

Please reclassify the Philippine *pamudpod*

Fredrick R. Brennan

(

 မြန်မာ့ သမ္မတမြန်မာနိုင်ငံတော်

 Psihedelista

 フレッド・ブレンナン

copypaste @ *kittens.ph*
)

23 September 2020

သမ္မတမြန်မာနိုင်ငံတော်
 သမ္မတမြန်မာနိုင်ငံတော်

သမ္မတမြန်မာနိုင်ငံတော်
 သမ္မတမြန်မာနိုင်ငံတော်

ACKNOWLEDGEMENTS

Antoon Postma, Mangyan Cultural Center

For popularizing the Mangyan script and for his efforts in preserving prehispanic Philippine culture which continue to resound today, far beyond Mindoro. *In memoriam*.

Dave Crossland, Google Inc.

For reminding me this issue exists.

Deborah Anderson, University of California @ Berkeley Script Encoding Initiative

For her tireless review of script proposals by n00bs like me.

Roozbeh Pournader, Unicode Consortium

For his review and helpful comments on an early version of this proposal.

Norman de los Santos, BHM Publishing Inc.

For his significant help in authoring § “On the other encoded Philippine scripts”.

Simon Cozens & Caleb Maclennan, S_IL^E

For their help with S_IL^E, an experimental typesetting engine and alternative to L^AT_EX.

Disclosure: This paper came about while I was working on a font project commissioned by Google Inc. on a contract basis, Noto Sans Tagalog. I was offered this contract because of my successful Unicode proposal which led to the pipelining of U+170D, [L2/19-258R](#). I am passionate about the subject and was not paid specifically to write this paper; this paper is not part of my contract.

Table of Contents

Quick summary	4
ScriptExtensions.txt	4
Why?	5
Why a <i>pamudpod</i> makes sense in <i>baybayin</i>	5
The <i>pamudpod</i> is, indeed, native	6
Why the <i>baybayin</i> community needs, once again, the Consortium’s intervention	8
Why not a new character?	9
Evidence of <i>pamudpod</i> in use in modern <i>baybayin</i>	11
Kalem (Anak Bathala), №1, 2013, BHM Publishing, graphic novel	11
Kristian Kabuay’s <i>Surat Magazine</i> , №1, December 2018	11
Wikipedia	12
Bohol–Panglao International Airport mural	13
On the other encoded Philippine scripts	14
Alternative solution: new character	15
Concerns about alternative solution	15
Existing fonts	15
Existing documents	15
Possible new character’s properties	15

0.1 Quick summary

If this proposal is accepted, the following change will be made:

- ◌, U+1734, HANUNOO SIGN PAMUDPOD, will receive character name alias PHILIPPINE PAMUDPOD, and be moved from the script Hano to the script Zyyy to match the Philippine punctuation //, U+1736, the PHILIPPINE DOUBLE PUNCTUATION.

I also argue that U+1734 should be reclassified as Mc—Mark, Spacing—but this is not critical for solving the user community’s woes.

0.1.1 ScriptExtensions.txt


I further suggest that the *pamudpod* be added to `ScriptExtensions.txt` with the following data:

```
# Script_Extensions=Hano Tglg
1734 ; Hano Tglg # Mc HANUNOO SIGN PAMUDPOD
# Total code points: 1
```

This is not the only way to solve the problem. I offer the alternative solution of a new character, should the UTC wish to avail of it, in § Alternative solution.


0.2 Why?

0.2.1 Why a *pamudpod* makes sense in *baybayin*

The *pamudpod* is a *virama*. *Baybayin*, the local name for what The Unicode Standard calls the Tagalog script, though, already has a *virama* encoded—, U+1714, TAGALOG SIGN VIRAMA, known in Tagalog as the *krus kudlit*.

So, why another *virama*? Well, first of all, it's being used, as I'll demonstrate, but the use of the *pamudpod* makes sense and is being enthusiastically adopted by a large portion of the *baybayin*-using community; several subject matter experts use it in their work, such as Norman de los Santos and Kristian Kabuay.

As far as I can tell, these are some of the reasons a *pamudpod* is preferred by many over the currently encoded *krus kudlit*:

- A reason to revive *baybayin* is to celebrate Filipino culture and the Tagalog language. The Spanish *krus kudlit*, however, is a symbol of colonization and Spanish administration of the country. Indeed, the very shape of the *krus kudlit* was meant to “civilize” the ancient Tagalogs: it is a cross, and this is not an accident; it was meant by its creator to symbolize Christianity. Even Christians, as most Filipinos are,¹ can find this colonial history problematic. The priest who first used it in 1620, Francisco Lopez, was perhaps unaware that native alternatives existed; or, seeing his chance to introduce a new letter to a “heathen” alphabet, chose a cross. The first book he printed with it, the *Doctrina Christiana*, was based on the teachings of Cardinal Roberto Bellarmine, one of Galileo's inquisitors.²
- While some fonts attempt to mask the symbolism of the *krus kudlit* by rendering it as an X, as , and some fonts such as [Noto Sans Tagalog v3+](#) indeed contain both selectable via OpenType features, in the case of that font, Stylistic Alternates (`sa1t`), it is impossible to convey such a shape reliably in plaintext. Furthermore, Unicode would never encode the X-shaped *krus kudlit* as a separate character, as it is so similar in function and semantics to the regular *krus kudlit*.

1. Roman Catholics aren't seen as Christian by all denominations, but such squabbles are not the point of this paper.

2. Morrow, Paul (2002). *Baybayin—The Ancient Script of the Philippines*. 14 July 2010 ed.

- The inventor of the *krus kudlit* disparaged the *baybayin* script (emphasis mine):³
- “ [T]he reason for putting the text of the Doctrina in Tagalog type...has been to begin the correction of the said Tagalog script, which, as it is, is so defective and confused (because of not having any method until now for expressing final consonants—I mean, those without vowels) that the most learned reader has to stop and ponder over many words to decide on the pronunciation which the writer intended. ”

This disparagement was enthusiastically repeated by Spanish scholars right up until the end of the Spanish presence in the Philippines.⁴

- From a mere design perspective, in most common fonts the *krus kudlit* is very difficult to distinguish from the regular *kudlit*. While in the beginning the cross shape was elongated,⁵ most fonts now render it as a mere plus sign, perhaps partially to combat the aforementioned perceived ills. Consider, if you would, $\text{—}\text{⋈}\text{—}\text{⋈}$ and $\text{—}\text{⋈}\text{—}\text{⋈}$. Which one is the Philippine dessert *halo-halo*, and which *halhal*, meaning “prone to laughing for no reason”? Certainly $\text{—}\text{⋈}\text{—}\text{⋈}$ is much easier to distinguish in small print.

0.2.2 The *pamudpod* is, indeed, native

Some would say that really the *pamudpod* is no better as it was “invented” by a Dutchman, Antoon Postma, who married into the Mangyan people and studied their language, Hanuno’o, during his life. Indeed, this argument was even tried against Postma’s Mangyan Cultural Center⁶ by an organization who claimed “*Surat Mangyan* (Hanuno’o) was changed to suit an anthropologist’s vision”.⁷

Even if this were true, it is certainly better to use a symbol originated in the independent Philippines which the Mangyan people voluntarily accepted rather than one which was forced upon the Tagalogs by a colonial dictatorship. However, it is not true; as the Mangyan Cultural Center notes...⁸

3. Scott, William Henry (1994). *Barangay: Sixteenth-Century Philippine Culture and Society*. Manila: Ateneo de Manila University Press—via [Internet Archive](#). Quoting and translating from Spanish an 1895 reprint of López, Francisco (1621). *Libro a naisurátan ámin ti bagás ti Doctrina Christiana nga naisúrat iti libro ti Cardenal a agnagan Belarmino* p. LXII.

4. Cipriano Marcilla y Martín (1895). *Estudio de los antiguos alfabetos filipinos*. Manila—via [Internet Archive](#). pp. 10, 57

5. Ibid. p. 20. “⋈”

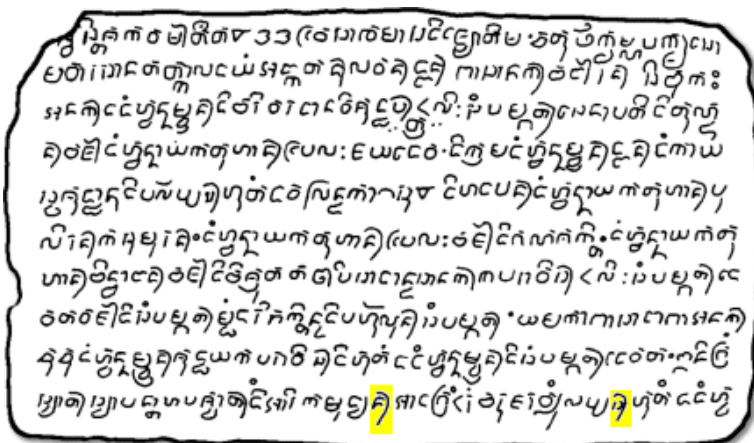
6. Postma is now deceased, but he founded the organization in 2000.

7. It seems that the organization who argued it, “Tsinelas Republic”, recanted their statement; their [Facebook page](#) no longer links to it, and the [page itself](#) returns an error. Internet Archive doesn’t have it; all that remains is the [Mangyan Cultural Center’s response](#).

8. Ibid.

“ Thank you so much Fr. Ferdinand Bajao for taking the time to explain so well Postma’s greatest contribution in preserving and keeping alive the rich Mangyan cultural heritage—the extensive documentation of the Mangyan script, and not to mention the over 20,000 *ambaban* he documented and [which are] now preserved [in] the Library of Congress in D.C. In fact, the vowel killer which is called *pamudpod* in the *Surat Mangyan* is not his invention. [It] was introduced in other writing systems as early as the 17th or 18th century for the Tagalog Baybayin. It was also used in the 900 AD Laguna Copperplate Inscription, the oldest Philippine document, which was discovered after Postma introduced the *pamudpod*. Even in the Tagbanwa script, a vowel killer was also introduced. Postma believed that the use of the *pamudpod* is necessary to keep the script alive and the Mangyans approved this for instruction purposes. Postma did not alter the original script. ”

Indeed, a *virama* with an awfully similar shape to the *pamudpod* can be seen in the mentioned inscription:⁹



A quick comparison of Morrow’s Latin transcription of the Laguna Copperplate Inscription with the inscription itself shows that the *pamudpod* serves the purpose of a *virama* here too.

Third party verification of this can be had by referring to L2/12-125.¹⁰ That paper recommends the encoding of Later Kawi, the script of the Laguna Copperplate Inscription, and specifically its *virama* at U+11D3F.¹¹ 11D3F ◌ KAWI SIGN VIRAMA ¹²

9. This version via Paul Morrow; I colored the backgrounds of two of the glyphs with a combining *pamudpod*-shaped *virama* yellow.

10. Pandey, Anshuman (2012). “Preliminary Proposal to Encode the Kawi Script”. pp. 1.

11. This codepoint now contains ◌ MASARAM GONDI VOWEL SIGN AU.

12. Ibid. p. 4.

Where then did this misunderstanding come from? It is indeed true that Postma introduced the *pamudpod* to the Mangyan people as part of his efforts to revitalize the Hanuno’o script. However, always the scholar, Postma was careful to borrow a construct from already existing Philippine scripts —not just simply draw a cross and call it a day.

0.2.3 Why the *baybayin* community needs, once again, the Consortium’s intervention

Of course, it is bad on its own that a character has a misleading name which does not reflect its use in real life. And it is bad on its own that a character so routinely used in *baybayin* (Tagalog)¹³ is marked as belonging to the Hanuno’o script.

I see how this mistake happened, of course. Most other languages, especially those like Devanagari and Bengali from which the word *virama* originates, cannot share a *virama*. The very idea is absurd; they don’t see the *virama* as a spacing mark, as Philippine script practitioners do.

And of course, due to systemic bias found in all tech, it’s mostly white Europeans and Americans engaging with the Unicode Consortium. And to top it all off, before my last Tagalog proposal, [L2/19-258R](#), a proposal had not been made since Michael Everson’s [L2/99-014](#). Obviously, fewer Filipinos were online then, fewer books were available (especially online), and so on.

However, due to the fact that some browsers¹⁴ rely on Unicode’s script determinations in ways I think are undue,¹⁵ the urgency is increased as we cannot simply work around the issue with clever OpenType tables. To explain...

Consider the sequence:

That is: PA RA PAMUDPOD KA KUDLIT_I NGA PAMUDPOD.

Assuming Unicode 14.0 is out and contains 𑀓 (RA),¹⁶ the scripts are divided thus, with Tagalog in blue and Hanuno’o in red:

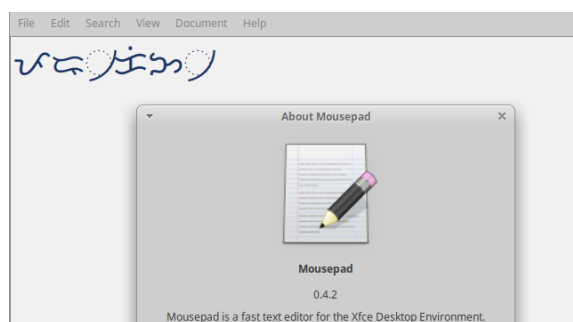
No problem, right? We see that S_IL^E is handling it fine. Alas, this is mere good fortune; behold Mousepad:

13. To the point where *pamudpod* appear multiple times in the first sentence of [Wikipedia’s baybayin](#) article, and have done so for years.

14. *cough* Chrome! *cough*

15. [Perhaps an issue for the legendary Text Shaping Working Group?](#)

16. Since in reality it isn’t, I have a whole universe of fun problems involving the 𑀓 in my fonts. See [Noto Sans Tagalog v3 documentation](#), pp. 14–15, for starters...



OpenType shaping expert Khaled Hosny puts it succinctly: there’s no standard way to chop glyphs into HarfBuzz (or whatever shaper) runs, a process known as “script itemization”.¹⁷

So, even if a font contains the needed OpenType Layout instructions, if the shaper refuses to ever put the glyphs in the same run, these instructions are worthless. Chrome therefore sees the *pamudpod* as the start of the run, and tries to be “helpful” by inserting a ◌, resulting in this derpy output:



The result is particularly broken because for some reason the *pamudpod* is considered by Unicode as a combining character; this makes little sense for *baybayin*, and in my fonts, I don’t implement the *pamudpod* with anchors and mark, but rather treat it as a punctuation mark that crosses its left bearing. In other scripts, I understand that the *virama* can ligate, and cause other transformations in the shapes of characters; this is not true for either Hanuno’o or *baybayin*, where it is static and has no effect on the shapes of the letters before or after it, and its shape does not change depending on where it is, nor does its attachment location.

Of course, it can be said that application-level bugs are not the Consortium’s problem, in which case I fall back on my trusty argument of correctness for why it needs to change! 🐛

It’s simply not correct to say that the *pamudpod* is only used in Hanuno’o...

0.2.4 Why not a new character?

The primary reason that I am not requesting a new character be encoded, but rather that U+1734 be repurposed, is that the *pamudpod* is already widely used how I describe as of this writing in 2020. Many of the fonts of Norman de los Santos, for example, contain the *pamudpod*.¹⁸ Many documents are already using U+1734 in *baybayin* text.

It is not abnormal for other scripts, even widely used scripts, to share marks, even marks that do not look identical. For example, both Cyrillic and Latin use U+0306, ◌̆, the COMBINING BREVE. We can see it in the example of the letter Ў, here encoded as U+0418 U+0306, and in the letter Ü, here U+0055 U+0306.

17. Hosny, Khaled. [Noto CJK Issue №172](#).

18. See full list in [L2/19-258R](#)

For clarity, here they are enlarged:

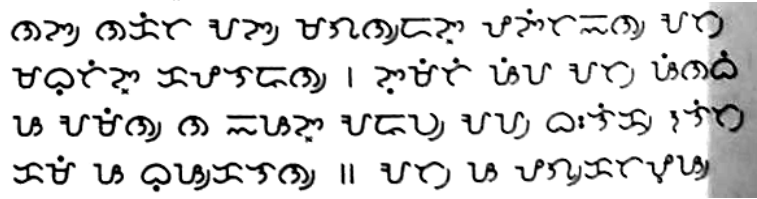
Йӱ Йӱ

It can be seen that the font I'm using, EB Garamond, displays the two breves quite differently, and it is not an outlier; most fonts do, for example, Libertinus Serif:

Йӱ Йӱ

It is my opinion that changing the encoding of the *рамудпод* at this late date would add considerable confusion to the encoding of *байбайин* with no discernible benefit.

The magazine is multilingual, however there is a long section written in the Tagalog *baybayin* script which uses a *pamudpod* everywhere a *virama* is called for. I will transcribe a line found on page 58 just as an example:



Transcription: 𑄆𑄗𑄗 𑄆𑄗𑄗 𑄆𑄗𑄗 𑄆𑄗𑄗𑄗𑄗 𑄆𑄗𑄗𑄗𑄗 𑄆𑄗𑄗
 𑄆𑄗𑄗𑄗 𑄆𑄗𑄗𑄗𑄗 | 𑄆𑄗𑄗𑄗 𑄆𑄗𑄗 𑄆𑄗𑄗 𑄆𑄗𑄗𑄗
 𑄆𑄗𑄗 𑄆𑄗𑄗𑄗 𑄆𑄗𑄗𑄗 𑄆𑄗𑄗𑄗 𑄆𑄗𑄗𑄗 𑄆𑄗𑄗𑄗 𑄆𑄗𑄗
 𑄆𑄗𑄗 𑄆𑄗𑄗𑄗𑄗𑄗 || 𑄆𑄗𑄗 𑄆𑄗𑄗𑄗𑄗𑄗

Tagalog: Nang nakita ang magandang pangitain at mabuting kapalaran, ngumiti siya at sinabi sa amin na isang araw at babalik ulit kami sa Biskalan. At sa pagkatapos...

English: When she saw a beautiful vision and a good fortune, she smiled and said to us that one day we would again return to Biskalan. After that...

0.3.3 Wikipedia

For years, *pamudpod* have been in use in Wikipedia’s *baybayin* article, and they are described as an alternate *virama*.

As of [revision 970344818](#), it is most prominently used in the lead sentence, as:

“ **Baybayin** (Tagalog pronunciation: [baɪˈbaːjm̄], 𑄆𑄗𑄗, virama-krus-kudlit: 𑄆𑄗𑄗𑄗𑄗, virama-pamudpod: 𑄆𑄗𑄗𑄗𑄗 ; also known as **alibata**) is an old writing system that was used in the Philippines. ”

Despite saying this, there is a whole section of examples of § “Modern usage” of the *baybayin* script, including on insignia of government agencies and the Philippine passport.

0.4 On the other encoded Philippine scripts

There are two other encoded Philippine scripts: Buhid, and Tagbanwa. Despite being on the surface similar to *baybayin*, they are really quite different. These other scripts are spoken by much smaller populations which are much more remote; besides Hanuno’o, their scripts are primarily of historical interest, although some modern use continues. They are not seeing modern revival on the same level *baybayin* is; *baybayin* appears in the insignia of many large Philippine organizations, including the Armed Forces of the Philippines and the National Museum of the Philippines; it also appears on all Philippine banknotes and passports.

Nevertheless, the question was posed in an e-mail thread about this proposal as to whether the other Philippine scripts use the *pamudpod*. Norman de los Santos answered thus:

“ The late Antoon Postma, Hanunuo leaders, and the Mangyan Heritage Center (MHC) tried to introduce the use of the *pamudpod* mark among the Buhid but because the tribe/group is inherently aloof and their villages more remote, continued contact and education attempts were not as easy as it was with the Hanunuo.

The University of the Philippines, SIL, Mangyan Mission, and the MHC recently concluded a years long data sampling/gathering and comparative study of Buhid scripts. I was privy to the data/samples from the surveyors and their expert consultant Christopher Ray Miller. Buhid now appears to have two distinct sets of scripts. One for Bansud Buhid, and another for Bongabong-Roxas. Many CV shapes are not shared. More importantly, they do not match many of the corresponding glyph shapes represented in the Buhid Unicode range. I don’t remember from whom or where the sample of Buhid Unicode typeface came from, but they are quite erroneous (if not highly stylized). The *pamudpod* has not been adopted by the writers. ...so there’s that.

As for Palawan scripts, there are emerging Tagbanwa variants from different Tagbanwa and Palawan tribes. But documentation is sketchy, places are too remote to study, and there is no official organization that is up to the task. We are left with what we already have in Tagbanwa Unicode, which if I recall correctly, came from the book *Alfabeto Tagbanwa con Algunas Reformas (Tagbanwa Alphabet with Some Reforms)* by Norberto Romualdez. As per Romualdez’s reform, he has suggested an *x-kudlit* type *virama* mark that precedes the final consonant (x before the final consonant if the parent syllable has an /a/ sound, superscripted x before the final consonant if the parent syllable has an /i/ sound, and subscripted x before the final consonant if the parent syllable has an /u/ sound.) ...there’s no evidence that his reforms were ever implemented beyond his book. ”

In sum, right now, we only have evidence of the *pamudpod* being used in *surat Mangyan* and *baybayin*. The problems de los Santos mentions regarding the encoding of Buhid are a possible avenue of future work.

0.5 Alternative solution: new character

An earlier version of this document was assessed by the Script Ad Hoc, and I received several comments from Roozbeh Pournader. The most important was that the SAH members thought it'd be better to have a new character. I'd like to present several arguments against this. However, I prefer a new character to the current situation, so for completeness have provided details for one if this is the UTC's ultimate preference despite my concerns.

My concerns can be classified in two broad categories:

0.5.1 Concerns about alternative solution

0.5.2 Existing fonts

As stated in the *Disclosure* section, I do have a conflict in that one of these fonts is my fonts. However, I am certainly not alone in including U+1734 in my fonts which do not otherwise contain Hanuno'o letters. Two recent *baybayin* fonts by Norman de los Santos include a *pamudpod* as well at U+1734 with no other Hanuno'o letters having glyphs:

Baybayin_Loc.otf






Baybayin_Doc.ttf



0.5.3 Existing documents

While there are not many documents that I know of which include the *pamudpod*, because of how long this situation has gone on, there are not zero. For example, the English Wikipedia as of the date of this document includes the *pamudpod* on its *baybayin* article encoded at U+1734.

Furthermore, many tweets include it:

- 
- 
- 
- ...etc.

0.5.4 Possible new character's properties

Nevertheless, I do understand that since we're already going to be adding TAGALOG LETTER RA and TAGALOG LETTER ARCHAIC RA the UTC might see fit to add another new character to the block. In that case I can provide a representative glyph and fill out a N4502-F for the concurrent ISO WG2 process.

So:

- , U+1717, TAGALOG SIGN PAMUDPOD

1717;TAGALOG SIGN PAMUDPOD;Mc;0;L;;;;;N;;;;;

If this is the avenue taken, I continue to recommend that U+1734 receive category Mc.