbourmasters) in several ports of the Malay archipelago.
- The following table compares Philippine lettershapes with old Devanagari lettershapes, early Gujarati shapes as they developed out of Devanagari, and modern handwriting shapes produced by similar structural changes.*
- Certain Sumatran lettershapes were added for comparison to illustrate relationships that would not be apparent without Gujarati shapes as plausible antecedents.
- A set of reconstructed intermediate proto-script shapes are illustrated, continuing the plausible evolution of lettershapes from informal Devanagari, without the further developments typical of the Philippine script itself.

* The fact that the Devanagari and developing Gujarati lettershapes come from manuscripts dating to no earlier than the 17th century (and especially from the early 21st century) may lead one to believe that they are a poor source of comparative data for pre-15th century informal Devanagari. However, since these lettershapes were considered appropriate enough to appear in the formal context of the Avestan scriptures, this can be taken as a clue that they would likely have existed for a considerable length of time as marginal and then low prestige informal variants before they moved into the mainstream, displacing the older variants. It is quite conceivable that it would have taken at least a couple of centuries for this to happen, which would place early informal variants in the appropriate timeframe (late 14th to early 15th centuries).

This comparative table arranges the original Devanagari letters and by form class (rather than the traditional classes based on letters' phonetic values). Each class is based on two criteria: 1) the presence or absence of a vertical stem in a letter; and 2) the starting and end positions and direction of the stroke sequence for the main body of the letter. Under these are arranged representative examples of corresponding Philippine script letters and, below them for comparative purposes, Batak letters from northern Sumatra and southern Sumatran letters as well as two letters from the Makassarese *Jangang-jangang* script of Sulawesi. Comparing letters across scripts by form class brings out regular structural changes and correspondences that relate directly to the structural features characteristic of each class.

Overall, there is a consistent correspondence between the body shapes of the Philippine letters (and Sumatran counterparts) and the form of Gujarati letters as they developed out of Devanagari. These and more idiosyncratic correspondences are summarised in the next slide.

Abstracting away the above-mentioned idiosyncratic changes, we can reconstruct a set of intermediate shapes or "proto-script" by applying the most consistent and regular correspondences; this will prove useful later for understanding the relationships of other scripts of Sumatra and Sulawesi.

*Departing from traditional practice for Indic scripts, I represent each letter only by its invariant value, rather than with a following vowel /a/, which is supplied by default by a reading rule — which does not apply in all contexts in some modern Indic scripts.

Nāgarī and Philippine scripts: regular correspondences

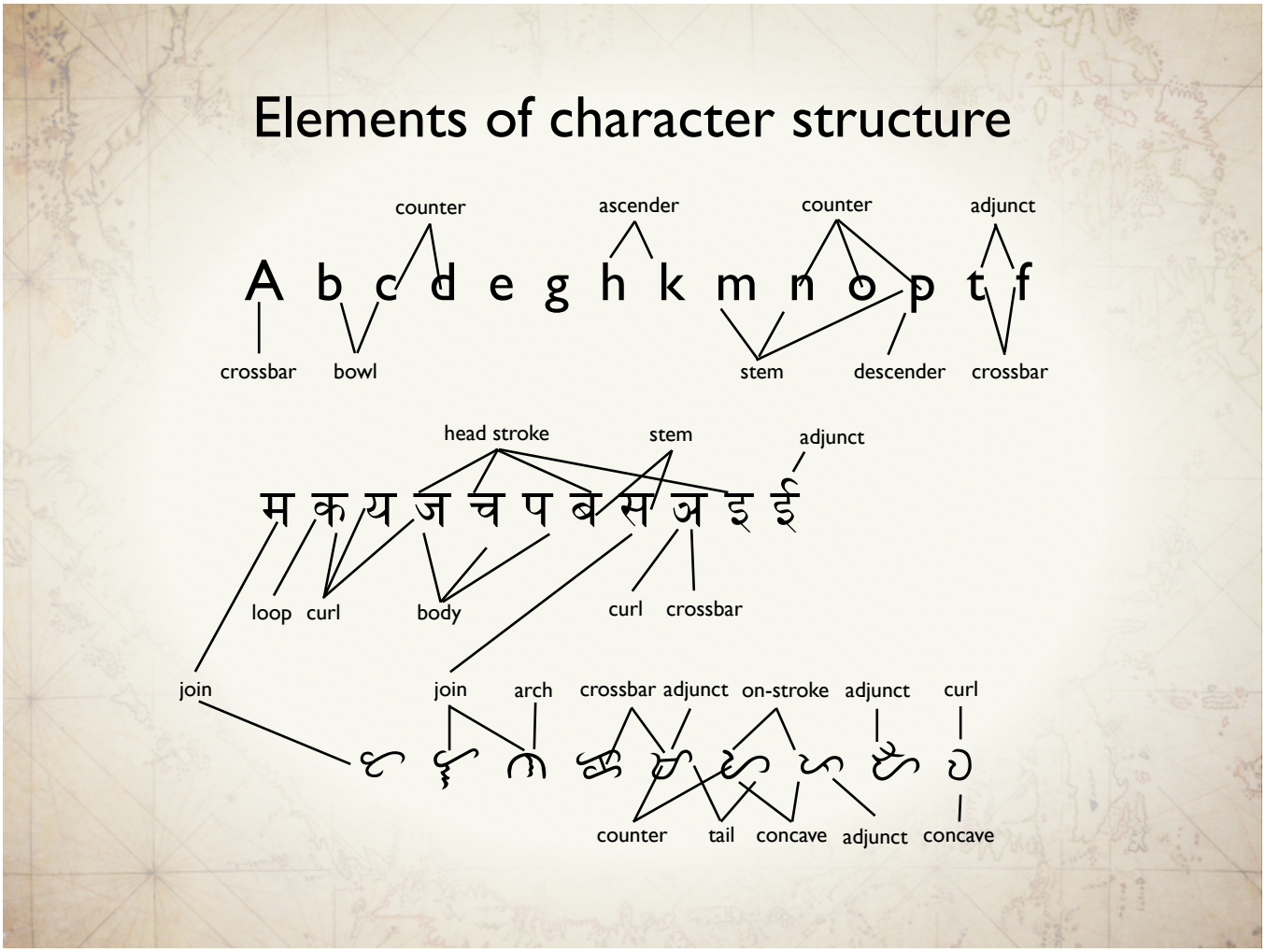
1. Body+stem join in formal Devanāgarī becomes looser and eventually a glide in informal Nāgarī.

2. Stem reduces to a short tail in Philippine script, usually downward but upward in ‹b› and ‹w›.

3. Base body shape corresponds closely and systematically between Nāgarī and Philippine scripts *for each form class*.

4. The Nāgarī headstroke is retained in Philippine ‹t›, ‹l›, ‹n›, curving directly into the short tail that corresponds to the Nāgarī stem.

5. Some Philippine letters undergo slight further changes:
   - ‹m› keeps the Nāgarī counter, but the loop shaped bowl of the base is reanalysed as the base body shape common to ‹a›, ‹p› and ‹y›, closed at the top by a crossbar.
   - ‹g› and ‹ng›, otherwise quite similar, undergo elaborations that increase their visual contrast.
   - ‹h› loses its final adjunct stroke while ‹p› gains one found in no other Indic script: a possible explanation for this later.
   - The initial three-stroke adjunct of ‹a› is directly reflected by the fork- or cross-like adjunct in the Philippine scripts: the Tagbanuwa position is probably the oldest.
   - The initial curl of Devanāgarī ‹y› is reflected in the (apparently displaced) extra curl that a appears as a “knot” at the end of some variants of the Philippine letter.



This slide illustrates a number of structural features of written characters that have been addressed informally in earlier slides.

Stereotypes

- One of the defining features of individual scripts is the presence of stereotypical structural elements that repeat themselves across a large proportion of a script's character inventory.
- These develop in the process of change from parent scripts, through a process of reanalysis of often unrelated earlier stroke combinations.
- The development of stereotypes may function to simplify the overall task of learning the character inventory of a script by providing easily learned basic shapes to which other graphomes (units of character structure) are added.

The idea of stereotypical body shapes in Philippine script has been raised earlier. This and the following slide expand on the notion.

Stereotypes

Latin script lower case: bowl + stem

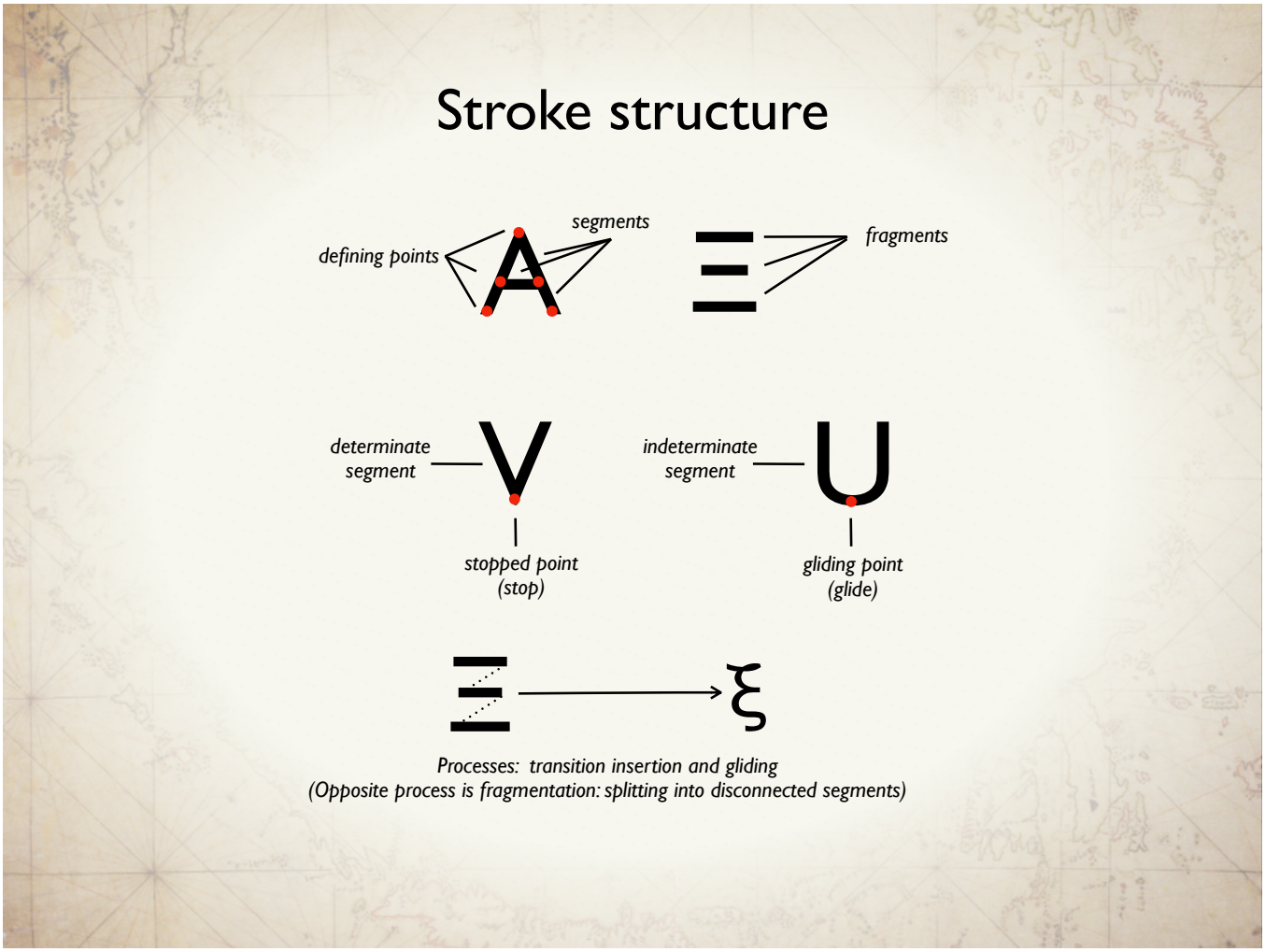
a b c d e f h j k l m n o p q r

Devanagari: headstroke + body + stem

म क य ज च प ब स ञ इ ई

Old Philippine script: (onstroke +) concave/curl + tail

Examples of stereotypes in three scripts.



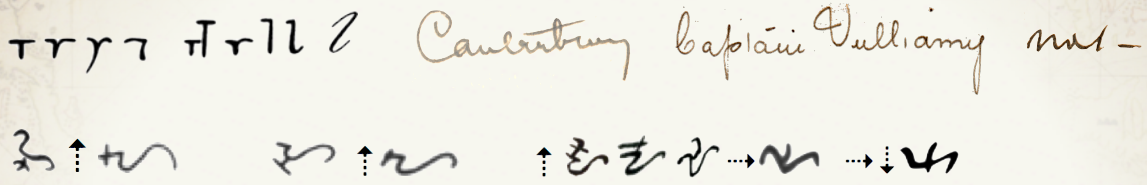
Below the level of graphomes, individual stroke segments can be described in terms of their connectivity or lack there of, and whether the transitions between segments are angular, with a clear-cut stop in of the writing implement before a change of direction, or curved, with the implement gliding through an approximate region or point as it gradually changes direction.

Varation and/or alternation between stops and glides is a very common process occurring across scripts.

Adjunct drift

• Greek tau and cursive variant (lower vertical stoke is adjunct)

• Adjunct bar on t, dot on i tend to drift rightward in handwriting:
"Canterbury" "Captain Vulliamy" "not"



Loop excrescence

• Instead of lifting the writing instrument, the writer continues beyond a stop and loops it back around to cross back over the preceding segment

Fusion and splitting at intersection points

• Distinct stroke segments join to form a single segment
• A single segment with an intersecting point defined by another segment splits at that point

Y y T *m* E ε

1) Latin, Cyrillic 2) Cyrillic 3) Greek

33

Several chirographic processes already observed in Philippine script (and observed later for other scripts) are illustrated here for more familiar scripts.

The discussion of adjunct drift for the evolution of lower case Greek *tau* and as a sporadic process in Latin script handwriting is followed by a recapitulation of the similar process as it applies to ‹a› in Palawan (first two examples) and in Luzon (right-hand sequence).

Loop excrescence is behind the development of lower case Greek *gamma* as well as cursive variants of Latin script letters.

Processes observed across scripts

- Rotation
- Reflection/mirroring
- Segment reordering
- Point reordering (direction change)
- Stopping/gliding
- Loop excrescence
- Fusion and splitting at intersection points
- Fragmenting at intersection points
- Transition insertion/excrescence
- Segment shortening
- Segment lengthening (swashing)
- Segment deletion
- Place assimilation
- Initial, final excrescence (serif formation)
- Adjunct drift
- Counter preservation
- Skewing (cursive)
- Stereotyping (reanalysis by analogy)

Each of these processes, found commonly in other scripts, can be observed in the scripts under discussion here or posited as an explanation for structural differences between otherwise similar corresponding letters in two scripts under comparison.

Intermediate origin of Old Philippine script?

- Philippine script was likely introduced indirectly via speakers of Bugis or Makassarese from South Sulawesi:
  - Old Philippine script did not spell coda Cs: it had a ‹$C^{(V)}$› syllabographic template.
  - This is the same as Bugis-Makassarese script, which also uses a ‹$C^{(V)}$› syllabographic template. This is natural for these languages because the range of coda (syllable-final) consonants is limited and relatively predictable:
    {-$N_{homorganic}$ /_$C_{homorganic}$, -ŋ/_#}
    {-C_{α} /_C_{α} , -ʔ/_{#,b,d,g}}.
  - Fox (1979): only plausible for the script to have this feature if it were adopted from Sulawesi users, not from other Indic scripts with a full complement of signs for spelling coda consonants overtly.
  - in Tagalog and other Philippine languages, there is a wide range of coda consonants and it is difficult to guess which one should appear in a given case; there would be no plausible functional motivation for not spelling these out overtly if the script already allowed for that.

This argument from orthographic structure, often known as the “Fox-Conklin hypothesis”, is one of the most convincing pieces of evidence for a specific origin for Philippine script.

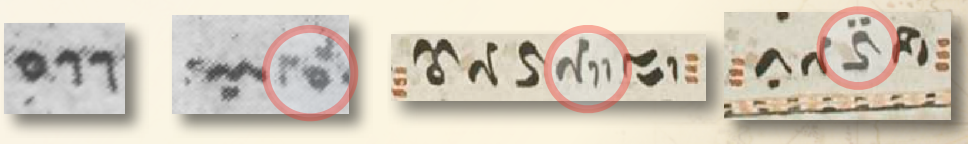
Interestingly, writers in Pampanga, a province northwest of Manila, had by the early 1600s settled on a somewhat makeshift and partial set of conventions for writing coda consonants while the rest of Luzon together with Palawan, Mindoro and the Visayas, continued to omit codas in spelling.

Vowel sign doubling

- Philippine and South Sulawesi scripts also share a unique abbreviation found only in these two scripts: two syllables with same consonant were often spelled with a single consonant letter bearing **two** vowel marks.

Bugis-Makassarese script

Makassarese *Jangang-jangang (bird)* script



soso 'asesĕ mapalapoporo tupalilika

(By permission, Noorduyn 1993)

(By permission, KITLV Tropenmuseum 668-216)

Data that have become available since the Fox-Conklin hypothesis was first proposed reinforce the evidence for a close genetic relationship between Philippine script and an antecedent from south Sulawesi.

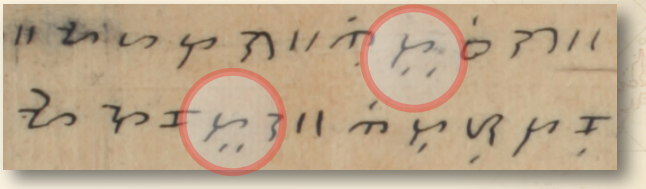
In a seminal 1993 paper on early variation in Bugis-Makassarese script, Jacobus Noorduyn noted (along with numerous observations on early variant letter shapes) the widespread phenomenon of vowel sign doubling. This could occur where two neighbouring syllables shared the same onset consonant, as long as both had vowels spelled with overt vowel signs.

As in other Indic scripts, the vowel /a/ is read off consonant letters by default and requires no overt vowel sign. As a result, vowel sign doubling never involves syllables with /a/ in BM script. In *Jangang-jangang*, this limitation was eventually overcome by borrowing the Arabic numeral ‹2› (illustrated below) to stand in for the second consonant and placing any overt vowel sign on this or the preceding consonant letter as appropriate.

Arabic numeral ‹2›: ٢

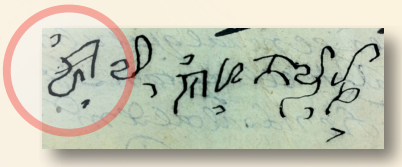
Vowel sign doubling in Philippine script

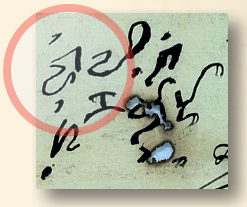
Nitong **tu**bigan

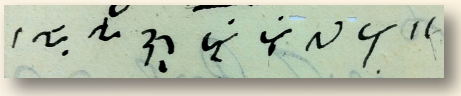


Ang pagka**tutu**o

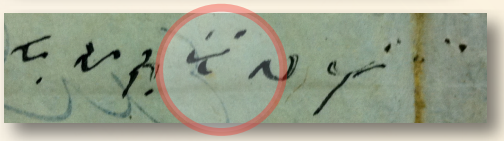
Don Dionisio Capolong







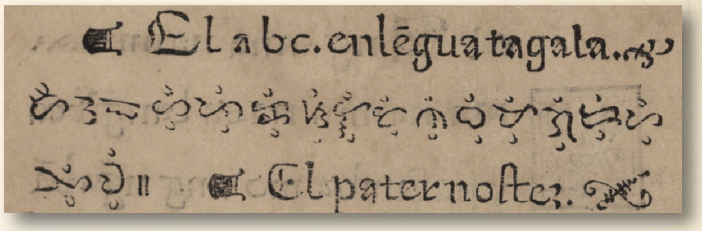
Don Agus**tín Ti**ualag



(University of Santo Tomas
Miguel de Benavides Library, Archives)

Although the evidence is much more sparse, it is clear from the handwriting of at least three different individuals in Philippine archives that the same convention, otherwise found in no other Indic script beyond Sulawesi, was known and used in early 17th-century Philippine script. Already noted by Villamor (1921) and Santamaría (1938), the relevant examples can be seen clearly in recent photographs of texts from the University of Santo Tomas archives, taken in 2011.

Didactic vowel sign combination: *Doctrina* and Tagbanuwa



Gardner 1943

Marcilla y Martín[88] says, "The expression or manner with which the Tagbanuas, when they read their alphabet, is to name with every sign or letter its three terminations in *a*, *e-i*, *o-u*, pointing with the finger on naming the termination in *e-i* the mark to the left; pointing with the finger to the left when naming the termination in *e-i*. When it is in *o-u*, he points to the right. If the reader will take any of the Tagbanua alphabets and hold it so that the vowel points are on the right and left sides, he will see that the writing is vertical." The same writer continues with the words of the chant or reading aloud of the alphabet—

"Gangairi Gangairo
Dandairi Dandairo
Kankairi Kankairo
Banbairi Banbairo."

Marcilla y Martín 1895

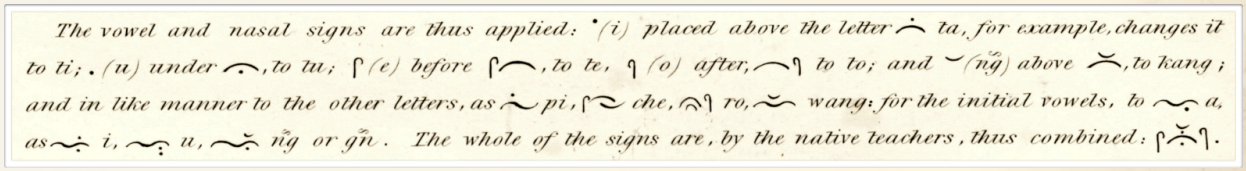
Conklin (1991), fieldwork observations:
"After pronouncing the three vocalic signs, a common way of reading this exercise aloud was to point to each basic or diacritic sign while chanting as rapidly as possible, *langláylu'-langláyli', mangmáymu'-mangmáymi',* etc., ending with *ngangngáyngu'-ngangngáyngi', wawa'*."

The most likely origin for vowel sign doubling can be found in the didactic combination of vowel signs on consonant letters. This is attested both for Luzon and for Palawan in the Philippines.

In the first example, from the 1593 *Doctrina Christiana*, the 17 Philippine letters are presented in conventional recital order (a u i h p k s l t n b m g d y ng w) followed by the double bar punctuation mark, and with the two vowel signs combined with each consonant letter (‹-i› above, ‹-u› below).

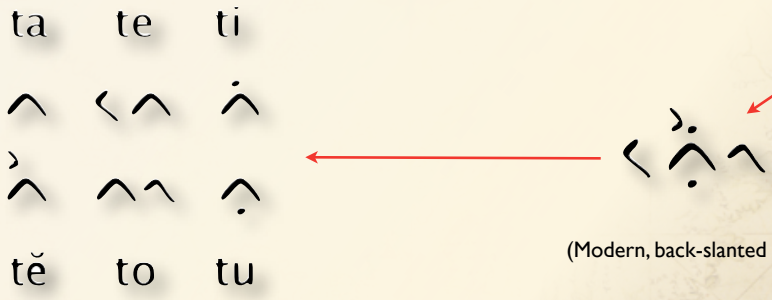
A similar didactic practice attested from Palawan by Marcilla y Martín (1895) and Conklin (1991) provides evidence that the visual combination of vowel signs on consonant letters in the 16th century was likely accompanied by a spoken recitation reading the vowel signs in combination with the letters. It is quite possible that the recitation was similar in form to the variants given by Marcilla y Martín and Conklin; however it is also possible that the signs may also have been read off in a simple fashion, one after the other, off an individual consonant letter. Thus, the ‹$^{i}d_{u}$› of the second-last CVV combination in the first line of "El abc. en lengua tagala" from the *Doctrina Christiana* could well have been read as "di du" or "du di" — the second reading being directly reflected in the initial abbreviation ‹$^{i}d_{u}$› in the signatures of Don Dionisio Capulong in the preceding slide.

Vowel sign combination in South Sulawesi



The vowel and nasal signs are thus applied: ˙(i) placed above the letter ta, for example, changes it to ti; ˍ(u) under, to tu; (e) before, to te, (o) after, to to; and ˘(ng) above, to kang; and in like manner to the other letters, as pi, che, ro, wang: for the initial vowels, to a, as i, u, ng or gn. The whole of the signs are, by the native teachers, thus combined:

Marsden (1834)



(Modern, back-slanted style)

Didactic use of ‹ĕ› in Makassarese to represent -N: Bugis has /a e ĕ i o u/; Makassarese has no /ĕ/. Possibility that this and the Tagbanuwa recitation are derived from "CaN ce Ci$^{?}$ C{a/ĕ}N Co Cu$^{?}$"?

William Marsden (1834) describes in abbreviated form a didactic practice similar to the one we have seen for the Philippines, with the difference that Bugis-Makassarese script uses five signs in total. Marsden describes the Makassarese use of the curved sign above a letter to represent ‹-ng›. This was essentially only used as a teaching aid for Makassarese, and not in everyday Makassarese writing. In Bugis, on the other hand, it is regularly used and stands for the mid-central vowel ‹-ĕ›.

If the letter-plus-vowel-signs combination given is read by supplying the Bugis values, beginning with the default /a/ and then reading the vowel signs off clockwise from the left, the recitation for the illustration provided by Marsden would be "ta te ti tĕ to tu", under the assumption that the ‹-i› sign, closer to the letter, would be read before the ‹-ĕ› sign that is written further out. This sequence clearly resembles the Palawan Tagbanuwas' recitation as it would be read for the corresponding letter: "tangtáyti$^{?}$ tangtáytu$^{?}$". Without going into the details here, properties of the Bugis and Makassarese sound systems in Sulawesi, and of Tagalog and Tagbanuwa in the Philippines, make it highly plausible that such a recitation was transmitted from Bugis and/or Makassarese speakers to Tagalog speakers, likely in the Manila are, and thence to Tagbanuwa speakers in Palawan — further evidence for transmission of the script from south Sulawesi.

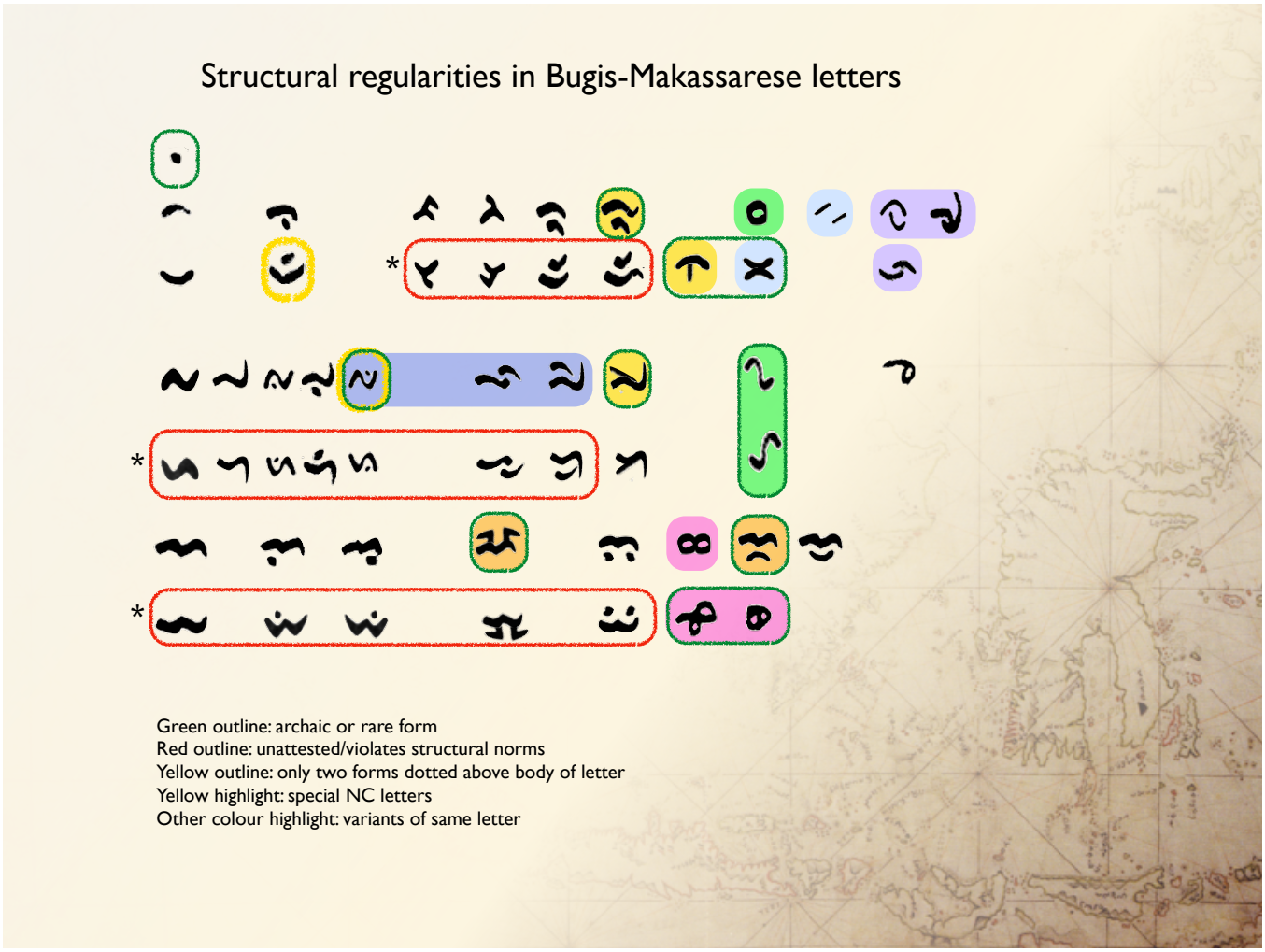
Bugis-Makassarese script has a remarkably minimalist form: Letters consist mainly of arches and rising strokes plus dots located (usually) underneath the arches, as well as flag-like adjuncts above certain letters and curved adjuncts above or below more complex letter shapes. The documents illustrated here display several different styles of BM script.

Top left: this book uses a style typical of the late 19th century and later. The back-skewed style is not found in texts datable to earlier periods, including those illustrated here and the lettering on the Bugis sea chart used as the background image for this presentation. One significant feature is that the arching "flag" adjunct has drifted to the right-hand side of the ‹b› letter (red circle), compared to its position atop the arch in the other three texts and standard Bugis type, which is based on mid-19th-century models.

Mid-left: this silver *kampu* box bears an inscription apparently made by beating an instrument with a rounded tip into the surface of the box. Apart from the peculiarities in shape due to the constraints of the writing instrument, the shapes of the letters are fairly consistent with the standard developed in the mid-1800s but for the ‹a› (green circle), in which the dot appears under the left-hand arch, an archaic feature from before the mid-1800s noted in Noorduyn (1993). The standard position under the right-hand arch — likely due to adjunct drift — can be seen in the top left and top right documents.

Bottom: this is an example of a method of writing on narrow strips of palm leaf sewn together in long ribbon-like strips apparently unique to Sulawesi but with possible antecedents in Borneo mentioned in an old Chinese chronicle. "Palm leaf" style (as named by Noorduyn) has a number of peculiarities compared with other styles: not only is the orientation of the letters resolutely vertical, but several letters have shapes rarely if ever seen in other kinds of texts. The ‹j› (orange circle) takes the form of three dots arranged in a triangle, unlike the standard shape (top right). The ‹a› letter (green circle) is a simple dot — compare its more elaborate form in the other three documents. As for ‹l› (blue circle), it has a dot inside the upper open counter of the right-hand side of the letter where it normally has an arch over the right-hand side (compare with the other three documents). This is an anomalous violation of the normal constraints on character structure described in the next slide.

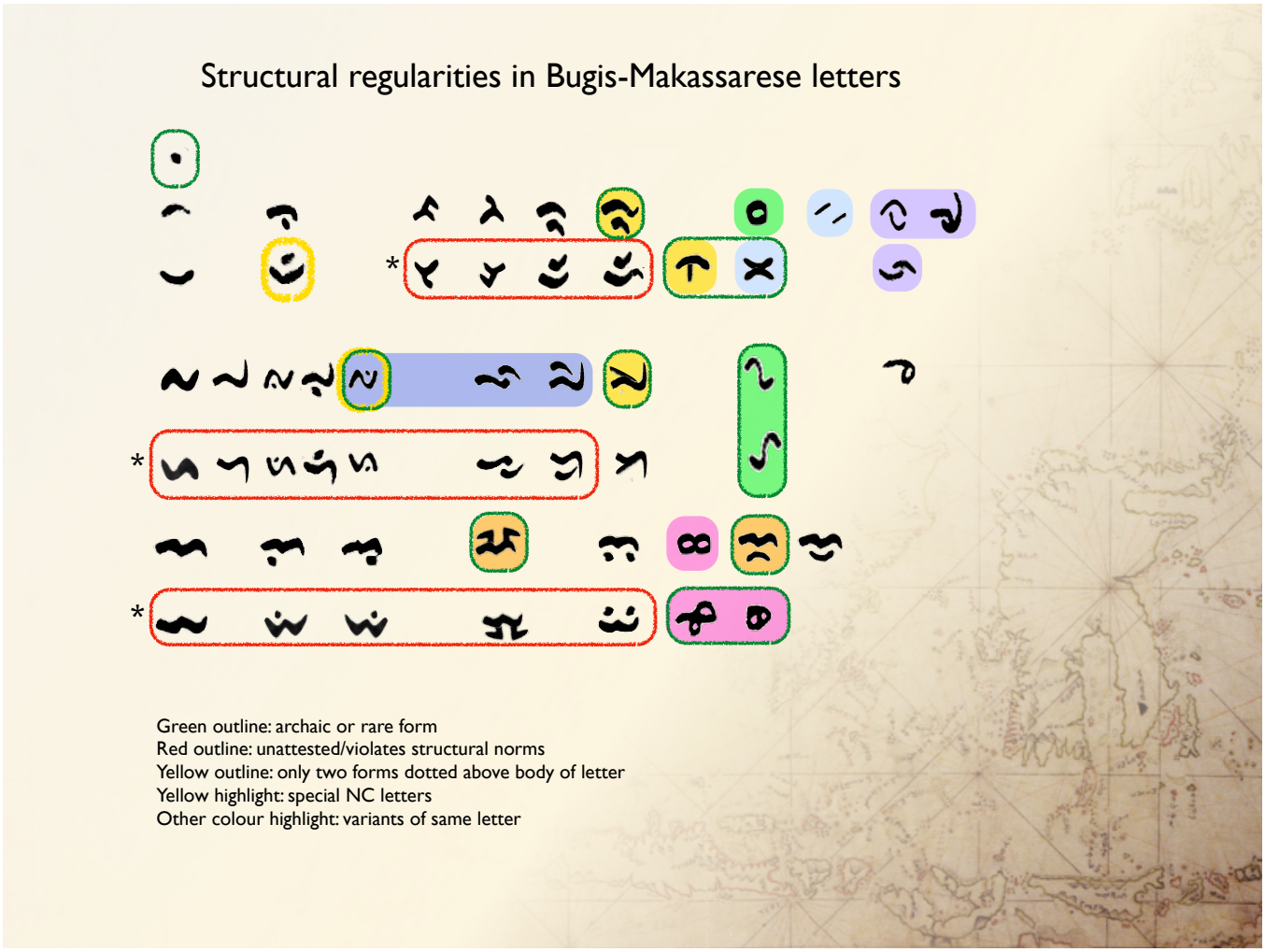
Top right: this is a fairly old manuscript, though the date of the specific text illustrated (the ms contains several) is not known to me. Although the letters are fairly close to the mid-19th-century standard, they lack the back-skewed feature characteristic of that period and, in particular, the extended upward-backward swashes on final upstrokes. In this, it is quite close to Palm-leaf style.



Examining regularities in the sound systems and grammars of spoken languages often helps to reveal patterns and anomalies that aid in reconstructing earlier stages of the language. Applying the same methodology to the structure of a script's inventory reveals some interesting asymmetries in the structure of BM letters that help to trace and ultimately confirm the hypothesis of its close relationship (despite the massive superficial differences) to Philippine script.

This slide examines most of the variant letter shapes found in different varieties of BM script, along with logically possible shapes constructed from similar elements but not otherwise attested. Two main patterns emerge from these data. First, letters not surrounded by any coloured outline (other than yellow) are the core inventory of the script; these are supplemented by letters outlined in green, which represent extra, rearely used letters, and variant shapes otherwise only rarely attested. Second, vertical mirror-images of these letters, outlined in red and preceded by an asterisk, are unattested: although logical possibilities, they appear to be excluded by structural constraints that appeared during the historical development of BM script.

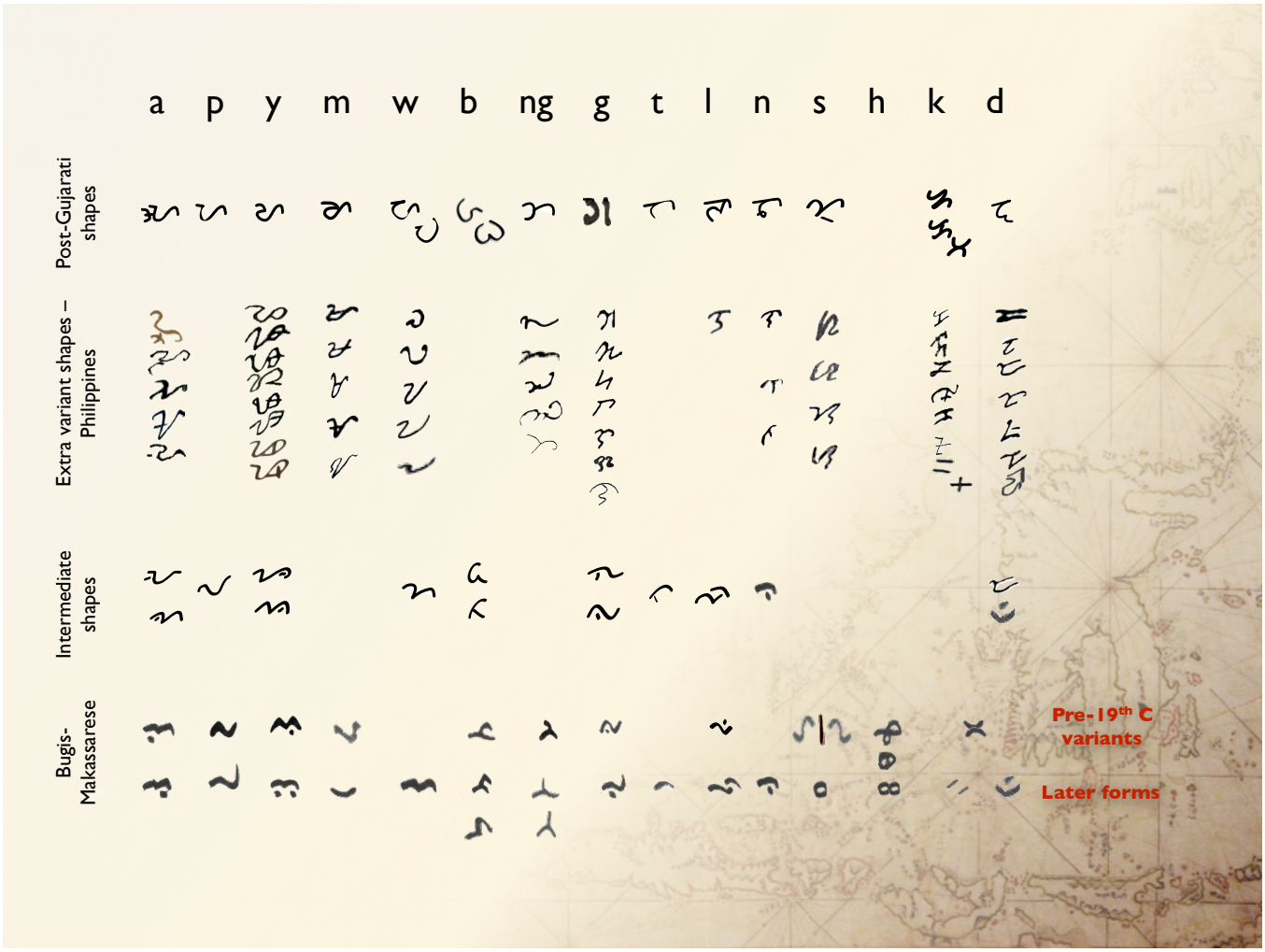
The first main pattern that emerges is that the preferred template for the body of letters is an arch on its own, followed by a final rising stroke, or a second arch. This arch is the core stereotype of BM script. In particular, their mirror image, an initial trough, is strongly dispreferred, being found only in two standard letter shapes (‹m› and ‹d›, beginning of the second full line) and one nonstandard variant of ‹c› (top right, lavender background) as well as a rare archaic variant of ‹k›, the 'X'-like trough-over-arch highlighted in pale blue on the second full line. More significantly, no trough shape is followed by a final *downward* stroke or a second trough: the structure of letters is biased in one direction along the horizontal and vertical dimensions.



The second main pattern is that dot adjuncts are normally found underneath arches, not above troughs. Apart for the defunct palm-leaf style variant of ‹l›, first of the three variants with a lilac highlight in the third full line, the only exception in the standard script is ‹d› in the second line. Third, a “flag” adjunct only appears above an arch, not below an arch or above or below a trough. However, arch or trough adjuncts show greater freedom and can appear either above or below the body of a letter. ‹s›, highlighted in green in the first line, is written as an arch with a trough directly attached to its underside.

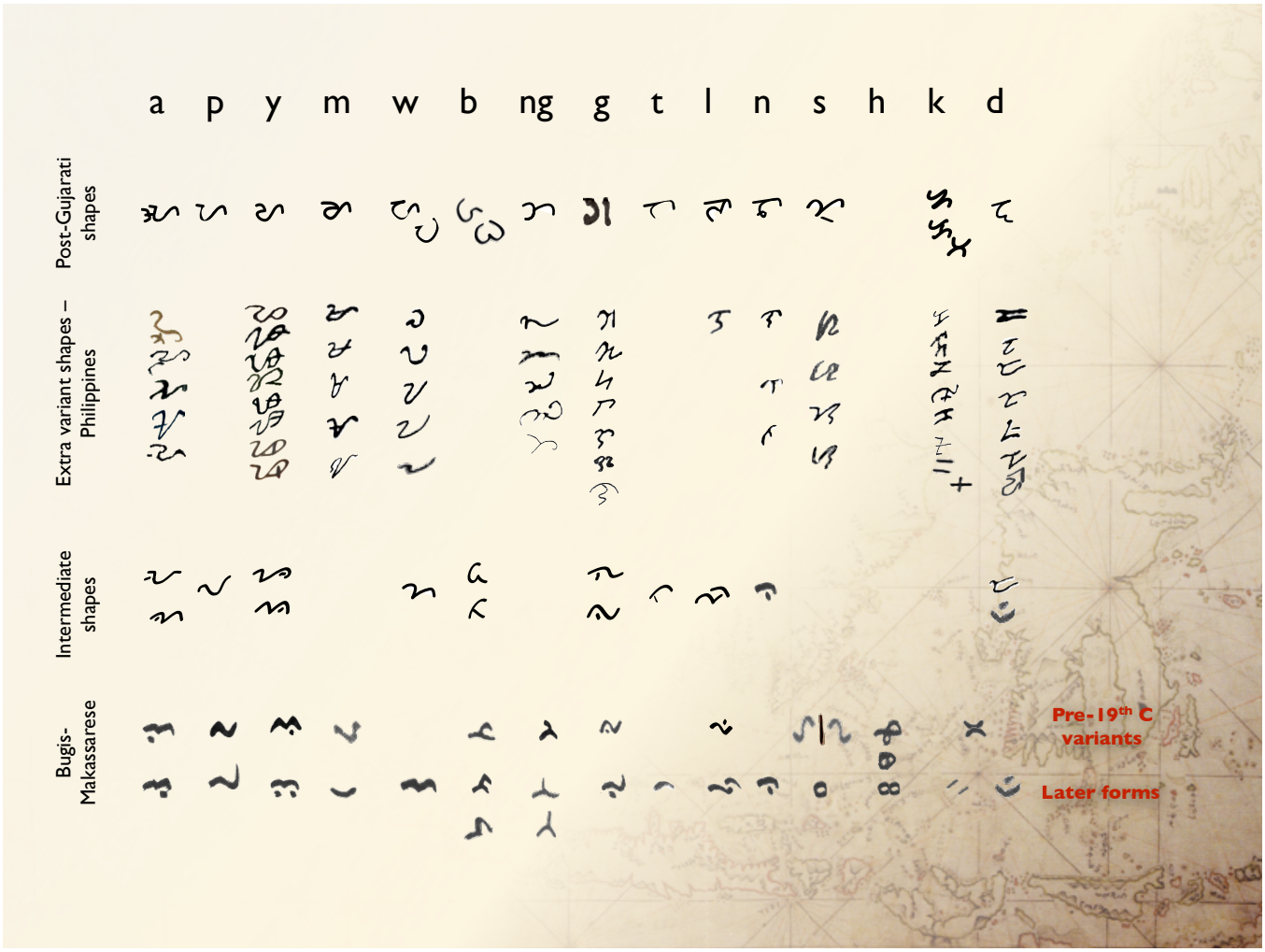
Two kinds of letters are to be set apart as later additions to the script. Those highlighted with a yellow background are special, rarely-used letters used to represent nasal +consonant sequences. As shown three slides below, these are almost certainly direct and/or structural borrowings from South Sumatran script models. Those highlighted in pink, variants of the letter ‹h›, are historically established as a borrowings from Arabic to wrote loanwords with /h/, a sound not originally present in South Sulawesi languages. The two at bottom right are actually two of several shape variants directly borrowed from Arabic script; the “lazy 8” shape above the leftmost of the two is the modern variant, which has conformed to the arch-trough stereotype that forms the core of the main body of most letters in the script, while preserving the two separate enclosed counters of the Arabic letter.

The core observation that letters tend to conform to a stereotypical shape beginning with an arch, with dot adjuncts confined to the counters of arches and flag adjuncts to the top of arches, leads to the hypothesis that attested letters conforming to these graphic structure constraints are probably modified versions of letters shaped somewhat differently before these constraints began to affect the latter-day form of the script.



Given the several strands of evidence that Philippine script was adopted from speakers of Bugis and/or Makassarese, it stands to reason that the forms of the Philippine letters would derive from shapes originally used by South Sulawesi literates. As we have already seen, these shapes are themselves simplifications and modifications of original informal Devanagari shapes as they developed in Gujarati script. It stands to reason, then, that the even simpler attested shapes of BM letters cannot be the antecedents of Philippine letter shapes. We should therefore examine the opposite possibility: that the Philippine shapes or intermediate reconstructed proto-script shapes represent the antecedents of the simpler shapes of BM letters as they appear in surviving texts.

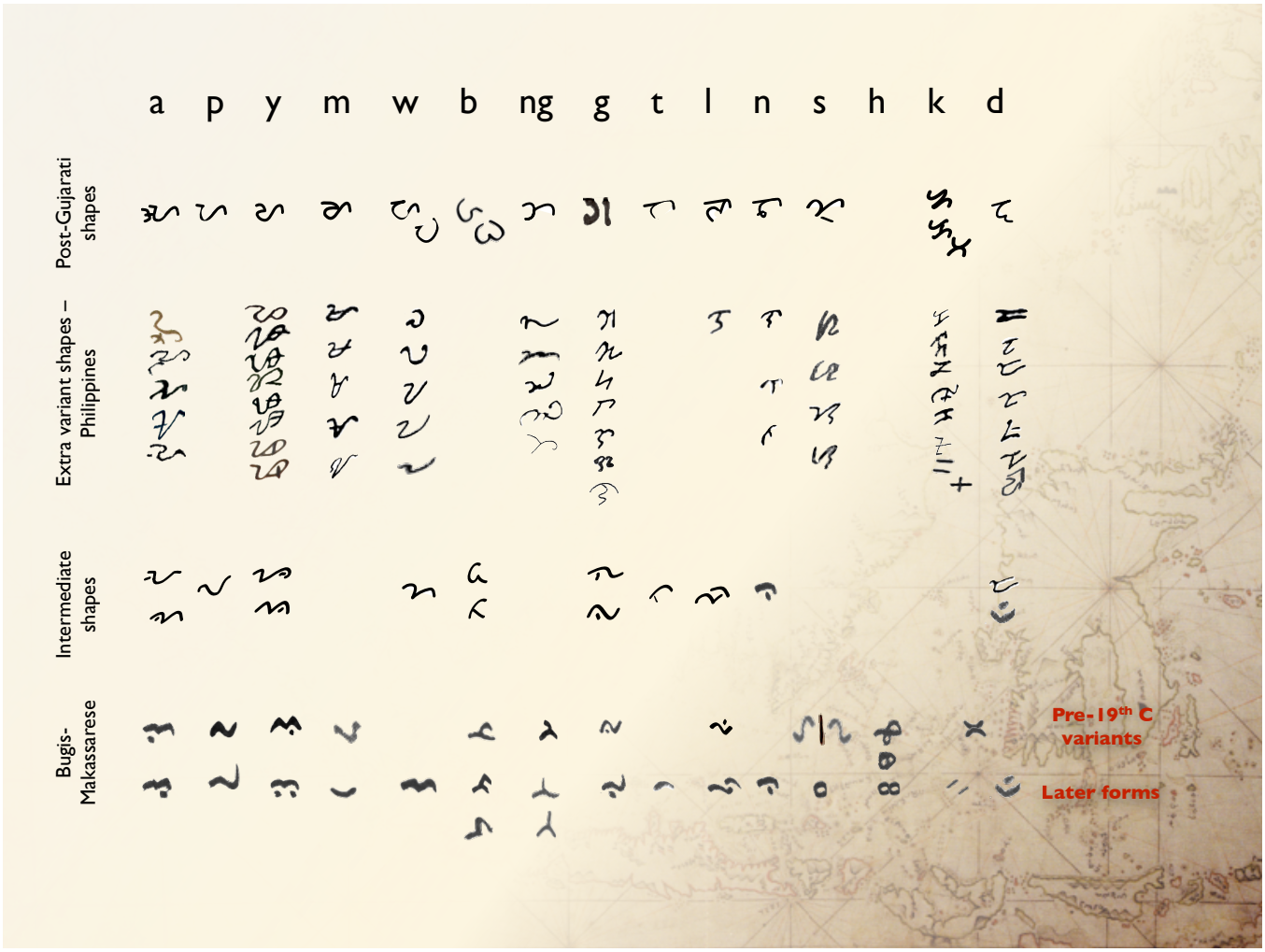
Two observations can be mentioned to begin. First, there is a certain tendency among some writers in the Philippines to reduce adjuncts to very short strokes that are often detached from the main body of their letter. This is an example of latent variation that can appear or reappear at different stages of a script's development. It stands to reason that this may well be the origin of the dot adjuncts in BM letters. Second, in connection with this, we can see that BM dots tend to be placed in positions corresponding to the adjunct in the corresponding Philippine letter. Clear examples of this are ‹a›, ‹n›, ‹d› and (possibly) an older form of ‹y›. If we allow for the possibility of reanalysis of a shape like the Philippine ‹g› along the lines of the Palawan Tagbanuwa shape with its continuous top stroke, then the resulting curved or squiggly adjunct on the underside, if reduced to a dot, corresponds directly to the position of the dot in the BM letter, with a main body structure analogous to several counterpart Philippine variants.



These two observations are complemented by the observation that the typical basic shape of many Philippine letters often takes on a pointy-shaped bottom rather than the standard curving trough-shaped bottom (similar to the relationship between Greek Υ and υ, but taken in the opposite direction as if the minuscule had changed into its capital counterpart). The left and/or right branches then take on a curved, extended arch shape. On the assumption this kind of variation was present in the earliest BM handwriting, the arch-point-arch variants would be the natural antecedents of the arch stereotype typical of the stages of BM script for which we have direct evidence. This is analogous to the change from older Greek Ω to the lower-case ω, but involving the vertical mirror images of the relevant letter shapes.

Returning to the question of adjuncts, we can observe that the non-dot adjuncts on BM ‹b› and ‹ng› correspond to an initual curved stroke in the reconstructed post-Gujarati protoscript shapes: counterclockwise for ‹b› and clockwise for ‹ng›. The arch underneath these BM adjuncts corresponds to the transition from the initial proto-script curve to the final tail. We can say for these letters that this transition was analogically reinterpreted as the arch stereotype, which thereby became the main body of the letter, and the previously initial curve was reinterpreted as an adjunct atop the letter's arch. The reanalysis in both cases is comparable to the one involving the relationship between Latin script Y and y.

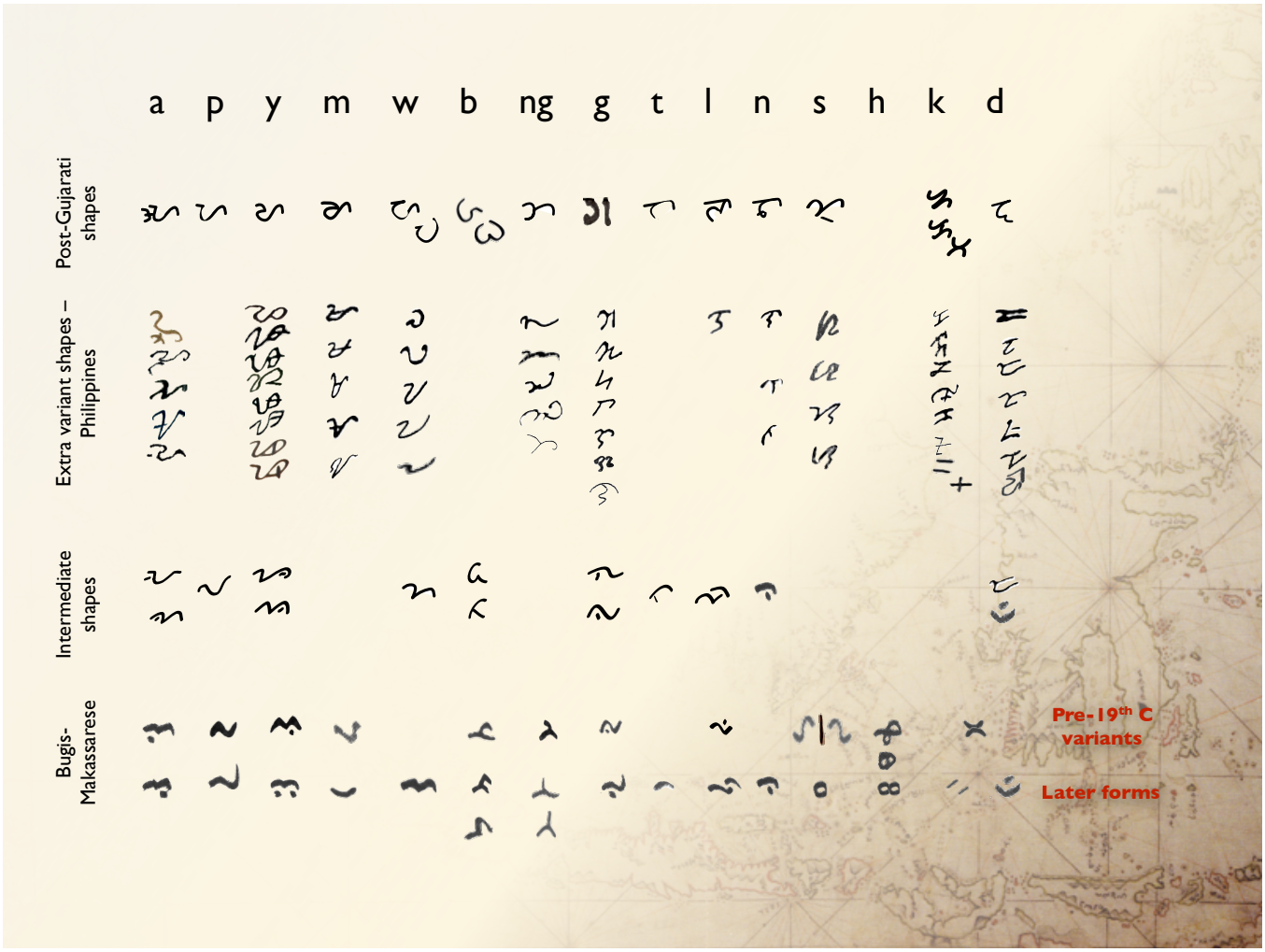
This brings us to BM ‹l›, which has the appearance of a vertically "squashed" Devanagari ‹l›, where the body of the original Devanagari letter (as in the proto-script letter as well) easily shifts to the arch-plus-upstroke of the BM letter, and the headstroke-plus-tail reduces to an arch adjunct. Interestingly, this adjunct reduces to a dot in palm-leaf style (upper image in the pre-19th century variants), corroborating the correlations obtaining with Philippine script letters, but otherwise the upper dot is excluded by the general constraint that they normally should appear under arches.



A related case is ‹t›, where the proto-script headstroke-plus-tail again corresponds to an arch in BM script, but the curl beneath is absent. Normally, it should have reduced to a dot as in ‹n› — however, that would have led to two different letters with identical shapes. It stands to reason that one of the two dropped the dot — ‹t›, as it turns out — to avoid conflict and ambiguity, and because there was no other letter with a simple arch shape.

This leads us to another dotted/dotless pair: ‹m› and ‹d›. These are puzzling because the proto-script (and Philippine) shapes ought normally to have changed to a double arch or arch-plus-upstroke like most other letters. However, the (anomalous) simple trough shapes of both BM correspond to a similarly anomalous feature of the corresponding Philippine letters: these are the only two (before the 17th-century adjunct drift in ‹a› took place) with an adjunct placed in or above the trough portion of a letter's body. In both letters, the adjunct serves to enhance the visual salience of the counter — fully enclosed in ‹m› and semi-enclosed in ‹d›, turning it into the visual focus of the letter. (Note the way the crossbar on ‹m› preserves the counter focus found in the Devanagari/Gujarati and Sumatran looped forms.) This perceptual focus on the trough counter in both letters, as opposed to other letters where the adjunct is on either side of the letter, is a likely factor that would counteract the general analogical tendency to conform to an arch-initial stereotype.

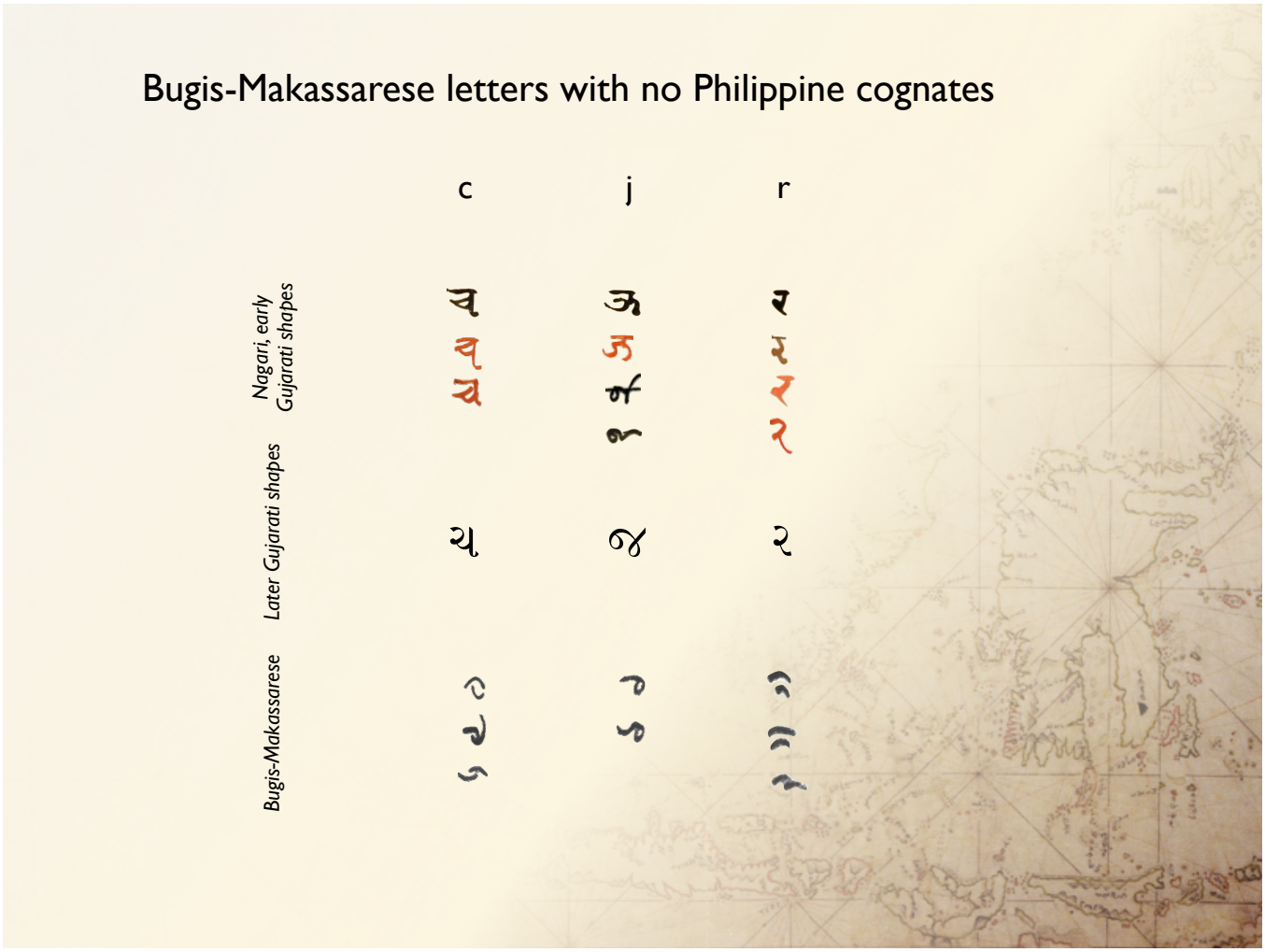
Simplifying these two letters into single trough shapes, with the adjuncts reduced (as expected) to simple dots, leads to the same problem of confusability as in the ‹n›/‹t› case. As in that case, one of the letters — here, ‹m› appears to have dropped the dot adjunct to preserve overall paradigmatic contrast in the letter inventory.



Apart from these more systematic cases, other BM letters with more unique shapes also show quite direct relationships to their Philippine and/or proto-script counterparts. ‹k›, with its parallel bars, relates directly to the common "rotated H" shape in the Philippines and especially to the tendency in some writers there to delete the crossbar, leaving only the parallel lines. A single, rare BM variant for this letter described by Noorduyn has a trough atop an arch, like a curved 'X' — which corresponds directly to the Palawan variant in the Philippines and its plausible antecedents, discussed earlier. Modern BM ‹s› has a circular shape, but in palm-leaf style this is not seen: instead, a shape similar to Latin "long ſ" (and its mirror image), as well as a simple vertical line (adopted in the Lota Ende offshoot of BM script) are used. It seems likely that the upper arch and lower trough of the "long ſ" shape were later realigned directly over each other to yield the now-standard circular or oblong shape. This "long ſ" shape relates directly to early variants of ‹s› in the Philippines, with only the final adjunct tail missing: given the direct equivalences established for other letters in the script, it is likely that the ancestor of both was similar in appearance to the curved Philippine shape, perhaps with the adjunct detached from the main body of the letter.

BM ‹p› is unusual with its final upstroke in not having a final downstroke corresponding to the final tail of the Philippine or proto-script letter. On the other hand, ‹w› does have a second arch with its final downstroke. In the Philippines, the corresponding letter either begins with a horizontal onstroke and ends in an upstroke, like BM ‹p› (in Luzon) or begins with a large counterclockwise curve and ends in a downstroke (Palawan). Each letter shape corresponds only partially to the BM shape, but reconstructing an intermediate shape with an initial onstroke analogous to the one still found in archaic Gujarati ‹b› and ‹d› (with their initial counterclockwise curl body strokes) furnishes a proto-script shape that would have easily reduced to the stereotypical double arch in the same way as other letters of a similar shape, as well as providing a plausible common ancestor for the two disparate Philippine shapes.

Although there have been attempts to compare the (modern) BM ‹h› with the corresponding Philippine shape (and a similar shape in other scripts of the archipelago), it is clear that this is actually a fairly late borrowing from Arabic script (cf. the two pre-19th-century variants illustrated here), later reanalysed, like ‹s› to conform to the arch-trough stereotype of BM script.



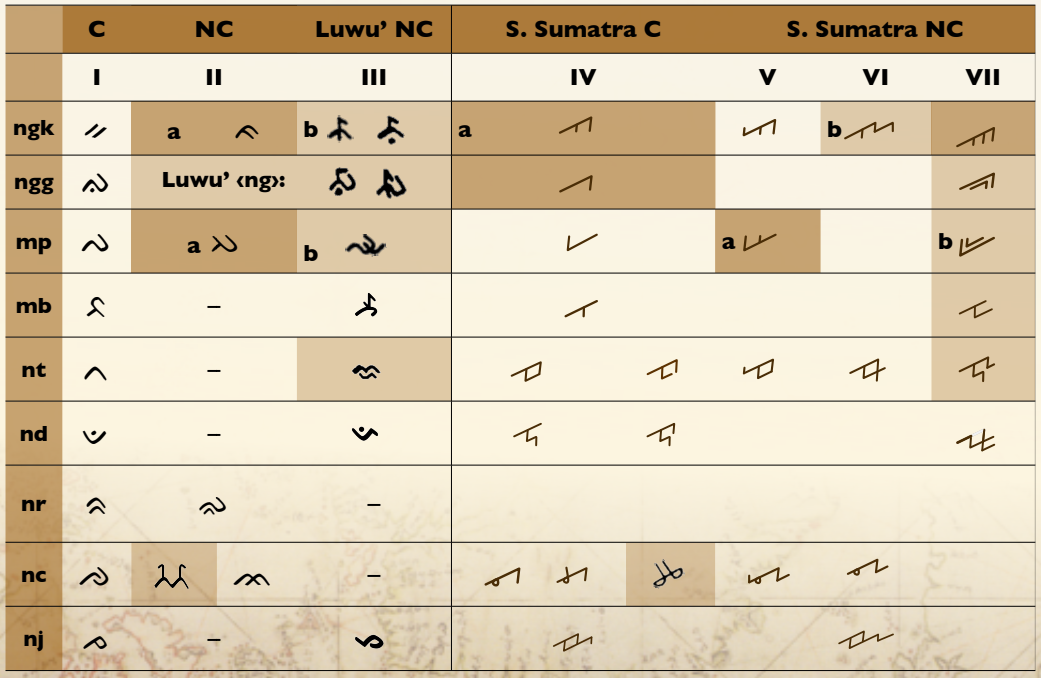
Additional nasal-consonant cluster letters not in common use and having no Philippine counterparts are discussed in the following slide.

Apart from these, four basic BM letters have no Philippine counterparts. This is explained — on the likely hypothesis that the script was first adopted by speakers of Tagalog and/or some other language(s) lacking the corresponding distinctive speech sounds — by the lack of a need for these letters in the languages first written in Philippine script.

Because of the peculiarities of its relationships and use, ‹ny› will be dealt with further below in conjunction with Sumatran scripts. This leaves ‹c›, ‹j› and ‹r›. For each of these three, a direct structural relationship with the early and modern shapes of the corresponding Gujarati letters is apparent. ‹c› and ‹r› are fragmented counterparts of the corresponding Gujarati letters (with a final upstroke instead of the expected downward tail on ‹c›, similar to BM ‹p› and Philippine ‹w›). Note that one rarer BM variant reverses the order and position of the arch and trough. BM ‹j› possesses a final loop, whereas archaic Gujarati has an initial loop. The looping stroke has (to borrow descriptive terms from phonology) “spread” or “copied” onto the lower right angled join in modern Gujarati script, and in loose modern handwriting, the initial loop is often reduced to an open curve or disappears, yielding a shape for all intents and purposes identical to the BM letter.

Taking Caldwell's (1988) dating of the beginning of literacy in Sulawesi to approximately the year 1400 together with the close formal correspondences with archaic Gujarati/informal Devanagari script, it seems likely the forerunner proto-script was introduced into the archipelago sometime during the 14th century — a time period consistent with early Gujarati activity in the region.

Bugis NC letters modelled on South Sumatran equivalents

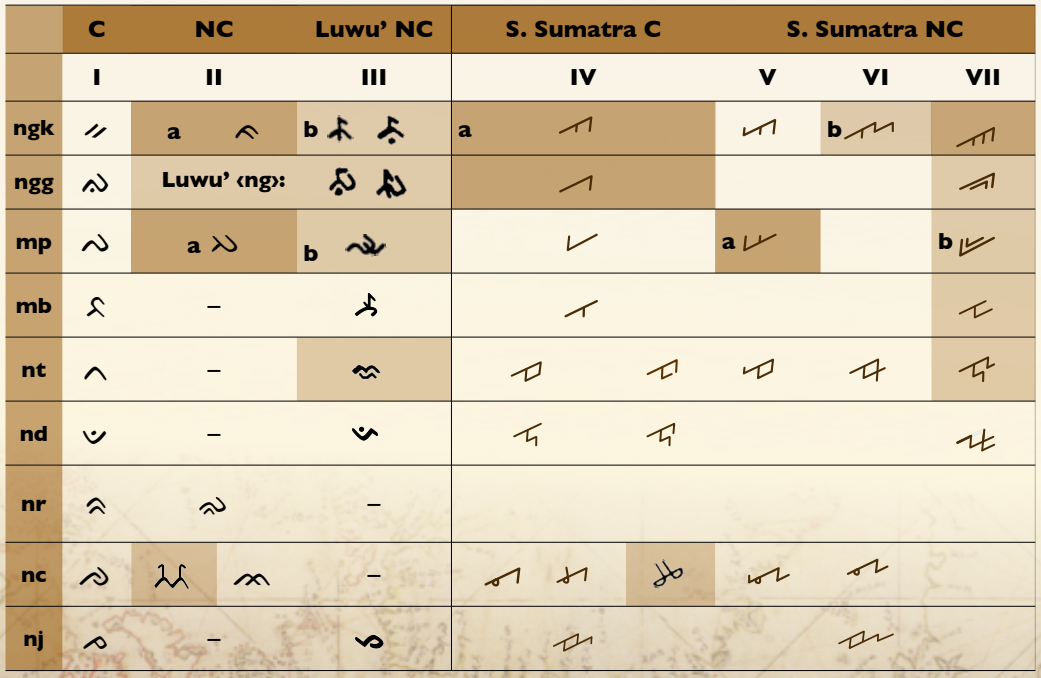


Noorduyn (1993) discusses the N(asal +) C(onsonant) letters used occasionally in BM script, arguing against the opinion expressed by some that they might be of Sumatran origin. In particular, he casts doubt on the hypothesis that BM ‹ngk› (column II, first line) is adopted from South Sumatran ‹k› (column IV) by pointing out that this particular BM shape is a later, back-skewed style whereas older attestations of the letter have a vertical adjunct line extending downwards on the underside of an arch. This still does not exclude the possibility that the vertical line is an adaptation to earlier BM norms of the South Sumatran diagonal stroke. Otherwise, the single added diacritic stroke in BM ‹mp› and ‹nr› (column II, lines 3, 8) are clearly comparable to common South Sumatran equivalent NC letters (including ‹mp›, line 3 of column V), and archaic BM ‹nc› (to the left of the more modern variant in column II), with its double flag adjuncts over arches with their (open) counters, shares most of its structure with a variant of ‹c› from Kerinci, South Sumatra (third of three, column IV).

Expanding on Noorduyn and Salim (1988), Noorduyn (1993) also briefly discusses a set of NC letters found in a single manuscript from Luwu', in the northeast of the Bugis-speaking region, but restricts himself to noting the shape modifications involved and the similarities to other BM letters. Like the above-mentioned "standard" set many of these, and a further set of letters representing doubled consonants, are transparently derived from the base BM letter by adding an extra diacritic stroke on the left or right side. This is clear, for example, of ‹mb›, ‹nd› and ‹nj› in column III (cf. the base letters in column I).

Other Luwu' NC letters do not relate in any systematic way to the corresponding base BM letter. However, comparison with their South Smatran equivalents shows not only systematic relationships but reveals that the specific choice of shape or stroke type added to the base BM letter, or the use of a suppletive shape, is borrowed directly from the corresponding South Sumatran NC letter — and in the case of more complex shapes, due to constraints on possible elaborated shapes that are specific to the structure of South Sumatran script.

Bugis NC letters modelled on South Sumatran equivalents



In the case of South Sumatran ‹g› (column IV line 2), adding a single perpendicular stroke on its left side would duplicate the existing letter ‹h› and on its right side, the letter ‹l›; adding a stroke inside the open counter in the bottom of the letter would duplicate ‹k› (cf. line 1). The next alternative was to add a double stroke, essentially a smaller angled arch, inside the counter. This is what is added to the BM base ‹g›, but above the arch in accordance with BM character structure constraints described earlier. Similar constraints act on South Sumatran ‹ngk› (line 1): adding a single perpendicular stroke on the right would duplicate the existing letter ‹ʔ(a)›, so the alternative was either to add one on the left (column V) or add an angled arch (VI, VII). As with Luwu' ‹ng(g)›, this appears as an angled arch "flag" atop the arch of the preexisting BM ‹ngk› letter. In the third line, South Sumatran ‹mp› has a variant with a perpendicular added stroke inside the counter (V) — corresponding to the standard BM ‹mp› — and one with a reduplicated angled trough inside the counter. These are the two main available alternatives, since adding a perpendicular stroke on the left or right encounters the same problems as mentioned for ‹g›. And the added reduplicate of the base shape is what is added to the base BM ‹p› shape in the special Luwu' ‹mp› letter, overlapping with that letter's counter.

A final mysterious Luwu' choice is again illuminated by referring to the corresponding South Sumatran NC letter. Noorduyn points out that Luwu' ‹nt› appears to be based not on BM ‹t›, but on ‹ny›, by adding a downstroke on the right-hand side of the adjunct trough beneath the counters of the letter's double arch body. Why a letter so completely unrelated in shape would have been chosen for this purpose remains quite a mystery. On the other hand, comparing the Luwu' ‹nt› with its South Sumatran equivalents immediately reveals that it is in all likelihood borrowed directly from South Sumatra, consistent with the other Luwu' and standard variants mentioned above. Taking the open-counter variant of South Sumatran ‹t› (second variant in column IV, line 5), the column VII ‹nt› is derived by adding a perpendicular stroke to the ends of the bottom diagonal and right-hand vertical strokes. The resulting derived shape corresponds to the column III Luwu' ‹nt›, with the single addition of another downstroke added at top right.

Bugis NC letters modelled on South Sumatran equivalents

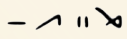
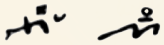
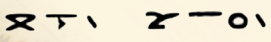
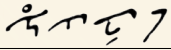
	C	NC	Luwu' NC	S. Sumatra C	S. Sumatra NC		
	I	II	III	IV	V	VI	VII
ngk		a	b	a		b	
ngg		Luwu' ‹ng›:					
mp		a	b		a		b
mb		–					
nt		–					
nd		–					
nr			–				
nc			–				
nj		–					

Taken together with the "basic" standard set of NC letters in BM script, these Luwu' variants are evidence for a detailed knowledge of South Sumatran script up to and before the 1850s on the part of BM writers in different regions of southwest Sulawesi, which testifies to the existence of contacts with users of South Sumatran script. Whether it was South Sumatran literates visiting Sulawesi who transmitted knowledge of their script or Sulawesi visitors to South Sumatran ports (Palembang and/or Jambi) who learned the script there remains an open question, but it seems clear that in earlier times this script was likely known on the Sumatran coast and not limited to the inland and western highlands of Sumatra as appears to be the case in descriptions coming to us from the early 1800s.

Scripts from Sumatra

Top: Batak script, north Sumatra (Pustaha *4301,* Logan Museum of Antropology, Beloit WI, USA)
Middle left: Kerinci *Surat incung* script, mid-south Sumatra. British Library Endangered Achives Project/Pusaka Kerinci web site.
Middle right: Tanjung Tanah book of laws, late 14th century or later, mid-south Sumatra. Uli Kozok, (2004). More typical of later writing from the Rejang-Middle Malay and Lampung regions further south.
Bottom: Lampung script variety, far south Sumatra. From H.N. van der Tuuk (1868), *Les Manuscrits Lampongs*.

Spelling rules in Sumatra: vowel sign shifted rightward or away

Mandailing, Toba Batak	*Kerinci*	*Lampung*
		
‹ $s^{ŋ}o$ › /soŋ/	‹ t_a ghi › /tagih/	‹ $p^{rə}$ › /pər/
		
‹ $p^{ŋ}i$ › /piŋ/	‹ si rhi › /sirih/	‹ s^{ni} › /sin/ ‹ $n^{ŋi}$ › /niŋ/
		
‹ $tp_{u}∅$ › /tup/ ‹ $sp_{i}∅$ › /sip/		‹ ^{ʔi} $kt_{u}∅$ › /ʔikut/

A curiosity of Batak writing systems described originally by van der Tuuk in the 1800s is shared with other scripts all the way to the southern end of Sumatra. In syllables closed by a coda consonant following the vowel, the vowel sign is not written (as in open syllables) on the letter for the onset consonant, but on the letter for the coda consonant, marked as such by a *virama* or "vowel killer" sign following the consonant letter and the syllable vowel letter if it is written to the right of its host as in the case of ‹-i› in Batak script and Kerinci *incung* script or Batak ‹-o/ĕ›. This shoft to the coda consonant letter is consistent and regular in Batak scripts in the north and Lampung scripts in the south, and found sporadically in Rejang-Middle Malay *Surat Ulu* script to the north of the Lampung region. In the case of the coda /ŋ/, /r/, /n/ or /h/, which depending on the specific script and variety are not written as independent consonant letters but as dependent signs above or (in the case of Kerinci *incung* ‹-h›) to the right of the syllable onset consonant letter, an accompanying ‹-i› or ‹-ĕ/o› will be written further away from the base letter, whether to the right of the coda consonant sign or above it. This second rule reflects the vowel sign displacement in the rule described above, but occurs more sporadically and/or variably in these scripts. Nonetheless, it is also found from the Batak script regions in the north to the Lampung script region in the south.

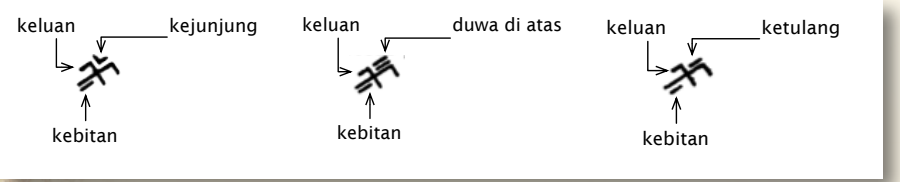
These two spelling conventions are rather counterintuitive in the way they rearrange the order of vowel and final consonant graphemes. Considering the vast difference in appearance between the Batak scripts and the ones further south on the island and the abstruse nature of these spelling conventions, the fact they are shared along the length of the island is almost certainly due to common inheritance from a shared ancestral script, rather than to borrowing (where there is little if any evidence of other borrowings between Sumatran script varieties).

Didactic vowel sign combination in South Sumatra

Bij het leeren spellen zet de inlander verscheidene teekens te gelijk bij de letter, en leest de leerling ꦏ *ka* kedjoedjoeng *kar*, keloean *kir*, kebitan *koer*;
ꦏꦶ *ka* doewa di atas *kan*, keloean *kin*, kebitan *koen*;
ꦏꦁ *ka* ketoelang *kang*, keloean *king*, kebitan koeng.

De Sturler (1843)

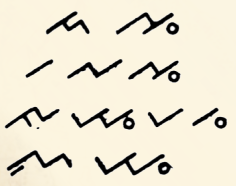
" When teaching spelling the native puts various signs together on the letter, and the learner reads ꦏ *ka* kejunjung *kar*, keluan *kir*, kebitan *kur*;
ꦏꦶ *ka* duwa di atas *kan*, keluan *kin*, kebitan *kun*;
ꦏꦁ *ka* ketulang *kang*, keluan *king*, kebitan *kung*. "



An invaluable clue to the likely origin of the conventions comes from de Sturler (1843), who describes a strategy used in South Sumatra for teaching the combinations of several vowel signs at a time with a syllable onset consonant plus a coda consonant letter or dependent coda consonant sign.

In this example de Sturler shows how learners would recite these combinations, placing vowel signs together with coda consonant signs on a base onset consonant letter. However, although the ‹-i› vowel sign here is written (iconically) to the left of the dependent coda consonant sign in the three examples given, it is recited *after* the consonant sign. It is plausible that this reversal of order in the recital eventually led to the vowel sign actually being written down *after* the consonant sign, no longer to its left but above it or to its right. This slight shift would lead to one of the two Sumatran conventions.

Didactic vowel sign combination, continued



da membunuh *wa:* dau (= one).
ta mungguh *te*, *la* membunuh *wa*, *lau:* telau (= three).
a membunuh *ma*, *am*, — *pa* membunuh *ta*, *pat:* ampat (= four).
na membunuh *ma*, *nam*, kebitan *num:* num (= six) [...].”

Van Hasselt (1881)

Quite likely that the two similar Sumatran rules derive from this practice:
a straightforward result of writing the characters and signs in the same order as the recitation.

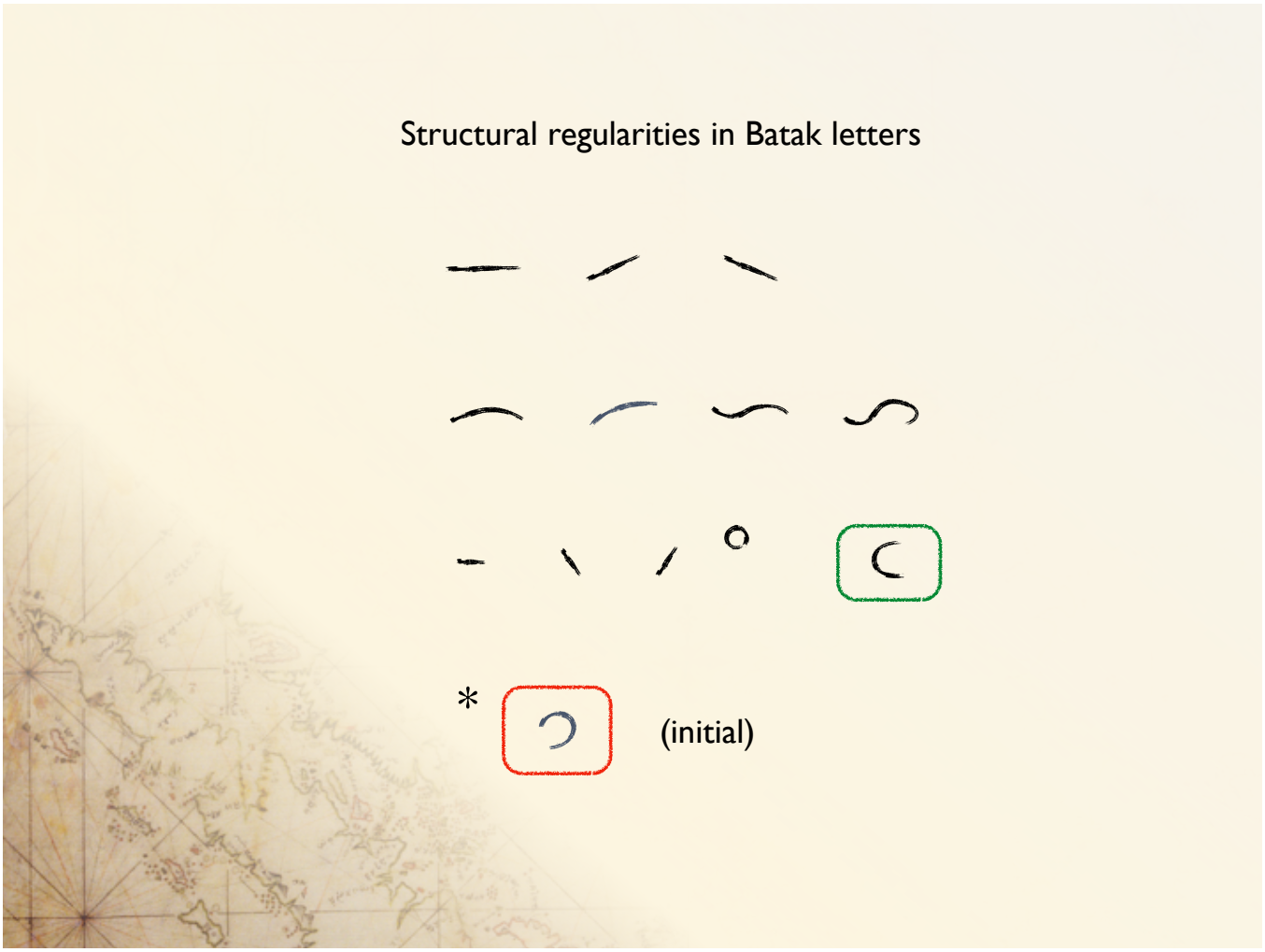
Also quite likely the Sulawesi-Philippine practice of combining (and reciting) vowel signs on base letters is a simpler version of the practice described by de Sturler, in which case it must be at least as old as the first date at which writing was adopted in Sulawesi — and likely older still.

Though he devotes less space to actually describing this, de Sturler also provides examples of similar recitations used to spell out the combinations of vowels with following coda consonants that are written out in full with an independent consonant letter followed by a “vowel killer” that signifies the default /a/ vowel (or equivalent, depending on the region) is not to be read off that consonant sign.

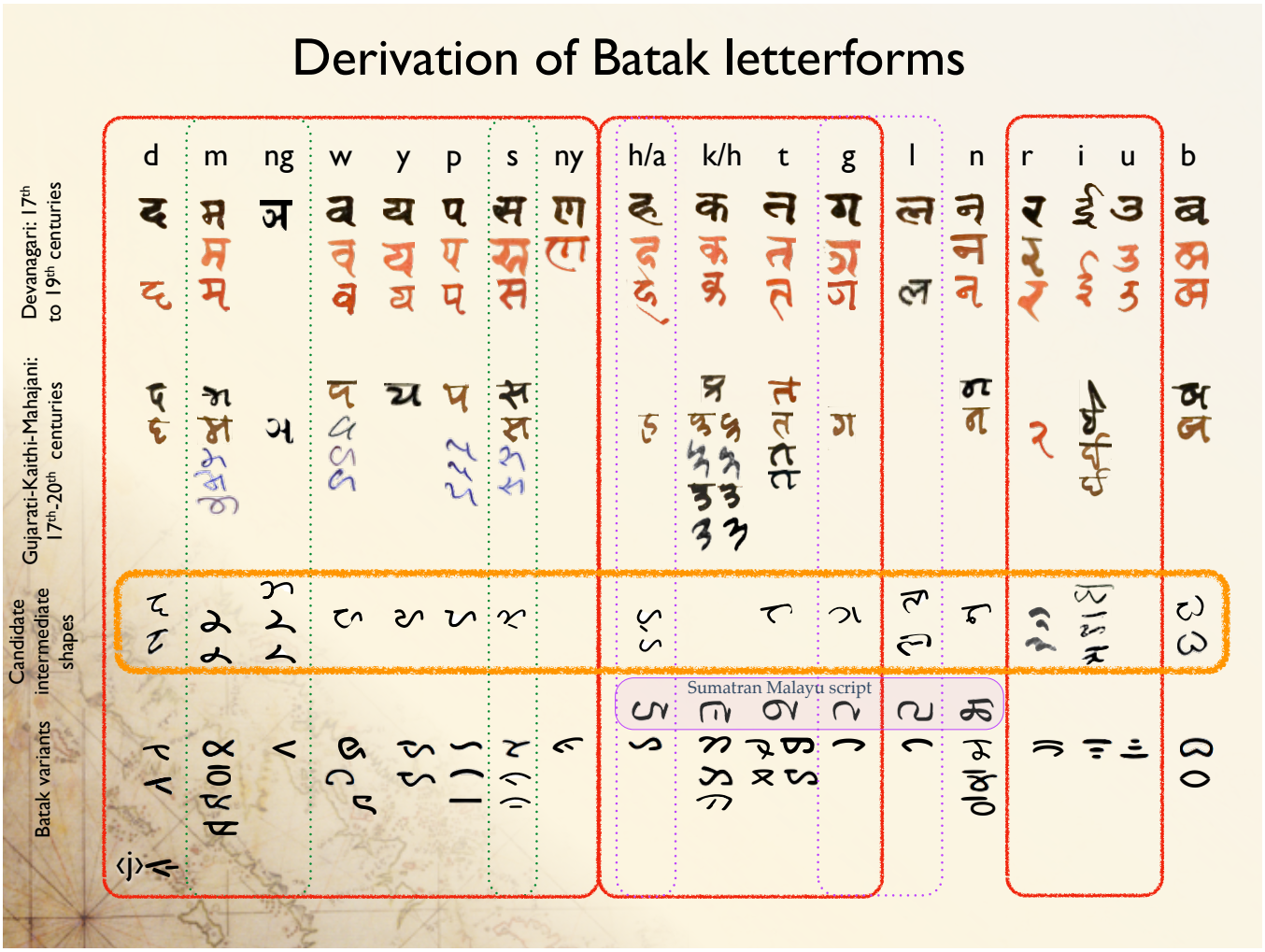
The vowel killer, here written in its later form as a small circle (almost certainly a borrowing of the Arabic *sukûn/jazma* which also marks the absence of a vowel) to the right of the coda letter, is not recited as a separate sign with its own name (variously attested as *tanda bunuh/tanda mati* ‘killer sign’ or *bunuhan*) after the consonant letter, unlike the coda consonant signs and vowel signs in the previous slide. Rather, the consonant letter with the vowel killer following is recited as a unit as “*membunuh* (+consonant letter name)” i.e. “kill (consonant letter name)”.

The first three examples show how this affects the second consonant letter in a syllable, blocking the default pronunciation of the following /a/ that is found in the letter’s name. The fourth example is interesting as it shows how (although it is written underneath the syllable-initial consonant letter here) *kebitan*, the name of the added vowel sign for ‹-u›, is recited after the *membunuh ma* ‘kill *ma*’: as in the examples on the preceding page, it is not a long step from reciting the *kebitan* out of order to actually writing it on the “killed” *ma* rather than on its expected host, the preceding onset consonant letter. This shift, like the one hypothesised on the previous slide, naturally gives rise to the second type of vowel sign displacement convention found in all Sumatran script regions, though less consistently in the middle of the island.

This practice of reciting combinations of signs on a single base letter in order to teach their use in spelling brings back to mind the similar practices attested for the Philippines and Sulawesi. To my knowledge, this particular teaching strategy is not used for other Indic scripts, whether Javanese or further afield. It seems most likely that the Sulawesi-Philippine practice is a simplified version of the one described for Sumatra, but without coda consonants since Bugis-Makassarese script had dispensed with them as unnecessary for spelling Bugis and Makassarese.



As with Philippine and BM scripts and others further afield, it is possible to isolate stereotypical structural elements as well as constraints on the form of letters. This slide illustrates the important stroke types found in Batak letters, as well as an important constraint: although a clockwise curve can be found in final position in several Batak letters, this particular stroke type is disallowed in initial position in letters (and dependent vowel signs). Where corresponding letters or other characters in other scripts have this kind of initial stroke, the Batak counterpart consistently replaces it with a related but shallower stroke, or in a couple of cases appears to have completely deleted it.



Given the extensive discussion for Philippine and BM scripts above, I will only add a few comments here.

The first grouping demonstrates how the final downward tail in reconstructed intermediate proto-script letter shapes (orange rectangle) corresponds in Batak script corresponds to a minimal rightward tail, downward-rightward sloping stroke or no tail at all if it there is a significant amount of preceding structure in the letter. In the second group, where the proto-script tail has very little preceding structure, this is reflected in the Batak counterpart letter by a pronounced leftward curve at the bottom of the final downstroke.

The deletion of a proto-script initial clockwise curve can be seen in ‹m›, ‹ng› and its replacement by a shallow descending stroke or fragmented (detached) shallow arch in ‹y›, ‹s›, ‹g›, ‹r› and ‹u›. This effect can also be seen in a couple of dependent signs including the vowel-killer originating in Javanese script, where it begins with a large clockwise curve: it appears in Batak as a short stroke sloping down to the right, or even a short horizontal dash in one variety.

Batak ‹h›, ‹g› and ‹l› (together with ‹k›) are the most obviously similar to old Javanese and later Sumatran Malayu scripts among Batak letters. Nonetheless, they also clearly related to plausible proto-script and informal Nagari variant antecedents. The first three have a simplified form in Batak, lacking the additional stroke in the proto-script letter. This trait, shared with their South Sumatran counterparts, might possible be due to the deletion of these strokes under influence from the form of the corresponding Javanese and/or Malayu script letters already known to early Sumatran adopters of the parent script.

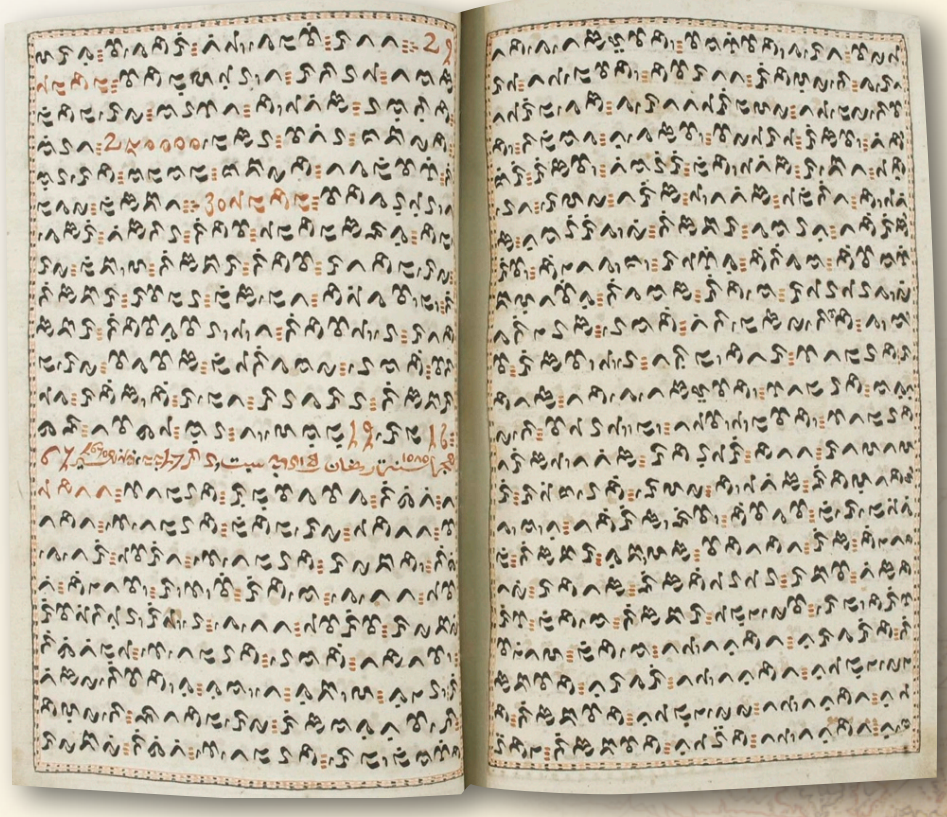
It is apparent on comparing South Sumatran script varieties (bottom four groups), that Rejang appears to conserve relatively more conservative shapes, from which (in most cases) both Kerinci and Lampung shapes appear to have derived through subsequent changes (and not vice versa).

One clear exception is in the second Lampung ‹b› (from which the third derives through straightforward processes). This shape relates directly to the North Sumatran Batak and Philippine shapes, through vertical inversion of its components. Otherwise, the more general South Sumatran ‹b› shape relates quite directly to the South Sulawesi letter. Apart from this, ‹ʔ(a)› has two unusual South Sumatran variants. One, resembling a '3v' sequence, (extreme bottom right image) is found in a particular Lampung variety; this appears to be a possible borrowing from Javanese script. The other, found in a single token in the Tanjung Tanah code of laws from Kerinci in northwestern South Sumatra, departs from the overall close resemblance of the South Sumatran text on the final two pages of that codex to Rejang-Central Malay and conservative Lampung letters. Though one hypothesis posits that it derives from the (improbable) combination of ‹h› with an older "vowel killer" shape, it seems more likely that it originates in the borrowing of a shape of Javanese ‹h› as it was developing toward its modern shape, sometime between the 15th and 16th centuries.

Comparing the more conservative shapes with the intermediate proto-script shapes (including two variants representative of specific changes drawn from Modern Gujarati and Pala'wán script from Palawan), it is overall quite clear (taking into account remarks already made regarding Batak) that despite the extremely angular and rectilinear form of these Sumatran shapes, they clearly reflect the shape of proto-script letter variants.

The ‹ng› letter has an added cross-stroke that appears inside or below the main body shape. This is the only thing that distinguishes the letter from ‹l›, abd without it the shape of ‹ng› is a fairly direct reflection of the proto-script letter. (Note how the variant with the stroke intersecting the middle of the letter takes on a fragmented two-part shape in less conservative forms of the script.) ‹c› relates straightforwardly to a form with an excrescent loop on the left-hand side, quite common in loose Gujarati handwriting and just as likely a variant in archaic Gujarati script. Although South Sumatran ‹r› has a shape that relates less obvious to the Gujarati shape (or Bugis-Makassarese and Batak shapes), a similar excrescent loop might be at the origin of the prominent closed counter in the South Sumatran shape. An intermediate shape is posited here for ‹y›, with the initial curl adjunct drifting and attaching to the right as in the Philippines: this is an alternative explanation for the shape of the South Sumatran letter, which is usually explained as deriving directly from Old Javanese ‹y› (which leaves unexplained why the Batak letter, for one, cannot easily be explained on the same basis). Finally, A modern development in which Palawan ‹w›, originally similar to the loose Gujarati shape and an archaic Batak shape, fragments into parallel strokes which turn downward at their ends, is a clue to how the apparently anomalous shape of the South Sumatran letter may have developed.

Makassarese *Jangang-jangang* (bird) script

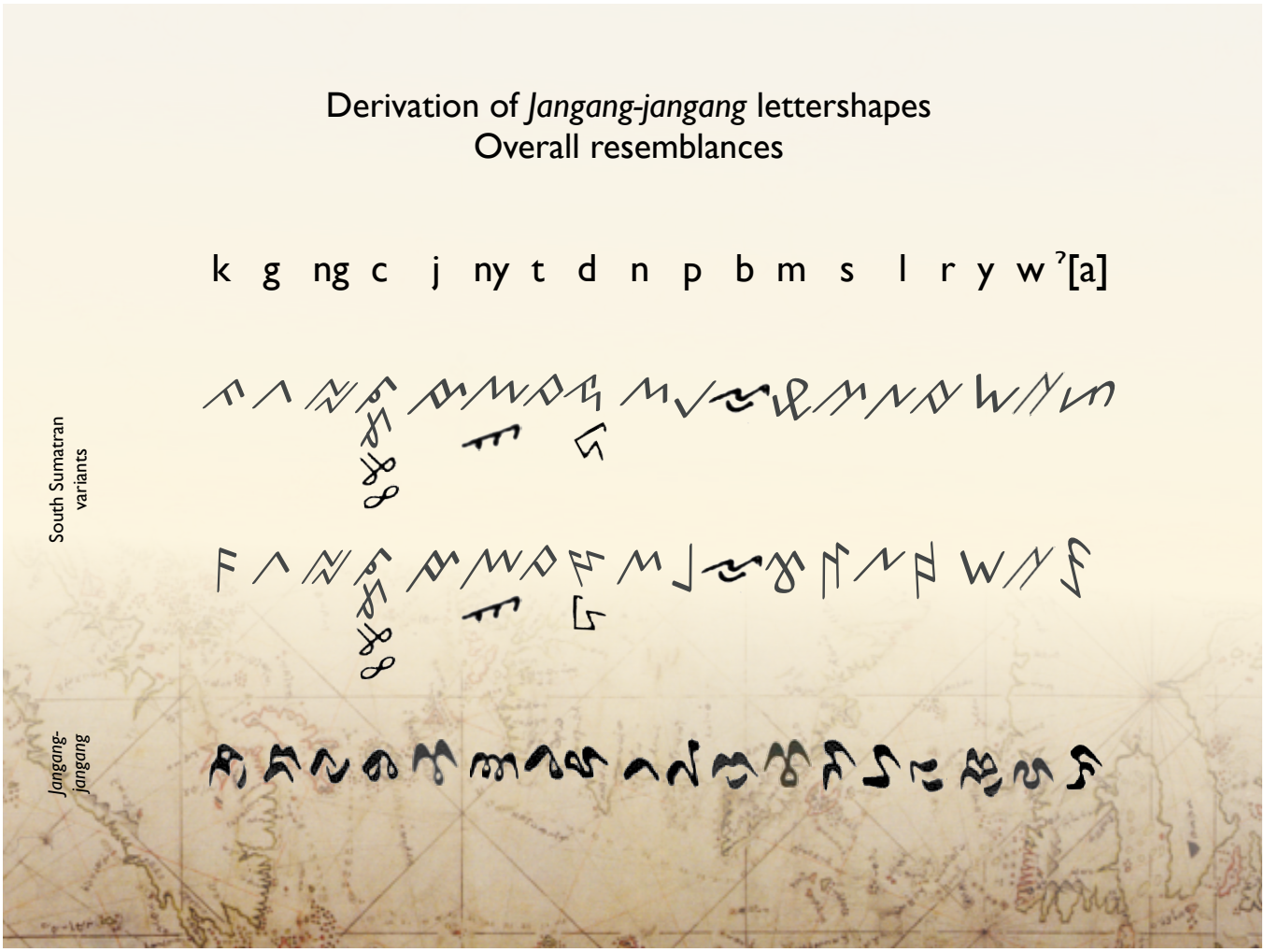


KITLV Tropenmuseum
668-216

Less well known than Bugis-Makassarese script is the *Jangang-jangang* 'bird' script used by the Makassarese but gradually replaced after the 17th century by BM script. Unlike BM script (apart from the palm leaf style), *Jangang-jangang* letters have a strong tendency to vertical orientation; as well, there is a greater tendency to complex structure than in the minimalist arch/cup-dot-flag stereotypes of BM script.

The origin of this particular script, and its exact relationship to BM script, has long been unclear. However, with the recent availability of a later copy of a document containing dated Makassarese historical records going back to the early 1600s and written in *Jangang-jangang* in styles that show gradual changes correlating with the increasing progression in the years of the relevant entries, it is possible to say that it is largely derived from South Sumatran script (with individual unusual letters drawn from different regional varieties) with some additions from Javanese script. An observation to retain is

The first page here contains part of a Makassarese translation of the treaty of Bungaya between the Makassarese and the Dutch and, half-way down the page, the date November 19, 1667 together with its Hijri equivalent written in the local Malay *Jawi* variety of Arabic script.



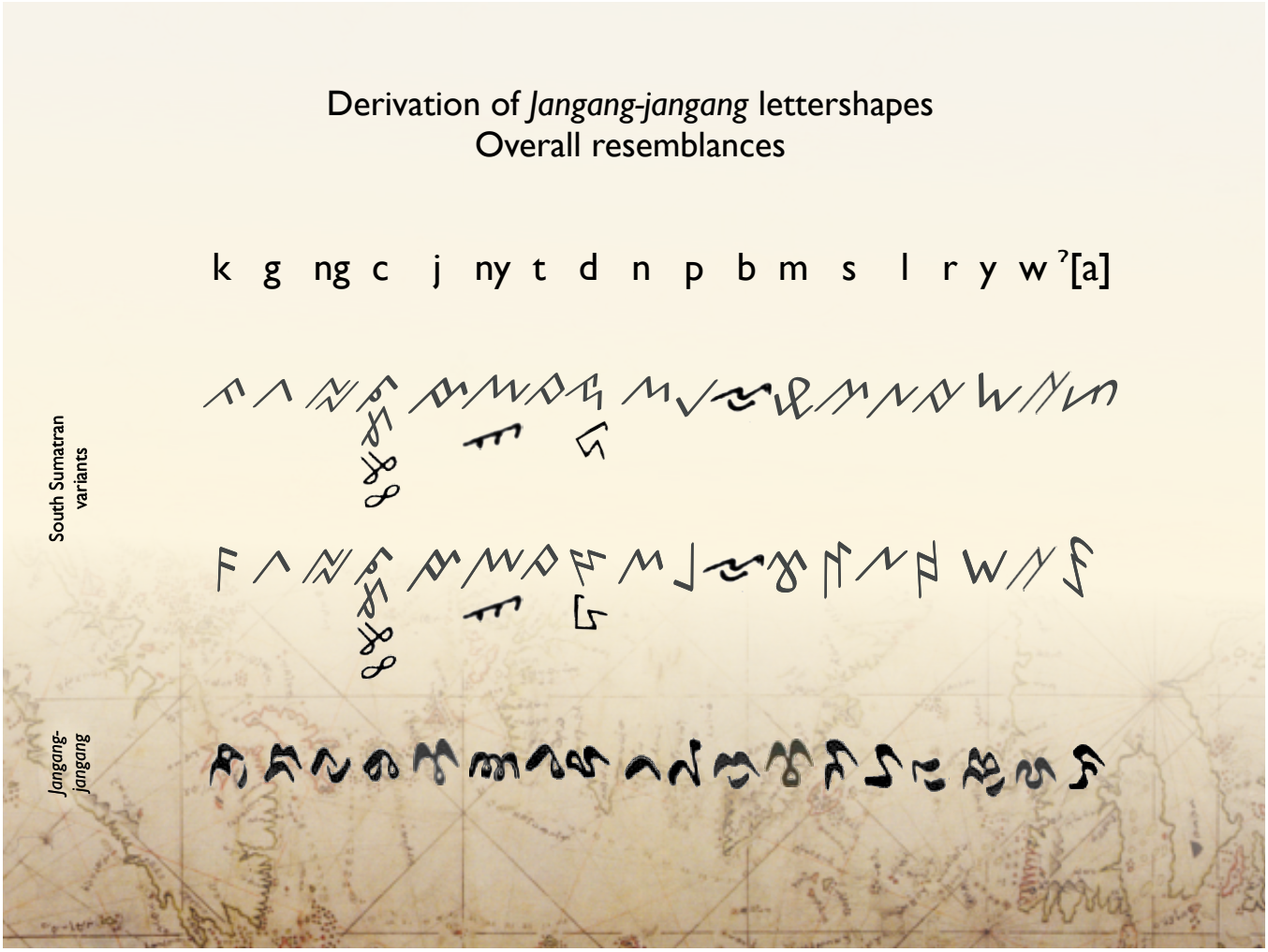
The first observation that reveals the direct relationship of *Jangang-jangang* to South Sumatran script varieties comes from recognising that the correspondences emerge when the South Sumatran letters are viewed rotated counterclockwise from their usual 30-45˚ slope (top line) to a vertical orientation (middle line), similar to the general tendency of *Jangang-jangang* letters (bottom line). As can be seen, this is a general tendency that does not apply equally to all letters.

One of the first questions that arises is how this unusual vertical orientation came about...

Gaja Mukur script play style (Van Hasselt 1881)

In an enlightening foldout plate, Van Hasselt (1881) illustrates a broad variety of observations on the usage of South Sumatran script in the Musi watershed between Palembang and the upstream inland region to the west. Among these are a group of playful manipulations of script style that he says were used by courting youth. A number of these are also described for Lampung script, to the south, by Pudjiastuti (1996, 2014) together with a number of others not shown in Van Hasselt's description.

One of these styles is what he describes as *Gaja Mukur* (*Gadja-Moekoer* in his older Dutch-based spelling) – which might be a distortion of the Malay phrase *gajah memukul* (striking elephant). In this style, all letters are rotated to the vertical from their normal 30-45˚ slope. At least one artifact written in this style exists: a fused seashell from Bengkehulu (Bencoolen) on the west coast in the collection of the KITLV Tropenmuseum, with the inscription *Kepada tuwan Sakarman dari Ahmad* 'For Mr. Sackerman/Zakerman from Ahmad'. (The personal name is unclear: Bencoolen was a British possession under the Bengal captaincy before its cession to the Dutch in 1825.)

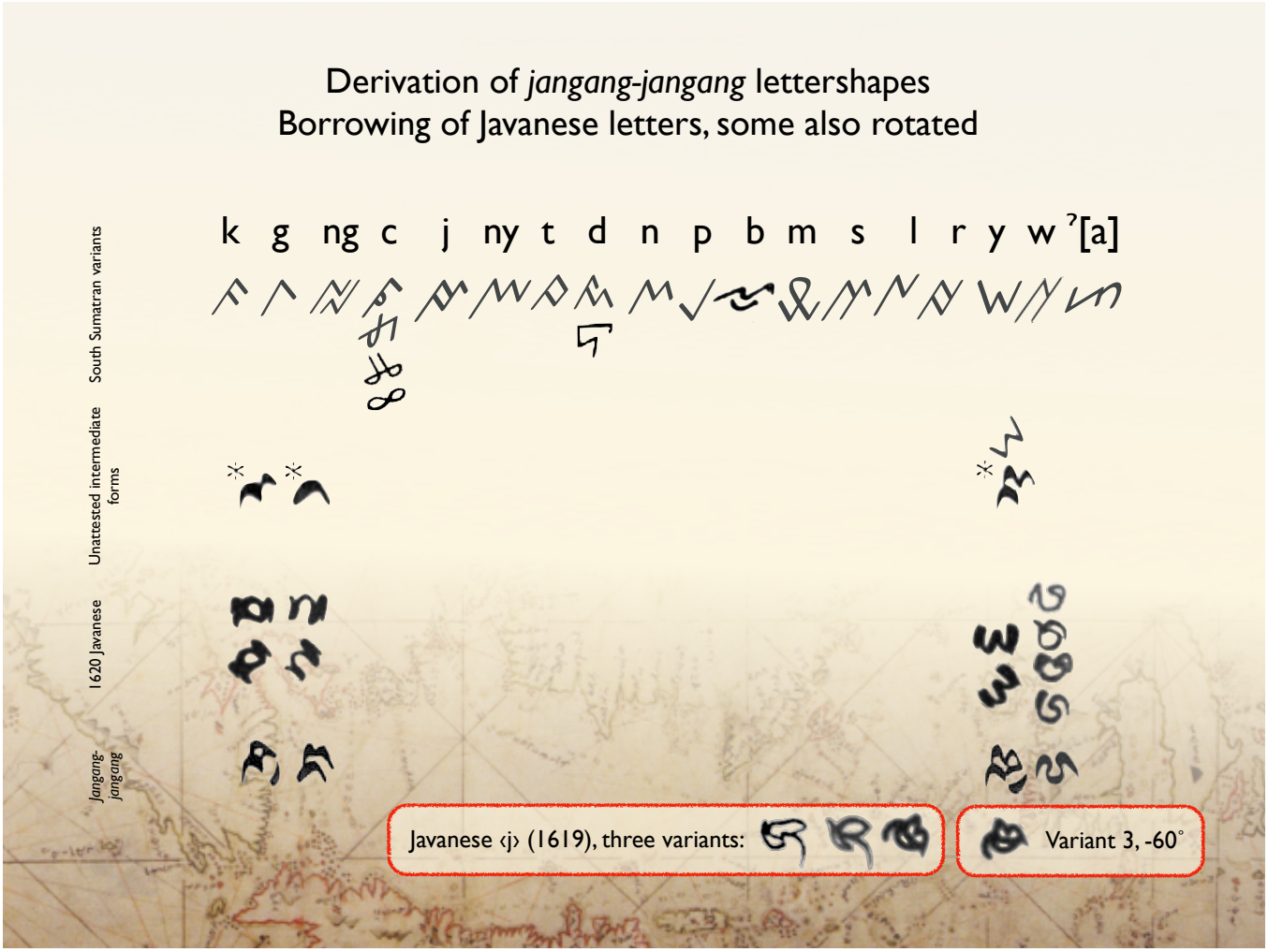


Returning to the comparison between *Jangang-jangang* letters and their South Sumatran counterparts, two important observations should be made.

First, for most the relationship is clear and direct, with only slight modifications in some cases: stroke simplifications in ‹ng›, ‹n›; horizontal mirror image in ‹l›, fragmentation in ‹r›; addition of loops in ‹ny›; and a general tendency to arching strokes in most, comparable to the Bugis-Makassarese stereotype.

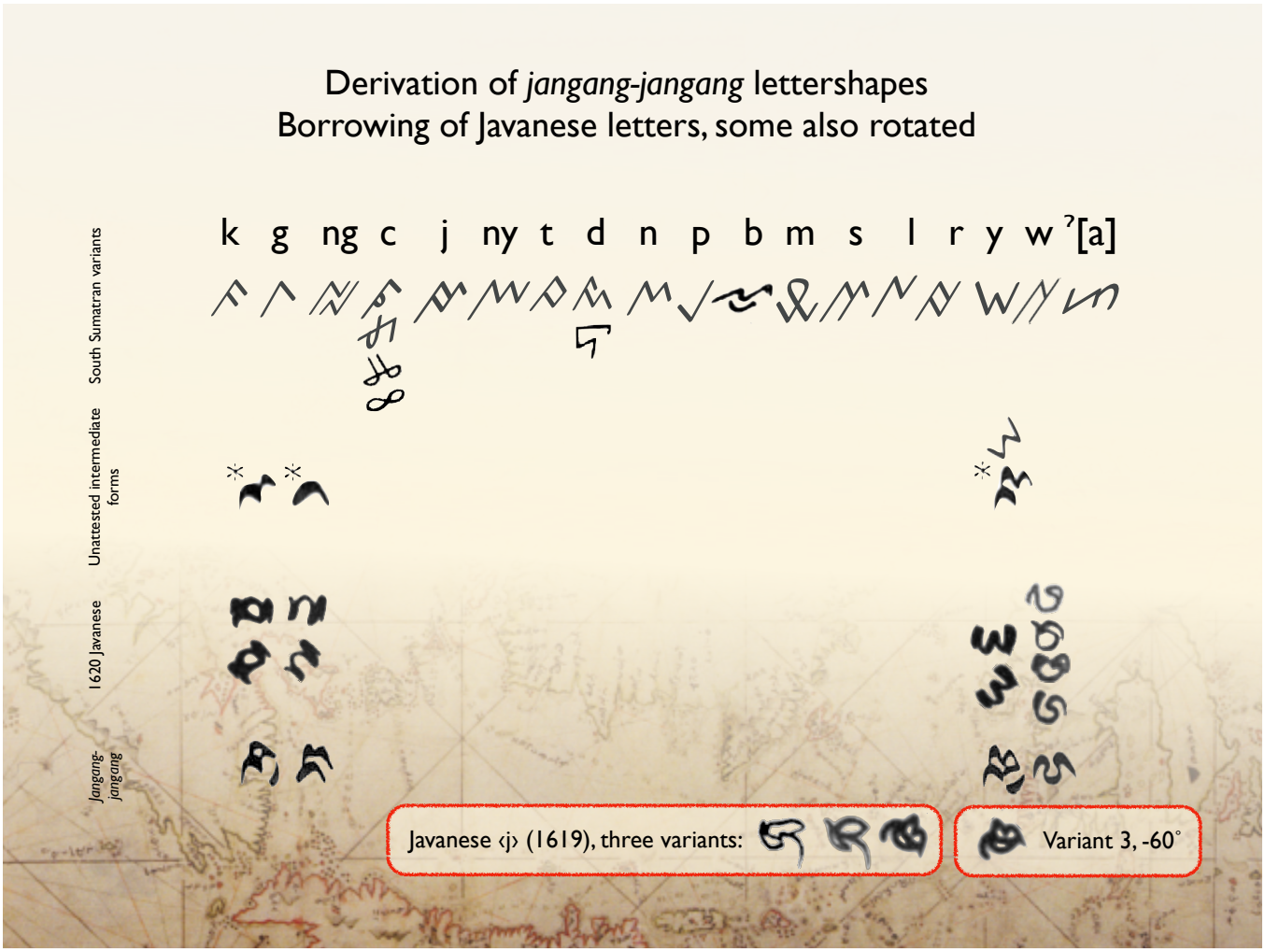
Second, two of the South Sumatran letters unambiguously related to their *Jangang-jangang* counterparts are relatively rare variants (in existing manuscripts). The ‹b› is typical of one set of particularly innovating Lampung varieties, and ‹ʔ(a)› is a hapax known (so far) only from a single token in a single text in the Tanjung Tanah code of laws. The presence of these two unusual letter shapes is evidence for two possibilities. The first is that there was a script variety including these two otherwise rare variants that was used in the east coast ports of Palembang and Jambi likely frequented by Makassarese seafarers (or alternatively, the home of Sumatran visitors to Makassar). The second is that Makassarese visitors were exposed to a range of South Sumatran varieties either in South Sumatra itself or through South Sumatran visitors staying in Makassar.

However, not all *Jangang-jangang* letters can be convincingly derived directly from any South Sumatran counterpart: this is most obviously the case for ‹k›, ‹g›, ‹y› and ‹w›.



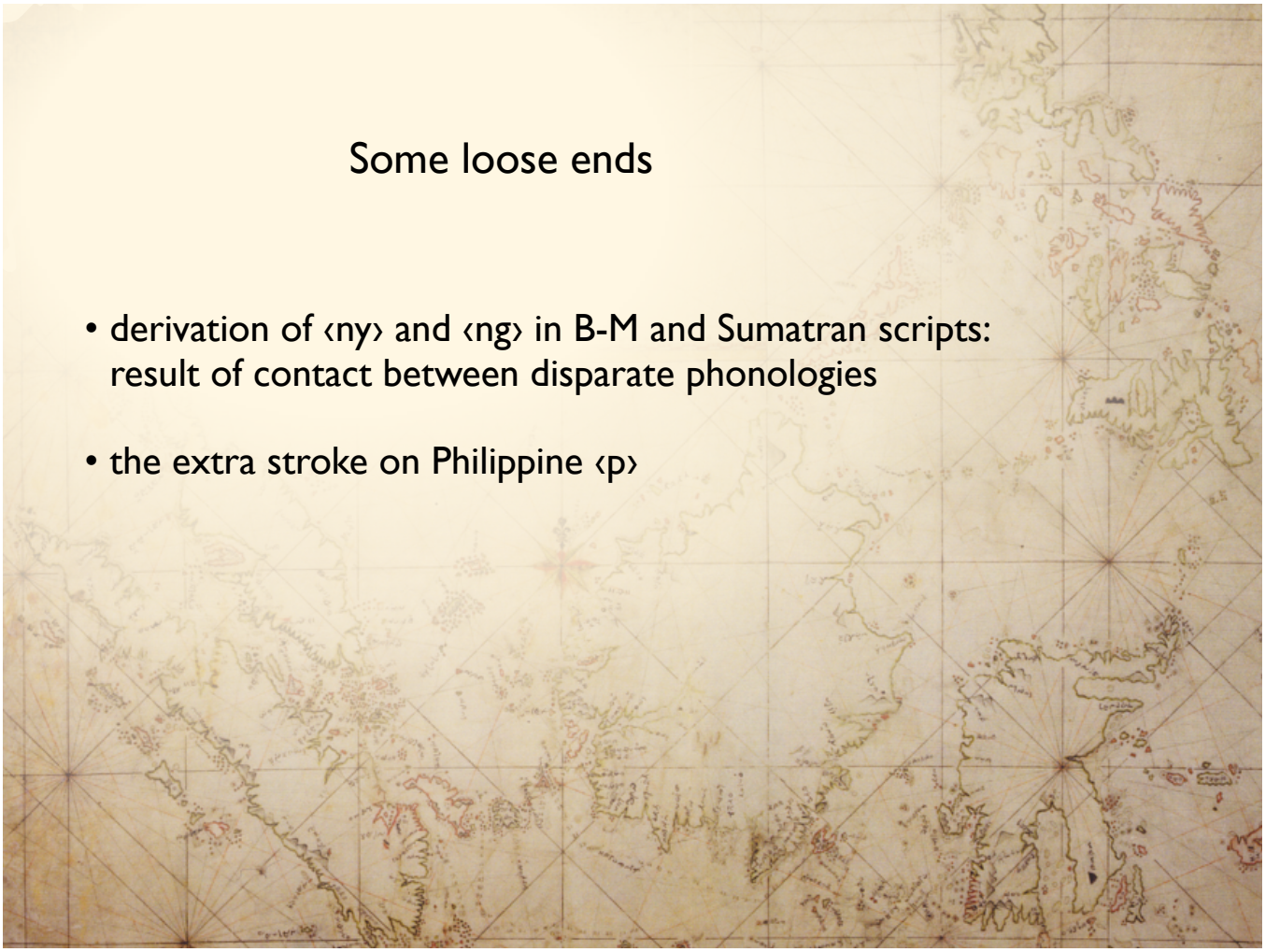
The letters with no clear derivation from any known South Sumatran script variety do, however, have convincing derivations from Javanese script. For comparison, we have available four hands from 1619 and 1620 letters from the Sultan of Banten in northwestern Java to the Dutch (Ricklefs 1976).

Although ‹k›, ‹g› and ‹y› *contain* substructures related to their South Sumatran counterparts, modified with the BM arch stereotype, their complete complex shapes can be related directly to variants of the corresponding Javanese letters illustrated in Ricklefs. The closest matches come from the hand of a single writer, found in a letter identified by Ricklefs as folio 327. If we rotate folio 327 ‹k› and ‹g› 60˚ counterclockwise, we can identify each component of that letter's structure (including the enclosed counter of ‹k›) directly with the corresponding components in the *Jangang-jangang* letter: the initial upward onstroke-plus-downstroke in the Javanese letter corresponds to the initial bottom arch in the *Jangang-jangang* letter and the remainder of the structure is placed atop this arch, the final downstroke of the Javanese letter being detached and moved rightward in *Jangang-jangang* ‹k›. In the case of ‹w›, it seems clear that the letter was borrowed without significant modification into *Jangang-jangang*.

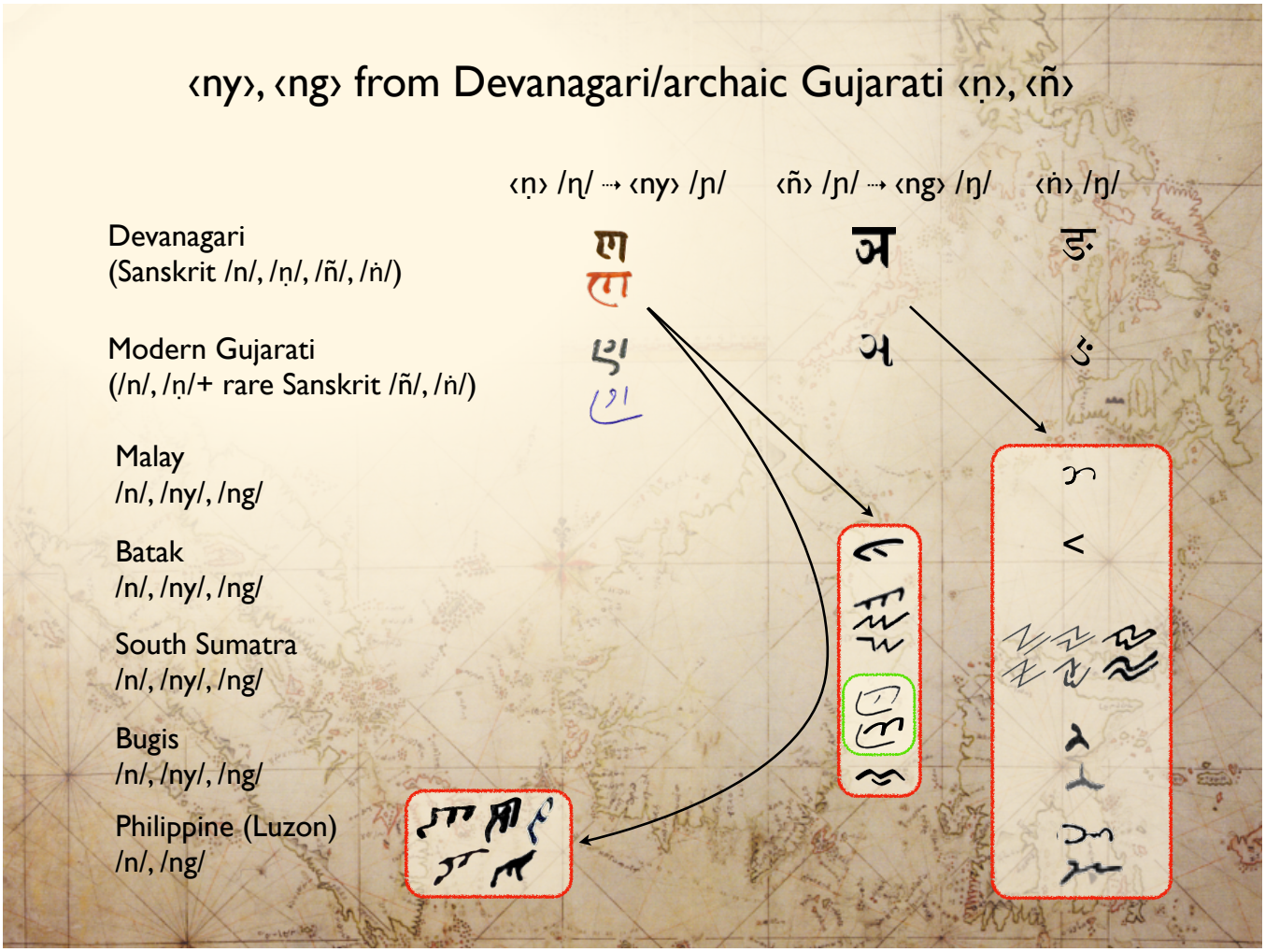


The case of ‹y› is more complex. As we see, the substructure of the *Jangang-jangang* letter relate, via the arch stereotype, to South Sumatran ‹y›. Unlike for ‹k› and ‹g›, rotating Javanese ‹y› does not yield an identifiable counterpart of the *Jangang-jangang* letter. However, *Jangang-jangang* ‹y› corresponds to a variant of Javanese ‹j› peculiar to the handwriting style in folio 327 when rotated 60˚ counterclockwise, with only minimal modification (detachment of the "belly" stroke, analogous to the case of ‹k›). Two more typical examples of ‹j› are taken from other folios for comparison: folio 327 shows a similar "bunching" of curved strokes upward to the left in several other letters. The correspondence to Javanese ‹j›, though initially unexpected, is not surprising given the phonetic similarity between the two sounds involved.

Given this evidence for borrowing and adaptation of specific shape variants of Javanese letters that can be traced directly to the hand found in a specific letter from 1619-20, two conclusions can be drawn. First, it is likely that the creation of *Jangang-jangang* — or at least the borrowing of these Javanese letters — can be dated to a time period extending to several decades before and after 1620, the likely productive lifetime of the folio 327 writer. Second, the clear evidence for borrowing from this specific writer's hand supports the hypothesis that *Jangang-jangang* was developed by eclectic borrowing of variants from a selection of South Sumatran scripts, rather than consistently reflecting a now lost east coast variety.



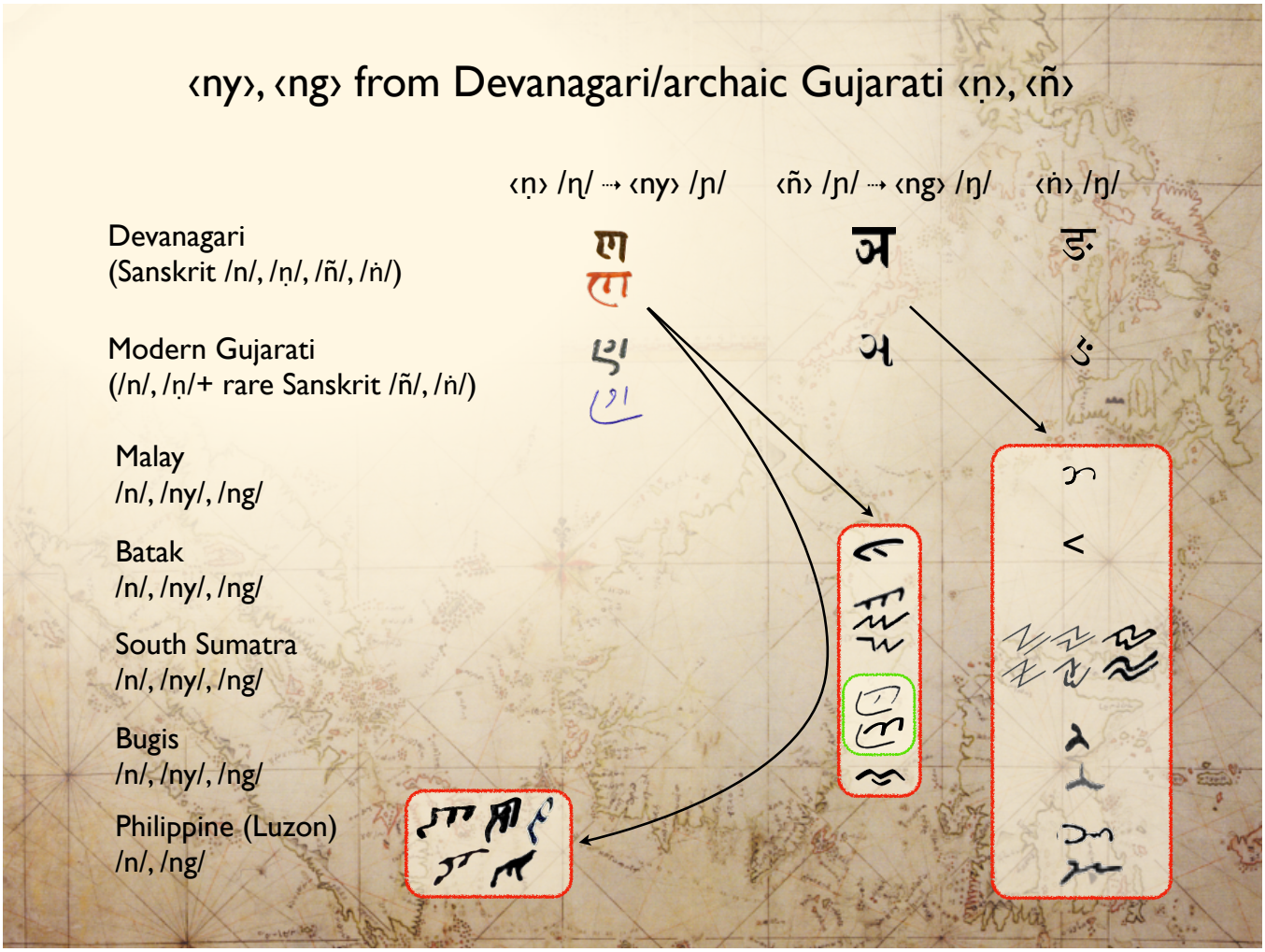
The next two slides deal mismatches between the shapes of certain letters and their expected regular antecedents in Nagari and archaic Gujarati scripts.



One point of mismatch between Sumatra-Sulawesi-Philippine letters and their expected regular correspondences in Devanagari and archaic Gujarati concerns SSP ‹ny› and ‹ng›. Rather than corresponding to the D/G ‹ny› and ‹ng› (conventionally represented as ‹ñ› and ‹ṅ›), they instead show systematic correspondences with the D/G nasal letters ‹ṇ› and ‹ny› (conventionally ‹ɳ› and ‹ñ›), respectively.

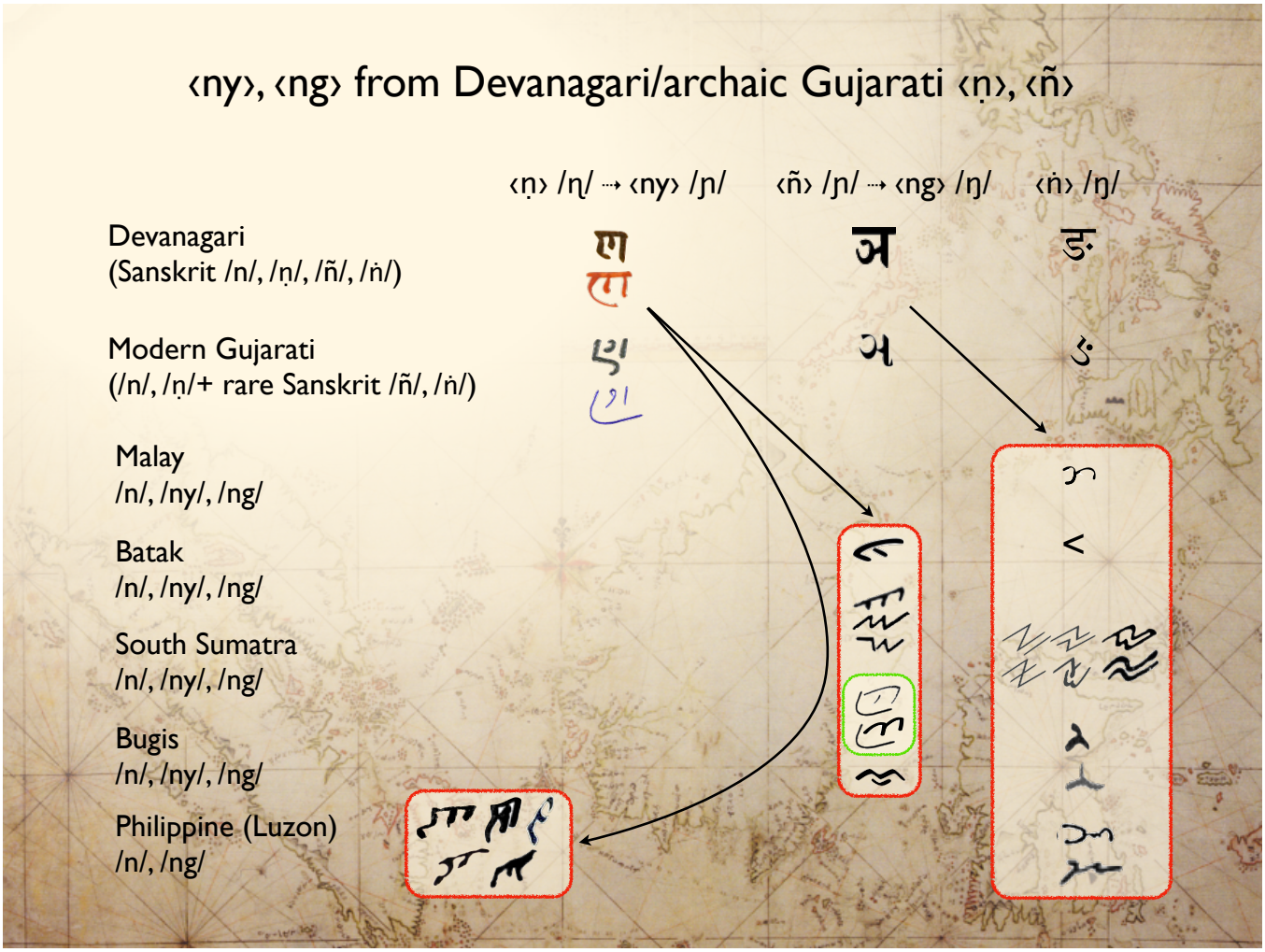
This can be explained as the product of contact between the modern Indo-Aryan phonology of the Gujarati language, with its basic native contrast between dental/alveolar /n/ and retroflex /ɳ/, and Malay, the indigenous language of the southeast Sumatran coast, which has a threeway contrast between dental/alveolar /n/, palatal /ɲ/ and velar /ŋ/ (n, ny and ng). Since the letters for dental/alveolar /n/ correspond fairly regularly, the mismatch to be explained concerns the SSP ‹ny› and ‹ng› letters.

Modern Gujarati includes ‹ñ› /ɲ/ and ‹ṅ› /ŋ/, but these are used only for Sanskrit loans and even in Sanskrit these have a marginal status, almost exclusively preceding other consonant letters in clusters (Mistry 2003). To write native Gujarati, only ‹n› and ‹ɳ› are needed. When archaic Gujarati script encountered Malay phonology, Malay writers were confronted with the problem of writing the /ɲ/ and /ŋ/ (ny and ng) sounds.



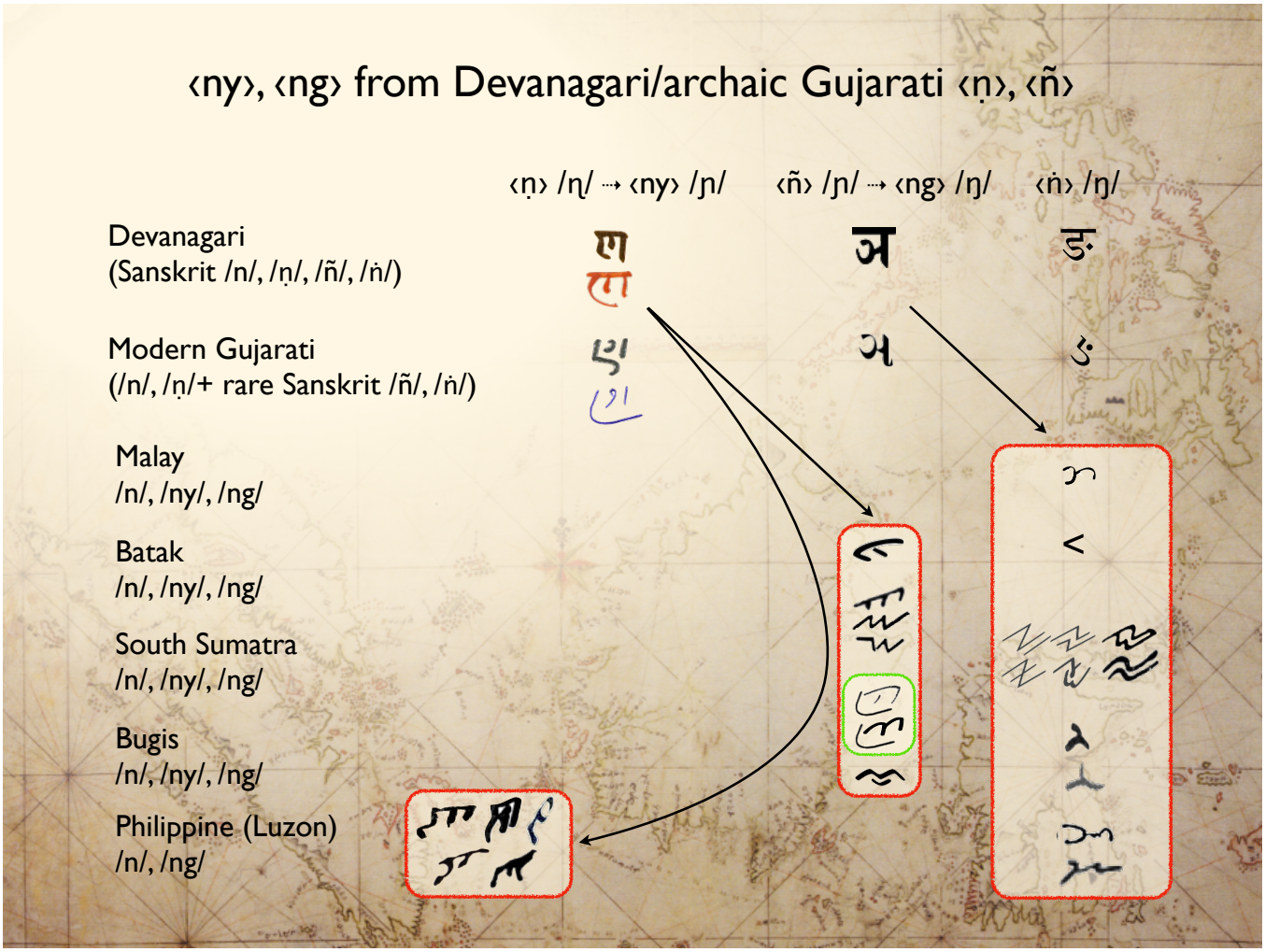
It appears that the first adaptation reused the common letter for retroflex /ɳ/ (the non-dental/alveolar nasal of Gujarati) to the most phonetically similar Malay sound, the palatal /ɲ/ (ny). Malay velar /ŋ/ (ng), pronounced farther back in the mouth, was a less likely candidate for this adaptation and was thus left without a letter. To fill this gap, two Sanskrit letters (‹ñ› and ‹ṅ›) were available. Since archaic Gujarati script was at the time essentially a low/intimate register script at the opposite end of the continuum from the high/formal register Devanagari of religious texts, it is not surprising that Gujarati merchants would have some familiarity with the extra Devanagari ‹ñ› and ‹ṅ› letters, if not with their values. (This is analogous to the familiarity of English-language literates with accented or otherwise derived letters used in various European languages (e.g. üöøçåæř) but not necessarily with their actual sound values.)

It is plausible that the choice of ‹ñ› may have been due in part to a poor awareness that ‹ṅ› was the “correct” equivalent — that ‹ṇ› was adapted to represent /ɲ/ (ny) instead of simply borrowing the direct equivalent ‹ñ› supports this hypothesis — but it also seems that the shapes of the two letters may have played a role in the choice made. Comparing the two, it is clear that the shape of ‹ñ›, with its right-hand stem, is more typical compared to ‹ṅ› with its right-hand dot. In the simplified form it would have taken on, ‹ṅ› would also be fairly easily confused with the ‘3’-like shape of ‹u› (conserved in Philippine script).



Taking into account these considerations, it appears that there is good evidence that archaic Gujarati ‹ṇ› and a loosely written variant of Devanagari ‹ñ› were adapted to represent the Malay non-dental/alveolar nasal consonants not through sophisticated phonetic analysis and comparison, but by makeshift adaptation via the perception of correspondences between Malay and Gujarati phonology and incomplete familiarity with marginal nasal consonant letters used to write Sanskrit. (In fact, more drastic adaptations of Bugis-Makassarese script were made to represent certain sounds of the Ende-Lio language not present in the simpler sound systems of Bugis and Makassarese.)

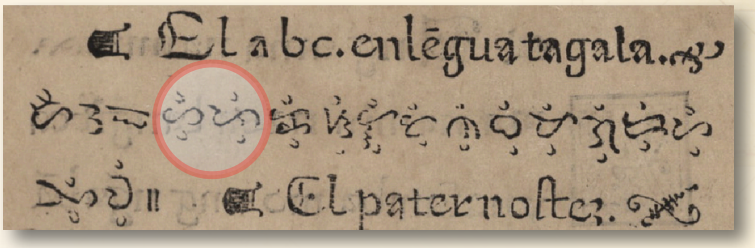
With this in mind, the shapes of ‹ng› in SSP scripts derive in expected ways as discussed in preceding slides; these shapes and the corresponding Devanagari and modern Gujarati letters are recapitulated on the right-hand side of this slide. The left-hand side illustrates the relationship between Devanagari/archaic Gujarati ‹ṇ› on the one hand and Sumatran and Sulawesi ‹ny› and a less common variant for ‹n› shape. The first shape is from Batak, which shows the kind of direct and regular changes discussed earlier, The next three are from South Sumatra; the first of these (a less common Lampung variant) shows a more direct shape relationship to the D/G letter. Despite the more innovative shapes of other letters in this Lampung script variety, it is a plausible hypothesis that this particular letter reflects an archaic South Sumatran shape that was reshaped in most varieties to the zigzag shapes below it.



The Philippine shape is especially interesting given its close (and in some hands) nearly exact correspondence to Devanagari/archaic Gujarati ‹ṇ›. Given its more complex, rectangular shape compared to the more usual arch-plus-adjunct variant, it stands to reason that the variant illustrated here may have originally been adopted from archaic Bugis-Makassarese script as its ‹ny›, but reinterpreted as ‹n› in Tagalog, which had no constrastive /ɲ/ palatal nasal phoneme.

This in turn is a clue that archaic BM ‹ny› likely had a shape very similar to these, consistent with the arguments presented above for the original letter shapes of the script in relation their to Philippine counterparts. Reconstructed transitional variants are illustrated in the green-outlined rectangle, above the modern standard shape. The first shows a plausible first stage in the development of the letter; below it, we see the shift to the general arch-based stereotype of the script. Following this reconstructed stage, the initial broad curving stroke of the letter would have been analysed as a secondary adjunct (cf. the discussion above of BM ‹b› and ‹ng›), which then reduced in size and underwent adjunct drift downward and to the right, beneath the main double arch body of the letter.

Extra stroke on Philippine ‹p›: miscopying from an early recital order



Early Philippine recital order in *Doctrina Christiana* (1593): ‹h›, ‹p› highlighted

Archaic Gujarati shapes

Reconstructed proto-script shapes

Shift of adjunct stroke from ‹h› to ‹p›

Philippine ‹h›, ‹p› in *Doctrina Christiana*

In the majority of Indic scripts, reflecting their Brahmi ancestor, ‹p› has a fairly minimalist shape consisting of an upward-concave “bowl” enclosing a counter open to the top. The added adjunct stroke on the right-hand side of the Philippine letter is unique among Indic scripts and has no counterpart in any of the Sumatran or Sulawesi scripts. This raises the question of how it originated. Related to this is the fact that the adjunct stroke on Devanagari and Gujarati ‹h› — which might be expected to be reflected in the Philippine letter given the overall close correspondences between the scripts — is absent on the Philippine letter.

A clue to resolving these two questions comes from the early recital order given in the 1593 *Doctrina Christiana*, where the two letters appear side by side. Taking the similarity of body shape in the reconstructions of the proto-script shapes (and in the Philippine shapes themselves), it seems plausible that at some point during the early transmission of the script, a writer may have miscopied the letters and mistakenly transferred the adjunct stroke of ‹h› onto ‹p›, as shown on this slide.

Devanagari/Gujarati vowel signs and Indonesian script counterparts

The regular and systematic structural correspondences reviewed above are convincing evidence that the base letters of the SSP scripts originate — apart from a very few later borrowings — in an early informal variety of Devanagari script introduced to Sumatran Malays by Gujarati merchants.

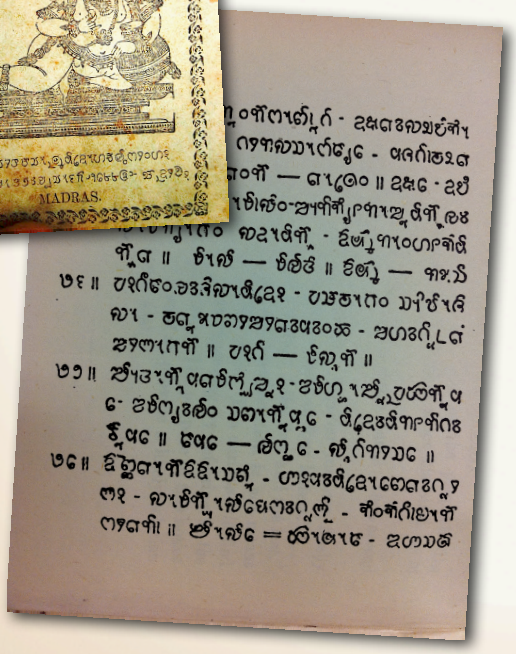
This is not the case for the inventory of dependent (bound) signs used to indicate vowels and certain coda consonants, and to indicate a consonant is to be read without any vowel, i.e. as a coda consonant.

As seen in the table of vowel and coda consonant signs shown at the beginning of this presentation and this table, although there are sporadic resemblances in some scripts with certain Devanagari/Gujarati dependent signs, it seems quite clear that across all the SSP scripts, the vowel/coda marking system is derived directly from the Old Javanese Kawi inventory. As discussed earlier, a signature feature of Batak scripts is a range of simple angled strokes corresponding to earlier clockwise curls, which came to be prohibited at some point in the development of the proto-Batak script. Given the otherwise clear relationship of the Batak vowel/coda marking system to the South Sumatran and Javanese inventories, it is certain that the similarity of Batak ‹-e› (via upward and rightward adjunct drift) and ‹-Ø› to their Gujarati counterparts is a chance result of the way the prohibition on initial curls restructured the relevant signs.

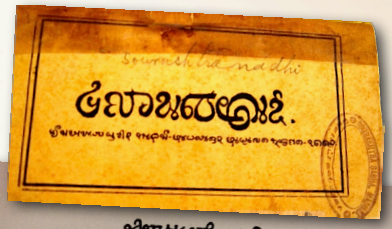
It is surprising on the face of it that despite the systematic evidence that their base letters derived from early informal Devanagari, the SSP scripts all use Javanese-derived vowel/coda marking systems. However, there are indications that the archaic Gujarati script transmitted by Indian merchants may have been a shorthand in which vowels were not normally written and that after its introduction, Malay users of the proto-script supplemented it with vowel and coda signs they were familiar with from Javanese script, resulting in a hybrid dichotomy between base letters with one origin and vowel/coda signs with another.

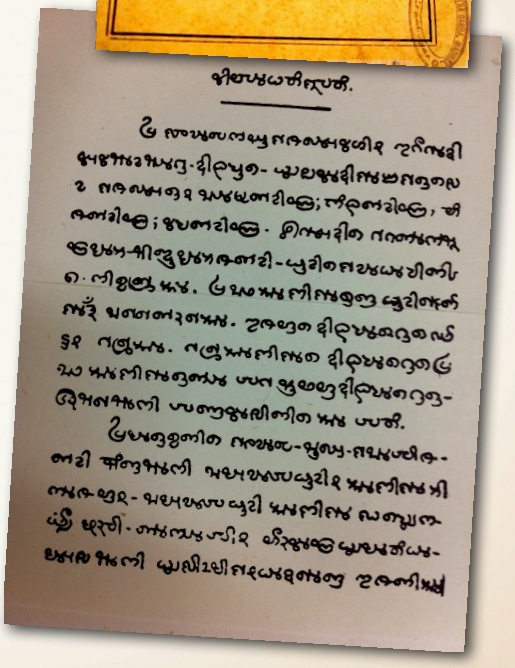
Sourashtra (Tamil Nadu)

Rama Rao
1902



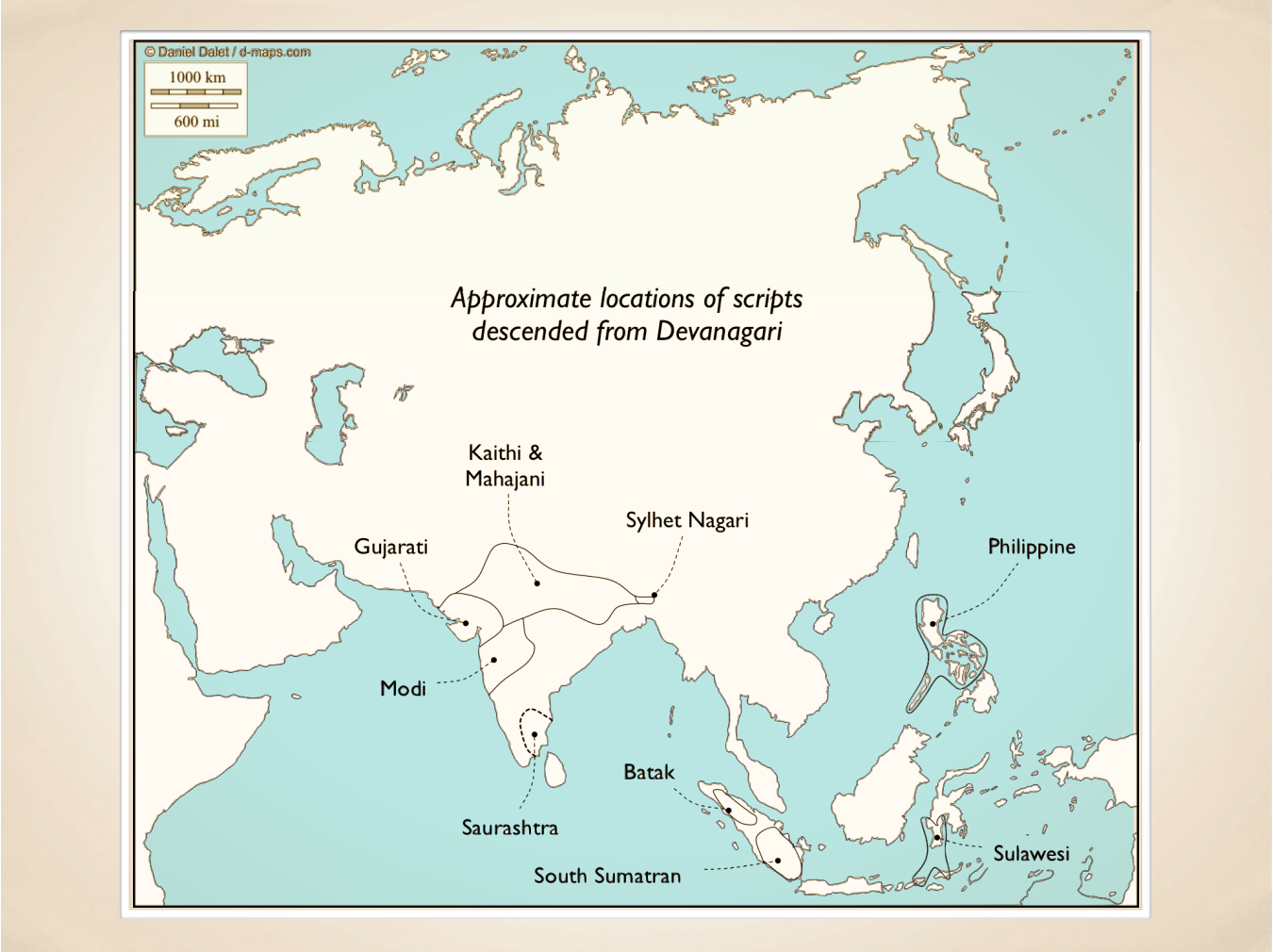
Halivi
1880





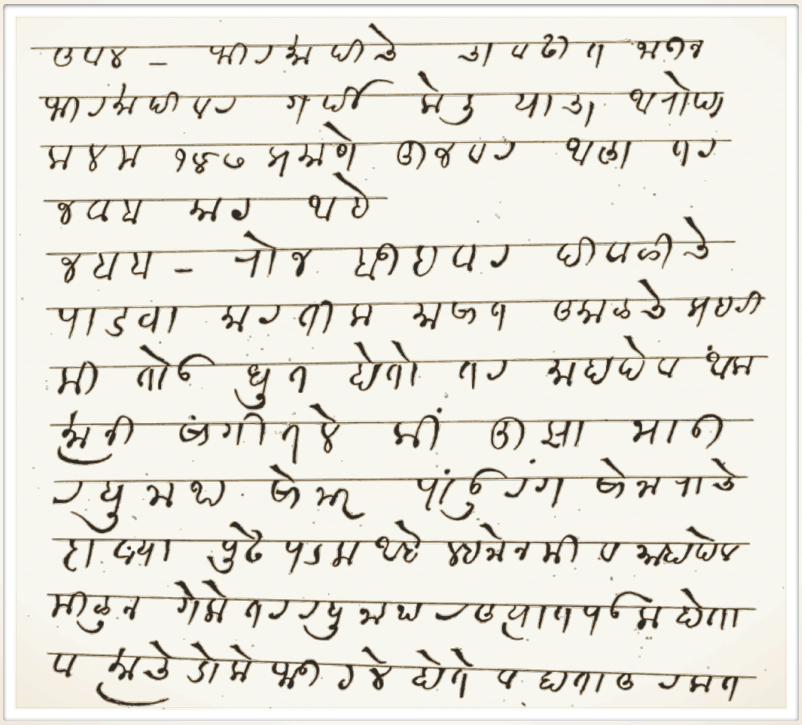
A similar case of hybrid structure can be found in two related scripts particular to the Sourashtran-speaking community (originally silk-weavers) of Tamil Nadu in southeast India. Sourashtrans speak an Indo-Aryan language related to Gujarati and Marathi in the northwest and have a tradition that they migrated from there between the 13th and 16th centuries.

This slide illustrates texts in the two known related Sourashtran scripts, reproduced from Randle (1944). The samples given are from the earliest source for each script known to Randle at writing. Above them are images of later book covers in the two scripts from the online image collection of Subramanian Obula.



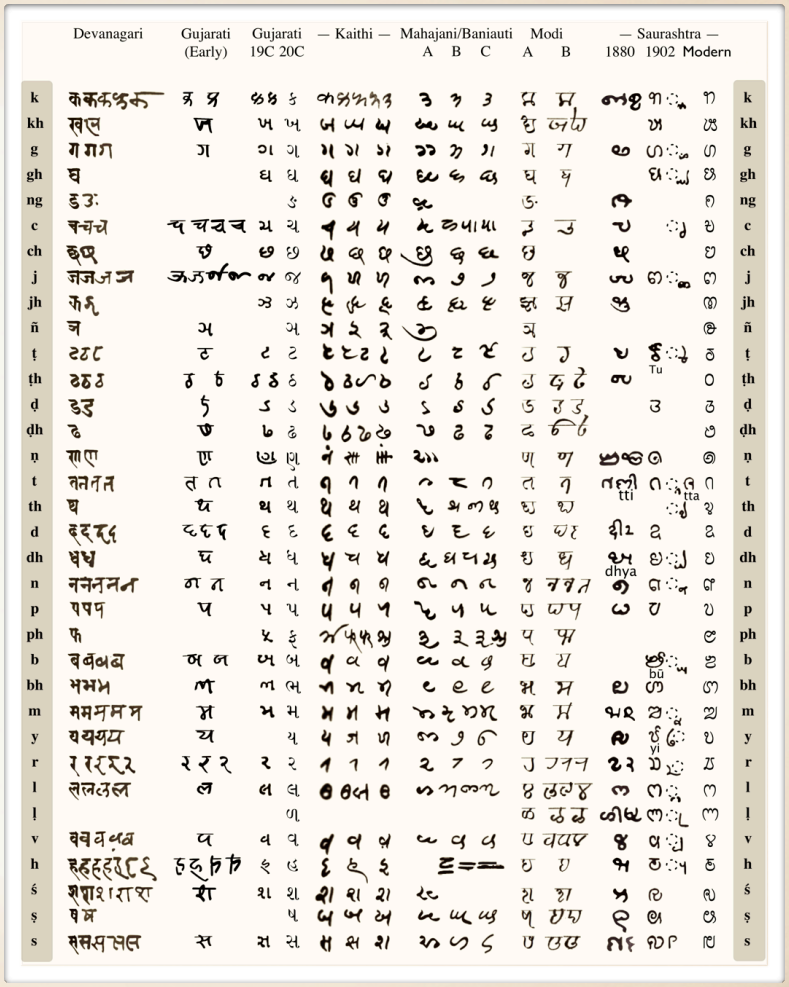
This slide illustrates the approximate locations of various scripts derived from Devanagari.

Modi (Maharashtra)



Moḍi was formerly the official court script of Maharashtra until its replacement by Devanagari, which is now the standard script for the Maraṭhi language. The letter shapes of Moḍi, like Gujarati, are derived from Gujarati and some of them are similar in shape to the corresponding Gujarati letters. The most characteristic features of Modi are the strong tendency for the body of letters to join the righ-hand stem at the bottom rather than at the top, and a similar tendency to develop excrescent loops at joins between stroke segments.

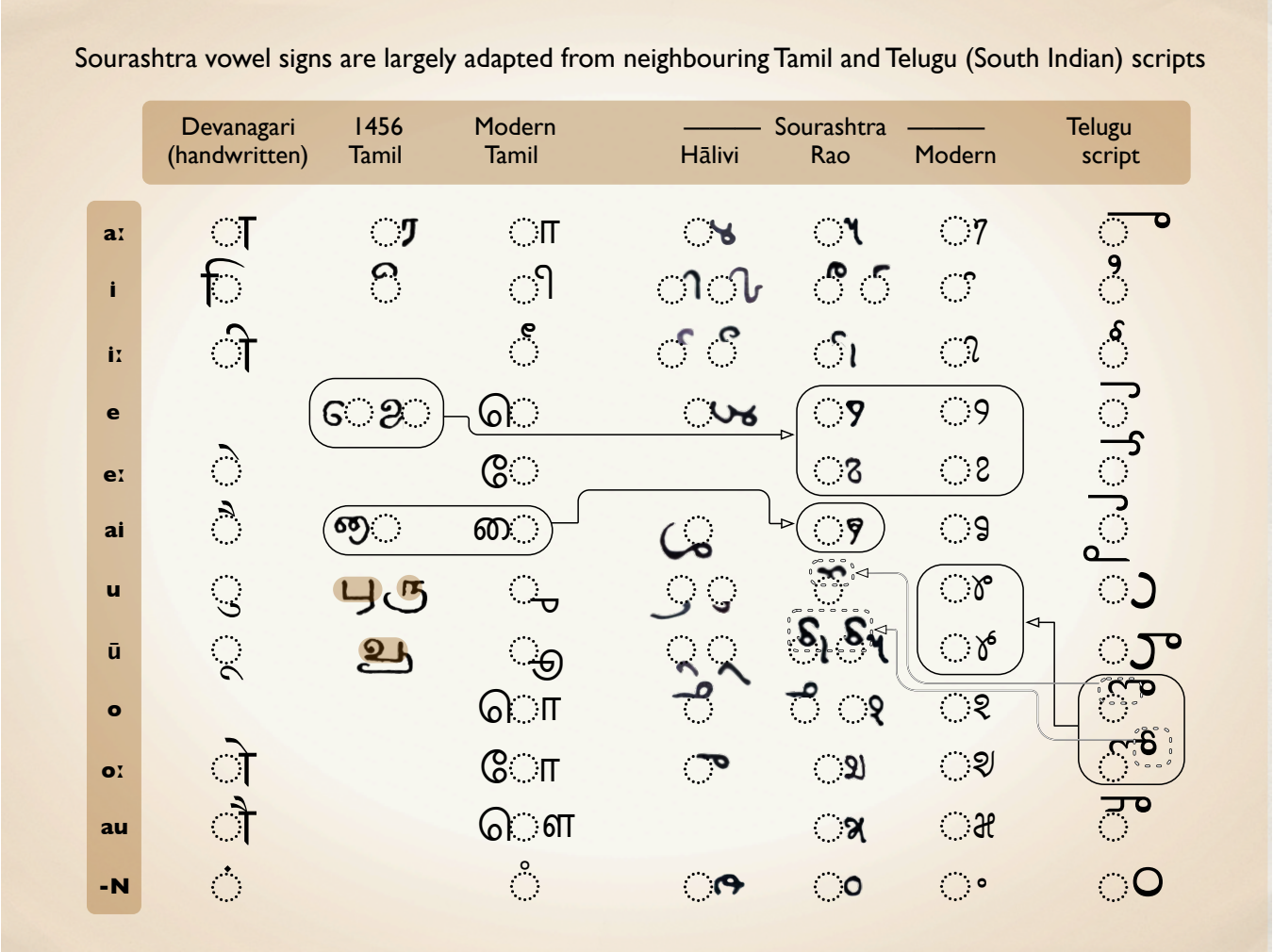
Sourashtran script varieties are related to informal North-Indian scripts descended from Nagari



Comparing the letters of the two Sourashtran varieties (vowel letters excluded because of space constraints) reveals numerous and systematic correspondences with North Indian scripts. Both scripts share with Modi the tendency to join the body to a rising stroke on the right, as well as to join stroke segments with excrescent loops. However, the similarities do not indicate that either Sourashtran script descends from Modi; rather, the three scripts, sharing similar features, appear to descend from a common predecessor in Maharashtra.

Otherwise, the letter shapes in the two scripts, though often different, relate directly to the informal shapes found in early Gujarati, Kaithi (the closely related script formerly widespread in northern India east of Gujarat) and the commercial Mahajani/Baniauti scripts. It can be concluded that both Sourashtran scripts developed in northwestern India, more specifically in Maharashtra during the period when the Sourashtran community, according to tradition, resided there.

The clear relation to the commercial scripts is interesting: these scripts (and similar scripts used in the old Northwest, now Pakistan) are known for their use as shorthands in which post-consonantal vowels were usually not written — much like Arabic script, whose use in the Northwest may have been an inspiration for the development of these scripts, not known from the south or east of India.



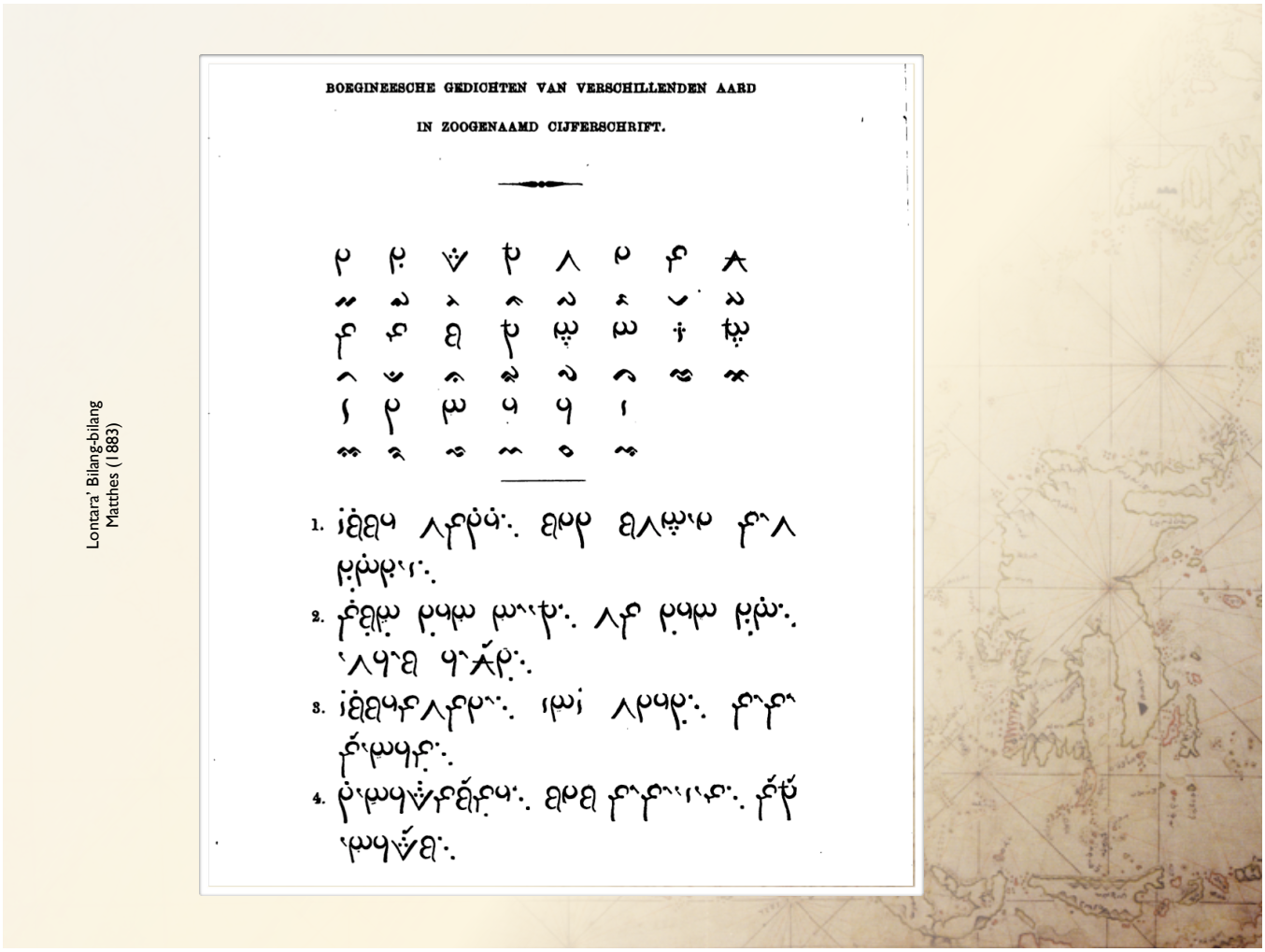
Given the consistent similarities to commercial scripts it is likely that the Sourashtran script varieties began as commercial shorthands, complete with the convention of not writing post-consonantal vowels. In fact, the scripts' appearance for formal purposes dates only to the late 19th and early 20th centuries. In this connection, Dave's (1976) observation that Tamils often refer to Sourashtrans as *ceṭṭi* (from Sanskrit *śreṣṭhī* 'merchant') is significant.

In this slide, we see not only that the vowel and coda consonant signs in the two scripts are for the most part clearly unrelated to those of Devanagari and related North Indian scripts, but they are also for the most part unrelated to each other. Only the signs for ‹-i› and ‹-ī› and a now archaic sign for ‹-o› are similar between the two scripts. The Hāḷivī script has signs for ‹-u› and ‹-ū› that are both similar to the corresponding Devanagari signs.

However, we do find clear similarities between the vowel and coda signs in both scripts and vowel and coda signs in Telugu and Tamil scripts. Before their migration came to its end in Tamil Nadu, the Sourashtrans stayed for a couple of centuries in Telugu-speaking regions and Dave observes that many still speak Telugu. Although various signs are either copied from one or the other, or adapted in shape, their values do not always correspond directly to the value in Tamil or Telugu.

The likely conclusion from all these observations is that at some point, in response to a perceived need to make them as easy to read as the surrounding Tamil and Telugu scripts, vowel and coda signs were added. These were not transferred wholesale from Tamil or Telugu, but adapted from the two scripts (together with Devanagari ‹-u› and ‹-ū› in Hāḷivī's script) in different ways.

The fact that this creation of hybrid scripts is observed for Sourashtran, whose introduction to southern India took place in a time frame similar to the likely 14th century introduction of archaic Gujarati script to southeastern Sumatra, corroborates the hypothesis that the hybrid structure of the Sumatran proto-script arose in a similar manner.



Lontara' Bilang-bilang
Matthes (1883)

The hybrid structure of the SSP proto-script is reflected in other scripts of the region. Makassarese *Jangang-jangang* script, which shows clear signs of being contrived rather than developing naturally from a single South Sumatran antecedent, does not use the South Sumatran vowel signs that might be expected. Instead, it uses the standard vowel signs of Bugis-Makassarese script.

Another script from South Sulawesi is *Lontara' bilang-bilang* 'Number script'. This is based on a cipher script of Indian origin that used Arabic numerals to represent the Arabic letters with corresponding numeral values in its earlier Aramaic-based letter order (similar to that of Hebrew). This was modified at one stage to represent the sounds of Malay absent in Arabic, and was then further adapted to Bugis by I Colliq Pujie, a 19th century Bugis princess and intellectual, according to Rahman (2014). Although the base letters of this script are all derived from Arabic numerals (in their usual shapes as used in the Malay archipelago), the vowel signs added to them are the standard vowel signs of Bugis-Makassarese script seen six slides above.

Gangga Malayu
Kern (1908)

Journal of the Royal Anthropological Institute, Vol. XXXVIII, 1908, Plate XVII.

ARABIC-MALAY ALPHABET AND GANGGA MALAYU COMPARED.

ا ب ت ث ج چ ح خ د ذ ر ز س ش

alif ba ta tha jim cha ha kha dal zal ra ze sin shin

ص ض ط ظ ع غ ڠ ف ق ک ڬ ل م

sod thod toh zoh ain rhin nga pa kop kap ga lam mim

ن و ه ي ڽ

nun waru ha ya nya

a i u an in un ba bi bu

LETTER IN GANGGA MALAYU WITH TRANSLITERATION.

di - ma'alumkan mĕngadap tuwan G. M. L(a)id - lāu yang muliya

di - ma'alumkan yang sahya - pun bĕsar duka chita dari

ḳal ḳurup gaṇga (= gangga) Malayu ini ka - rana tiyada-lah

di - dapāti bageimana mu - la - nya yang bulih di - dapati ulih

ḥapa sahya itu mĕl(a)inkan yang sahya dapat itu dengan penga-

jaran dariyada bapa sahya itu juwa ada-nya.

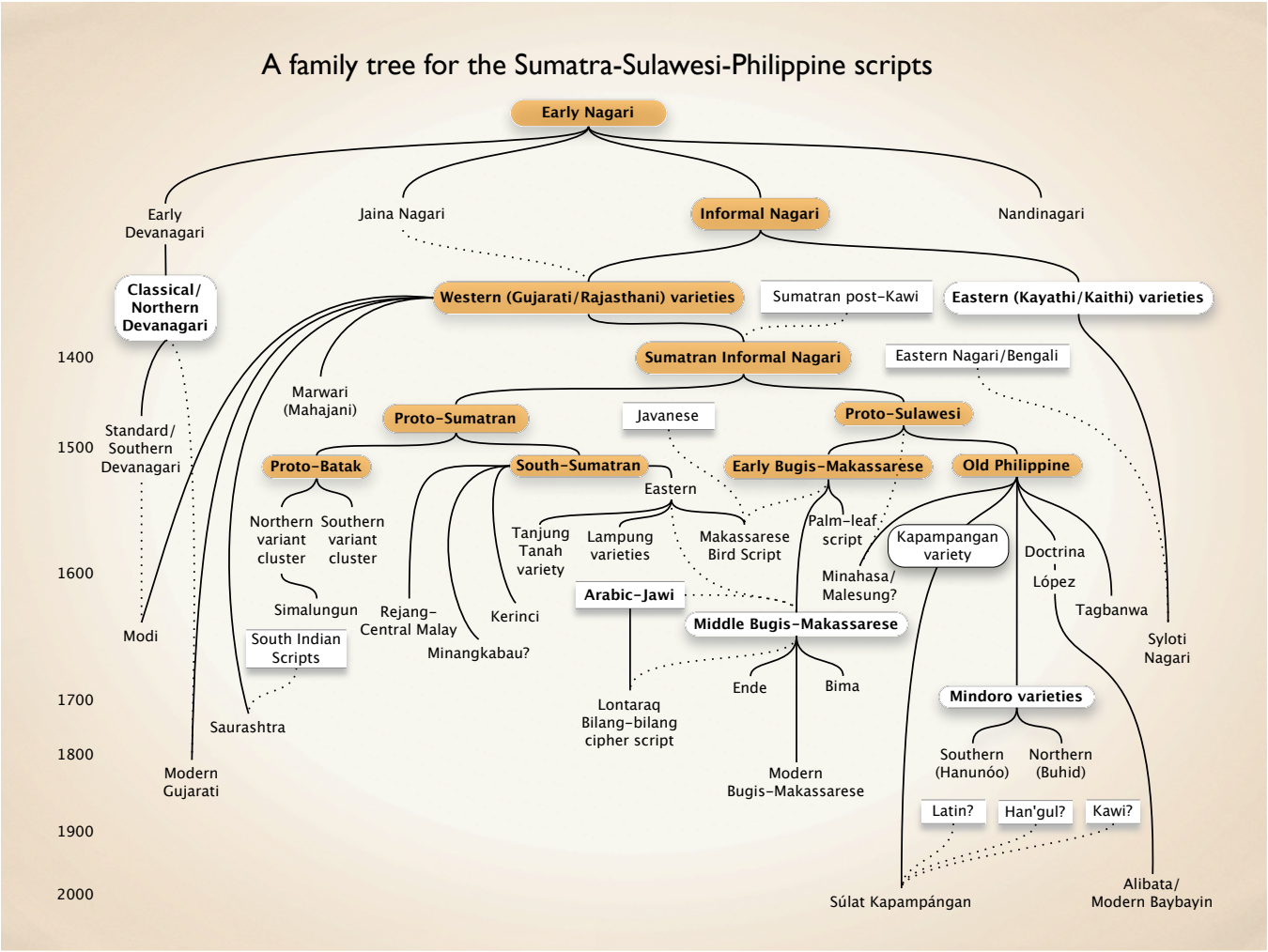
dariyada Muḥammad Sirajud'in.

1 2 3 4 5 6 7 8 9 10

A MALAY CIPHER ALPHABET.

A third case of a similar hybrid script in the archipelago is *Gangga Malayu*, reported by Kern (1908). This uses apparently invented letter shapes; some are loosely based on the corresponding letter in Jawi, the Malay version of Arabic script, while others with no clear shape relationship nonetheless reflect the paradigmatic relationships between their Jawi counterparts. The vowel signs, however, are derived directly from the shapes of corresponding vowel signs in modern Javanese script.

Gangga Malayu, *Jangang-jangang* and *Lontara' bilang-bilang* are all contrived to one degree or another, and all three share a hybrid structure pairing letters from one source with vowel (and in GM, vowel-plus-coda) signs from a completely different source or sources. This consistent hybrid structure is puzzling; however, given the similar hybrid structure of the Sumatran proto-script reconstructed here, it seems at least plausible that a memory of the hybrid genesis of that script survived in the lore surrounding writing in the archipelago, together with the didactic combination of dependent vowel and/or coda signs with base consonants and letter recitation orders surviving in Batak and Philippine scripts as well as the Javanese *Hanacaraka* order, all ultimately based on an alternative phonetic grouping of letters different from the standard Indic place of articulation-based *varṇamāla* (Miller 2014).



Given the preceding arguments, it is now possible to construct a family tree to elucidate the relationships between the Sumatra-Sulawesi-Philippine scripts and between these and other scripts descended from Devanagari. Although some details need to be updated (among them the existence of two distinct Sourashtran scripts and their origin in a single ancestor shared with Moḍi), this tree gives a good initial idea of the variety of descendants of Nagari or early Devanagari and the complex interrelationships in the genesis of its descendants in the Malay archipelago.