Merap Historical Phonology in the Context of a Central Bornean Linguistic Area

Alexander D. Smith

UNIVERSITY OF HAWAIʻI AT MĀNOA

Merap (Mbraa/Mpraa) is a language spoken by a small community in North Kalimantan, Indonesia. It is remarkable in its striking phonology, despite being part of the fairly conservative Kayanic group of languages. This paper describes both the synchronic and diachronic phonology of Merap, as spoken in the village of Langap along the upper Malinau river. It is argued that Merap, despite being highly innovative, is not only a Kayanic language, but that it subgroups specifically with Ngorek, a phonologically conservative Kayanic language of Sarawak and North Kalimantan. The arguments for subgrouping Ngorek and Merap together apart from all other Kayanic languages are purely phonological, as both languages devoiced *-b, but nasalized *-d, a combination found nowhere else in Borneo. Other evidence involves reflexes of nasal-obstruent clusters *mb, *nd, *nj, and *ŋg, and a small list of exclusively shared lexical replacement innovations. Merap is only one of several languages in the area of central Borneo, south of Sabah and north of the equator, that have independently undergone drastic changes. This paper also discusses some of the other phonologically aberrant languages of this linguistic area, including Sa'ban, Modang, Gaai, Kelai, Kiput, and Berawan. Although many of these languages have undergone sound changes that are quite similar to those found in Merap, they are the product of convergence rather than inheritance. Many of the languages in this area have independently shifted stress to the final syllable and expanded the inherited Proto-Malayo-Polynesian vowel inventory.

1. THE ABERRANT LANGUAGES OF NORTHERN KALIMANTAN AND SARAWAK.[1] The languages of Borneo, particularly those in the area of Borneo just south of Sabah, have been widely reported as having striking, difficult phonologies (Blust 1974, 2001, 2002b, 2003, 2007; Soriente 2006a; Guerreiro 1983, 1989, 1996; and Revel-Macdonald 1982). Blust (2007:1) may have put it best when he described the area as being a "hot spot" for linguistic change. The most widespread feature of this area is stress shift, from the penultimate to the final syllable, which in turn has led to massive

1. I would like to first thank Ingan Faunaliwan, a university student from the Merap village of Langap who was kind enough to teach me about his language, and Edena, who's help was instrumental in the success of my fieldwork in Kalimantan. The advice of two anonymous reviewers helped a great deal in making the presentation of data more complete, although I am responsible for any errors that remain.

Oceanic Linguistics, Volume 56, no. 1 (June 2017)

changes in the phonologies of these languages. Unlike most Austronesian languages of island Southeast Asia, the languages of this part of Borneo tend to be monosyllabic or sesquisyllabic, they often have relatively large vowel inventories including expanded diphthongs and triphthongs, and their historical phonologies are particularly opaque; and unlike the Chamic languages, whose word-final stress arose through influence from Mon-Khmer languages (Thurgood 1996, 1999), the languages of central Borneo cannot be shown to have had any contact with non-Austronesian stress-final languages.[2]

Several of these aberrant languages are discussed in detail below, where I elaborate on the specific properties of the central Borneo linguistic area. These language are Sa'ban (Dayic), Kiput (and other languages of the Berawan-Lower Baram group), Gaai (Kayanic), Kelai (Kayanic), Modang (Kayanic), Long Gelat (Kayanic), Hliboi Bidayuh (Land Dayak), and Òma Lóngh (Kenyah). Because this is a linguistic area, and not a genetic subgroup, these languages do not have a special relationship to one another. Also, many of the languages in the geographic area of Borneo did not shift stress to the final syllable. Rather, a small number of languages exhibit these changes, while most others do not. Among these languages, Merap, spoken in Langap village of Malinau Regency Indonesia, is particularly aberrant and is one of the more historically interesting languages of this area.

The remainder of this paper consists of two parts. The first part, the remainder of section 1 and section 2, describe the features of a central Bornean linguistic area. Specific member languages are briefly described below, with analysis from published data and field notes. The second part, sections 3 through 5, describe the phonology, history, and linguistic position of Merap, one of the more historically interesting languages of the central Bornean linguistic area. The reason for including the first two sections is to place the historical phonology of Merap into context. It is not a language that has changed in complete isolation, but rather it presents a particular case of a much larger trend in central Borneo.

1.1 SA'BAN. Sa'ban is a Dayic language spoken in the upper Baram river area, near the border of Sarawak and Kalimantan, and is most closely related to Kelabit. Blust (2001) produced a fairly robust description of the synchronic and diachronic phonology of Sa'ban, and Clayre (1972, 1992) provides earlier references, but few other works have been written on this language. Blust reports twenty six vowels, which includes ten pure vowels (/i ɪ e ɛ a u ʊ o ɔ ə/), nine diphthongs (/iə̯ eə̯ əy əw ay aw uə̯ oə̯ oy/), and seven triphthongs (/iə̯w eə̯w aə̯y aə̯w oə̯y oə̯w uə̯w/).

1.2 KIPUT. Kiput is part of the larger Berawan-Lower Baram group, found in the lower Baram river area, from the coastal town of Miri to around the junction of the Tinjar and Baram rivers. These languages in general have undergone a number of interesting historical changes. Long Terawan is reported as having 16 vowels in Burkhardt (2016), but Kiput appears to be the most aberrant in this group. It has a vowel inventory similar to that found in Sa'ban; Blust (2002b, 2003) reports a typical consonant inventory, but 24 vowels including eight pure vowels (/i ɪ e a u ʊ o ə/), 14 diphthongs (/iw iə̯ ew eə̯ əy əw ay aw a:y a:w oə̯ oy uə̯ uy/), and two triphthongs (/iə̯y iə̯w/).

2. Adelaar (1995) has proposed a link between Land Dayak and Aslian (Mon-Khmer) because of similarities in the words 'to die' and 'to bathe', but this lexical similarity does not exist for other central Bornean languages.

1.3 SEGAI-MODANG. Although there are several works discussing the Segai (Gaai and Kelai) and Modang (Modang and Long Gelat) languages, few contain substantial wordlists or in-depth discussion of the phonology. My own field notes, which were collected in East Kalimantan in 2015 and 2016, contain vocabularies of nearly 1,000 items for Modang (of the Kelinjau river), Long Gelat, Gaai (of Long Laai), and Kelai (of Long Lamjang) along with sentence lists of various lengths. Those notes will be the source of data used in this paper. Gaai and Kelai are mutually intelligible, and both have similar phonologies. Modang is more diverse, and the two dialects of Kelinjau and Long Gelat described below, although mutually intelligible, are different from one another in important ways (see 2.2 on reverse umlaut).

1.3.1 Gaai. Gaai (also, Mengga'ai) is spoken along the Segah river in the Berau Regency of East Kalimantan. The phoneme inventory below is from the dialect of Long Laai. Gaai has a typical inventory of consonants, but 15 vowels including seven pure vowels (/i e a ə u o ɒ/), seven diphthongs (/iw uy əw oy ay aw ɑ:y/), and a single triphthong (/e̯w/).

1.3.2 Kelai. Kelai (also referred to as Punan Kelai, although it is not a Punan language), is spoken along the Kelai river in the Berau Regency of East Kalimantan. The following data are from the dialect spoken at Long Lamjan, sometimes labeled Long Palai on maps of the area. Kelai has a typical consonant inventory, and 17 vowels including nine pure vowels (/i e ɛ æ u o ɒ a ə/) and eight diphthongs (/iw ew æw a̯o a̯i a̯e oy uy/).

1.3.3 Kelinjau Modang. Kelinjau is a major dialect of Modang, and is spoken along the Kelinjau river, in East Kutai, East Kalimantan. It is often referred to as Long Wai, and is distinct from the Wahau and Long Gelat dialects. Data for this language are from Woeq Helaq, a settlement along the Kelinjau river. Modang has 16 vowels including seven pure vowels (/i e a a: u o ə/) and nine diphthongs (/ia̯ ea̯ ey əw oy uy ay aw aɵ/[3]). The low vowel appears short and long.

1.3.4 Long Gelat Modang. Long Gelat is most closely related to the Modang dialects of East Kutai, but is spoken quite a distance from Modang, in the upper courses of the Mahakam river. Because of its location, it is under fairly heavy influence from the Busang dialect of Kayan, although it maintains its distinctively "Modang" features. Long Gelat has thirteen vowels including six pure vowels (/i e a u o ə/) and seven diphthongs (/ia̯ ey əw oa̯ uy ay aw/).

1.4 HLIBOI BIDAYUH. Hliboi is a Land Dayak language, and belongs to the Bidayuh subgroup of Land Dayak.[4] It has strong word-final stress and has deleted or reduced penultimate vowels, giving it a monosyllabic canonical word form with initial consonant clusters, geminates, and voiceless sonorants. The vowels are phonetically complex (there are few pure vowels, as most are phonetically diphthongs), but the vowel inventory is phonemically not large (Smith 2017 reports /i u e o a ə əw ay aw ey uy/).

1.5 ÒMA LÓNGH. Òma Lóngh is an aberrant Kenyah language spoken in the Malinau Regency of North Kalimantan. Data on the language were first published in the form of texts and a 3,000-item wordlist in Soriente (2006a), with Blust (2007) offering a

3. The symbol *ɵ* represents a mid-central rounded vowel. Here, it is a nonsyllabic off-glide.

historical analysis based on Soriente's publication. Soriente reports nine pure vowels (/i e ɛ a u o ɔ ɯ ə/) and, surprisingly, no diphthongs.

2. PARALLEL SOUND CHANGES IN LANGUAGES OF CENTRAL BORNEO. The languages of central Borneo have undergone numerous sound changes that define a linguistic area that occupies roughly the interior of Borneo south of Sabah and north of the equator. Languages of this area share several convergent sound changes that cannot be argued to have been inherited from a single common ancestor. This section defines and exemplifies these parallel sound changes, particularly the reduction and deletion of penultimate vowels, reverse umlaut, the development of word-initial consonant clusters, voiceless sonorants, palatalization of word-final stops, and strengthening of intervocalic glides. The last sound change, glide strengthening, is widespread in Borneo, and is included here because Hudson (1978) suggests several genetic relationships on the basis of this change, although cases where glides do strengthen are all almost certainly convergent. Abbreviations used in this section are PKAY (Proto-Kayanic), PDAY (Proto-Dayic, which includes Kelabit, Lun Dayeh, and Sa'ban), PMP (Proto-Malayo-Polynesian), and PKEN (Proto-Kenyah).

2.1 REDUCTION/DELETION OF PENULTIMATE VOWEL. Throughout Borneo, stress-final languages have deleted or reduced penultimate vowels in a number of ways. Most typically, reduction to schwa occurs across the board. Full deletion of the penultimate vowel is also widespread, but less broad in its application. A number of languages delete inherited nonlow vowels, *i, *u, and *ə, but retain *a. Kelai and Gaai have deleted penultimate vowels in the vast majority of cases, but retain some penultimate vowels both irregularly (where one expects deletion) and regularly (typically to avoid a disallowed consonant cluster).

Modang and Long Gelat have reduced all penultimate vowels to schwa, and in cases where words were historically vowel initial, have deleted the penultimate vowel altogether. Some examples from both are shown below:

(1) MODANG

PKAY *kitan 'binturong'	> kətin
PKAY *biʔil 'difficult'	> bəʔen
PKAY *bulu 'body hair'	> bəlun
PKAY *kulih 'leopard'	> kəlih

4. The internal subgrouping of Land Dayak used here is from Smith (2017), and differs from the recent proposal in Rensch et al (2012). In Smith (2017), Land Dayak is split into two branches, as follows:

1. Land Dayak
   a. Banyadu-Bekati
   b. Bidayuh-Southern Land Dayak

Bidayuh-Southern Land Dayak is further split into a Bidayuh group (the self-identifying Bidayuh of Sarawak, plus Sungkung and Hliboi in Indonesia) and a Southern Land Dayak group (Golik, Jangkang, Ribun, Sungkung). Land Dayak is the accepted term for the entire Land Dayak subgroup, while Bidayuh, or Bidayuhic, only applies to the Bidayuh subgroup itself. In Sarawak, Bidayuh is often used to refer to all Land Dayak languages, but this is only because the majority of Land Dayak speakers in Sarawak are in fact Bidayuh. In Kalimantan, the term Bidayuh is nearly absent.

PKAY *saləŋ ‘earthworm’	> səlaŋ
PKAY *paku ‘fern’	> pəkaə̯
PKAY *atay ‘liver’	> tey

(2) LONG GELAT

PKAY *siŋaw ‘shade’	> həŋiə̯
PKAY *kitan ‘binturong’	> kətün
PKAY *kulih ‘leopard’	> kəleh
PKAY *puluʔ ‘ten’	> pələwʔ ‘multiples of ten’
PKAY *mata ‘eye’	> mətin
PKAY *palad ‘palm’	> pəlin
PKAY *atay ‘liver’	> tɛy

Sa’ban deleted nonlow vowels in the penultimate syllable, but *a was not deleted. Some examples are reprinted below from Blust (2001).

(3) SA’BAN

PDAY *bədʰuk ‘pig-tailed macaque’	> ssuə̯k
PDAY *tuba ‘derris root’	> bbəh
PDAY *kini ‘this way’	> hnay
PDAY *paʔit ‘bitter’	> paʔet
PDAY *takut ‘afraid’	> taʔəwt

Hliboi Bidayuh deleted nonlow vowels in the penultimate syllable. *a and in some cases *ə (where it had merged with *a) was not deleted but changed regularly to *i* in the penultimate syllable. This condition directly parallels that described in Sa’ban above.

(4) HLIBOI BIDAYUH

PMP *bulan ‘moon’	> blatn
PMP *lubaŋ ‘hole’	> bbakŋ
PMP *bibiR ‘lips’	> bbitn
PMP *silu ‘finger nail’	> hlutn
PMP *daRaq ‘blood’	> iaʔ
PMP *nanaq ‘pus’	> ninaʔ
PMP *aku ‘I; me’	> ikuʔ

PMP *ə merged with *a in the penultimate syllable (and eventually became *i*) after nonlabial consonants in the following examples, which bled the deletion of nonlow vowels.

(5) HLIBOI BIDAYUH

PMP *təbəŋ ‘to fell a tree’	> tibokŋ/nibokŋ
PMP *təbuh ‘sugar cane’	> tibuh

2.2 REVERSE UMLAUT. Umlaut in Germanic is a historical change whereby penultimate vowels assimilated to vowels in the final syllable. The historical development of pluralizing ablaut in Germanic languages is commonly used to exemplify this change. Briefly, the long *u of Proto-Germanic *muːsiz ‘mice’ was fronted due to the umlaut effect of *i in the final syllable. This gave rise to Old English *müs* [myːs], which in turn became Middle English *mis,* and Modern English *mice* [maə̯s]. Stated as a single change, *muːsiz > [maə̯s]. Thus, umlaut is defined as a process whereby a vowel becomes more like a following vowel. In Borneo, reverse umlaut follows similar changes but is distinct in that it is the vowel of the *final* syllable that assimilates to the pre-

ceding vowel (thus, it is a regressive assimilation, where prototypical umlaut is progressive). While the exact motivation remains unclear, it is apparent that the qualities of stressed high vowels in the penultimate syllable "followed" stress shift, where *úCaC > *uCúC > *əCúC > CúC, and *íCaC > *iCíC > *əCíC > CíC. Kelai, Merap, Sa'ban, Long Gelat, Modang, and Gaai all have some form of reverse umlaut and the first four cases are examined in more detail below.

2.2.1 Kelai. The reverse umlaut pattern in Kelai is the most transparent for analysis. Wherever a high vowel occurred in a penultimate syllable with *a in the final syllable, it transferred to the final syllable, replacing *a. (6a) shows *..iCaC > ..CiC, and (6b) shows *..uCaC > ..CuC.

(6) KELAI

a. PKAY *jəpitan 'nine' > jəptin
PKAY *kitan 'binturong' > ktin
PKAY *siap 'chicken' > jip

b. PKAY *kulat 'mushroom' > klut
PKAY *bəhuaŋ 'bear' > wahguŋ
PKAY *suŋay 'river' > ŋuy

2.2.2 Merap. The pattern of reverse umlaut in Merap is similarly transparent, but with an additional environment. High vowels in Merap transferred to the final syllable regardless of the shape of the vowel in the final syllable. Additionally, where Kelai deleted the original final vowel, Merap maintains it as an off-glide. In the case of *a or *ə, *uCaC/*uCəC > Cu̯C and *iCaC/*iCəC > Ci̯C. High vowels in final syllables became their corresponding glides, *uCiC > CuyC and *iCuC > CiwC.

(7) MERAP

PKAY *ñipa 'snake' > pi̯ʔ
PKAY *lima 'five' > mi̯ʔ
PKAY *kitan 'binturong' > tĩ̯
PKAY *siku 'elbow' > kiwʔ

PKAY *udaŋ 'shrimp' > roy̯
PKAY *bulan 'moon' > mblũ̯
PKAY *uləd 'maggot' > lu̯n
PKAY *putiʔ 'white' > tuy

2.2.3 Sa'ban. A similar type of vowel transfer is reported for Sa'ban in Blust (2001). But here again, the details differ in important ways from both Kelai and Merap. Reverse umlaut in Sa'ban appears to be more related to the spread of labiality across consonants than stressed vowel transfer to final syllables associated with stress shift. Only *u causes reverse umlaut (8a), while *i in the penultimate syllable does not (8c).

(8) SA'BAN

a, PDAY *tuqəlaŋ 'bone' > hlo̯ŋ
PDAY *uRat 'vein' > ro̯t
PDAY *kulat 'mushroom' > lo̯t
PDAY *uta 'vomit' > to̯

b. PDAY *upan ‘bait’ > pan
PDAY *upa ‘splitting’ > pa

c. PDAY *Ribu ‘thousand’ > pə-lbəw
PDAY *lipən ‘tooth’ > lepən
PKDAY *lima ‘five’ > emah

Blust (2001:275) notes that labial consonants block rounding of the final vowel (see [8b]), so while a sequence *uta became *to̯a̯*, *upa did not become **po̯o̯. Compare *kulat > *lo̯a̯t* ‘mushroom’ with *upaʔ > *paʔ* ‘yam’. Blocking labial consonants provide additional evidence that the changes observed in Sa’ban were caused by coarticulation of lip rounding across consonants, as labial consonants force the closure of the lips, physically interrupting rounding.

2.2.4 Long Gelat Modang. Reverse umlaut in Long Gelat is perhaps the most opaque. In all of the cases outlined above, there was a clear transfer of either the penultimate vowel itself, or a feature of that vowel, to the final syllable. Long Gelat shows clear differential reflexes where penultimate high vowels have influenced the forms of modern final vowels, but there appears to be no straightforward vowel or feature transfer.

To begin, *a in the final vowel is regularly reflected as *i* in the modern language:

(9) LONG GELAT MODANG
PKAY *mata-n ‘eye’ > mətin
PKAY *palad ‘palm’ > pəlin
PKAY *panas ‘hot’ > pənih
PKAY *ləpaw ‘field hut’ > piə̯
PKAY *laŋaw ‘house fly’ > ləŋiə̯

If *a was preceded by a high vowel (*i or *u), then the reflex is *u*.

(10) LONG GELAT MODANG
PKAY *ibah ‘saliva’ > wuh
PKAY *siap ‘chicken’ > jup
PKAY *suŋay ‘river’ > həŋuy
PKAY *udaŋ ‘shrimp’ > luŋ

If *a was preceded by a high vowel and closed with an alveolar, then the reflex is *ü* [yɪ̯].

(11) LONG GELAT MODANG
PKAY *jəpitan ‘nine’ > səptün [ˌsəpˈtyɪ̯n]
PKAY *kitan ‘binturong’ > kətün
PKAY *kuman ‘to eat’ > kün
PKAY *kulat ‘mushroom’ > kəlüt

It is difficult to find a satisfactory phonetic motivation that explains why *i in the penultimate vowel would result in a *u* reflex of *a in the final vowel (as exemplified by *ibah > *wuh* ‘saliva’ above). The allophone *ü*, which appears only before *-n*, can be explained as a result of anticipatory fronting, as *-n is produced with a front tongue position. Regardless of how this tricky historical sound change is eventually explained, it demonstrates that reverse umlaut has proceeded in different manners in different languages, and is not likely to have been inherited from any common ancestor. Example (12) directly

compares reflexes of words that show a reverse umlaut pattern in three of the languages discussed above.

(12)

	KELAI	MERAP	LONG GELAT
PKAY *kitan 'binturong'>	ktin	tĩə̯	kətün
PKAY *udaŋ 'shrimp'>	yuŋ	royə̯	luŋ
PKAY *ibah 'saliva'>	wih	ɓiə̯h	wuh

2.2.5 Reverse umlaut or metathesis? Descriptively, at least some of these patterns might be analyzed as a metathesis of the penultimate vowel with the onset of the final syllable, where *VCV became *CVV. There are certain aspects of reverse umlaut that argue against this interpretation. As noted above, the Sa'ban examples are best analyzed as historical progressive (left-to-right) coarticulation of lip-rounding. The fact that labial consonants, which physically alter lip position, blocked coarticulation in Sa'ban provides good phonetic evidence for this analysis. There is, on the other hand, no phonetic motivation for why labial consonants would block metathesis. In Kelai and Merap, metathesis seems ill-equipped to explain why inherited vowel clusters are treated different from the vowel clusters that would have arisen through metathesis (where *VCV became *CVV).

In Merap, inherited vowel clusters are retained as clusters, with stress on the final vowel and no coalescence. For example, PKAY *kiuŋ 'myna', *bəhuaŋ 'bear', and *dua 'two' became *kiawə̯* [kiˈawə̯], *bueə̯* [buˈeə̯], and *waʔ,* where the final vowel retained its syllabicity. If metathesis were used to explain observed reverse umlaut, it would predict similar treatment, as there must have been an intermediate stage where *VCV became *CVV. However, there does not appear to have ever been an intermediate stage, as all examples of reverse umlaut result in single vowels with complex articulations: PKAY *siku 'elbow', *qudaŋ 'shrimp', and *lima 'five' became *kiwʔ* (not ***kiawʔ*), *royə̯* (not ***rueə̯* as predicted by *bəhuaŋ > *bueə̯*), and *miə̯ʔ* (not ***miaʔ*). Metathesis also fails to explain cases in Merap where reverse umlaut occurs and the penultimate vowel does not delete. Examples include PKAY *tiruh 'sleep' > *tɛrɛwh,* *lubaŋ > *loɓoyə̯,* *duman > lumũə̯ 'year', and *miʔaŋ 'to split' > *miʔiə̯.* In these cases especially, metathesis cannot fully explain the data.

In Segai-Modang the observation that reverse umlaut results in single syllable segments while inherited vowel clusters remained two syllables does not apply, as transition glides strengthened from *ia to *ji* and *ua to *gu*. Thus, in Kelai, *dua 'two' became *agoʔ* and *duhian 'durian' became *lahjin.* Reverse umlaut, if analyzed as metathesis, would have resulted in additional cases of glide fortition if one assumes a feeding relationship (for example, *iCa > *Cia, which fed glide fortition, *Cji*). Evidence from Segai-Modang however, does not support a feeding relationship, and thus cannot be used to argue for or against the metathesis. The unique reflexes of reverse umlaut in Long Gelat, however, strongly argue against metathesis. Note that *a in the final syllable is reflected as *u if the penultimate vowel was either *u or *i. Metathesis, where *iCa became *Cia, does not explain how *Cia might become *Cu. One can imagine multiple scenarios, however, where *a become more like *i and *u through raising. Thus, *a in *iCa and *uCa may have raised to an intermediate high vowel, *iCɨ and *uCɨ, only after a high vowel, and later backed to *u. This is a phonetically more plausible scenario than metathesis. Although I hesitate to posit intermediate stages without direct evidence, reverse umlaut

remains a more robust and phonetically motivated explanation for the above patterns than metathesis, which ignores the relationship between the quality of the penultimate vowel and the reflex of the syllable-final vowel.

2.3 WORD-INITIAL CONSONANT CLUSTERS. In most languages where word-initial consonant clusters are found, stress shift followed by penultimate vowel deletion can be identified as the cause. Although some secondary consonant clusters are common throughout western Indonesia and Malaysia (*s* + voiceless obstruent in Malay three-syllable words—*sekali~skali* 'very' for example) the following data only deal with languages that have developed a robust set of consonant clusters in the onset of the initial syllable of carefully pronounced words.

2.3.1 Kelai and Gaai. Kelai and Gaai are quite remarkable in that they allow a wide range of word-initial consonant clusters.[5] These clusters arose through the regular deletion of penultimate vowels, resulting in the sound change *C_1VC_2VC > * C_1C_2VC. Interestingly, C_1 and C_2 have largely resisted any form of assimilation, resulting in a number of relatively uncommon word-initial consonant clusters. Tables 1 and 2 display all observed word-initial consonant clusters in Gaai and Kelai. The vertical column represents possible first segments in consonant clusters, and the horizontal row represents possible second segments. Each table is followed by several examples of words with initial consonant clusters and a description of observable phonotactic constraints.

The following constraints on Kelai consonant clusters are observable in table 1:

(i) Palatal stops are banned as the first segment in a word-initial-consonant cluster.
(ii) Glottal stop is banned as the first segment, and *h* is only found in one example, *hŋ*.
(iii) Glottal stop appears as the second segment only after sonorants and fricatives. There are no obstruent-glottal stop clusters.

TABLE 1. KELAI WORD-INITIAL CONSONANT CLUSTERS

	p	**t**	**c**	**k**	**ʔ**	**b**	**d**	**j**	**g**	**s**	**h**	**l**	**m**	**n**	**ñ**	**ŋ**
p		pt	pc	pk			pd		pg	ps	ph	pl		pn	pñ	
t				tk						ts	th	tl	tm	tn		tŋ
c																
k	kp	kt				kb	kd			ks	kh	kl		kn		
ʔ																
b							bd	bj				bl				bŋ
d												dl				
j																
g									gg			gl				
s	sp	st		sk	sʔ	sb	sd					sl	sm			sŋ
h																hŋ
l				lk	lʔ	lb				ls	lh		lm	ln		lŋ
m		mt		mk	mʔ		md		mg	ms	mh	ml			mñ	mŋ
n	np	nt		nk	nʔ		nb				nh	nl	nm		nñ	nŋ
ñ																
ŋ					ŋʔ							ŋl			ŋñ	

5. Guerreiro (1996) briefly mentions the consonant clusters in these languages.

(iv) There are no voiced obstruent-voiceless obstruent clusters.
(v) *c* and *j* are rare in any consonant clusters, and where they do occur, it is only after *p* (as *pc*) or *b* (as *bj*).

Some Kelai lexical items that show these clusters are presented below:

(13) KELAI

ptaệ 'corpse'	(< *patay)	pñat 'to tamp soil'	
pciə̂k 'thigh'		bloə̂n 'body hair'	(< *bulu-n)
pkaw 'fern'	(< *paku)	glæʔ 'fishing net'	
kbaʔ 'rattan bag'	(< *kiba)	sʔaệ 'frog'	(< *saʔay)
kdiŋ 'goat'		lʔæn 'branch'	(< *daʔan)
phaʔ 'firewood storage area'	(< *paRa)	ŋʔæŋ 'horned toad'	
thuə̂n 'collapse'		ŋñæn 'name'	(< *ŋajan)[6]
khas 'traditional skirt'	(< *taRas)	nhaə̂ 'eagle'	(< *ñahu)
tmæŋ 'thick'	(< *təmaŋ)	bŋeʔ 'rice wine'	(< *bəŋiʔ)

Table 2 shows Gaai word-initial clusters. The following constraints on consonant clusters are observable in that table:

(i) Palatal consonants are banned from initial position.
(ii) Glottal consonants are banned from initial position.
(iii) *l* is banned from initial position.
(iv) Geminate consonants can only be formed with the voiced obstruents, and in the data, only *bb* and *dd* are attested.
(v) Obstruents must match in voice in order to form a cluster: for example, *tg* is banned, but *tk* is allowed; *dk* is banned, but *dg* is allowed.
(vi) Voiced obstruents cannot be the initial segment in a cluster with any nasal.

Some lexical items that show these clusters are presented below:

TABLE 2. GAAI WORD-INITIAL CONSONANT CLUSTERS

	p	**t**	**c**	**k**	**ʔ**	**b**	**d**	**j**	**g**	**s**	**h**	**l**	**m**	**n**	**ñ**	**ŋ**
p		pt		pk	pʔ					ps	ph	pl		pn	pñ	
t	tp		tc	tk	tʔ						th	tl	tm	tn		tŋ
c																
k	kp	kt			kʔ					ks	kh	kl	km			
ʔ																
b						bb	bd	bj	bg			bl				
d						db	dd		dg			dl				
j																
g						gb		gj			gh	gl				
s	sp	st		sk	sʔ	sb			sg			sl			sñ	sŋ
h																
l																
m		mt		mk	mʔ	mb	md			ms	mh	ml		mn	mñ	
n	np			nk	nʔ	nb	nd				nh		nm			nŋ
ñ																
ŋ		ŋt							ŋg	ŋs	ŋh	ŋl			ŋñ	

6. *ñ* irregular.

(14) GAAI

ptəwŋ ‘aflame’		khuə̯ŋ ‘dry’	
pʔoə̯n ‘base’	(< *puʔun)	bjiw ‘deer’	(< *payaw)
phaw ‘grasshopper’	(< *pahu)	dbəwʔ ‘to grow, as a plant’	
pñek ‘broom’		dlam ‘deep’	(< *daləm)
tcəwʔ ‘seven’	(< *tujuʔ)	gjaw ‘tree; wood’	(< *kayu)
thuə̯n ‘person’		sʔaə̯h ‘shy’	(< *səʔih)
tlay ‘rope’	(< *tali)	mtaʔ ‘raw; unripe’	(< *mataq)
tŋeə̯w ‘cat’		nʔoə̯n ‘carry on the shoulder’	(< *ñuʔun)
ktal ‘itchy’	(< *gatəl)	nmas ‘tomorrow’	
kʔɒk ‘dull’		ŋlak ‘overflow’	

There are differences between Kelai and Gaai regarding the phonotactics of initial consonant clusters. Although palatal consonants are barred from initial position in both languages, Kelai has one example of *h* as the initial segment, while Gaai does not appear to allow *h* in initial position. Gaai allows glottal stop as the second segment in a cluster in a range of environments, and it is only banned after voiced obstruents, where Kelai bans glottal stop after both voiced and voiceless obstruents. Kelai allows *l* to be the initial segment in clusters, but Gaai innovated an initial *a,* which created a syllable boundary between *l* and the following consonant. For example, PKAY *daʔan ‘branch’ became Kelai *lʔœn,* but Gaai *alʔan,* where *l* and *ʔ* are separated by a syllable boundary: [al.ʔan]. Finally, while Gaai requires obstruent clusters to match in voice (only voiced-voiced and voiceless-voiceless obstruent clusters), Kelai allows voiceless-voiced obstruent clusters. This difference is apparent in reflexes of *kayu ‘tree’, where Gaai voiced *k- in accordance with its phonotactics, giving *gjaw,* but Kelai did not, giving *kjaw.*

2.3.2 Sa’ban. As noted above, Sa’ban deleted nonlow penultimate vowels. A quick glance at the appendix in Blust (2001) reveals the following consonant clusters: *pl-, plok* ‘slack, as a rope’; *pr-, prək* ‘crowded’; *pw-, pwət* ‘navel’; *bl-, bləy* ‘purchase’; *br-, breə̯* ‘husked rice’; *ml-, mləy* ‘to buy’; *mr-, mray* ‘dried up, as a stream’; *mp-, mpəŋ* ‘to accuse’; *ns-, nsəw* ‘to push’; *nj-, njəwʔ* ‘ashamed’; *nc-, nceə̯t* ‘expensive’; *nt-, ntəp* ‘to cut’; *ŋk- ŋkoə̯ʔ* ‘to allow to fall’. Constraints on initial consonant clusters are: (i) only nasals and noncoronal voiceless obstruents can be the first segment of a consonant cluster; (ii) only the sonorants *l, r,* or homorganic voiceless nonsonorant consonants can appear as the second element. Additionally, *pw-* is allowed as an apparent exception to the second rule.

2.3.3 Hliboi Bidayuh. Hliboi allows a range of initial consonant clusters. All nasals can form a word-initial consonant cluster with a homorganic stop: *mb-, mbot* ‘canine teeth’; *mp-, mpuə̯kŋ* ‘to float’; *nd-, ndaat* ‘monitor lizard’; *nt-, ntiə̯ʔ* ‘pregnant’; *ns-, nsey* ‘paddle’; *ŋg-, ŋgaʔ* ‘to see’; *ŋk-, ŋkuʔ* ‘mine’. The alveolar nasal also forms consonant clusters with palatal stops: *nj-, njug* ‘to stand’ and *nc-, nciə̯t* ‘to squeeze’. Any noncoronal stop can form a consonant cluster with *l* as the second element: *pl, pləwʔ* ‘multiple of ten’; *bl, bləwtn* ‘body hair’; *ml-, mlan* ‘just; recently’; *kl-, kluə̯h* ‘egg’; *ŋl-, ŋləwk* ‘gather’; *gl- glaʔ* ‘tongue’. There are a large number of additional consonant clusters that involve *h-.* These are enumerated in the discussion of voiceless sonorants in 2.5.

2.3.4 Merap. There are 20 attested word-initial consonant clusters in Merap (see section 3 below). This includes seven that involve three consonants: *mbl-, mpr-, ŋgl-, ŋkr-, hmp-, hmn-, hŋk-*.

2.4 VOICELESS SONORANTS. In all of the languages in Borneo where voiceless sonorants are present, and where etymologies are available, it appears that they arose through consonant clusters where CVS (Consonant Vowel Sonorant) became CS through deletion of the penultimate syllable. Later, the consonant cluster was simplified to *h*S, which is perceived in the modern languages as a voiceless sonorant. Three languages, Merap, Sa'ban, and Hliboi, have developed voiceless sonorants. All three belong to separate subgroups (Merap to Kayanic, Sa'ban to Dayic, and Hliboi to Land Dayak) and have no history of contact, so the changes are parallel and independent. Despite this fact, reflexes of *pənuq 'full' provide strikingly similar results from sonorant devoicing. Compare directly Hliboi *hnuə̲ʔ,* Merap *hnaų̲* and Sa'ban *hnoʔ* in the examples from each language listed below:

(15) HLIBOI BIDAYUH

PMP *silu-n 'finger nail'	> *slun	> hluə̲tn
PMP *pənuq 'full'	> *pnuʔ	> hnuə̲ʔ
PMP *suŋut 'black'	> *sŋut	> hŋuə̲t

(16) MERAP

PKAY *kuman 'eat'	> *kmuan	> hmūə̲
PKAY *pənuʔ 'full'	> *pnu	> hnaų̲
PKAY *sulu-n 'finger nail'	> *slun	> hləwŋ^{w}
PKAY *siŋət 'sting'	> *sŋiət	> ŋiə̲t

(17) SA'BAN

PDAY *kini	> *kni	> hnay 'this way'
PDAY *pənuʔ 'full'	> *pnuʔ	> hnoʔ
PDAY *turun	> *trun	> hruə̲n 'descend'
PDAY *tulaŋ	> *tluŋ	> hloə̲ŋ 'bone'

2.5 PALATALIZATION OF WORD-FINAL STOPS. Generally speaking, word-final palatal stops are rare in Austronesian languages. PMP did not allow any palatal consonants in word-final position.[7] Where they are found, they are almost always the product of conditioning: specifically, final palatals have been innovated as the result of high front vowels influencing final stops. Languages that have final palatals are listed below, with statements on their historical developments and examples.

2.5.1 Òma Lóngh. Of the many remarkable aspects of the historical phonology of Òma Lóngh are word-final palatals, *-c* (written as *-j* in Soriente 2006a) and **-ñ* (written as *-ny* by Soriente). The development of final palatals occurred as a result of assimilation to preceding vowels, where front and high or mid vowels triggered palatalization.

7. PMP *j, which was phonetically a palatalized velar stop [gʸ], was allowed in final position, but the true palatals *ñ and *z (phonetically [ʤ]) were not.

(18) ÒMA LÓNGH

PKEN *liʔip		> liʔic ‘shoulder’
PKEN *taʔat	> *taʔɛt	> taʔɛc ‘to see’
PKEN *paʔit		> faʔic ‘bitter’
PKEN *səŋim		> səŋiñ ‘cold’
PKEN *uman	> *ɔmɛn	> ɔmɛñ ‘year’
PKEN *buʔin		> buʔiñ ‘domesticated pig’

As the above data suggest, historically final labial and alveolar stops were palatalized, but final velars (*-k and *-ŋ) were not.

2.5.2 Long Jegan Berawan. Data from Robert Blust’s field notes (Blust n.d.) shows historical palatalization of word-final consonants in Long Jegan Berawan. Both velars and alveolars (*t, *k, *n, and *ŋ) were targeted for palatalization. Labials, however, do not appear to have been affected. According to Burkhardt (2014:75), *c* is a conditioned allophone of *k* (an alternation that also exists in Merap). Some examples are:

(19) LONG JEGAN BERAWAN

PMP *kulit ‘skin’	> kolayc
PMP *sakit ‘sick’	> cakayc
PMP *pawit ‘wing’	> pawayc
PMP *bətik ‘tatoo’	> bətiəyc
PMP *tumid ‘heel’	> tomayñ
PMP *məriŋ ‘hard’	> mərayñ
PMP *lamin ‘room’	> lamayñ ‘house’
PMP *diŋdiŋ ‘wall’	> jicayñ[8]

2.5.3 Modang. As part of a synchronic process, velar nasals in Modang are regularly palatalized in palatal environments, that is, where adjacent to either a high or mid front vowel or a vowel with a high-mid front off-glide. Some examples from Modang that show this are: /mənhaɐ̯ŋ/ > [ˌmənˈhaɐ̯ɲ] ‘red’, /kəlaːŋ bəkoɐ̯ŋ/ > [kəˈlaːŋ bəˈkoɐ̯ɲ] ‘a rattan basket’, /həŋet/ > [həˈɲet] ‘spicy’, and /ŋəwaɐ̯k/ > [ŋəˈwaɐ̯c˺] ‘to shout’. Note that in some cases, consonants are palatalized before high front vowels, not just after.

2.5.4 Merap. In Merap, as discussed at length in 4.1 below, the stops *-p and *-t are reflected as *-c* while *-m and *-n are reflected as *-ñ* when preceded historically by *i. In the modern language, *-c* and *-ñ* are still found only in predictable environments, and are not phonemic. Native speakers perceive *-c* as an allophone of *k,* while *-ñ* is perceived as *ŋ*. Word-final *-ñ* does not contrast with any other nasal, but *-c* does contrast with *-ʔ* where *-it or *-ip > *-ayc* and *-i > *-ayʔ,* through regular sound changes.

2.6 STRENGTHENING OF INTERVOCALIC GLIDES. This is one of the most widespread sound changes found in Borneo, but is less common outside of Borneo. The number of languages where this change is found is large, and it would take too much space to give examples for each language. I have instead listed each language where glide strengthening is found, followed by the reflex of ‘tree; wood’ (*kayu) and ‘two’ (*dua),[9] which is sufficient to establish glide fortition in each language.[10]

8. Burkhardt (2014) has *dicəyñ* ‘wall’ where Blust recorded *jicayñ.*

(20)	‘wood; tree’	‘two’
Modang	kəjəo̯	əŋgao̯ʔ
Long Gelat	kəjəo̯	əŋgao̯ʔ
Gaai	gjaw	agoʔ
Kelai	kjao̯	agoʔ
Beketan	kaju	lugo
Kadorih	kacu	duoʔ (*w did not strengthen in Kadorih)
Tunjung	kajuʔ	rəgaʔ
Sekapan	kaʐəw	dəgʷa
Kejaman	kaʐəw	dəgʷah
Lahanan	kaʐəw	ləgʷah
Kiput	kacəw	dufih
Long Jegan	kajəw	duvey
Long Terawan	kajuh	ləbih
Bintulu	kazəw	ba
Miri	ajuh	dəbeh
Rungus	kazu	duvo
Kadazan	kazu	duvo

2.7 CONVERGENT SOUND CHANGES AND SUBGROUPING. The six sound changes discussed above—final palatalization, reverse umlaut, penultimate vowel deletion/reduction, development of voiceless sonorants, development of initial consonant clusters, and glide strengthening—are not common at all in the Austronesian family, but abound in Borneo. They are all, however, convergent in the languages where they appear. They are for the most part phonetically motivated, and caused primarily by stress-shift (although glide fortition does not correlate with word-final stress). It is necessary to point out these convergent changes, as the peculiarity of some of them both within Austronesian (voiceless sonorants, initial consonant clusters, penultimate vowel deletion) and more generally (reverse umlaut) may be misinterpreted as subgrouping evidence when signs otherwise point to convergence. The goal of any historical linguist working with these languages must be to locate particular sound changes that are not motivated by stress shift but are nevertheless of high quality. As a general rule, the sound changes listed above cannot form the basis for any subgrouping argument. This is the challenge one faces when working with Merap. Although the language has much in common with its neighbors, it is unlikely that those similarities were inherited. It is necessary to look beyond the six sound changes listed above when attempting to place Merap in a larger subgroup.

3. MERAP PHONOLOGY. Little information is available for Merap, save for a presentation given at the 13th International Conference on Austronesian Linguistics by Antonia Soriente (2015).The nature of the presentation, being only 20 minutes with little time for presenting data, did not lend itself to a thorough classification or analysis. How-

9. PMP ‘two’ was *duha, but this is reflected as *dua in all languages of Borneo. *dua, in turn, had a phonetic transition glide and was pronounced [duwa]. The glide was then strengthened in languages that also strengthened phonemic glides.
10. Data are primary for Modang through Lahanan, and secondary for Kiput through Kadazan. Burkhardt (2014) recorded Long Jegan *kajjəw, duβiəy* and Long Terawan *kajjŭh, ləbbĭh*

ever, one can identify some minor differences between the data in that presentation and the data being presented below:

(i) It was reported that *-ŋ had become *h,* but I recorded it as Ø. It is not difficult to imagine that *h from *ŋ would further reduce to Ø. We were either recording different dialects or the speech of different generations (my speaker was a university student).

(ii) Soriente also recorded *mbr* where I recorded *mpr.* The name of the language contains this example: /mbra:/ in Soriente (2015), where I recorded /mpra:/ Again, devoicing in this environment is not out of the question.

There are no major differences between the data being reported here, and those found in Soriente's earlier presentation.

There is also a short mention of Merap in Guerreiro (1996), who cites unpublished data from Sellato. These data appear to show the nasal distinction in the vowels, but little else is available, as the publication did not deal primarily with Merap.

3.1 MERAP PHONEMES. Merap has a typical inventory of consonants (see table 3), as do most languages of Borneo. Complexity appears in its vowel system, which is shown in table 4 below.

As table 4 makes clear, the number of diphthongs and triphthongs is very large when compared to other western Austronesian languages. There are 26 vowels—seven pure vowels, 15 diphthongs, and four triphthongs—although an accurate count is difficult given the complexity of the language. Of the pure vowels, only *a* and *a:* appear in the final syllable, while *i, ɛ, u, o,* and *ə* are confined to the unstressed penultimate vowel. Three-syllable words are rare, but where they do occur, only schwa is found in the prepenultimate syllable. Diphthongs and triphthongs are confined to final syllables, and are interpreted as single, complex vocalic units. All of the vowels shown above are derivable from the PMP vowels *i, *a, *u, *ə, and the diphthongs *aw, *ay, *iw, and *uy.

TABLE 3. MERAP CONSONANTS

	LABIAL	ALVEOLAR	PALATAL	VELAR	GLOTTAL
VOICELESS PLOSIVE	p	t	c	k	ʔ
VOICED PLOSIVE	b*	d	j	g	
NASAL STOP	m	n	ñ	ŋ	
FRICATIVE		s			h
LATERAL		l			
TRILL		r			

* /b/ is realized as an implosive stop [ɓ] in the onset of final syllables.

TABLE 4. MERAP VOWELS

	FRONT	CENTRAL	BACK
HIGH	i ĩə̰, iə̰, iw		u ũə̰, uə̰, uy
MID	ɛ ɛy, ɛyə̰	ə əw	o oyə̰, oy
LOW		a, a: ãə̰, aə̰, aọ, aụ, awə̰, aị, aɛ̰, ayə̰	

Justifying the above vowel inventory is dependent on one's analysis of stress. In Merap, stress appears regularly on the final syllable. In every vowel phoneme presented above, stress falls on the initial segment. Thus, *a* is stressed [ˈa], *əw* is stressed [ˈəw], *ĩə̯̃* is stressed [ˈĩə̯̃], and *aɣə̯* is stressed [ˈajə̯]. The stress pattern is the basic motivator for interpreting diphthongs and triphthongs as single segments. If, for example, one were to interpret *aɣə̯* as *aɣə*, that is, with the schwa as a separate syllable, one would have to explain why stress is penultimate [ˈa.jə] and not final [a.ˈjə]. Some words do have legitimate vowel clusters, like *kiawə̯* 'the Myna bird', which is stressed [kiˈawə̯] and reflects PKAY *kiuŋ. During the preparation of this paper, multiple alternative analyses were proposed to remove triphthongs from the vowel inventory. Many were proposed by an anonymous reviewer, in an attempt to simplify the vowel inventory. Below, the basic reanalysis of triphthongs into two syllables is considered using *kiawə̯* as an example. The result maintains the original analysis, that Merap does have triphthongs, and that reanalyzing triphthongs as two syllables causes unnecessary irregularities in both canonical syllable shape and stress.

It was pointed out that there are no stressed schwas in Merap (other than stressed syllable nuclei in diphthongs like *əw*). With this in mind, one could assume that schwa in the final syllable prevented stress shift. However, the larger picture reveals that schwa never altered the strictly word-final stress pattern of the language. Rather, where schwa appeared in the final syllable, it received stress, but lowered to *a* (see 4.14 for more discussion). Thus, if one were to assume that *kiawə̯* be reanalyzed as *ki.a.wə,* one would also have to assume irregular treatment of schwa in these words only, since schwa in *salən 'earthworm', and all words where schwa was historically found in the final syllable, failed to prevent stress shift. The result is modern Merap *halaŋ* [haˈlaŋ] 'earthworm' rather than **halən* [ˈhaləŋ], because schwa was the nucleus of the final syllable, but *kiawə̯* [kiˈawə̯] rather than **kiawa* [kiaˈwaː] because schwa is not the nucleus of the final syllable.

3.2 WORD-FINAL VELARS *k* AND *ŋ*. All stops are found in word-final position in Merap, including the palatal stops *c* and *ñ* and labialized velar consonants k^w and $ŋ^w$ (found *only* in final position). However, *-c, -ñ, -*k^w*,* and -$ŋ^w$ are conditioned by the quality of the preceding vowel, and appear to be allophones of *-k* and *-ŋ*:

- *-c* only appears in a palatal environment, which is to say, only when immediately preceded by either the vowel *i* or the glide *y* [i̯]. As noted above, all voiceless consonants appear in word-final position after a short *a*, but after *i/y* only *c* and *ʔ* occur.
- *-ñ* also only appears in a palatal environment. Like the voiceless stops, all nasals (except for *ñ*) appear in word-final position where they historically followed a schwa. After other vowels, however, all final nasals either deleted (after *a*), merged as $ŋ^w$ (after *u), or merged as *ñ* (after *i). The result is that there is no distinction in word-final nasals after vowels other than schwa.
- -k^w only appears in a labial environment, which is to say, only when immediately preceded by either *u* or *w.* It is in complementary distribution with *-c* and both are allophones of *k.*
- -$ŋ^w$ also only appears in a labial environment. As noted above, it is in complementary distribution with *-ñ,* and both are allophones of *ŋ.*

3.3 CONSONANT CLUSTERS IN INITIAL POSITION AND VOICELESS SONORANTS. Merap allows several word-initial consonant clusters, and is unique in Borneo in that it allows clusters of three consonants in this position. There are some constraints on the shape of clusters:

(i) no clusters begin with *d*, *g*, or *t*;
(ii) an obstruent can only be immediately followed by *r* or *l*;
(iii) obstruents can only be preceded by nasals;
(iv) clusters with three consonants can only be of two types: noncoronal nasal + homorganic obstruent + *l/r*, or *h* + nasal + any consonant.

All of the attested consonant clusters in Merap are listed below, with an example word. Voiceless sonorants that are formed with *h*+sonorant clusters are discussed later. Clusters like *mp-* and *mb-* are interpreted as belonging to the same syllable. However, additional fieldwork is needed to provide experimental evidence for this interpretation. It is quite possible that nasal-initial clusters will be shown to have an internal syllable boundary.

(21)

pl	pləo̯ ‘ten’		ŋk	ŋkayə̯ ‘to bring’
pr	prawə̯ʔ ‘feelings’		ŋg	ŋgua ‘fruit’
kl	klanao̯ ‘finger’		ŋl	ŋlatawkʷ ‘to float’
kr	kraę̯ ‘now’		mbl	mblayʔ ‘to buy’
bl	bluhuə̯ʔ ‘to stab from below’		mpr	mpruyc ‘alive’
mp	mpau̯ ‘smelly’		ŋgl	ŋgla ‘sticky’
mt	mtɛyh ‘to lean’		ŋkr	ŋkrawə̯ʔ ‘to snore’
mj	mjah ‘waterfall’		hmp	hmpiə̯n ‘to sneeze’
nc	ncɛy ‘one’		hmn	hmnɛyə̯ ‘bright’
nt	ntao̯h ‘bathe’		hŋk	hŋkiə̯t ‘every’

All sonorants in Merap can be pronounced without voice when immediately preceded by *h.* Historically, these consonant clusters have several origins, but in the modern language the following voiceless sonorants have been observed:

(22)

hlawʔ ‘they’	hnau̯ ‘full’
hləw:ŋʷ ‘fingernail’	hñãə̯ ‘tortoise’
hmũə̯ ‘to eat’	hŋkiə̯t ‘every’
hmneyə̯ ‘bright’	hŋiə̯t ‘bee’
hmayh ‘sweet’	hrĩə̯ ‘when’

Whether these are true phonemic voiceless sonorants, or allophones where voicing was lost next to *h*, is still problematic. In this paper, they are treated as consonant clusters, but further testing is needed to come to a definitive conclusion. A possible avenue of research is secondary stress testing in words with clusters in medial position. For example, words with CVCCVC structures receive a secondary stress in Merap: ˌCVC'CVC. If a word like *ŋahŋaə̃*˜‘to breathe’ receives secondary stress, it would indicate a consonant cluster, not a single voiceless sonorant. However, it remains to be seen if this observation also holds for initial voiceless sonorants, as stress cannot be used as a test in this environment.

3.4 3.4 INSERTION OF *b* AND *p* BETWEEN *ml-* AND *mr-*. There are no examples of *ml* or *mr* in Merap. In any case where *ml* or *mr* is expected, either through

synchronic alternations or historical sound changes, one finds *mbl* instead of *ml* and *mpr* instead of *mr.* A diachronic example of *p* insertion is *mprəwa̱* 'to snap' where PKAY had *məruŋ, and a synchronic example is *ruyc* 'life', which becomes *mpruyc* 'alive' after the verbal prefix *m-* is attached to the base. Additionally, some examples are historically misleading, as in *mbləwŋ*^{w} 'body hair' from *bulu-n. In this example, it appears that *b was retained, and *m* was innovated, but other examples where *b > *m* in initial position offer an alternative analysis. For example, *batu 'stone' is reflected as *mataq̱ʔ,* and *bəRsay 'paddle' with *mahaq̱.* Considering these examples, *mbləwŋ*^{w} most likely went through the following ordered sound changes:

(23) *bulu
*bulun
*buləwn
*buləwŋ^{w}
*bələwŋ^{w}
*məlәwŋ^{w}
*mlәwŋ^{w}
mbləwŋ^{w}

In this scenario, *mbləwŋ*^{w} offers diachronic evidence for synchronic *-b-* and *-p-* insertion between *m* and *l/r:*

3.5 THE STATUS OF /g/. Examples of *g* are difficult to come by in Merap, and the only case where *g* is found alone (not in a consonant cluster) is a suspected loan, *gawa̱* 'gong' (Indonesian *agung*). There are, however, several examples where *g* is found as part of a homorganic nasal-obstruent cluster *ŋg,* and these examples are given below:

(24) ŋgueq̱ 'bear'
ŋgua: 'fruit'
ŋguaʔ 'traditions; customs'
ŋgla: ayc 'sticky'
liŋgaya̱ 'dragonfly'
paŋga:h 'swamp'
pluŋguq̱ʔ 'civet cat'

Because Merap also has *ŋk* clusters—such as *ŋkuaŋn* 'parang', *ŋkaya̱* 'to gather things together', and *taŋka:ʔ* 'the dry season'—and *ŋ* as a stand-alone consonant, allophony can be ruled out. It appears, then, that *g* occurs only in homorganic nasal-obstruent clusters in native vocabulary, and as a stand-alone only in loan words.

3.6 LONG VOWELS. All vowels in word-final position automatically lengthen and all vowels in closed final syllables (except *a*) are phonetically long. Before word-final glottal stop and *h,* however, there is a phonemic distinction between regular and long *a*. Before other consonants, historical changes have prevented distinction (all final voiceless stops merged as glottal stop after *a, while *a became nasalized before nasals). Some examples, including minimal or near minimal pairs, are:

(25) paʔ 'fathom; length of outstretched arms'
pa:ʔ 'four'
prah 'sick; painful'
taʔa:h 'type of traditional long skirt'
lata:h 'flat; plain'

3.7 RISING DIPHTHONGS. A rising diphthong is any complex vowel where the initial segment has a lower sonority than the second segment. Because high vowels have a lower sonority than low vowels, any diphthong that starts high and ends low is referred to as a rising diphthong. In Merap there are four such diphthongs: *iə̯, ĩə̯̃ uə̯, ʉə̯̃* Only the oral rising diphthongs appear in closed final syllables (exemplified with words that are closed with a glottal stop below), but the nasals are confined to open syllables.

(26) miə̯ʔ 'five' tuə̯ʔ 'blind'
hiə̯ 'he/she/it' kuə̯ 'the two of you'
jəʔtĩə̯̃ 'nine' hmũə̯̃ 'eat'

3.8 FALLING DIPHTHONGS. A falling diphthong is any complex vowel where the initial segment has a higher sonority than the second segment. Any diphthong that starts low and ends high or mid, or a diphthong that starts mid and ends high, is thus referred to as a falling diphthong. Merap has eight such diphthongs: /ɛy əw aə̯ ao̯ au̯ ai̯ ae̯ oy/. Each is exemplified below.

In final position, *oy* and *ɛw* are realized as [oe̯] and [eo̯], but this is not phonemic.

(27) mɛrɛyh 'to slice' blalayʔ 'thunder'
pləwk^{w} 'tree sap' paraę̯ 'field rice'
ŋaə̯ 'scaly ant-eater' toyʔ 'banana'
klaʔawh 'barking deer' ŋoy [ŋoe̯] 'rattan'
ncau̯ 'to lift; carry'

The diphthongs *au̯, ao̯, ai̯,* and *ae̯* are only distinctive in final position. The following minimal and near minimal pairs demonstrates their distinction:

(28) mpao̯ 'tall'
mpau̯ 'smelly'
pai̯ 'fire'
mpae̯ 'why'

Because these four diphthongs are only distinctive in final position, *aw* and *ay* are written in closed syllables, but *ao̯, au̯, ae̯,* and *ai̯* are written in open final syllables.

The rising diphthongs *ow* and *ɛw* are only present in words with *o* and *ɛ* in the penultimate syllable. For example, *kotowʔ* louse', *kocowʔ* 'heron', *tɛrɛwh* 'to sleep', *lɛbɛw* [lɛbɛo̯] 'shallow'. Thus, both *ɛw* and *ow* are considered allophones of *əw*, and are conditioned by the quality of the preceding vowel.

3.9 LEVEL DIPHTHONGS. A level diphthong is any complex vowel that does not change height between segments, such as a diphthong that begins and ends high, or begins and ends mid. Merap has two such diphthongs: *iw,* and *ʉy.* Note that *oy* (see 3.8) also appears in final position, and is thus distinguished from *ʉy.*

(29) ñəliw 'swallow'
luy 'return home'

3.10 TRIPHTHONGS. A triphthong is a complex vowel with three places of articulation, typically one that starts high, moves towards the center, then ends high, or one

that starts low, moves towards the center, and ends low again. In Merap four such triphthongs have been identified: /ayə̯ awə̯ ɛyə̯ oyə̯/.

(30) hayə̯ʔ ‘ripe’
maɓawə̯ʔ ‘drunk’
kloyə̯ ‘bone’
kɛyə̯ʔ ‘the two of us (exclusive)’

3.11 NASAL VOWELS. Three diphthongs in Merap contrast with a nasal counterpart, /aə̯/ and /ãə̯̃/, /iə̯/ and /ĩə̯̃/, and /uə̯/ and /ũə̯̃/. The distinction only occurs in open final syllables. It is difficult to find an example where *aə̯* is distinguished from *ãə̯̃* in final position, as historical processes have confined *aə̯* almost exclusively to closed final syllables. However, /ŋaə̯/ ‘scaly anteater’ does have an oral diphthong, which appears to form a near minimal pair with /hŋãə̯̃/ ‘breath; the act of breathing’.

The nasal diphthongs /ũə̯̃/ and /ĩə̯̃/ have a greater number of examples of distinction, and minimal pairs are easy to come by. Two minimal pairs are given below:

(31) kũə̯̃ ‘scabbard for a machete’
kuə̯ ‘the two of you’
ɓĩə̯̃ ‘parent-in-law; child-in-law’
ɓiə̯ ‘underneath’

4. MERAP SOUND CHANGES. In this section, all reconstructions are to Proto-Kayanic (PKAY), which includes all groups that identify as Kayan, Bahau, Busang, Ngorek (or Murik),[11] Merap, Segai, and Modang, unless otherwise indicated. A further internal subgrouping of Kayanic languages follows this section. Before that, however, the interesting and at times difficult historical phonology of Merap is discussed. Although the modern language is quite different from its closest relatives, it is possible to explain every difference in cognates in terms of the historical developments of Merap. This is an extraordinary example of the usefulness of the comparative method, as many of the cognates have changed to the point where they are difficult to recognize. For example, Merap *yãə̯̃* regularly reflects *duRian ‘durian’ through a series of sound changes including deletion of the initial syllable that produced *Rian, loss of *R which produced *ian, stress shift to the final syllable resulting in loss of syllabicity of word-initial *i which produced *yan, and finally, word-final nasal deletion with residual nasalization of *a which produced the modern form *yãə̯̃*

Throughout this section, statements of sound changes are immediately followed by several examples. Due to the sheer number of sound changes, I have tried to make these statements as concise as possible.

11. The name of this language is somewhat inconsistently reported in the literature. Blust (1974) reports “Murik” from Long Semiyang, Sarawak, but during my own research in Long Semiyang I was told by my consultant that Ngorek was the more proper term. Soriente (2008:59) lists Ngorek as the name of the larger subgroup that contains Pua’, Hueng Ba, and Murik. Both words, Ngorek and Murik, reflect PMP *udahik ‘upriver’ (Iban *mudik* ‘to go upriver’, Kayan *hudik* ‘upriver areas’, Punan Bah *muriəʔ* ‘to go upriver’). Because my consultant claimed that Ngorek was the preferred term, I am inclined to use it when referring to the language of Long Semiyang. For the larger subgroup, I use Murik after Blust (1974) who reconstructed “Proto-Kayan-Murik.”

4.1 WORD-FINAL VOICELESS STOPS. Word-finally, Merap has reduced the environments where *-p, *-t, and *-k are distinguished from one another, but it has not completely merged them, and distinctions remain where they occurred after schwa.

Word-final voiceless stops *-p, *-t, *-k merged as *-ʔ* if preceded by *a.

(32) *siap 'chicken' > hɛaʔ
*əpat 'four' > paːʔ
*anak 'child' > nayə̲ʔ

Word-final *k also became glottal stop after *i and *u, although not before altering the quality of the vowel.

(33) *buk 'head hair' > ɓawə̲ʔ
*baʔik 'short' > maʔayə̲ʔ

Final *-t and *-p are reflected as an unreleased voiceless palatal stop *c* if they were preceded by *i.

(34) *murip 'alive' > mpruyc
*laʔip 'shoulder' > laʔayc
*hakit 'raft' > kayc
*laŋit 'sky' > laŋayc

Final *-t and *-p became a voiceless labialized velar stop k^w when preceded by *u.

(35) *kaʔub 'lie prone' > *kaʔup > ŋaʔawk^w
*pulut 'sap' > pləwk^w
*takut 'afraid' > takawk^w

All three, *-p, *-t, and *-k did not change if preceded by schwa.

(36) *kələb > PMUR[12] *kələp 'turtle' > klap
*siŋət 'bee, to sting' > hŋiə̲t
*utək 'brain' > tuə̲k

Glottal stop in word-final position was deleted in all environments.

(37) *puluʔ 'ten' > pləw
*utaʔ 'vomit' > toa
*piliʔ 'choose' > mblɛy

4.2 NONFINAL VOICELESS STOPS. In general, the voiceless stops did not change in nonfinal position, with some exceptions. Where initial *t preceded an *l in the following syllable (*tVl-), it became *k* when the vowel deleted.

(38) *tulad 'animal' > klũə̲
*tulaŋ 'bone' > kloyə̲
*təlu 'three' > klawʔ
*təlis 'squirrel' > klayh

Where a voiceless stop in initial position became adjacent to a nasal, it is reflected as *h*.

(39) *pənuʔ 'full' > hnau̯
*kuman 'to eat' > hmũə̲

12. PMUR = Proto-Murik-Merap.

4.3 WORD-FINAL VOICED STOPS. Merap is one of only a handful of Kayanic languages that reflect voiced word-final stops with both a nasal (*n*) and a voiceless stop (*p*). In final position, *b > *p (*p in turn was further altered depending on the quality of the preceding vowel). Examples of this change were difficult to come by, but two were located.

(40) *kaʔub 'lie prone' > *kaʔup > ŋaʔawk^{w13}
*kələb 'turtle' > *kələp > klap

*-d on the other hand became *n* (which often surfaces as *ŋ*^{w}, ñ, or ṽ depending on the quality of the preceding vowel).

(41) *maʔud 'wake up' > *maʔun > maʔawŋ^{w}
*uləd 'worm' > *ulən > luə̯n
*tulad 'animal' > *tulan > klũə̯̃
*apid 'twins' > *apin > kapayñ

4.4 NONFINAL VOICED STOPS. *b- is reflected as *m*

(42) *baya 'crocodile' > maya:ʔ
*batu 'stone' > matawʔ
*basay 'paddle' > maha:e̯

If *b- occurred before medial *-l-, it still became *m,* but a secondary *b* was inserted between *m* and *l*, as in the following examples:

(43) *buluʔ 'bamboo' > mbləw
*bulan 'moon' > mblə̃wə̯̃

*b- is reflected as *ɓ* where it occurred as the initial segment in a monosyllable or in medial position.

(44) *bi 'carry on back' > ɓayʔ
*buk 'head hair' > ɓawə̯ʔ
*təbu 'sugarcane' > təɓawʔ
*nubus 'to plant' > loɓowh
*mabuk 'drunk' > maɓawə̯ʔ

In initial position *d became *l.*

(45) *daʔan 'branch' > laʔãə̯̃
*dahaʔ 'blood' > la:
*duman 'year' > lumũə̯̃

In medial position, PMP *d became *r,* which is likely a retention from PKAY, as all Kayanic languages reflect *-d- > *-r-, except Segai-Modang, where *-d- > *-r- > *-l-*.

(46) *tiduR 'sleep' > *tiruh > terewh
*paday 'field rice' > *paray > parae̯

*z- (which is written as Proto-Kayanic *j) is reflected as *ñ* in three examples (47a), but also as *t* (47b) and *j* (47c) in one example each.

13. With nasal substitution on initial k-.

(47) a. *jaʔa-n ‘chin’ > ñaʔãə̰
*jalan ‘road’ > ñalãə̰
*jihi ‘house post’ > ñɛːy
b *japitan ‘nine’ > jəʔtĩə̰
c. *jaʔak ‘bad’ > taʔayə̰ʔ

In medial position *z became *c*.

(48) *kajəl ‘dull’ > kacan
*ujan ‘rain’ > cə̃wə̰

In two cases, however, *-z- > *ʃ*, a voiced palatal implosive.

(49) *nəpujuk ‘jump’ > ləʔʃəwə̰ʔ
*təkəjət ‘surprised’ > ʃat

4.5 NASAL-OBSTRUENT CLUSTERS. Nasal-obstruent clusters underwent devoicing, *mb > *mp,* *nd > *nt,* *nj > *nc,* and *ŋg > *ŋk.*

(50) *ləmbaw ‘tall’ > ləmpaọ
*lindəm ‘dark’ > ntiə̰m
*mənjuʔ ‘lift’ > ncaụ
*ñaŋgu ‘fry’ > ñaŋkawʔ

4.6 *s. In all positions *s became *h.*

(51) *saləŋ ‘earthworm’ > halaŋ
*siŋəhət ‘sting’ > hŋiə̰ʔ
*pusəd ‘navel’ > puhuə̰n
*isi-n ‘flesh’ > hiːñ
*pərəs ‘sick’ > prah
*təlaʔus ‘barking deer’ > klaʔaọh

4.7 *l. In initial and medial positions, *l did not change.

(52) *lubaŋ ‘hole’ > loboyə̰
*laŋaw ‘a fly’ > laŋaọ
*bulu-n ‘body hair’ > mbləwŋʷ
*tulad ‘animal’ > klũə̰

In final position, however, *l became *n, which has various reflexes in the modern language (see reflexes of *-n).

(53) *kajəl ‘dull’ > *kajən > kacan
*gatəl ‘itchy’ > *gatən > gatan
*kapal ‘thick’ > *kapan > kapãə̰

4.8 NASALS. The nasals in Merap are fairly stable, but in word-final position, as discussed earlier, their reflexes depend on the quality of the preceding vowel. In final position, the nasals delete after *a, and nasalized the preceding vowel.

(54) *maram ‘rotten’ > marãə̰
*jalan ‘road’ > ñalãə̰
*lubaŋ ‘hole’ > loboyə̰

*-ŋ also deleted after *i and *u, but did not nasalize the preceding vowel.

(55) *uruŋ 'nose' > ruə̯
*laʔuŋ 'back' > laʔawə̯
*ikiŋ 'pinky' > kiə̯
*mariŋ 'new' > marayə̯

*-n, however, became ñ after *i and *ŋ^{w} after *u. Unfortunately, the data do not contain reflexes of *-um and *-im, so their reflexes are unknown.

(56) *lamin 'room' > lanayñ 'floor'
*ləbin 'to wring' > plaɓayñ
*sulu-n 'finger nail > hləwŋ^{w}
*puʔun 'base' > puʔowŋ^{w}

4.9 DEVELOPMENT OF *h*-SONORANT CLUSTERS. Merap has developed voiceless sonorants from the underlying consonant clusters *hm, hn, hñ, hŋ, hr,* and *hl.* The modern clusters are derivable through processes where more complex consonant clusters were reduced: for example, *pənuq 'full' became pre-Merap *pnu, and the cluster *pn- was simplified to *hn-* in *hnau̯.* Voiceless sonorants only developed from clusters of voiceless stops and sonorants, such as *km, *pn, *sl. More examples are listed below:

(57) *pənuʔ 'full' > *pnuʔ > hnau̯
*kuman 'eat' > *kmuan > hmũə̯
*sulu-n 'finger nail' > *slun > hləwŋ^{w}
*siran 'when' > *srian > hrĩə̯
*siŋət 'sting' > *sŋiət > hŋiə̯t

Voiceless sonorants did not develop from clusters that involved voiced stops, but the only example involve a reflex of a three-syllable word, which result in a modern *bl-* cluster in *blawaŋ* 'empty' (compare with Ngorek *bəlawəŋ*). In other cases, *b- became *m-* and *d- became *l-,* which bleeds any further development of voiced stop-sonorant clusters.

4.10 REFLEXES OF *a. *a has numerous reflexes (/a iə̯ ə̯ uə̯ yə̯ ãə̯ ĩə̯ ũə̯ aː/) that occur under specific conditions that are explained in further detail below:

Word-final *-a is reflected as either *a, iə̯,* or *uə̯,* all with an innovated glottal stop closing the syllable, giving *aʔ, iə̯ʔ,* and *uə̯ʔ*; *aʔ* occurs where *-a is preceded by *a, *ə, or Ø in the penultimate syllable; *iə̯ʔ* occurs where preceded by *i; and *uə̯ʔ* occurs where preceded by *u.

(58) *ata 'water' > kataʔ
*dəpa 'fathom' > paʔ
*ñipa 'snake' > piə̯ʔ
*buta 'blind' > buə̯ʔ

Before a velar consonant, *a in the final syllable is reflected as *yə̯* as part of a triphthong, with the nucleus dependent on the quality of the preceding vowel. Thus, *aCak > *Cayə̯ʔ,* *uCak > *Coyə̯ʔ,* and *iCak > *Ciyə̯ʔ (*iyə̯ here became *iːə̯* and further contracted to *iə̯* in a closed syllable, while it remained *iːə̯* in an open syllable). In all cases, it is assumed that final velars created a schwa off-glide after first causing *a to front: *ak > *eə̯k*, *aŋ > *eə̯ŋ*. Later, strong word-final stress caused the vowel nucleus to break: *eə̯k > *ayə̯k, *eə̯ŋ >

*ayə̯ŋ. Velars caused *a-fronting with schwa off-glides in many Bornean languages, including Melanau, Kajang, Punan Bah, and several Kayan dialects (Blust 1977; Smith 2017), so there is strong phonetic motivation to assume the same history in Merap.[14] This ordering (off-gliding before velars, followed by vowel breaking) is also assumed for triphthongs that developed from *-uk, *-uŋ, *-ik, and *-iŋ (4.11 and 4.12 below).

(59) *lubaŋ 'hole' > loɓoyə̯
*miʔaŋ 'split' > miʔiːə̯
*məndaŋ 'fly' > məntayə̯
*anak 'child' > nayə̯ʔ

If *a was followed by *n in the final syllable, *n was deleted and *a > *ãə̯̃* [ə̃ə̯̃] (the vowel is centralized as a result of nasalization but speakers perceive is as /ã/ and not as /ə̃/). *ĩə̯̃* and *ũə̯̃* also occur as a result of influence from the penultimate vowel.

(60) *jaʔa-n 'chin' > ñaʔãə̯̃
*daʔan 'branch' > laʔãə̯̃
*kitan 'binturong' > tĩə̯̃
*ujan 'rain' > cũə̯̃

*a became *iə̯* and *uə̯* when preceded by *i and *u and when followed by a nonvelar consonant.

(61) *kulat 'mushroom' > kluə̯ʔ
*ibah 'saliva' > ɓiə̯h

Otherwise, *a was lengthened to *aː* in a closed final syllable with either *a or *ə (or Ø) in the penult.

(62) *datah 'plain; flat' > lataːh
*əpat 'four' > paːʔ
*taʔas 'trad. skirt' > taʔaːh
*baʔ 'mouth' > ɓaː

4.11 REFLEXES OF *u. In the final syllable, *u is reflected as *aə̯, əw, ow, au̯, awə̯, uːə̯,* and *iw* under specific conditions that are elaborated on below.

Word-final *-u is reflected as *aw* [aʊ̯], and closed with a glottal stop, giving *awʔ* where it was preceded by *a, *ə, or Ø in the penultimate.

(63) *alu 'large pestle' > lawʔ
*təbu 'sugarcane' > təɓawʔ
*tu 'ghost' > tawʔ

If word-final *u was preceded by *i in the penult, it is reflected as *iw.*

(64) *siku 'elbow' > kiwʔ
*ñilu/ñəlu 'to swallow' > ñəliw

If *u occurred in the final syllable, after *u in the penult, then it is reflected as *ow,* unless *u was subsequently deleted, in which case it is reflected as *əw.* This is true regard-

14. Reflexes of *tuqəlaŋ 'bone' provide good examples: Mukah (Melanau) *tuleəŋ,* Sekapan (Kajang) *tuʔɻeəə̯,* Kajaman (Kajang) *tuʔɻeəŋ,* Punan Bah *toleəŋ,* Data Dian Kayan *tuleəŋ.* Also, in the Iban of the upper Kapuas in West Kalimantan, *a fronted to *æ* only before *-k, but there is no secondary off-glide: *anak > anæʔ.

less of whether *u was in an open final syllable or a closed final syllable, unless it is closed with a velar, in which case the reflexes are different.

(65) *kutu ‘louse’ > kotowʔ
*kuju ‘heron’ > kocowʔ
*tujuʔ ‘seven’ > tosow
*nubus ‘to plant’ > loɓowh
*pulut ‘tree sap’ > pləwkʷ
*uru ‘grass’ > rəwʔ
*tujuʔ ‘seven’ > cəw

In two cases where *u occurred word-finally after *h, a glottal stop was not innovated, and *u became the off-glide of a diphthong, thus *ahu# > *au̯#*.

(66) *pahu ‘grasshopper’ > pau̯
*ñahu ‘eagle’ > ñau̯

If *u occurred in a final syllable closed with *ʔ and was preceded by *a, *ə, or Ø, then it is reflected as *au̯*.

(67) *mənjuʔ ‘lift’ > ncau̯
*juʔ ‘far’ > cau̯
*aruʔ ‘long’ > rau̯

Where *u occurred in the final syllable before a velar consonant, and was preceded by any vowel other than *u, it is reflected as *awə̯*.

(68) *manuk ‘chicken’ > manawə̯ʔ
*təkuk ‘nape’ > kəkawə̯ʔ
*buk ‘head hair’ > ɓawə̯ʔ
*laʔuŋ ‘back’ > laʔawə̯
*kətuŋ ‘porcupine’ > tawə̯
*tiuŋ/kiuŋ ‘myna’ > kiawə̯

If *uk/uŋ was preceded by *u in the penult, it became *u:ə̯*. Length in this case is phonetic.

(69) *lunuk ‘banyan tree’ > lunu:ə̯ʔ
*tutuŋ ‘aflame’ > tutu:ə̯
*uruŋ ‘nose’ > ru:ə̯

Elsewhere, *u in a closed final syllable is reflected as *ao̯,* which is written simply as *aw* in closed syllables. Note that while *aw* is phonetically [ao̯], it is distinct from *au̯* [au̯] in word-final position (see above).

(70) *takut ‘afraid’ > takawkʷ
*put ‘blowpipe’ > pawkʷ
*təlaʔus ‘barking deer’ > klaʔawh
*maʔud ‘wake up’ > maʔawŋʷ

4.12 REFLEXES OF *i. *i has the following reflexes in Merap, *ay, ɛy, ayə̯, i:ə̯, uy, oy.* The conditions for these various reflexes are stated below.

Typically, word-final *-i became *-ay* with a glottal stop closing the vowel, giving *-ayʔ*. If *-i occurred after a penultimate *u, then it is reflected as *oyʔ*. If it followed *i, it is reflected as *ɛyʔ*.

(71) *taʔi ‘excrement’ > taʔayʔ
*laki ‘man’ > lakayʔ
*məli ‘to buy’ > mblayʔ
*punti ‘banana’ > toyʔ
*ini ‘this’ > nɛyʔ

In one case, where *i was preceded by *i in the penult, and separated by *h, the *h deleted, and *ii became *iy, which is reflected as *ɛy*. Note that in this form there is no final glottal stop, presumably because final *i had become a glide.

(72) *jihi ‘house post’ > ñɛy

In a closed final syllable, *i is reflected as *ay, ayə̯, i:ə̯, ɛy,* and *ɯy* under specific conditions, which are elaborated upon below. Where *i was preceded by *a, *ə, or Ø and was closed with a nonvelar consonant, it became *ay*.

(73) *kapid ‘twins’ > kəpayc
*təlis ‘squirrel’ > klayh
*paʔiʔ ‘bitter’ > paʔay

Before *k and *ŋ, *i is reflected as *ayə̯*.

(74) *baʔik ‘short’ > maʔayə̯ʔ
*mariŋ ‘new’ > marayə̯

Before a velar, but if preceded by *i, it became *i:ə̯*. Length in this case is phonetic.

(75) *dindiŋ ‘wall’ > rintiːə̯
*kisiŋ ‘laugh’ > kihiːə̯

Where *i was preceded by *i in the penultimate syllable and closed with a nonvelar consonant, it became *ɛy*.

(76) *isiʔ ‘snail’ > hɛy
*miliʔ ‘choose’ > mblɛy
*ñipis ‘thin’ > lɛpɛyh
*miris ‘slice’ > mɛrɛyh

If *i was preceded by *u, then it is reflected as *ɯy*.

(77) *murip ‘alive’ > mpruyc
*putiʔ ‘white’ > tuy
*kulih ‘leopard’ > kluyh

4.13 REFLEXES OF WORD-FINAL DIPHTHONGS. *ay is typically reflected as *ae̯,* unless preceded by *u, where it became *oy* [oe̯], or **i* where it became *ɛy*.

(78) *atay ‘liver’ > tae̯
*matay ‘die’ > matae̯
*dahulay ‘left’ > loy [loe̯]
*inay ‘mother’ > nɛy

*aw is reflected as *ao̯,* unless preceded by *i, where it is reflected as *ɛw* [ɛo̯]. There are no examples in the data where *aw was preceded by *u.

(79) *panaw ‘walk’ > panaọ̯
*kayaw ‘headhunting’ > kayaọ̯
*ñibaw ‘shallow’ > lɛbɛw

*uy had merged with *i as *i and is reflected as *aị̯* as in *apuy > *paị̯* ‘fire’. The only two reflexes of final *iw give conflicting evidence: *bahiw ‘strong wind’ had apparently become *bayu, as evidenced by Ngorek *bayu* and Merap *mayaụ̯;*[15] but *baliw ‘to become, transform’ is reflected as *malaị̯,* which suggests that *iw and *uy had both merged with *i.

4.14 REFLEXES OF SCHWA. In a final syllable, schwa is reflected as a short *a,* which contrasts with long *a:* from *a (see 4.10 above).

(80) *salәŋ ‘earthworm’ > halaŋ
*tәbәŋ ‘to fell trees’ > ɓaŋ
*nәm ‘six’ > nam
*pәrәs ‘sick; pain’ > prah

If schwa was preceded by *i or *u in the penultimate syllable, however, it coalesced with the preceding vowel through reverse umlaut.

(81) *utәk ‘brain’ > tuọ̯k
*uləd ‘worm’ > luọ̯n
*siŋәt ‘to sting’ > hŋiọ̯t

In the penultimate syllable, schwa was deleted.

(82) *pәru-n ‘gall’ > prәwŋʷ
*sәŋәm ‘cold’ > hŋam
*kәra-n ‘neck’ > krãẽ̯

In two cases, however, schwa apparently did not delete:

(83) *bәsuR ‘satiated’ > məhəwh
*tәbu ‘sugar cane’ > tәɓawʔ

4.15 ORDERING RELATIONSHIPS. Many of the sound changes that have altered word-final syllables, particularly those that have resulted in Merap’s expanded vowel inventory, the reduction of final voiceless stops, and nasalization of vowels before final nasal stops other than *ŋ, can be shown to have occurred after several earlier sound changes.

***-d > *-n, *-b > *-p.** Word-final *d became *n before *n was altered in any way. Where *-n became a labial velar after *u, *d had already merged as *n: *maʔud ‘wake up’ > *maʔun > *maʔawŋʷ. Where *-n was palatalized, *d had also already merged as *n: *apid ‘twins’ *apin > *kapayñ*. Where *-n was deleted (resulting in nasalization of the vowel), again *d had already merged as *n: *tulad ‘animal’ > *tulan > klũẽ̯. The same holds true for reflexes of *-b. That is, *-b became *-p before final consonants were altered, as evidenced by *kaʔub ‘lie prone’ > *kaʔup > *ŋaʔawkʷ*.

15. The likely explanation for this disparity is that PMP *baRiw became PKAY *bahiw. In Proto-Murik-Merap, *-h- deleted, which produced *baiw. This created an environment similar to that where PMP *kahiw ‘tree’ became *kaiw through the same processes. *kahiw is reflected as *kayu throughout Borneo, and the same probably holds true for *baiw > *bayu. Thus, Merap *mayaụ̯* reflects *bayu.

Velar consonant deletion. Word-final *k and *ŋ were deleted after all vowels other than schwa. However, the velar series had a specific effect on preceding vowels that must have occurred before they were lost (*-k became *-ʔ* and *-ŋ was deleted altogether). For example, a triphthong formed where *-k became glottal stop in *manuk 'chicken' > *manuə̯k > *manawə̯ʔ*, but not where inherited glottal stop was deleted (*aruʔ 'long' > *aru > *araɰ*) or where glottal stop was innovated (*təbu 'sugar cane' > *təbuʔ > *təbawʔ*). Had *-k first become glottal stop, *manuk would have gone through the following stages, *manuk > **manuʔ > ***manawʔ*.

Penultimate vowel deletion. In some cases, the penultimate vowels were deleted. Where the penultimate vowel was historically high, it had reverse umlaut effects on the following vowel before it was deleted. Thus, reverse umlaut preceded vowel deletion. If, for example, the penultimate vowel in *ujan 'rain' (Merap *cuã̯*) had deleted before reverse umlaut, the following stages would have occurred, *ujan > **jan > **cãə̯.

Vowel lengthening. All vowels lengthen in word-final position. It is also true, however, that vowels in closed syllables are phonetically long, but do not form contrasts except where *a > *a:* and *ə > *a*. *a must have lengthened before schwa lowered, otherwise schwa would have become *a, then both *a (from *a) and *a (from schwa) would have lengthened.[16] For example, if *ənəm 'six' (Merap *nam*) had lowered schwa first, the following feeding order would have occurred, *ənəm > *nəm > **nam > ***na:m*.

Schwa lowering. As noted above, schwa lowered to *a only after *a had lengthened to *a:. It must also be the case that schwa lowered only after word-final voiceless stops reduced to glottal stop (a change that occurred only after *a). If schwa lowered first, then the following order would be observable for *utək 'brain' (Merap *tuə̯ʔ*): *utək > **utak > **uteə̯k > ***toyə̯ʔ*.

The following four rounds of sound changes must have occurred, in the specific orders shown, given the data presented in this paper. However, the history of Merap is especially complex and the following ordering relationships will almost surely have to be adjusted as more research is undertaken. The first round of sound changes are inherited, as evidenced by identical changes in Ngorek, Merap's closest relative.

(84) **First Round (inherited from an immediate ancestor)**
nasal-voiced obstruent devoicing
*-d > *-n
*-b > *-p

(85) **Second Round**
stress shift to final syllable
off-gliding before velars
reverse umlaut

(86) **Third Round**
velar reduction/deletion
assimilation of final consonants to the color of preceding vowels
reduction of word-final voiceless consonants to glottal stop (after *a)
penultimate vowel deletion
vowel lengthening

16. A parallel change has been posited for Proto-Berawan (Burkhardt 2014, 2016)

(87) **Fourth Round**
schwa lowering
simplification of consonant clusters to voiceless sonorants

5. MERAP AND NGOREK. Kayan has been the topic of several studies, including basic wordlists (Barth 1910; Burns 1849; Douglas 1911; Rousseau 1974), historical descriptions (Blust 2002b), synchronic descriptions (Blust 1977; Cubit 1964; Clayre and Cubit 1974; Effendy 1989; Guerreiro 1983; Rousseau 1983; Soriente 2013), a dictionary (Effendy et al. 2006), and collections of texts (La'ing 1968; Rubenstein 1973). Ngorek, on the other hand, has received sparse attention. There is an early wordlist (Douglass 1912), which was updated and expanded by Blust (1974), and a more recent use of Murik data in a historical discussion on final glottal stop innovation in Kayanic languages (Blust 2002b). Additionally, Ngorek data can be found in the appendix of Soriente (2003) and Smith (2015). Soriente's dissertation is particularly useful, as it is the only source of data for Murik communities in Kalimantan. Smith (2017) also contains a large amount of data not only on Ngorek and Merap, but on several Kayan and Segai-Modang communities. With this material, it is possible to discuss the relatedness of Merap and Ngorek, while appealing to specific sound changes found in only these two Kayanic languages. All data for the subgrouping arguments below are from my own field notes.

5.1 REFLEXES OF WORD-FINAL VOICED OBSTRUENTS IN KAYANIC. Proto-Kayanic retained voicing in reflexes of PMP *-b, *-d, *-j, and *-g, as evidenced by reflexes of word-final voiced stops in modern languages, where *-b is reflected as *-p, -m, -v,* and *-w,* and *-d and *-j are reflected as *-l, -r, -n.* Most Kayanic languages either reflect the final voiced stops as nasals, or as other, nonnasal segments. Ngorek, however, has been viewed as unique among Kayanic languages in that it reflects *-b with a nonnasal, *p,* but *-d and *-j with a nasal, *n* (Blust 1974). The data presented in this paper show that Merap also reflects *b with *p* and *-d and *-j with *n.* Table 5 summarizes reflexes of final voiced obstruents in all Kayanic languages for which primary data are available.

5.2 REFLEXES OF NASAL-OBSTRUENT CLUSTERS IN KAYANIC LANGUAGES. Most Kayanic languages simplified nasal-obstruent clusters, and reflect them with simple voiced stops, thus, *mb > *b,* *nd > *d,* *nj > *j,* and *ŋg > *g.* Only

TABLE 5. KAYANIC REFLEXES OF WORD-FINAL VOICED OBSTRUENTS

PKAY	*-b	*-d
Long Naah Kayan	-m	-n
Data Dian Kayan	-m	-n
Modang	-w	-n
Gaai	-w	-l
Balui Liko Kayan	-v	-r
Busang	-v	-r
Bahau	-p	-l
Ngorek	-p	-n
Merap	-p	-n

two groups did not undergo this change, Ngorek and Merap. Rather, these two groups devoiced the obstruent, but maintained the cluster, thus *mb > *mp,* *nd > *nt,* and *ŋg > *ŋk.* See table 6.

TABLE 6. KAYANIC REFLEXES OF NASAL-OBSTRUENT CLUSTERS

PKAY	*mb	*nd	*nj	*ŋg
Long Naah Kayan	b	d	j	g
Data Dian Kayan	b	d	j	g
Balui Liko Kayan	b	d	j	g
Busang	b	d	j	g
Bahau	b	d	j	g
Modang	b	d	s	?
Gaai	b	d	c	?
Kelai	b	d	c	?
Ngorek	mp	nt	nc	ŋk
Merap	mp	nt	nc	ŋk

5.3 MERAP AND NGOREK SHARED LEXICON. Phonological evidence is the most important type of evidence for forming a subgrouping hypothesis. However, a short list of Murik-Merap lexical innovations can be constructed based on the words provided in this study. That list follows:

(88) PKAY *masak > PMUR *maru ‘cook’
PKAY *bahat > PMUR *ləmən ‘heavy’
PKAY *məgan > PMUR *təʔuh ‘dry’
PKAY *halah > PMUR *bəlawaŋ ‘empty’
PKAY *mitah/mindah > PMUR *nəŋaw ‘to wait’

Although the above list is short, it is important to keep in mind that despite the phonological changes that have taken place in Merap, both Merap and Ngorek are lexically conservative. Few lexical replacement innovations have taken place: Merap, Ngorek, and in fact, most Kayanic languages show the same tendency for slow lexical replacement. Ngorek and Merap share 90 percent of their vocabulary with each other, while Merap and Data Dian Kayan share 81 percent, and Merap and Modang share 57 percent.[17] Clearly, the lexicons of Merap and Ngorek have diverged only slightly, despite the numerous phonological changes.

5.4 MERAP AS AN ABERRANT DIALECT OF NGOREK. The phonological evidence for calling Merap an aberrant dialect of Ngorek is quite strong. There are no other languages in Borneo that reflect *-b with *p* but *-d with *n.* The typical pattern is devoicing, like Kenyah *-b > *p* and *-d > *t,* nasalization like Long Naah Kayan *-b > *m* and *-d > *n,* or a combination of lenition without nasalization such as Balui Liko Kayan *-b > *v* and *d > *r;* Ngaju Dayak *-b > *p* and *-d > *r;* or Bahau Kayan *-b > *p* and *-d > *l.* It is thus highly unlikely that both Ngorek and Merap show the changes *b > *p* and *-d >

17. Cognate percentages do not imply any type of subgrouping relationship, as clearly demonstrated in Blust (2000). Segai-Modang languages in general show a much higher tendency for lexical replacement than other Kayanic languages, but this does not mean that Murik-Merap is any more closely related to Kayan than it is to Segai-Modang. The percentages stated were calculated for Smith (2017) from a 200 item wordlist.

n as the result of chance or parallel sound change, since they are found nowhere else in Borneo. In addition, the devoicing of obstruents in nasal-obstruent clusters provides a second piece of evidence linking Merap and Ngorek. Although certain dialects of Highland Kenyah languages also devoiced obstruents in this environment (Lepo' Tau and Badeng, for example, from Smith 2015), the changes appear unrelated, since other dialects of Kenyah did not devoice. Such devoicing, because it is also found in Kenyah, is thus weaker evidence for subgrouping, but it is still rather uncommon. Combined with reflexes of *-b and *-d, the evidence for Ngorek-Merap is strong.

Murik and Merap thus form one subgroup of the much larger Kayanic group. Several publications have claimed that there is a special relationship between Kayanic and Kenyah (Hudson 1978; Soriente 2003, 2006b, 2008;, Dyen 1965:43), but this paper assumes that Kayan is not closely related to Kenyah, or any other language of North Sarawak (see Smith 2015 for a recent argument). Kayanic is internally diverse, and in addition to Murik-Merap include the phonologically conservative Kayan group and the aberrant Segai-Modang group. A hypothesis on Kayanic internal subgrouping, from Smith (2017) is reprinted as figure 1.

FIGURE 1. KAYANIC SUBGROUPING HYPOTHESIS

KAYANIC

Kayan-Murik-Merap

Kayan

Including the Kayan languages of the Baram river, the Rejang river, Balui Liko, Uma Juman), the Apo Kayan area (Data Dian), the Upper Mahakam (Busang and Bahau), and the upper Kapuas

Murik-Merap

Various widely dispersed languages including Ngorek in Sarawak, Pua' and Huang Bau in Kalimantan (Soriente 2003), and Merap

Segai-Modang

The languages of Berau Regency in Kalimantan, including Segai (Gaai), Punan Kelai, Modang dialects of the Wahau area in East Kutai, Kalimantan, and Long Gelat

The Murik-Merap subgroup includes Pua' and Huang Bau. Data on these two Murik dialects can be found in Soriente (2003), and they appear to be quite similar to the Ngorek of Long Semiyang in Sarawak. This is one of three primary groups in Kayanic, and the divisions are based largely on a set of sound changes (for Segai-Modang) and lexical innovations (for Kayan). For further arguments on Kayanic internal subgrouping, Smith (2017) provides a recent reference.

The immediate question, then, is why did Merap run wild with sound changes while Ngorek remained relatively unchanged? The Merap-Ngorek situation directly parallels the relationship of Sa'ban and Kelabit. Blust (2001) showed that, lexicostatistically, Sa'ban is nothing more than a Kelabit dialect. The phonology of Sa'ban, however, is drastically different from Kelabit, with a plethora of sound changes similar in type to those found in Merap. Merap and Sa'ban both shifted stress to the final syllable, where word-final stress wreaked havoc on the conservative phonologies of Proto-Kelabit as well as Proto-Ngorek-Merap. It is not within the scope of the present study to take this question further, that is, *why* did these languages shift stress to the final syllable in the first place? There are numerous possible explanations:

(i) There was a non-Austronesian language or group of languages with word-final stress whose speakers shifted to Proto-Kelabit and Proto-Ngorek.
(ii) An Austronesian language independently shifted stress to the final syllable and that change influenced other languages of the area.
(iii) There is an area of intense cultural and linguistic malleability in central Borneo, and social pressures favored stress shift in genetically unrelated but socially intertwined languages.

None of these hypotheses is testable, and we may never be able to do anything more than make an educated guess as to what *might* have happened. But the fact that Merap and Ngorek are so closely related but so different, and that in the same general area Sa'ban and Kelabit have a similar conflicting relationship, is tantalizing. What is more, Segai-Modang languages are also Kayanic and are more like Merap in their synchronic phonology than any other Kayanic language, but Merap is a Ngorek dialect, not a Segai-Modang dialect. There certainly does seem to be "something in the air" that has caused languages to recently change direction towards increased phonetic complexity as the result of strong word-final stress.

6. CONCLUSION. The languages of Borneo, despite decades of study, continue to surprise. Kalimantan in particular has been shrouded in mystery, as few historical linguistic works have focused on the languages of this area. Only recently has the veil been lifted, and true to the spirit of Borneo the languages of eastern Kalimantan present unique challenges in documentation, classification, and analysis as the result of riotous sound changes. This paper has addressed several topics. There is a linguistic area located approximately south of Sabah, in the area of northern Sarawak and North and East Kalimantan, where strong word-final stress has caused impressive changes to the phonologies of many of the languages spoken in the area. Sound changes that characterize this linguistic area are stress shift to the final syllable, penultimate vowel reduction/deletion, reverse umlaut, final consonant palatalization, word-initial consonant clusters, and voiceless sonorants. Languages that show such changes include North Sarawak languages like Sa'ban, Kiput and Berawan, and Kayanic languages including Merap, Modang, Gaai, and Kelai, and the Land Dayak language Hliboi Bidayuh. Merap stands out as particularly innovative, and this paper has given a wider audience its first detailed look into the phonology (synchronic and historical) of this highly aberrant dialect. Merap has 26 vowels, according to the analysis presented in this paper (a number that may very well change as more work is done on the language). These vowels include diphthongs and triphthongs. Historically, numerous changes have taken place that can account for modern Merap's vowel inventory, which is ultimately an expanded version of PMP's four-vowel system (*a, *i, *u, and *ə). Certain sound changes, including the reflexes of word-final voiced obstruents and homorganic nasal-obstruent clusters, provide fairly strong evidence that Merap is most closely related to Ngorek, and has only recently undergone the changes that make it so distinct.

APPENDIX. SUPPORTING DATA FOR ALL RECONSTRUCTIONS USED IN THIS PAPER.

Data Dian is a dialect of Kayan spoken in the Apo Kayan highlands of North Kalimantan, around the headwaters of the Kayan river and the Boh tributary of the Mahakam river. Busang is a dialect of Kayan spoken in the upper Mahakam area of East Kalimantan.

PKAY	Merap	Ngorek	Data Dian	Busang	Kelai
*alu ‘rice pestle’	lawʔ	lu	aloʔ	aloʔ	alaw
*anak ‘child’	nayə̲ʔ	anak	aneə̲k	anak	ənæk
*apid ‘twins’	kapayñ	(pit)	papin	apir	pəkpel
*apuy	paɰ̲	api	apuy	apuy	əpoy
*aruʔ ‘long’	raɰ̲	aru	aruʔ	aruʔ	(kjah)
*ata ‘water’	kataːʔ	ataʔ	ataʔ	ataʔ	(ŋuy)
*atay ‘liver’	taə̲	ate	atay	ate	ataə̲
*babuy ‘wild boar’	maɓaɰ̲	mabiʔ	bafuy	bavuy	awoy
*bahat ‘heavy’	man	ləmən	bahat	bahat	bəhæt
*bahiw ‘strong wind’	mayaɰ̲	bayu	–	buhuy	uhiw
*baliw ‘to become; transform’	malaɰ̲	–	–	baluy	blew
*basay ‘paddle’	mahaə̲	bahe	bəsay	bəse	psaə̲
*batu ‘stone’	mataw̲ʔ	batoʔ	batoʔ	batoʔ	(sutaw)
*baya ‘crocodile’	mayaːʔ	bayaʔ	bayaʔ	bayaʔ	wajɛʔ
*baʔ ‘mouth’	ɓaː	ba	baʔ	baʔ	(gulɛŋ)
*baʔik ‘short’	maʔayə̲ʔ	baʔek	baʔiə̲k	biʔik	(guiʔ)
*bəli ‘to buy’	mblayʔ	bəleʔ	bəleʔ	bəleʔ	(nlew)
*bi ‘to carry on the back’	ɓayʔ	beʔ	beʔ	beʔ	(bɛə̲n)
*buk ‘head hair’	ɓawə̲ʔ	bok	buk	buk	wok
*bulan ‘moon’	mblũə̲	bulan	bulan	bulan	ulun
*buluʔ ‘bamboo’	mbləw	bulu	buluʔ	buluʔ	ləwʔ
*buta ‘blind’	tuə̲ʔ	botaʔ	butaʔ	butaʔ	(pset)
*dahaʔ ‘blood’	laː	raa	dahaʔ	dahaʔ	lhæʔ
*dahulay ‘left side’	loɛ̲	ole	ulay	ule	(manlis)
*datah ‘flat; a plain’	lataːh	latah	datah	datah	(diʔ)
*daʔan ‘branch’	laʔãə̲	laʔan	daʔan	daʔan	lʔæn
*dəpa ‘a fathom’	paʔ	ləpaʔ	dəpaʔ	dəpaʔ	paʔ
*dindiŋ ‘wall’	rintiə̲	lentiŋ	lidiŋ	lidiŋ	diŋ
*duman ‘year’	lumũə̲	loman	duman	duman	(tʔon)
*əpat ‘four’	paːʔ	pat	pat	pat	pæt
*gatəl ‘itchy’	katan	katən	katən	katəl	ktal
*hakit ‘raft’	kaɰ̲ic	–	akit	akit	(stoŋ)
*halah ‘empty’	blawə̲ŋ	bəlawəŋ	halah	halah	(anæ cɛʔ)
*ibah ‘saliva’	ɓiə̲h	ebah	(ataʔ luran)	iwah	wih
*ikiŋ ‘pinky finger’	kiːə̲	–	(jahaŋaw uk)	ikiŋ	(kəŋget)
*inay ‘mother’	nɛy	nay	inay	ine	neə̲ʔ
*ini ‘this’	nɛyʔ	neʔ	ani	aniʔ	nɛh
*isi-n ‘flesh’	hɛyñ	ihin	sin	sin	sen
*isiʔ ‘snail’	hɛy	ihi	si	si	seə̲ʔ
*jalan ‘road’	ñalãə̲	jalan	alan	alan	gulæn
*jaʔa-n ‘chin’	jaʔaʔ	ñaʔãə̲	jaʔan	jəʔaʔ	cəʔæn
*jihi ‘house post’	ñɛy	ji	jahe?	jiheʔ	(skol)
*juʔ ‘far’	caɰ̲	su	suʔ	suʔ	(kjah)
*kajəl ‘dull’	kacan	kasən	kasən	kasəl	(khɒt)
*kapal ‘thick’	kapãə̲	kapan	kapan	kapal	(tmæŋ)
*kayaw ‘headhunting’	kayaɷ̲	kayo	kayaw	kayo	ŋñiw
*kaʔub ‘to lie prone’	ŋaʔawkʷ	–	ŋaʔum	–	kʔəwp
*kələb ‘turtle’	klap	kələp	kələm	kələβ	klaw
*kətuŋ/təhətuŋ ‘porcupine’	tawə̲	kətoŋ	kətuŋ	kətoŋ	təhtɒŋ
*kisiŋ ‘to laugh’	kihiə̲	–	kəsiə̲ŋ	kəsiŋ	lawʔ ŋan

PKAY	Merap	Ngorek	Data Dian	Busang	Kelai
*kitan ‘binturong’	tĩə̃	ketan	(bunin)	kitan	ktin
*kuju ‘heron’	kocowʔ	kuju (L)	kusoʔ	kusoʔ	(jɛʔ)
*kulat ‘mushroom’	kluaʔ	kolat	kulat	kulat	klut
*kulih ‘clouded leopard’	kluyh	koleh	kuleh	kuleh	kleh
*kuman ‘eat’	hmũə̃	koman	kuman	kuman	mun
*kutu ‘louse’	kotowʔ	kotoʔ	kutoʔ	kutoʔ	ptao
*laki ‘man’	lakayʔ	lakeʔ	lakeʔ	lakeʔ	maŋkay
*lamin ‘house; room’	lamayñ	amin	amin	amin	–
*laŋaw ‘a fly’	laŋao	laŋo	laŋaw	laŋo	lŋaw
*laŋit ‘sky’	laŋayc	laŋit	laŋit	laŋit	lŋet
*laʔip ‘shoulder’	laʔayc	laʔip	laʔip	liʔip	ləʔep
*laʔuŋ ‘back’	laʔawə	laʔoŋ	laʔuŋ	laʔuŋ	(koʔ)
*ləbid ‘to wring’	plaɓayñ	–	jələfin	kəlvir	–
*ləmbaw ‘tall’	mpao	mpo	baw	bo	kəmbao
*lindəm ‘dark’	ntiə̃m	lintəm	lidəm	lidəm	ŋəldam
*lubaŋ ‘hole’	loɓoyə	lobaŋ	lufeə̃ŋ	luvaŋ	guə̃ŋ
*lunuk ‘banyan tree’	lunuːə̃ʔ	lunok	lunuk	lunuk	lnuk
*mabuk ‘drunk; intoxicated’	maɓawə̃ʔ	mabok	mafuk	mavuk	(nuok)
*mandaŋ ‘to fly’	məntayə̃	mantaŋ	madeə̃ŋ	madaŋ	mñæŋ
*manuk ‘chicken’	manawə̃ʔ	manok	manuk	manuk	mnok
*maram ‘rotten’	marãə̃	maram	–	(butuŋ)	mñæm
*mariŋ ‘new’	marayə̃	mareŋ	mariə̃ŋ	mariŋ	məlhiŋ
*masak ‘to cook’	marau	maru	pakseə̃k	paksak	(mtəwk)
*matay ‘to die’	mataẹ	mate	matay	mate	(lwas)
*maʔud ‘wake up’	maʔawŋʷ	maʔun	kiniə̃ŋ	muʔur	mʔɒl
*məgan ‘to dry’	təʔəwh	toʔoh	məgeə̃ŋ	pakgaŋ	pəkkhuə̃ŋ
*məŋjuʔ ‘lift’	ncau	məncu	məjuʔ	juʔ	(gɒh/læʔ)
*miris ‘slice’	mɛrɛyh	ñereh	(mutun)	miri	(dɛə̃ʔ)
*mitah/mindah ‘to wait’	ŋao	nəŋo	pidah	(kave)	dah
*miʔaŋ ‘to split’	miʔiə̃	meʔaŋ	miʔeə̃ŋ	miaŋ	(tæʔ)
*m-urip ‘alive’	mpruyc	murip	murip	murip	(blɒm)
*ñahu ‘eagle’	ñau	ñau	ñihoʔ	ñhoʔ	ñahoʔ
*ñaŋgu ‘to fry’	ñaŋkawʔ	ñaŋkoʔ	ñagah	ñagah	nkæh
*nəpujuk ‘to jump’	ləʔɟəwə̃ʔ	nəpəjok	napjuk	nəpujoʔ	napcɒh
*ñĩbaw ‘shallow’	lɛɓɛo	ñebo	ñifaw	ñivo	–
*ñĩpa ‘snake’	piə̃ʔ	ñepaʔ	ñipaʔ	ñipaʔ	pɛʔ
*ñĩpis ‘thin’	lɛpɛyh	ñepeh	ñipih	ñipi	pes
*ñĩluʔ/ñəluʔ ‘to swallow’	ñəliw	ñəlu	ñəluʔ	ñəluʔ	(nal)
*nubus ‘to plant; sow’	loɓowh	(toŋkan)	tufuh	nuvu	wus
*pahu ‘grasshopper’	pau	pauʔ	pahoʔ	pahoʔ	phau
*panaw ‘to walk’	panao	pano	panaw	pano	pnæw
*paray ‘field rice’	paraẹ	pare	paray	pare	plaẹ
*paʔiʔ ‘bitter’	paʔai	paʔi	paʔiʔ	paʔiʔ	(paʔet)
*pərəs ‘sick; in pain’	prah	pərah	pərah	pərah	(ak)
*piliʔ ‘choose’	mblɛy	mili	miliʔ	miliʔ	leh
*pulut ‘sap; sticky substance’	pləwkʷ	pulut	pulut	pulut	ploə̃t
*puluʔ ‘ten’	pləw	pulu	puluʔ	puluʔ	(suaŋ)
*pusəd ‘navel’	puhuə̃n	pohən	(ubut)	(ubut)	(guboʔ)
*put ‘blowpipe’	pawkʷ	put	hiput	hamput	poə̃t
*puti ‘banana’	toyʔ	poteʔ	puteʔ	puteʔ	ptay
*putiʔ ‘white’	tuy	puti	putiʔ	putiʔ	(maslɛt)
*puʔun ‘base’	puʔowŋʷ	puʔun	paʔun	puʔun	–
*saləŋ ‘earthworm’	halaŋ	aləŋ	haləŋ	haləŋ	(kəlgæt)
*siap ‘chicken’	hɛaʔ	yap	hiap	həñap	jip
*siku ‘elbow’	kiwʔ	ikoʔ	hikun	hikoʔ	(paŋgok)
*siŋət ‘bee; to sting’	hŋiə̃t	iŋət	hiŋət	hiŋət	ŋat
*siran ‘when’	bəhrĩə̃	meran	hiran	hiran	(dao moh)
*sulu-n ‘finger nail’	hləwŋʷ	(ulok)	hulun	hulun	sloə̃n

PKAY	Merap	Ngorek	Data Dian	Busang	Kelai
*takut ‘afraid’	takawkʷ	takut	takut	takut	tkuə̯t
*taʔas ‘a traditional skirt’	taʔaːh	taʔah	taʔah	taʔah	khas
*taʔi ‘excrement’	taʔayʔ	taʔeʔ	taʔeʔ	taʔeʔ	–
*təbu ‘sugarcane’	təɓawʔ	təboʔ	təfoʔ	təvoʔ	(kuaw)
*təkəjət ‘surprised’	ƒat	(məlak)	nakjək	təkəjət	(hawan)
*təkuk ‘nape’	kəkawə̯ʔ	təkok	təkuk	təkuk	patok
*təlaʔus ‘the barking deer’	klaʔawh̯	təlaʔoh	təlaʔuh	təlaʔu	ɒs
*təlis ‘squirrel’	klayh	təleh	təlih	təli	(liə̯ŋ)
*tiruh ‘sleep’	tɛrɛwh	teroh	(tuduʔ)	–	(dəwʔ)
*tiuŋ/kiuŋ ‘myna bird’	kiawə̯	kioŋ	kiuŋ	tiuŋ	kjɒŋ
*tu ‘ghost’	tawʔ	toʔ	toʔ	toʔ	(wəlguə̯n)
*tujuʔ ‘seven’	tosow	tusu	tusuʔ	tusuʔ	cuʔ
*tulad ‘animal	klũə̯̃	tolan	tulan	tular	(dakot)
*tutuŋ ‘to be aflame’	tutuə̯	totoŋ	tutuŋ	tutuŋ	təwŋ
*ujan ‘rain’	cũə̯̃	usan	usan	usan	ciə̯n
*uləd ‘worm; maggot’	luə̯n	ulən	ulən	ular	laɪ̯
*uru ‘grass’	rəwʔ	oroʔ	uroʔ	uroʔ	law
*uruŋ ‘nose’	ruə̯	oroŋ	uruŋ	uruŋ	yuŋ
*utaʔ ‘vomit’	toa	nuta	nutaʔ	nutaʔ	tuʔ
*utək ‘brain’	tuə̯k	utək	otak	utək	tok

REFERENCES

Adelaar, K. Alexander. 1995. Borneo as a cross-roads for comparative Austronesian linguistics. In *The Austronesians: Historical and comparative perspectives*, ed. by Peter Bellwood, James J. Fox, and Darrell Tryon, 75–95. Canberra: Department of Anthropology, Research School of Pacific and Asian Studies, The Australian National University

Barth, J. P. J. 1910. *Boesangsch–Nederlandsch woordenboek.* Batavia: Government Printer.

Blust, Robert. 1974. A Murik vocabulary, with a note on the linguistic position of Murik. In *The peoples of central Borneo. Special issue of the Sarawak Museum Journal*, 22.43 (NS), ed. by Jérôme Rousseau, 153–89.

———. 1977. Sketches of the morphology and phonology of Bornean languages 1: Uma Juman (Kayan). *Papers in Bornean and western Austronesian languages* no. 2, 7–122. Canberra: Pacific Linguistics.

———. 2000. Why lexicostatistics doesn’t work: The “universal constant” hypothesis and the Austronesian languages. In *Time depth and historical linguistics,* vol. 2: *Papers in the prehistory of languages,* ed. by C. Renfrew, A. McMahon, and L. Trask, 311–31. Cambridge: The McDonald Institute for Archaeological Research.

———. 2001. Language, dialect and riotous sound change: The case of Sa’ban. In *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society*, ed. by Graham W. Thurgood, 249–359. Tempe: Program for Southeast Asian Studies, Arizona State University.

———. 2002a. Kiput historical phonology. *Oceanic Linguistics* 41:364–418.

———. 2002b. Formalism or phoneyism? The history of Kayan final glottal stop. In *Between worlds: Linguistic papers in memory of David John Prentice*, ed. by K. Alexander Adelaar and Robert Blust, 29–37. Canberra: Pacific Linguistics.

———. 2003. *A short morphology, phonology and vocabulary of Kiput, Sarawak.* Canberra: Pacific Linguistics.

———. 2007. Òma Lóngh historical phonology. *Oceanic Linguistics* 46:1–53.

———. n.d. Field notes on 41 languages of northern and central Sarawak.

Blust, Robert, and Alexander D. Smith. 2014. *Bibliography of the languages of Borneo (and Madagascar)*. Phillips: Borneo Research Council.

Burns, Robert. 1849. A vocabulary of the Kayan language of the northwest of Borneo. *Journal of the Indian Archipelago and Eastern Asia* 3:182–92.

Burkhardt, J.M. 2014. The reconstruction of the phonology of Proto-Berawan. PhD. diss., Johann-Wolfgang-Goethe-University.

———. 2016. How did Long Terawan Berawan develop sixteen vowel phonemes? *Oceanic Lingistics* 55:588–619.

Clayre, Beatrice. 1972. A preliminary comparative study of the Lun Bawang (Murut) and Sa'ban languages of Sarawak. *The Sarawak Museum Journal.* 20 (40-41 NS):319–29.

———. 1992. Sa'ban: A case of language change. In *Shifting patterns of language use in Borneo: Papers from the Second Bi-Ennial International Conference, Kota Kinabalu, Sabah, Malaysia, July 1992*, ed. by Peter W. Martin, 209–25. Borneo Research Council Proceedings Series, vol 3. Williamsburg, VA: Department of Anthropology, The College of William and Mary.

Clayre, Beatrice, and L. Cubit. 1974. An outline of Kayan grammar. *Sarawak Museum Journal* 22 (43 NS):43–91.

Cubit, L. E. 1964. Kayan phonemics. *Bijdragen tot de Taal-, Land- en Volkenkunde* 120(4):409–23.

Douglas, R. S. 1911. A comparative vocabulary of the Kayan, Kenyah, and Kelabit dialects. *Sarawak Museum Journal* 1(1):75–119.

Dyen, Isidore. 1965. *A lexicostatistical classification of the Austronesian languages*. Indiana University Publications in Anthropology and Linguistics, and Memoir 19 of the International Journal of American Linguistics. Baltimore: The Waverly Press.

Effendy, Chairil. 1989. *Morfologi Bahasa Kayan.* Pontianak: Departemen Pendidikan dan Kebudayaan: Universitas Tanjungpura, Falkutas Keguruan dan Ilmu Pendidikan.

Effendy, Chairil, A. R. Muzammil, Firman Susilo, Dedy Ari Aspar, Agus Syahrani, Hangga Dwitika, and Ediyanto. 2006. *Kamus Bahasa Kayan–Bahasa Indonesia.* Pondianak: STAIN Pontianak Press untuk Pusat Pendidikan dan Kebudayaan, Universitas Tanjungpura.

Guerreiro, Antonio J. 1983. A note on pronouns in the Long Gelat and Busang languages (Upper Mahakam, Kalimantan Timur). *Borneo Research Bulletin* 15(2):98–105.

———. 1989. Éntités, rhétorique et intention dans le discours ritual Modang Wehèa (Bornéo). In *Anthropologie de la prière: Rites oraux en asie du Sud-Est*, ed. by Stephen C. Headley, 89–124. Paris: Études du Centre de l'Asie du Sud-Est.

———. 1996. Homophony, sound changes and dialectal variation in some central Bornean languages. *Mon-Khmer Studies.* 25:205–26. (Special volume dedicated to Professor André-Georges Haudricourt.)

Hudson, Alfred B. 1978. Linguistic relations among Bornean peoples with special reference to Sarawak: An interim report. *Studies in Third World Societies* 3:1–44.

La'ing, Joseph Anyie. 1968. *Na'an balei lung Kayan* [Some Kayan stories]. Kuching: Borneo Literature Bureau.

Revel-Macdonald, Nicole. 1982. Synchronical description at the phonetic and syllabic level of Modang (Kalimantan Timur) in contrast to Kenyah, Kayan, and Palawan (Philippines). In *Papers from the Third International Conference on Austronesian Linguistics*, ed. by Amran Halim, Lois Carrington, and Stephen A. Wurm, 321–31. Canberra: Pacific Linguistics.

Rensch, Calvin R., Carolyn M. Rensch, Jonas Noeb, and Robert Sulis Ridu. 2012. *The Bidayuh language: Yesterday, today, and tomorrow* (revised and expanded). SIL E-Books 33. Kuching: SIL International.

Rousseau, Jérôme. 1974. A vocabulary of Baluy Kayan. In *The peoples of central Borneo. Sarawak Museum Journal* 22.43 (NS): 93-152 (special issue).
———. 1983. Kayan personal names. *Sarawak Museum Journal* 32:53 (NS), ed. by Jérôme Rousseau, 251–72.
Rubenstein, Carol. 1973. *Poems of indigenous peoples of Sarawak: Some of the songs and chants*. 2 vols. Kuching: Sarawak Museum.
Smith, Alexander D. 2015. On the classification of Kenyah and Kayanic languages. *Oceanic Linguistics* 53:333–57.
———. 2017. The languages of Borneo: A comprehensive classification. PhD diss., University of Hawai‘i at Mānoa.
Soriente, Antonia. 2003. A classification of the Kenyah languages in East Kalimantan and Sarawak. Ph.D. diss., Universiti Kebangsaan Malaysia.
———. 2006a. *Mencalèny and Usung Bayung Marang: A collection of Kenyah stories in the Òma Lóngh and Lebu' Kulit languages*. Jakarta: Atma Jaya University Press.
———. 2006b. Uma' Kulit: A Kenyah or Kayan language? Linguistic classifications and local epistemology. *Linguistik Indonesia* 24(1):71–81.
———. 2008. The classification of Kenyah languages: A preliminary assessment. In *SEALS XIV(2): Papers from the 14th meeting of the Southeast Asian Linguistics Society (2004)*, ed. by Wilaiwan Khanittana and Paul Sidwell, 49–62. Canberra: Pacific Linguistics.
———. 2013. Undergoer voice in Borneo Penan, Punan, Kenyah, and Kayan languages. In *Voice alternations in Austronesian languages*, ed. by K. Alexander Adelaar. *NUSA* 54:175–203.
———. 2015. Mbraa: A Modang-Bahau language? Paper presented at 13-ICAL, Academia Sinica, Taipei, Taiwan, July 18–31.
Thurgood, Graham. 1996. Language contact and the directionality of internal drift: The development of tones and registers in Chamic. *Language* 71(1):1–31.
———. 1999. *From ancient Cham to modern dialects: Two thousand years of language contact and change, with an appendix of Chamic reconstructions and loanwords*. Oceanic Linguistics Special Publication, no. 28. Honolulu: University of Hawai‘i Press.

smithad@hawaii.edu