SPELLING ISSUES IN WEST COAST BAJAU: A COMPARATIVE LOOK

Mark Miller SIL Malaysia (mark_miller@sil.org) DOI: https://doi.org/10.22452/brj.sp2019no1.12

Abstract

A number of values inform people's decisions on how to spell their language. Sometimes these values conflict. For example, some may hold strongly to the value that a language should be spelled according to its sound system, while others prioritize the value that their language should be spelled as much as possible like the national language. Both values have been instrumental in shaping the orthography used in West Coast (WC) Bajau [bdr], spoken on the west coast of Sabah, Malaysia. In 2004, a Bajau orthography workshop was held in Kota Kinabalu, involving some 30 WC Bajau leaders and educators. The decisions reached at this workshop have formed the basis of the current WC Bajau orthography. With the publication of the Bajau Sama – Malay - English Basic Dictionary last year (2017), more people are being exposed to the spelling system than perhaps ever before. In the present paper, some of the key features of the WC Bajau orthography will be discussed. These include: the representation of the glottal stop; the representation of the pepet vowel; the use of semivowels ('y' and 'w') vs vowels ('i' and 'u') in vowel glides; and the representation of long vowels. We shall consider how the Malay sound system and orthography have influenced the development of the WC Bajau orthography. We shall also examine how the WC Bajau orthography differs from the spelling systems used in other Sama Bajau languages such as Mapun, Central Sama, and Southern Sama, languages which are much more influenced by Filipino/Tagalog than Malay since they are spoken mostly in the Philippines.

Keywords: orthography, West Coast Bajau, Mapun, Central Sama, Southern Sama

Introduction: The WC Bajau language and its orthography development

The West Coast (WC) Bajau people number approximately 210,000 and are found mostly on the western and northern coasts of Sabah (Simons & Fennig, 2018). Their cultural homeland is in the area surrounding Kota Belud town, but they have substantial populations in Kota Marudu District to the north, and in Tuaran, Kota Kinabalu, Putatan, and Papar to the south. The WC Bajau language (ISO code: [bdr])

is still spoken widely in the Kota Belud area, but its vitality is weakening as people migrate to the cities seeking employment, and as parents increasingly use Malay with their young children to prepare them for education in that medium.

While the language vitality of WC Bajau is an ongoing concern, in the last twenty years a concerted effort has been made to create and eventually to standardise a spelling system for the language. A major step was achieved in February 2004, when a Bajau orthography workshop was held, involving some 30 WC Bajau community leaders and educators, with the goal of reaching consensus on the spelling system to be used for the language. During the workshop I presented the spelling issues that needed to be resolved, based on previous testing I had done in the community and in consultation with members of the PSBB (Persatuan Sendi Budaya Bajau). While I facilitated the discussion, decision making was done by the group. Subsequent to the workshop, a little booklet was produced which summarised the spelling decisions made during the workshop. Using this approved orthography, in 2006 a series of 25 WC Bajau language learning lessons was published by the Sabah State Museum. Then, just last year, a Bajau Sama – Malay – English Basic Dictionary (Kamus Asas) was published. The dictionary contains over 5,000 entries and uses substantially the same orthography that was agreed upon in 2004. It is important to note that the orthography is based on the Kota Belud dialect of WC Bajau. The question remains whether the present orthography will adequately serve speakers of other varieties of WC Bajau, such as Tuaran or Putatan.

Values which help shape orthographies

In an ideal (linguist's) world, the orthography developed for a given language would be that which most succinctly captures all the necessary sound distinctions in the language. In other words, the linguist strives for the most economical and elegant spelling. However, othographies are not the exclusive property of linguists! Rather, a spelling system reflects the choices and preferences of many people in the language community, and multiple factors inform those choices and preferences. One might think of these factors as **values** held (consciously or unconsciously) by speakers of the language, values which inform their spelling decisions. In a 1998 address given to an audience of WC Bajau educators and leaders, Dr. C. Henry Bradley (who mentored me during the early years of my research on WC Bajau) identified five such values. An orthography should:

- Be based on the sound system of the language
- Take into account grammatical and semantic information
- Be acceptable to speakers of the language
- Be approvable by those who have oversight of national language policy
- Conform totally to the writing system of the national language

These values will conflict at times. Choices must be made as to which value(s) are prioritised in deciding how to handle a given spelling decision. In the sections that follow, some of the key orthography issues are described for WC Bajau. They are "issues" because they represent conflicts (or potential conflicts) between competing values. The issues to be examined include (1) spelling the glottal stop; (2) spelling the pepet (schwa) vowel; (3) spelling word-initial consonant clusters; (4) spelling long vowels; and (5) spelling using vowels vs semivowels at the end of the word.

As these spelling issues are presented, I will also consider how the same issues have been treated in the orthographies of other Sama-Bajau (S-B) languages. A total of nine Sama-Bajau (S-B) languages have been identified (Grimes, 2003, pp. 496-498). Six of these are located in the Sulu Archipelago of the southern Philippines, including Mapun [sjm], Balangingih Sama [sse], Central Sama [sml], and Southern Sama [ssb]. People speaking these languages have also settled in Sabah, most of them on the east coast. By far the largest S-B group on Sabah's east coast are the Southern Sama, including the Bajau Kubang of Semporna, who have long dwelled in Sabah and are generally considered to have indigenous status. Other varieties of Southern Sama include Sibutu', Simunul, and Obian/Ubian. Another S-B language, Indonesian Bajau [bdl], is spoken in Sulawesi and neighbouring islands. A comparative look at the orthographies of these S-B languages will suggest that, apart from differences in their sound systems, the national language (whether Malay, Indonesian, or Tagalog) also plays an important role in how these different languages are spelled.

The WC Bajau alphabet

The WC Bajau language has 16 consonants and 6 vowels, to comprise a total of 22 significant sounds, or phonemes. Normally a phoneme will possess its unique symbol or 'letter' in the alphabet, but there are exceptions. In the WC Bajau orthography, the two vowels /e/ and /ə/ (called 'taling' and 'pepet', respectively) are symbolised using the same letter 'e', even though they are separate phonemes. That is why the WC Bajau orthography has 22 phonemes but only 21 letters. The following table shows each phoneme in the WC Bajau language, together with its corresponding symbol or letter:

Significant Sounds (= Phonemes)	Written as:
/a/	а
/b/	b
/d/	d
/e/	e
/ə/	e

Table 1: Phonemic inventory and alphabet for WC Bajau

g
i
j
k
1
m
n
ng
0
р
r
S
t
u
W
у
'

The Malay language has certain phonemes which WC Bajau does not have: these are /c/, /f/, /h/, and /z/. Therefore the WC Bajau spelling system does not require the letters used to represent these sounds in Malay.

Spelling Issues

Spelling the glottal stop

The glottal stop (Malay *hentian glotis*) is a phoneme in WC Bajau as it is in many local languages of Sabah. In the Kota Belud dialect of WC Bajau, the glottal stop occurs only at the end of the word. Examples include: [mu.le?] 'go home', [so.o?] 'to command', and [tu.nu?] 'burn'. The glottal stop sound also occurs in Malay between two 'a' vowels in words like *perkataan* where the *-an* suffix has been added, and in some loan words from Arabic (such as *maaf*). These occurrences are usually predictable, meaning that the glottal stop does not need to be written in Malay. However, in WC Bajau, the occurrence of the glottal stop at the end of a word can change the meaning of the word: compare [bo.o] 'bring' with [bo.o?] 'bamboo'; and [ŋə.ga] 'deny' with [ŋə.ga?] 'startle (s.one)'. Thus, the glottal stop in WC Bajau should be written. The symbol chosen to represent this sound is the simple apostrophe ('). An apostrophe-type symbol is used to represent this sound in several local languages including Kadazan Dusun, Lotud, Lundayeh, and Murut Timugon. Also, the simple apostrophe is commonly used as the Latin transliteration of the glottal stop in Arabic (the *hamzah*).

While there is ample precedent for choosing the apostrophe to symbolise glottal stop, people who write in Bajau informally (such as when texting friends) often

do not know how to write it, and may neglect to write it at all. When they do write it, they sometimes use 'k' to represent the glottal stop. This preference probably owes to the influence of Malay as spoken in Peninsular Malaysia, where the final 'k' in a word (such as *tepuk* or *kakak*) is pronounced as a glottal stop. So, why not use 'k' to spell the word-final glottal stop in WC Bajau? In Bajau, there is a contrast in word-final position between the [k] sound and the [?] sound. Examples of [k] in word final position include: [kak'] 'older sibling'; [bu:k'] 'worn, spoiled'; and [tiko:k'] 'head'. Minimal pairs such as [pak'] 'frog' vs [pa?] 'uncle' confirm that the [k] and [?] sounds are contrastive, that is, they determine the meaning. If the letter 'k' is used to represent both word-final /k/ and word-final /?/, then there is under differentiation and confusion may arise, especially for those who are learning the language and would not be able to tell from context alone which sound is intended.

How is the glottal stop represented in other Sama-Bajau (S-B) languages? While Central Sama uses the apostrophe, the Southern Sama orthography uses 'h' to represent the glottal stop. Using the letter 'h' to represent glottal stop might also be a solution for WC Bajau, because the [h] sound is produced at the same point of articulation as the glottal stop, and it does not occur as a phoneme in WC Bajau. However, here again the influence of Malay is important. In Malay, /h/ is phonemic and it occurs at the end of many words: *buluh, marah, sedih,* etc. The 'h' in Malay words is always pronounced as [h], never as [?]. If the WC Bajau orthography attempted to ascribe a different sound (= the glottal stop) to 'h', this might confuse people accustomed to Malay spellings.

Spelling the pepet (schwa)

Like Malay, West Coast Bajau has the /e/ taling vowel and the /ə/ pepet (= schwa) vowel. The old Malay orthography distinguished these two vowels by using 'e' to spell the /e/ taling vowel and 'ë' (with the diacritic) to spell the /ə/ pepet vowel. Thus: *semak* 'edit' vs *sĕmak* 'undergrowth'. However, the modern Malay orthography does not distinguish these sounds, even though they make a difference in meaning. The modern Malay orthography spells both of these words as *semak*, so one must rely on the context to know which word is meant.

As in Malay, these two vowels—/e/ taling and /ə/ pepet—are also phonemic in WC Bajau. One option is to write them using two different symbols (or a diacritic with one of the symbols, as did the old Malay orthography). It turns out, though, that in WC Bajau, the /ə/ pepet vowel never occurs in the last (= ultimate) syllable of a word, whereas in most cases the /e/ taling vowel <u>only</u> occurs in the last syllable, as shown in Tables 2A and 2B below.

Phonemic word	Spelled as	Gloss
/bə.toŋ/	betong	'stomach'

Table 2A: /ə/ pepet in first syllable

/lə.ma?/	lema'	'soft'
/bə.ge/	bege	'like that'
/sə.le?/	sele'	'take turns'

Phonemic word	Spelled as	Gloss
/mu.le?/	mule'	'go home'
/bu.e?/	bue'	'water'
/bə.ge/	bege	'like that'
/sə.le?/	sele'	'take turns'

Table 2B: /e/ taling in last syllable

For the words shown in Tables 2A and 2B, there is no real difficulty to use 'e' to symbolise both /ə/ pepet and /e/ taling, as we can tell from word position alone which vowel sound occurs where. However, there are a handful of words in which /e/ taling occurs in both the last syllable <u>and</u> the second to last syllable, as shown in Table 2C below.

Phonemic word	Spelled as	Gloss
/pe.ne?/	pene'	'choose'
/ke.ket/	keket	'bite'
/sem.bet/	sembet	'chase'

Table 2C: /e/ taling in both syllables

This creates a potential problem because 'e' in the second to last syllable can symbolise either /e/ taling (as with **pene**' and **keket** in Table 2B) or /ə/ pepet (as with **bege** and **sele**' in Table 2A). Thus, when 'e' appears in both the second to last and the last syllable of a word, we cannot predict which vowel sound occurs in the second to last syllable. So then, is it necessary to distinguish /ə/ and /e/ in the WC Bajau spelling system? If one's overriding value is to base the orthography on the sound system, then 'yes'. But since the occurrence of such words is relatively rare, it was acceptable to WC Bajau speakers to use the one symbol 'e' for the two phonemes. Also, the fact that these two vowel sounds are no longer distinguished in the Malay orthography is an important factor.

While Proto Sama-Bajau had pepet as a phonemic vowel, it has dropped out of some of the present S-B languages, including Southern Sama in most of its dialects. For those languages where it has been retained, different orthographic solutions are proposed to represent the pepet vowel. In Mapun, the same letter 'a' is used for both the pepet sound and [a], since there is almost always a doubled (= geminate) consonant following the pepet sound, enabling the two to be differentiated (WC Bajau does not have geminate consonants, except for some words that begin with a lengthened nasal consonant or lengthened /l/, as further described in Section 4.3. So the Mapun solution

would not be viable for WC Bajau). In Central Sama, at least two solutions have been proposed. One solution is to use the simple apostrophe—also used for glottal stop—for the pepet, since the two sounds (pepet and glottal stop) never occur in the same environment. The second solution, advocated by the Tawi-Tawi Department of Education, is to place a diacritic over the other 5 vowels (e.g. *ä*, *ü*, *ö*) to represent the pepet (Schroeder, 2017). As far as I am aware, WC Bajau and (possibly) Indonesian Bajau are the only S-B languages where the pepet vowel is represented by 'e'. Chandra Nuraini-Grangé, in her doctoral dissertation (2008) on the Kangean variety of Indonesian Bajau, appears to spell /a/ pepet with the letter 'e' and /e/ taling with the diacritic 'é'. It is not clear whether this was her personal preference or whether she was working from a working or provisional orthography already in place. Since Malay and Indonesian Bajau to be represented the same way as in Malay and Indonesian. By way of contrast, Tagalog lacks the pepet vowel, so there is no strong precedent for S-B languages in the Philippines for how to symbolise it.

The following table depicts how the pepet vowel sound (in bold font) is spelled in WC Bajau and a few neighbouring S-B languages (Southern Sama lacks the pepet vowel, so the replacement vowel is shown instead).

	West Coast Bajau (uses 'e')	Mapun (uses 'a')	Central Sama (uses ' or a diacritic)	Southern Sama (Sibutu' dialect) no pepet
'gully'	l e bak	l a bbak	l'bbak / l ä bbak	labbak
'sugarcane plant'	t e bu	t a bbu	t'bbu / t ü bbu	tubbu
'stomach'	betong	b a ttong	b'ttong / b ö ttong	bottong
'lower leg'	b e tis	b a ttis	b'ttis / bïttis	bittis

Table 3: Spelling the pepet vowel in some of the S-B languages

Spelling word-initial consonant clusters

A fair number of words in WC Bajau begin phonetically with a consonant cluster, that is, a sequence of two consonants. Most of these consist of a nasal consonant followed by a (voiced or voiceless) stop at the same point of articulation, e.g. <u>mb</u>, <u>mp</u>, <u>nd</u>, <u>nt</u>, <u>ng</u>, <u>nk</u>. Some words begin with a lengthened nasal consonant or lengthened /l/, which are interpreted as a sequence of two like consonants. In careful speech, a pepet vowel can be heard to precede the consonant cluster. Examples of word-initial consonant clusters in WC Bajau are shown below.

[m'beren] ~ [əm'beren] 'when?	21
[n,'do?] ~ [ən'do?]	'take'
[mˌˈmaʔ] ~ [əmˈmaʔ]	'father'
[nˌˈnaʔ] ~ [ənˈnaʔ]	'to place'
[lˌˈlum] ~ [əlˈlum]	'alive; to live'

The initial consonant is best interpreted as syllabic, with an underlying (phonological) VC syllable (Miller, 2007, p. 44). Note that, although the initial pepet vowel can sometimes be detected in careful or emphatic pronunciation of these words, usually the consonant cluster sounds the beginning of the word. How should these words be spelled? If we follow their usual pronounciation, we would simply start the word with the consonant cluster, e.g. **mberen** 'when?', **ndo**' 'take', **mma**' 'father', **llum** 'live, alive'. However, this creates consonant sequences not found word-initially in Malay, and they look strange to people accustomed to Malay spelling. Therefore it was decided to spell such words inserting an 'e' (representing the pepet /ə/) prior to the nasal, even though it is not usually pronounced. Thus, the above words in the present WC Bajau orthography are spelled: **emberen**, **endo**', **emma**', **enna**', and **ellum**.

Other S-B languages contain these word-initial nasal consonant clusters, but there is no pressure (e.g. from Tagalog or other wider languages of communication) to insert a vowel prior to the consonant cluster in spelling these words. Table 4 shows the spelling of some words containing word-initial consonant clusters in a few S-B languages.

	West Coast Bajau	Mapun	Central Sama	Southern Sama (Sibutu' dialect)
'wife'	endo	nda	nda	nda
'grandchild'	empu	mpu	mpu	mpu
'hold'	entan	ntan	ntan	
'NEG'	enggai		ngga'i	nggey
'to place	enna'	nna'	nna'	nnah
'father'	emma'	mma'	mma'	mmah
'sun; day'	ellaw	allaw / llaw	llaw	llow

Table 4: Spelling word-initial consonant clusters in some S-B languages

Spelling long vowels

Quite a number of words in WC Bajau are pronounced with a long vowel sound. Examples include such common words as [la:n] 'road, way', [tu:t] 'knee', [se:?]

'companion', [di:ŋ] 'fish', and [so:] 'snake'. The long vs. short vowel sound can make a difference in the meaning of the word, as shown by the following minimal pairs:

'thin'
'to scoop'
1
'at / in'
'long'

Therefore, it is necessary to represent the long vowel sound in WC Bajau. In the orthography, the long vowel sound is represented by doubling the vowel letter. So, the spelling of the words introduced above would be:

[la:n]	laan	'go'
[tu:t`]	tuut	'knee'
[se:?]	see'	'companion'
[di:ŋ]	diing	'fish'
[so:]	SOO	'snake'

One might raise the question, whether the reader might want to insert a glottal stop when he or she reads a sequence of two like vowels. Lengthened vowel sounds do not occur in Malay. Adding the -an suffix to root words ending in 'a' creates a sequence of two 'a' vowels in the spelling, but this sequence is always pronounced with an intervening glottal stop (e.g. *keadaan, kerajaan*). Also, Malay contains some loan words from Arabic in which a sequence of two 'a' vowels occurs, again with an intervening glottal stop (e.g. *maaf*, *Jumaat*). So far, however, this does not seem to create a problem for WC Bajau readers in the Kota Belud area. Note that, in other WC Bajau dialects spoken along the western coast, sequences of like vowels are sometimes pronounced with an intervening glottal stop: for example, [poon]~[po?on], 'tree'. In the Banggi and Sandakan dialects of WC Bajau, an /h/ may intervene. For example, [boo] 'shoulder' in the Kota Belud dialect is pronounced [boho] in the Banggi and Sandakan dialects (Craig Soderberg, p.c.). The nature and extent of such variations, and their potential implications for the spelling of long vowel sounds in these dialects, requires further research.

Lengthened vowels occur in all of the S-B languages. In most S-B languages, including Mapun and Central Sama, a diacritic over the vowels (\bar{a} , \bar{e} , \bar{i} , \bar{o} , \bar{u}) is used to represent long vowels. Even though use of a diacritic is problematic in that it requires extra effort to produce the character, the alternative may be worse. Readers influenced by the Tagalog orthography may want to automatically insert a glottal stop when they see a duplicated vowel, as this is how they would pronounce the word with such a spelling in Tagalog. The Southern Sama orthography sidesteps this problem simply by

not distinguishing vowel length. Table 5 below shows how words containing lengthened vowel sounds are spelled in the neighbouring S-B languages:

	West Coast Bajau	Mapun	Central Sama	Southern Sama
'sea cucumber'	baat	bāt	bāt	bat
ʻlight, lamp'	suu'	sū'	sū'	suh
'head'	tikook	takōk	kōk	kok
'money'	siin	sīn	sīn	sin

 Table 5: Spelling long vowels in some S-B languages

Spelling using vowels or semivowels at the end of the word

WC Bajau, like Malay, has the vowel glides *au*, *ai*, and *oi* at the end of many words. These vowel glides might best be analyzed as containing a vowel + semivowel: thus, /aw/, /ay/, and /oy/. As such, each glide comprises one syllable, not two. While in Malay these word-final glides are written as two vowels, in WC Bajau they are written as a vowel + semivowel, according to the sound system of the language.

Some examples:

langkaw	(not langkau)	'long'
kerabaw	(not kerabau)	'water buffalo'
dangay	(not dangai)	'how many'
rungay	(not rungai)	'lost'

It is true that some people prefer the vowel endings rather than the consonant endings because they are used to the Malay spelling. In my observation, most people informally use vowel endings when texting on their phones. However, the consonant endings were chosen for the Bajau Sama orthography, because they accurately represent the Bajau sound system, and because the consonant endings serve as an identity marker to distinguish WC Bajau from Malay. Here we see the value 'an orthography should be based on the sound system of the language' overrides the value that 'an orthography should conform totally to the writing system of the national language'. As more people use the WC Bajau orthography, we shall see which value prevails. For those S-B languages spoken in the Philippines, it is unproblematic to use 'y' and 'w' to represent vowel glides at the ends of words. This follows the convention used for spelling vowel glides in Filipino: *ay*, *oy*, *aw*, *ow*, *uw*, etc.

Note that some common words in WC Bajau end in the vowel combinations /au/ and /ai/ rather than the vowel glides /aw/ and /ay/. These are spelled with final 'u' or 'i' accordingly. In this case, each vowel comprises a separate syllable. For examples:

ba.u	(not baw)	'new'
da.u	(not daw)	'before'
ga.i	(not gay)	'they'
du.wa.i	(not du.way)	'get down'

Conclusion

In this paper, I have introduced some of the key features of the WC Bajau spelling system, 'key' because they often involve a clash of values. Of the many values that play into decisions for how to spell a newly written language, we have seen that the value 'an orthography should be based on the sound system of the language' has been instrumental in shaping the WC Bajau orthography. We have also seen the importance of the value 'an orthography should conform totally to the writing system of the national language' whether that national language is Malay (for West Coast Bajau) or Tagalog (for most other S-B languages).

As we have looked at how different S-B languages handle similar spelling issues, we might also consider the potential influence of the orthographies of other S-B languages on WC Bajau, and particularly those S-B languages spoken on Sabah's east coast. The WC Bajau community might want to consider the context of other S-B languages when making decisions on its orthography, in view of a wider "Bajau Sama" identity, and also to make it easier to read literature in these other languages. This could be an important factor, for example, in whether to continue using semivowels to represent vowel glides at the end of words, and whether to write consonant clusters at the beginning of words. We might see a new value emerge to influence how to spell in WC Bajau: 'Write the sound as it is generally written in other S-B languages'. This may result in greater uniformity between the orthography used in WC Bajau and those used for neighbouring S-B languages. At the same time, of course, some variation between orthographies is necessary, because of differences in the sound systems between the different S-B languages. For example, some S-B languages have the schwa as a phoneme, and others do not. Some allow for the glottal stop between vowels, and others do not.

It is hoped as well that speakers of S-B languages on Sabah's east coast will familiarize themselves with the orthographies developed for their languages in the Philippines. Then they might consider, when producing their materials, whether they

wish to conform their spelling to those already established orthographies, or whether they want to adopt a more customised (localised) orthography showing greater adaptation to Malay.

References

- Grimes, B. F. (2003). Sama-Bajaw Languages. In W. J. Frawley (Ed.), International Encyclopedia of Linguistics, Volume 3 (2nd ed.) (pp. 496-498). Oxford: Oxford University Press.
- Miller, M. (2007). A Grammar of West Coast Bajau. Doctoral dissertation, Faculty of the Graduate School, University of Texas at Arlington, United States of America. Retrieved from http://hdl.handle.net/10106/577
- Nuraini-Grangé, C. (2008). Langue et production de récits d'une communauté Bajo des îles Kangean (Indonesié). Doctoral dissertation, Université de La Rochelle, France. Retrieved from https://www.theses.fr/130045055
- Schroeder, L. (2017, 23-26 March). Popularizing a Practical Sinama Orthography Based on Its Unique Language Features. Paper presented at the second International Conference on Bajau/Sama Diaspora & Maritime Southeast Asian Cultures (ICONBAS-MASEC 2017), Semporna, Sabah, Malaysia.
- Simons, G. F., & Fennig, C. D. (Eds.). (2018). Ethnologue: Languages of the World, Twentyfirst Edition. Dallas, Texas: SIL International. Retrieved from http://www.ethnologue.com

Date of Received: 15 January 2019