

Southwest Sabah Revisited

Jason William Lobel

UNIVERSITY OF HAWAII AT MĀNOA

This paper presents the results of a new innovation-based subgrouping argument for the Southwest Sabah languages of northern Borneo, including over 60 languages in Sabah, northern Sarawak, Brunei, and northern Kalimantan Timur. Data for many of these languages have never appeared in the literature before, and few of them have been included in previous studies except those based on lexicostatistics and intelligibility testing. In contrast with previous works, the current study is based on phonological and functor innovations. Discrepancies between the findings of the current study and previous studies are explained as the effect of contact and borrowing, which lexicostatistics and intelligibility testing cannot account for.

1. INTRODUCTION.¹ The languages of northern Borneo have been the subject of scholarly attention since at least Beech (1908) and Ray (1913), yet, as Blust points out, “the subgrouping of the languages of Sabah remains rudimentary” (1998:32) and “published data on the Sabahan languages remain discouragingly sparse” (1998:34). Moreover, with regard to subgrouping hypotheses that have been put forth regarding these languages, Blust notes that “all . . . are offered simply as conclusions, without support of any kind (viz. evidence of exclusively shared innovations in the case of qualitative arguments, explicit information on cognate decisions in the case of lexicostatistical arguments)” (1998:32). In fact, no subgrouping argument about the internal structure of the Southwest Sabah (SWSAB)² subgroup has ever been presented based on phonological and morphological

1. Thanks are due to my hundreds of informants in Sabah, Sarawak, Brunei, and Kalimantan Timur; to the staff of SIL Malaysia; to Paul Kroeger, Yabit Alas, and Robert Blust; to Sanen Marshall and Mika Okushima for indispensable discussions on north Bornean history; and especially to Dave Moody for access to the SIL survey data that enabled me to refine the itinerary for my longest fieldwork trip to north Borneo from 2010 to 2011. Special thanks also to two anonymous reviewers, whose feedback helped improve the final version of this paper. All mistakes are mine alone.

Abbreviations of the names of subgroups are given in the text where the subgroup is first mentioned. The abbreviation for the relevant protolanguage simply adds P to the subgroup abbreviation: thus GMUR is given as the abbreviation of the Greater Murutic subgroup at the end of section 1, and, therefore, PGMUR refers to Proto-Greater Murutic.

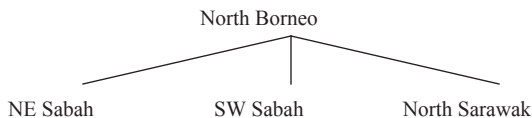
2. The name “Southwest Sabah” was coined by Blust (2010), and is a slightly more appropriate name for what King (1984) calls the “Bornean Stock,” reflecting the hypothesis that the homeland of these languages is southwestern Sabah, where the largest number of representatives of the primary branches of the component subgroups can be found (Greater Murutic, represented by the Tatana and Papar languages in Kuala Penyu, and by various conservative Murutic languages just uphill in Keningau and Tenom; and Greater Dusunic, represented by Sabah Bisaya and a number of varieties of Dusun and Kadazan in Beaufort, Membakut, Kimanis, and Papar).

innovations. Blust himself (1998), lacking any significant amount of data for most of its individual member languages, only discusses the Southwest Sabah subgroup in the context of its relationship to languages elsewhere in Borneo and the Philippines, without consideration of the internal structure of the Southwest Sabah subgroup. More recently, Blust (2010:62) simply argues that the immediate external relationships of Southwest Sabah are to Northeast Sabah and North Sarawak (cf. figure 1), and that “NE Sabah and SW Sabah may share a Sabahan node not far below North Borneo, with NE Sabah and SW Sabah branching off from it shortly after North Sarawak had separated from them.”

Utilizing phonological and functor innovations, this paper will demonstrate that, although the traditional terms “Dusunic” and “Murutic” can still be used to define valid subgroups, their membership differs from that indicated by the less reliable methods used in previous studies, the most comprehensive of which are the various papers in King and King (1984), on whose subgroupings the *Ethnologue* relies (Lewis 2009 in its most recent edition).³ First, Lotud, classified by King and King (1984) as Dusunic, should be subgrouped with Sabah Bisaya, Limbang Bisaya, and Brunei Dusun (the last two not in King and King 1984) in a small “Bisaya-Lotud” subgroup that is distinct from the core Dusunic group. Second, two other languages—Papar and Tatana—were classified as Dusunic, whereas the comparative evidence suggests that they are more closely related to the Murutic languages. Similarly, the Gana language has been listed in Lewis (2009) as Dusunic, even though Smith (1984) classified it as Murutic, and in this case the data and analysis presented here also support an assignment to Murutic. Furthermore, three languages in the Indonesian province of Kalimantan Timur near the border with Sabah—Abai Sembuak, Abai Tubu,⁴ and Bulusu—are also Murutic, although they were outside the geographically based scope of King and King (1984). Although Bulusu or “Burusu” was misclassified as part of the Rejang-Sajau branch of “North Borneo,” and the two Abai varieties were erroneously listed as dialects of “Putoh,” an alternate name in Kalimantan Timur for Lun Bawang (~ Lun Dayeh) (Lewis 2009), all three are clearly Murutic languages.

The conclusion of the present study, as will be discussed in Section 4, is that four lower-level subgroups can be defined based on phonological and morphological innovations: Dusunic (DUS), Bisaya-Lotud (BISLO), Paitanic (PAIT), and Murutic (MUR). The first three

FIGURE 1. NORTH BORNEO (Blust 2010:62)



3. However, it must be noted that the *Ethnologue* listing of linguistic relationships has acquired a number of errors over the years, most problematic being the listing of (1) Yakan (a Sama-Bajaw language) as a member of the Paitanic subgroup; (2) Bonggi (a member of Blust’s Northeast Sabah and the current author’s Molbog-Bonggi subgroup) as a member of the Bisaya branch of the Dusunic subgroup; and (3) Bulusu/Burusu as a Rejang-Sajau language.

4. These two languages should not be confused with Abai Sungai, a Paitanic language spoken in eastern Sabah. Abai Sungai is spoken by a tribe called “Sungai” located in the town of Abai; Abai Sembuak and Abai Tubu, on the other hand, are spoken by a tribe called “Abai” residing in the villages of Sembuak and Tubu, respectively.

of these can be grouped together in a Greater Dusunic (GDUS) subgroup, while the Murutic subgroup, Tatana, and Papar form three branches of a Greater Murutic (GMUR) subgroup.

1.1 METHODOLOGY AND PREVIOUS STUDIES. Previously, the only study of the internal relationships of the Southwest Sabah languages was King and King (1984), which was based on wordlists representing hundreds of language communities in Sabah surveyed by various members of the Summer Institute of Linguistics (SIL) in Sabah. While impressive in its scope and groundbreaking at the time, it is also highly problematic for two major reasons. First, it was limited geographically to the Malaysian state of Sabah, even though closely related languages are also located in nearby northern Sarawak (Malaysia), northern Kalimantan Timur (Indonesia), and Brunei. The only languages from outside Sabah that were included were those for which speakers were found in Sabah. King and King (1984), therefore, failed to include Limbang Bisaya in northern Sarawak, the Dusun dialects in Brunei, and several Murutic languages spoken on the Indonesian side of the Sabah-Kalimantan Timur border.⁵ In all fairness, King and King (1984) was a survey of languages located in Sabah, and did not claim to be an all-inclusive study of the Southwest Sabah subgroup. As such, while it did not include Southwest Sabah languages spoken outside Sabah, it did include more distantly related languages spoken in Sabah, like Iranun, Sama-Bajaw, Idaan, Bonggi, Bugis, and Tausug (or “Suluk” as it is called in Sabah), as shown in table 1.⁶ Furthermore, the languages—especially those in Kalimantan Timur—were much less accessible decades ago than they are today, since many communities have now moved (or been moved) downriver closer to more central towns.

The second and more serious problem with King and King (1984) is that neither of the methodologies that it utilized—lexicostatistics and intelligibility testing—is known to be a reliable basis for subgrouping. Lexicostatistics does not differentiate between retentions, borrowings, and shared innovations, and has been largely discredited by a large number of linguists—including many Austronesianists such as Grace (1964,

**TABLE 1. THE “NORTHWEST AUSTRONESIAN SUPERSTOCK”
(after Smith 1984:41ff, adapted by Blust 2010:53)**

1. Lundayeh language	8. Bajaw family
2. Bonggi language	9. Bornean stock
3. Iranun language	9.1 Tidung
4. Suluk language	9.2 Paitanic family
5. Bugis language	9.3 Dusunic family
6. Ida'an language	9.4 Murutic family
7. Malayic language	

5. Note that use of the term “Murut” appears to be limited to the Sabah side of the border, while on the Indonesian side, speakers of Murutic languages are included in the generic term “Dayak,” which can be applied to any of the majority of ethnolinguistic groups in Kalimantan Timur, regardless of their linguistic affiliation. Note that the term “Murut” should not be confused with the Lun Dayeh (~ Lun Bawang) language, which is often called “Murut” in Brunei but belongs to the North Sarawak subgroup (Blust 2010).

6. Smith’s spellings “Banggi” and “Tidong” in table 1 have been updated to “Bonggi” and “Tidung” to more accurately reflect the native speaker pronunciations of these language names.

1992), Blust (1981, 2000), Ross (1991, 2005), and Pawley (1999)—as a basis for subgrouping arguments.⁷

Intelligibility testing, on the other hand, has never been demonstrated to provide reliable subgrouping results, nor has it been widely accepted and/or utilized for this purpose; instead, its use has been limited to studies by certain SIL members. This method was developed to aid the SIL in determining where two or more communities could be served by a single set of literacy materials and other vernacular publications, or whether separate materials would be necessary. For its intended purpose, the importance of intelligibility testing should not be underestimated, since incorrect decisions about how well various communities understand each other could result in decades of extra work by linguists, translators, educators, supervisors, and proofreaders, and large amounts of additional money spent. For any organization with limited human and financial resources, this ability is indescribably valuable. However, to use intelligibility testing as a basis for subgrouping is highly problematic, as there are a number of uncontrollable factors that can affect whether speakers of one speech variety may understand speakers of another, including geographic proximity, trade, political or military power of one community over another, historical lexical borrowing, and even the relative conservativeness of the lexicon and phonology of each language. At best, intelligibility testing, like lexicostatistics, can only be used to provide a numerical basis for describing the lexical similarities between languages, with the caveat that *similarity* is not a generally accepted basis on which to subgroup languages. As such, while it is undoubtedly a useful tool for applied linguistics, intelligibility has little if any reliable application for historical-comparative linguistics.

In contrast to King and King (1984), the methodology in the current study consists of the analysis of phonological and functor innovations.

While not completely problem-free, phonological innovation is one of the most widely accepted bases for subgrouping. Care must still be taken to ensure that shared phonological correspondences are really the result of exclusively shared innovations, and that it is not the case when one language has adopted pronunciation patterns from another more influential or prestigious language, or, as Blust (1992) demonstrated, when it is the result of heavy lexical borrowing from a different language with different reflexes of one or more protophonemes.⁸ Still, in response to the question of whether it can be determined that there is evidence that two languages form a group as against a third, Blust (1998:31) observes “more than a century of scholarship has shown that reliable answers ... can only be based on exclusively shared innovations.”

On the other hand, functors are a less-commonly utilized basis for subgrouping, not because of any known problem inherent in their use, but because the focus of most large-scale language surveys has been almost exclusively on eliciting lists of individual lexical

7. The arguments need not be repeated here, but Blust (2000:327), for example, summarizes the argument against the use of lexicostatistics for subgrouping Austronesian languages as follows: “we cannot tell when it gives valid results and when it does not. ... Since we now know that languages vary widely in retention rate of basic vocabulary over lengthy intervals of time, lexicostatistics must be seen as an unreliable foundation for subgrouping hypotheses that are not independently confirmed by the evidence of exclusively shared innovations.”

8. As in places such as northeastern Mindanao, where languages belonging to three separate subgroups (Bisayan, Manobo, and Mamanwa) share two phonological innovations (*y > /j/ and *l > /y/), which are not otherwise found in the closest relatives of each of these languages.

items, often for lexicostatistical comparison and/or as a basis of searching for phonological correspondences in basic vocabulary. Unfortunately, with a few noteworthy exceptions (for example, McFarland 1974, Zorc 1974 and 1977, Yamada and Tsuchida 1983, and my own fieldwork), little if any attention has been paid to the systematic collection of complete inventories of functors (like pronouns, case markers, demonstratives, adverbial particles, basic verbal morphology, negators, and adverbs of time) across a wide variety of languages.⁹ Importantly, however, without such sets being elicited in sentence context, complete and accurate elicitation of functors is rendered virtually impossible, as is functor analysis itself.

In its quantitative incarnation, functor analysis has been championed by Philippinists such as Zorc (1977, 1978) and McFarland (1974), based on the assumption that closed grammatical sets are the backbone of a language and are much less prone (although not completely immune) to borrowing than open-class lexicon. Observing that “a language is more readily defined by its grammar than by its lexicon,” Zorc (1978:510) also points out that functors have “obvious importance within any given speech variety,” “high text frequency,” and a “tendency towards stability and a low rate of replacement.” Although few attested examples of this had been documented in the Philippines at that time (Wolff 1967 being one of the earliest), a number of examples have since emerged of languages that have borrowed lexicon heavily from their neighbors while retaining their substrata largely intact—for example, Inati (Pennoyer 1986–87), Tiruray (Blust 1992), Utudnon/Baybayanon (Rubino 2005), and Manide and Inagta Alabat (Lobel 2010)—in addition to other languages known to have similar histories of heavy lexical borrowing, such as Mamanwa in Mindanao, Kinamiging on Camiguin Island, Remontado Dumagat in central Luzon, Kinabalian in Southern Leyte, and Kasigurarinin in northeastern Luzon. As will be demonstrated in this paper, the functor evidence is, in fact, very important and agrees almost perfectly with the phonological evidence in Sabah.

The evidence presented here will be limited to phonological and functor innovations. A comprehensive, qualitative study of the lexicons of these 60 or so speech varieties is beyond the scope of this paper, as it would require many times more space, and would necessarily include a thorough assessment of lexical strata and the direction and sources of borrowing, both from each other and from external languages such as various dialects of Malay and Philippine languages like Tausug, similar to works published by Burton (1996) and Pallesen (1985) for the languages of the southern Philippines. Also, this paper will concentrate on the defining features of the major branches of the Southwest Sabah macrogroup, while the lower-level relationships of the individual languages will be dealt with in a future paper.

9. In fact, in many of the SIL wordlists elicited in Sabah, the handful of sentences at the end of the earliest version of the elicitation list were either only partially filled out or completely left blank, and later versions of the list omitted the sentences altogether. Elsewhere, where pronouns, deictics, negatives, and so on are included in wordlists taken out of sentence context, it is difficult to determine which form the researcher has elicited, for example, whether a pronoun response in a Philippine language is a nominative, genitive, or oblique pronoun.

1.2 A NEW SUBGROUPING PROPOSAL. Based on phonological and functor innovations, it is argued that a Dusunic language can be defined as one that descends from a protolanguage in which:

1. The innovated pronouns **ya* ‘1EXCL.GEN’, **dəJə(n?)*¹⁰ ‘1SG.OBL’, and **daJay* ‘1EXCL.OBL’, replaced PSWSAB **mai*, **d[i]-ak(əi)(n?)*, and **d[i]-am(əi)(n?)*, respectively, which themselves reflected Proto–Malayo–Polynesian (PMP) **mami*, PPHNB¹¹ **d[i]-akə(nq)*, and PPHNB **d[i]-amə(nq)*, respectively (cf. 2.1).
2. PMP/PSWSAB **R* became **w* in word final position (cf. 2.2).
3. PMP/PSWSAB **R* became **g* before **i*, but became **h* in other intervocalic environments (except after **ə*, where it had already shifted to **g* in PSWSAB) (cf. 2.2).
4. The initial consonant of the PMP/PSWSAB adjectival and abilitative/accidental prefix **ma-* (but not **ma[R]-* or **maN-*) was dropped when prefixed to consonant-initial roots (cf. 2.3).
5. PMP/PSWSAB **b-* split to **b* and **w* (cf. 2.4).
6. PSWSAB **d-* (< PMP **d*, **j*, and **z*) split to **d* and **r* (cf. 2.5).
7. Schwa was retained as **ə* in nonfinal syllables unless the following syllable contained **a*, in which case it became **a* (cf. 2.6).
8. Penultimate **a* was neutralized to **ə* in PSWSAB **aku* ‘1SG.NOM’, **ə-kai* ‘1EXCL.NOM’ (< PMP **kami*), and **ə-kau* ‘2PL.NOM’ (< PMP **kamu*) (and its short form, **kau*, as well as **takau*, ‘1INCL.PL.NOM’ of which the final sequence **-kau* originates from the 2PL.NOM pronoun **kau*) (cf. 2.7).
9. Although not an innovation, it is worth noting that PDUS reflected **nu* ‘2SG.GEN’ instead of the more widespread form **mu* that is reflected in Bisaya-Lotud, Paitanic, and Murutic languages. The only other languages in northern Borneo to reflect **nu* are Bonggi, and a Paitanic language known as Dusun Puawang,¹² which may have borrowed it from contact with Dusunic languages. Within Dusunic, only Dumpas reflects **mu*, but this is likely a loan resulting from its intense contact with Sungai Beluran (a Paitanic language) and other Paitanic languages, Tausug, Tidung Bangawong, and Tidung Sambal, and other languages that reflect **mu* (cf. section 2.8).

Note that all of these innovations except 1 and 9 also apply to the Bisaya-Lotud and Paitanic subgroups, which, together with the Dusunic subgroup, likely form three branches of a larger subgroup referred to herein as “Greater Dusunic” (cf. 4.2).

10. The convention **J* represents an unexplained correspondence set found only in these two innovated pronouns, and nowhere else in the lexicon or functors. **J* is reflected as the affricate /*j/* in Dusun Membakut and Dusun Kimanis; as /*h/* in Dusun Papar, Kadazan Ovai Kambizaan, Rungus, Dusun Tambunan, Dusun Kiulu, Dusun Tindal, and Kujau; and as /*g/* in Dusun Tambahlu, Dusun Tamparuli, Dumpas, Mangkak, Minokok, Sonsogon, Tinagas, Talantang, Kadazan Penampang, Dusun Klias, and Kimaragang Dusun.

11. Reconstructions attributed to Proto-Philippines and Northern Borneo (PPHNB) may, in fact, turn out to be reconstructible to PMP or Proto-Western Malayo-Polynesian (PWMP), but lacking evidence outside of the Philippines and northern Borneo, a more cautious approach is taken here, since the forms are only found in these two geographical areas. Further study will hopefully determine whether the lowest level protolanguage shared by the languages of the Philippines with the languages of northern Borneo is PMP, PWMP, or some other yet-unidentified lower-level node within Malayo-Polynesian.

12. A Paitanic language, in spite of being spoken by a population whose members self-identify as “Dusun.”

On the other hand, a Murutic language is defined as one that descends from a proto-language in which:

1. PMP/PSWSAB *R became *h before any vowel (except after *ə, where it had already shifted to *g in PSWSAB), then PGMUR *h later shifted to zero in all daughter languages except Papar (cf. 2.2 and section 3).
2. PMP/PSWSAB *R became *g word-finally (cf. Sections 2.2 and section 3).
3. PMP/PSWSAB *aw and *ay shifted to *ow and *oy, respectively (cf. 3.1).
4. PMP/PSWSAB *iw became *uy (cf. 3.2).
5. PGMUR *g- > zero after adjectival prefix *ma- (cf. 3.3).
6. PMP/PSWSAB *ə became *a in nonfinal syllables, except in the environment *_Cə, where it is reflected as /o/ (cf. 3.4).

The absence of innovations 4 and 5 in Tatana, and the absence of the *h > Ø shift in Papar, indicate that Tatana and Papar belong to a node above the core Murutic languages, forming two of three branches of a “Greater Murutic” subgroup (cf. 4.1).¹³

It should be pointed out that only a few innovations link the Greater Murutic and Greater Dusunic subgroups together, none of which is especially high quality, as all are also found in a number of other subgroups:

1. PMP *h > PSWSAB Ø. PMP *h is only retained in the Batanic/Bashiic, Manide-Alabat, and Central Philippine subgroups, and reflected in “a handful of other languages,” including Kayan, Malay, and Soboyo (Blust 2009:579).
2. PMP *a > PSWSAB *ə / _# (also found in some dialects of Peninsular Malay, standard Javanese, Gorontalo, Buol, and the Idaanic languages; as this is also found in Idaanic, it could possibly be an areal feature in Sabah or northern Borneo).
3. PMP *R > PSWSAB *h / (aiu)_ (aəu): the *R > /h/ shift is also found in Samal, Kayan, Ngaju Dayak, and Kove (Blust 2009:582), as well as in Gorontalo and Ponosakan (both via *g), and in some Central and Western Manobo languages (via *g > *y).
4. PMP *R > PSWSAB *g / ə as the result of fortition after schwa, which is “extra-short, and cannot hold stress without compensatory lengthening of a following pre-vocalic consonant” (Blust 2009:245).
5. PMP *-m- > Ø in PSWSAB reflexes of the PMP pronoun forms *kami ‘1EXCL.NOM’, *mami ‘1EXCL.GEN’, and *kamu ‘2PL.NOM’, an innovation that is also found in some members of the Manobo, Palawan, and Molbog-Bonggi subgroups, and is generally widespread in Malayo-Polynesian languages (Dyen 1974).
6. Reduction of most PMP consonant clusters to either singletons or prenasalized clusters, also characteristic of various subgroups in Mindanao (for example, Danao, Subanen) and northern Sulawesi (for example, Mongondow-Gorontalo), as well as in most areas to the south.

1.3 THE DATA. The primary source of the data in this paper is the writer’s field-work conducted during more than a dozen trips to Borneo between April 2008 and October 2012, totaling seven months. Well over 100 language communities were visited in Sabah, Brunei, northern Sarawak, and northern Kalimantan Timur. In fifty communities,

13. As an anonymous reviewer correctly suggests, at least one morphological innovation might also be mentioned: *-in ‘Location Focus’, which is found throughout the core Murutic languages, although not in Tatana and Papar.

an 800-item wordlist was elicited along with a hundred or so sentences covering functor subsystems and verbal morphology.¹⁴ In over 50 other communities, only the sentences were elicited.

In early 2010, David Moody of SIL Malaysia generously shared an invaluable database containing hundreds of wordlists (approximately 200 to 400 items each) elicited by various members of SIL Malaysia in the 1970s and 1980s. These cover virtually every known speech variety in northern Borneo, and so allowed me to survey the full extent and distribution of phonological innovations in Sabah, and to consolidate plans for my longest field trip to northern Borneo, from October 2010 to February 2011. Access to these SIL wordlists no doubt saved me years of work and an unimaginable amount of money, allowing me to prioritize the speech varieties that had the most important and widest range of features. Pronominal data for Dusun Kimaragang (from Kroeger 2005) and Dusun Tobilung (from Buck 2009) were also very generously provided by Paul Kroeger and David Moody in 2009, and these data were later supplemented by my own fieldwork on these two languages.

The approximate distribution of the languages included in this survey is illustrated in map 1.

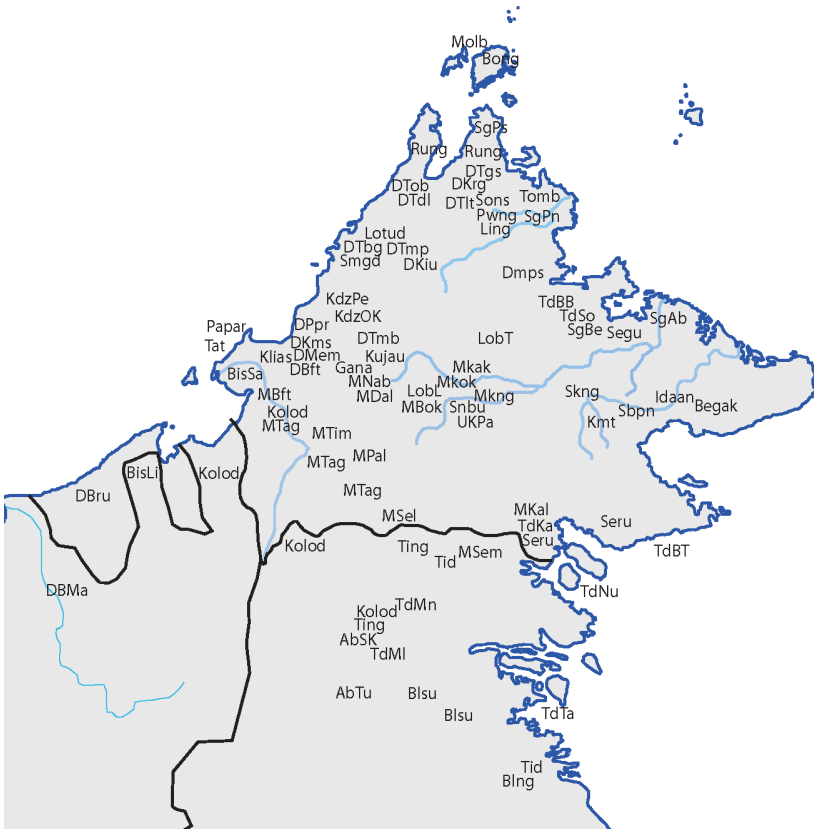
2. EVIDENCE FOR THE DUSUNIC SUBGROUP. As stated in 1.2 above, there are nine phonological and functor innovations that define the Dusunic subgroup. In order to be able to consider these (and the Murutic innovations in section 3) in context, a tentative reconstruction of the Proto-Southwest Sabah phonological system is given in table 2. Note that the status of PSWSAB *j and *ñ is uncertain. A phoneme /j/ occurs only sporadically in a handful of forms in a few Dusunic, Bisaya-Lotud, and Murutic languages, and never appears to be a direct continuation of PMP *j or *z, both of which merged with PMP *d as PSWSAB *d. Likewise, /ñ/ appears in scattered languages in a small number of forms with no known etymologies, but also appears in the *mañ- allomorph of *maN- when prefixed to *s-initial roots in Sabah Bisaya, Tatana, and the Murutic languages spoken in Kalimantan Timur (Kolod, Tingalan, Abai Sembuak, Abai Tubu, Bulusu, and the various varieties of Tidung). In any case, the phonemic status of *j and *ñ is not relevant to the innovations discussed in this paper, and will not be considered further.

2.1 DUSUNIC PRONOUN INNOVATIONS. Three replacement innovations can be identified in the pronouns of Proto-Dusunic, and are absent from the Bisaya-Lotud, Paitanic, and Greater Murutic languages, as illustrated in tables 3 and 4. The first two, *daJə(n?) ‘1SG.OBL’ and *daJay ‘1EXCL.OBL’, replaced PSWSAB *d[i]-ak(əi)(n?) and *d[i]-am(əi)(n?), respectively,¹⁵ and are reflected in all of the core Dusunic languages. The third, *ya ‘1EXCL.GEN’, replaced PSWSAB *mai (< PMP *mami), and is reflected in all core Dusunic languages except Dumpas (which has borrowed the pronoun *may* under influence from various Paitanic languages and Tidung) and various lan-

14. This set of elicitation materials is a slightly revised version of materials elicited by the current author for around 200 speech communities in the Philippines and northern Sulawesi.

15. The Proto-Southwest Sabah forms are also supported by evidence from Idaanic, Molbog-Bonggi, and Proto-Philippines, and from a top-down perspective by Proto-Malayo-Polynesian reconstructions such as those by Ross (2006) and Blust (1977).

MAP 1. THE SOUTHWEST SABAH LANGUAGES (with key to abbreviations)



- | | | |
|----------------------------------|--------------------------------------|---------------------------|
| AbSK = Abai Sembuk | AbTu = Abai Tubu | Begak = Begak |
| BisLi = Limbang Bisaya | BisSa = Sabah Bisaya | Blng = Bulungan |
| Blsu = Bulusu | Bong = Bonggi | DBft = Dusun Beaufort |
| DBMa = Dusun/Bisaya Marudi | DBru = Brunei Dusun | DKiu = Dusun Kiulu |
| DMem = Dusun Membakut | DKms = Dusun Kimanis | DKrg = Dusun Kimaragang |
| DTbg = Dusun Tambalugu ("Lotud") | Dmps = Dumpas | DPpr = Dusun Papar |
| DTmp = Dusun Tamparuli | DTdl = Dusun Tindal | DTgs = Dusun Tinagas |
| Idaan = Idaan | DTlt = Dusun Talantang | DTmb = Dusun Tambunan |
| Klias = Dusun/ "Kadazan" Klias | DTob = Dusun Tobilung | Gana = Gana |
| Kujau = Kujau | KdzOK = Kadazan Ovai Kambizaan | KdzPe = Kadazan Penampang |
| LobL = Lobu Lanas | Kmt = Sungai Kuamut ("Dusun" Segama) | Kolod = Murut Kolod |
| MBft = Murut Beaufort | LobT = Lobu Tampios | Ling = Lingkabau |
| Mkak = Mangkak | MBok = Murut Bookan | Lotud = Lotud |
| Mkok = Minokok | MKAl = Murut Kalabakan | MDal = Murut Dalit |
| MPal = Murut Paluan | MNab = Murut Nabaay | Mkng = Sungai Makiang |
| MTag = Murut Tagol | MSel = Murut Selungai | Molb = Molbog |
| Pwng = Dusun Puawang | MTim = Murut Timugon | MSem = Murut Sembakung |
| | Rung = Rungus | Papar = Papar |
| | | Sbpn = Subpan (cont'd) |

Segu = Sungai Seguluid	Seru = “Murut” Serudung	SgAb = Abai Sungai
SgBe = Sungai Beluran	SgPn = Sungai Paitan	SgPs = Sungai Pitas
Skng = Sukang	Smgd = Sumagid	Snbu = Sungai Sinabu
Sons = Sonsogon	Tat = Tatana	TdBB = Tidung Bangawong
TdBT = Tidung Bangawong	TdKa = Tidung Kalabakan	Beluran/Labuk
Tawau	TdMI = Tidung Malinau	TdMn = Tidung Mansalong
TdNu = Tidung Nunukan-	TdSo = Tidung Sambal/	TdTa = Tidung Tarakan
Sembakung	Sombol	Tid = Tidung (unspecified
Ting = Tingalan	Tomb = Tombonuwo	dialect)
UKPa = Upper Kinabatangan Paitan		

TABLE 2. THE TENTATIVE PROTO-SOUTHWEST SABAH PHONOLOGICAL SYSTEM

CONSONANTS				VOWELS		
*p	*t		*k	*ʔ	*i	*u
*b	*d	(*j)	*g		*ə	
	*s			*R	*a	
*m	*n	(*ñ)	*ŋ			
	*l					
	*r					
*w	*y					

TABLE 3. DUSUNIC PRONOMINAL INNOVATIONS

	1EXCL.GEN	1SG.OBL	1EXCL.OBL
PDUS	*ya	*dɔJə(nʔ)	*daJaj
DMEM	za	dojoʔ	dajaj
DKMS	ja	dojoʔ	dajaj
DPPR	za	dohoʔ	dahay
KDZOK	za	dohoʔ	dahay
KUJAU	za	dəhəʔ, dəhən	dahay
DTMB	ya	dohoʔ	dahay
DTDl	(dahay, < OBL)	dohoʔ	dahay
RUNG	(dahay, < OBL)	dohoʔ, dohon	dahay
KDZPE	za	dogoʔ	dagay
KLIAS	za	dogoʔ	dagay
SKNG	ja ~ ya	dogoʔ	dagay
MKAK	ja	dogoʔ	dagay
MKOK	za	dəgəʔ	dagay
DTLT	ja	dogoʔ	dagay
DKRG	ya	dogoʔ, *dogon	dagay
DMPS	may (<PAITAN)	dogoʔ	dagay
DTOb	(dagay, < OBL)	dogoʔ, dogon	dagay
DTGS	(dagay, < OBL)	dogoʔ	dagay
DTBG	ya	dogoʔ	dagay
DTMP	(dagay, < OBL)	dogoʔ	dagay
SONS	?	dogoʔ	dagay
PMUR	*may	*dak(oi)(nʔ)	*dam(oi)(nʔ)
PBisLO	*ni-amiʔ	*jəkiʔ	*jamiʔ
PSWSAB	*mai	*d[i]-ak(əi)(nʔ)	*d[i]-am(əi)(nʔ)
PMP/PPHNB	*mami	*d[i]-akə(nq)	*d[i]-amə(nq)

* This form was contributed by an anonymous referee to this paper, and credited to “Janama Lontubon, native speaker,” personal communication.

**TABLE 4. DUSUNIC PRONOMINAL INNOVATIONS
(NON-DUSUNIC EQUIVALENTS)**

	1EXCL.GEN	1SG.OBL	1EXCL.OBL
PBisLo	*ni-ami?	*joki?	*jami?
DBRU	(jami?, < OBL)*	jai?	jami?
BISLI	(jami?, < OBL)†	jai?	jami?
BISSA	(jami?, < OBL)	joki?	jami?
LOTUD	nyami?	joki?	jami?
PGMUR	*may	*dako(n?), *daki?	*damo(n?), *dami?
TAT, GANA, MNAB, MBOK	may	daki?	dami?
PAPAR	may	daki?	(dakay)†
MTIM	may	raki?	ramon
MPAL, KOLOD, MSEM, ABSK, TING, BLSU	may	dakon	damon
MTAG	may	dakon, rakon, dako?, rako?	damon, ramon, damo?, ramo?
MKAL	may	rako?, rakon	ramo?, ramon
ABTU	?	doxon	domon
TDBB, TDSo	may	dakon	damon
TDBT	may	dako?	damo?
TDNU	may	dakon	damo?
TDKA	(damo? < OBL)	dako?	damo?
TDMN	(damo? < OBL)	dakon	damon
TDML	(domo? < OBL)	doko?	domo?
TDTA	(damo? < OBL)	dako?	damo?
PPAIT	*mai	— (*sa+*aku < NOM)	— (*sa+*kai < NOM)
LOBL, MKNG	mai	saaku	sakai
LOBT, UKPA, KMT, SNBU	mai	saaku	sakai
TOMB	me?e	so aku	so ke?e
PWNG	mae	saaku	sokae
SGBE, LING	may	saaku	sakai
SGPN	may	səaku	səka?i
SGPS	ma?ay	saaku	sakai
SERU	mee	saaku	sekee
PIDAAAN	—	*akon	*amon
IDAAN	— (= NOM)	engkon	ngamon
BEGAK	— (= NOM)	nakon	namon
SEGU	— (= NOM)	akon	ngamon
PMoBo	*may	*di-ak(əi)(n?)	*dikay
MOLB	may	yahi?, nahi?	dikay
BONG	mi	diyadn	dih

* In these languages, Oblique forms have replaced the earlier Genitive and Nominative forms.

† An innovation derived from oblique formative *d(a)- + *akay '1EXCL.NOM', and not derived from PDUS *daJay.

guages where it has been replaced as the result of the widespread Dusunic phenomenon of Oblique pronouns replacing their Genitive counterparts, as in Sonsogon, Tinagas, Rungus, Dusun Kiulu, Dusun Tindal, and Dusun Tobilung.

Note that the innovated PDUS *ya '1EXCL.GEN' should not be confused with the PPHNB pronoun *ya '3SG.GEN', which had shifted to *yə in PSWSAB due to an innova-

tion in which word-final PMP *a became *ə. The shift of word-final *a > *ə in PSWSAB clearly occurred prior to the innovation of PDUS *ya ‘1EXCL.GEN’.

The full sets of pronoun reconstructions for Proto-Dusunic, Proto-Bisaya-Lotud, Proto-Murutic, and Proto-Paitanic are included in tables 5–8 for reference. From the pronouns alone, it can be shown that the Bisaya-Lotud languages are not Dusunic, as they do

TABLE 5. PROTO-DUSUNIC PRONOUNS

	TOP	NOM	GEN	OBL
1SG	*i-əku	*əku	*ku	*dəJə(n?)
2SG	*i-ika[w]	*kə	*nu	*d-ika[w]
3SG	*i-siyə	*i-siyə	*yə	*di-siyə
1EXCL	*i-(iə)əkəy	*[(iə)]kəy	*ya	*d-aJəy
1INCL.DU	*i-kitə	*kitə	*tə	*datə?, *di-kitə
1INCL.PL	*i-təkə[w]	*təkə[w]	*təkə[w]	*datən, *dati?, *di-təkə[w]
2PL	*i-kə(w)y	*kəw	*yu, *muyu, *nyuy	*di-kə(w)y
3PL	*i-sidə, *yə-sidə	*i-sidə, *yə-sidə	*ni-sidə, *nə-sidə	*d(iə)-sidə

TABLE 6. PROTO-BISAYA-LOTUD PRONOUNS

	NOM	GEN	OBL
1SG	*əku	*ku	*jaki?
2SG	*ik(aə)w, *=kə	*mu	*[d]ijun
3SG	*iyə	*nyə	*[di]siyə
1EXCL	*i(iə)kəy	*nyami?	*jami?
1INCL.DU	*[k]itə	*[ki]tə	*di[ki]tə
1INCL.PL	*[i]təkəw	*təkəw	*jati?
2PL	*i(iə)kəw, *=kəw	*muyu[n]	*[di]jamuyu[n]
3PL	*idə	*də	*[di]sidə

TABLE 7. PROTO-MURUTIC PRONOUNS

	NOM	SHORT NOM	GEN	OBL
1SG	*a[k]u	(*a[k]u)	*ku	*dak(oi)(n?)
2SG	*oko[w]	*=ko[w]	*mu	*di[]un
3SG	*iso, *(io)yo	*iyə	*no	*di[s]o
1EXCL	*akay	(*akay)	*may	*dam(oi)(n?)
1INCL.DU	*ito	*=to	*to	*dito
1INCL.PL	*itaka[w]	*=taka[w]	*taka[w]	*ditaka[w]
2PL	*aka[w]	*=ka[w]	*muyu[n] ~ *mi[n]	*damuyu[n]
3PL	*iro	(*iro)	*niro	*di[si]ro

TABLE 8. PROTO-PAITANIC PRONOUNS

	NOM	GEN	OBL
1SG	*aku	*ku	*sa-aku
2SG	*i(ə)kaw, *=kə	*mu	*sə-(ə)kaw
3SG	*iyə	*niyə	*sə-iyə
1EX	*kai	*mai	*sə-kai
1INCL.DU	*kitə	*tə	*sə-kitə
1INCL.PL	*təkə	*təkə	*sə-təkə
2PL	*kau	*muyu	*sə-kaw
3PL	*sirə	*nə-sirə	*sə-sirə

not share any of these Dusunic innovations. In addition, they also lack the Topicalized Nominative forms with formative *[y]i-, and clearly preserve the set of Oblique bases much like those of the Murutic subgroup (cf. table 7), inherited from PMP/PPHNB (table 9), and also reflected in various Philippine subgroups, as well as in Proto-Idaanic (table 10). Note that the Paitanic languages are unique among the Southwest Sabah languages in having completely replaced the Oblique bases with forms from the Nominative set preceded by case-marking formative *sa ~ *sə (not attested in any of the other Southwest Sabah subgroups), as shown in table 8. As a result, there is no way of determining whether the Paitanic languages shared the two Dusunic Oblique pronoun innovations prior to the replacement of the Paitanic Oblique pronoun set.

Based on the reconstructions in tables 5–8 for the four Southwest Sabahan subgroups, the Proto-Southwest Sabah pronominal system can be reconstructed as illustrated in table 11.

TABLE 9. PROTO-PHILIPPINES AND NORTH BORNEO PRONOUNS

	NOM	GEN	LONG GEN	OBL	OBL-2
1SG	*aku	*ku	*naku	*akə(nq)	*d[i]-akə(nq)
2SG	*ikaw, *=ka[w]	*mu, *nu	*nimu	*imu, *iu[n]	*dimu, *diu[n]
3SG	*[si]-ia	*na, *ni-ia, *ya	*nia	*ia	*dia
1EXCL	*kami	*mi	*nami, *mami	*amə(nq)	*d[i]-amə(nq)
1INCL	*kita	*ta	*nita	*atə(nq)	*d[i]-atə(nq)
2PL	*kamu, *kayu	*niu, *muyu	*namu	*inyu, *imuyu	*dinyu, *d[i]-amuyu
3PL	*sida	*da	*nida	*ida	*dida

TABLE 10. A VERY PRELIMINARY RECONSTRUCTION OF THE PROTO-IDAANIC PRONOUN SYSTEM*

	NOM	GEN	OBL
1SG	*aku	—	*noŋ nakon
2SG	*ik(ao)w	—	*noŋ niyun
3SG	*rumo	—	*noŋ rumo
1EXCL	*k(əu)m̄mi	—	*noŋ namon
1INCL	*kito	—	*noŋ naton
2PL	*muyu	—	*noŋ muyon
3PL	*[]iro	—	*noŋ iro

* An anonymous referee points out that Goudswaard (2005:127) notes distinct genitive forms *ku* '1SG.GEN' and *mo* '2SG.GEN' for Begak, although it is unclear whether these are retentions from PMP or borrowings from Malay, Tausug, Sama-Bajaw, or any other language with which Begak has been in contact historically. Regardless, this does not affect the arguments being made in this paper.

TABLE 11. PROTO-SOUTHWEST SABAH PRONOUNS

	NOM	GEN	OBL
1SG	*aku	*=ku	*d[i]-ak(əi)(n?)
2SG	*[əi]-ka[w], *=kə	*=mu, *=nu	*d[i]-iyun
3SG	*[s]iyə	*=yə, *=nə, *nyə	*di[s]iyə
1EXCL	*ə-kai	*=mai	*d[i]-am(əi)(n?)
1INCL.DU	*[k]itə	*=tə	*d[i]-at(əi)(n?)
1INCL.PL	*[ki]ta-kau	*=ta-kau	*d[i]-at(əi)(n?)
2PL	*ə-kau, *=kau	*=muyu[n]	*d[i]-amuyu[n]
3PL	*[s]idə	*=[ni-]də	*di[s]idə

2.2 REFLEXES OF *R IN DUSUNIC vs. MURUTIC. Unlike most Philippine languages, which have only one inherited reflex of *R (cf. Blust 1991:90, 2009:582–83),¹⁶ or two reflexes, one of which is due to borrowing or an unconditioned split, reflexes of PMP *R in the descendants of Proto-Southwest Sabah involve fairly complex conditioning. These reflexes have not been previously discussed in the literature—which is not surprising considering the lack of accessible lexical materials available for most of these languages. However, they quite strikingly define the boundary between the Greater Dusunic and Greater Murutic subgroups.

The Greater Dusunic and Greater Murutic languages share two environmentally conditioned reflexes of *R that can be attributed to Proto-Southwest Sabah. First, after schwa, *R is reflected as *g, probably due to compensatory lengthening to *-gg- after the phonetically short schwa, then subsequent shortening of *-gg- to *-g-. This series of shifts is also found in some languages in the northern and southern Philippines and has been discussed by Blust (2010).¹⁷

The second reflex of *R shared by Proto-Dusunic and Proto-Murutic is *R > *h in the environment {aiu}_ {auə}, the result of intervocalic lenition as opposed to the fortition that occurred after *ə. Note that PGMUR *h was lost in all of the Greater Murutic languages except Papar; and PGDUS *h was likewise lost in (1) a number of Dusunic languages except Rungus, Kujau, and the Dusun dialects of Papar, Kota Belud, and some parts of Tambunan; (2) all Paitanic languages; and (3) all Bisaya-Lotud languages except Lotud.

However, there are also at least two environments where the Greater Dusunic and Greater Murutic languages have different reflexes of *R. The first is word-finally, where PSWSAB *R became *g in Proto-Greater Murutic, but *w in Proto-Greater Dusunic. The second is before *i, where *R became *g in Proto-Dusunic and Proto-Bisaya-Lotud, but *h in Proto-Murutic (that is, the reflex of *R before *i in Proto-Greater Murutic is the same as that before *a, *u, or *o, unless the preceding vowel is *ə). Paitanic has a zero reflex of PMP *R before *i, probably via PSWSAB *R > Pre-Paitanic *h, with subsequent loss of *h in Proto-Paitanic.

It should be noted that the lenition of *R > /w/ word-finally is rare; Blust (2009:582), in his survey of the entire 1,200-member Austronesian family, lists only “Bisaya”¹⁸ (apparently referring to data collected on the Limbang Bisaya and another group of Bisaya living upriver from Miri in the vicinity of Marudi) as having a /w/ reflex of *R. Therefore, the /w/ reflex of word-final *R, shared by the Dusunic, Bisaya-Lotud, and Paitanic languages, is considered a highly diagnostic innovation indicating membership in the Greater Dusunic subgroup (cf. section 4).

Table 12 outlines the reflexes of *R in the various intermediate protolanguages of Southwest Sabah, while table 13 illustrates the reflexes of some PMP forms with *R in the various subgroups of Sabah, and tables 14a–b illustrate the reflexes of *R in individual

16. Umiray Dumaget is one such language with two environmentally conditioned reflexes of *R, namely, /g/ and zero (Himes 2002:279–80).

17. This same process of fortition also affected *b and *d. Note that gemination after schwa is a process that is also reflected in a number of Northern Luzon (Cordilleran) languages, a few Manobo languages, and phonetically in Maranao (cf. also Blust 2009:219).

18. Blust (pers. comm., November 14, 2011) notes that in 1971 he elicited data for Limbang Bisaya, and for another group that self-identified as “Bisaya Bukit” or “Sang Bukit,” which apparently was spoken in the border area between interior Sarawak and Brunei.

TABLE 12. REFLEXES OF *R BY ENVIRONMENT

ENVIRONMENT	PDUS	PBISLO	PPAIT	PGMUR	PAPAR	TAT	PMUR
*ə_	g	g	g	g	g	g	g
(aiu)_(auə)	h	h	∅	h	h	∅	∅
#	w	w	w	g	g	g	g
u_i	g	g	∅	h	h	∅	∅
i_i	g	g	∅	h	h	∅	∅

TABLE 13. REFLEXES OF *R IN SOUTHWEST SABAH

PMP or PSWSAB	PDUS	PBISLO	PPAIT	PGMUR	PIDAAAN
*zaRami 'straw'	*rahami	*rahami	*raami	—	*dami
*baRa 'coals'	*baha	—	*waa	*baha	*əbba?
*daRaQ 'blood'	*raha?	*raha?	*raa?	*daha?	*ədda?
*daRat 'sea'	*(dr)jahat	*rahat	*raat	*da[]at	—
*tiŋaRaQ 'look up'	*tiŋaha?	*tiŋaha?	*tiŋaa?	*tiŋaha?	—
*duRi 'thorn'	*rugi	*rugi	*duwi	*duhi	*duwi
*hadiRi 'post'	*ərigi	*ərigi	*əndii	*ad[]i	—
*linsəR 'seed'	*linsəw	*linsaw (L)	*linsəw	—	*lissog
*luaR 'loose'	*luwaw	*luwaw	*luwaw	*luwag	*luwag
*ibəR 'phlegm'	*iwəw	*iwəw	*iwəw	*iwog	—
*siaR 'brave'	*siyəw	*siyəw	—	*siyog	—
*liqəR 'neck'	*li?əw	*li?əw	*li?əw	*li?əg	*lig
*laməR 'slippery'	*laməw	*laməw	—	*lamog	*lamog
*dəŋəR 'hear'	*rəŋəw	*rəŋəw	*rəŋəw	*rəŋəg	*kiŋog
*bəsuR 'full'	—	—	*wasug	*[w]asug	—
*sandiR 'lean on'	*səndiw	*səndiw	*səndiw	*sandig	*sandig
*suliR 'floor'	*suliw	*s(ui)liw	—	*sulig	—
*bihaR 'alive; full'	*wiyaw	*[w]iyaw	*iyaw	*biyag	*biag
*bakaR 'sweet potato'	*wakaw 'rattan'	*wakaw	—	*bakag	—
*pəRəQ 'squeeze'	*paga?	*paga?	*pəga?	*paga?	*[]əgka?
*[]əRis 'sand'	*əgis	*əgis	əgis	*əgis	bəris
*bəRaS 'uncooked rice'	*wagas	*wagas	*wəgas	*bagas	*bəgkas
*bəRəqat 'heavy'	*wagat	*wagat	*wəgat	*bagat	*bəgkat
*baqəRu 'new'	*wagu	*wagu	*wagu	*bagu	*bəgku
*uRaT 'vein'	*uhat	*uhat	*uwat	*uhat	*uwat
*dapuRaN 'stove'	*rəpuhan	*rəpuhan	—	*dapu[]an	—
*suRaT 'wound'	*suhat	*suhat (L)	—	*su[]at	—
*duRaŋ 'add'	*ruhaŋ	*ru[]aŋ (B)	*ruwaŋ	*du[]aŋ	—
*hiRup 'slurp'	*ihup	*ihup	*iyup	—	—
*Ratus 'hundred'	*hatus	—*	*atus	*[]atus†	—
*baRəqaŋ 'molar'	*wiyaŋ‡	*bagaŋ	*bagaŋ	*bagaŋ	*bagaŋ
*duRay 'short time'	*ruhay	*ruhay (L)	*ruwəy	*ru[]oy	—
*tiRəb 'burp'	*tihəb	*tihəb	*tiəb	*ti[]ob	tigob/sigob/ igkab
*suRuq 'order'	*suhu?	*su[]u?	*suu?	—	*su?

* Evidence for *h in this form in PGMUR could only come from Papar, but unfortunately, all Papar numbers above ten are Malay borrowings.

† Note that *Ratus rarely occurs without being preceded by a number or a prefix, so its initial *R rarely occurs in true word-initial position.

‡ Proto-Dusunic *wiyaŋ 'molar' is irregular, for expected **wagaŋ.

TABLE 14. REFLEXES OF *R IN DUSUNIC AND MURUTIC LANGUAGES

Table 14a. *R medial before *i, and final

	MEDIAL_i		FINAL					
	‘thorn’	‘post’	‘neck’	‘alive, full’	‘spit, saliva’	‘hear’	‘floor’	‘slippery’
PSWSAB	*duRi	*hadiRi	*liʔəR	*biyaR	*iwəR	*dəŋəR	*suliR	*laməR
PBisLO	*rugi	*arigi	*liʔəw	*[w]iyaw	*iwəw	*rəŋəw	*suliw	*laməw
DBRU	rugi	—	liyow	miyaw	iwow	-oŋow	siliw	lamow
BisLI	rugi	rigi	liyaw	mayaw	—	-roŋow	siluy	lamaw
BisSA	rugi	rigi	liyow	iyow	—	-roŋow	saliw	-lamu
LOTUD	rugi	origi	liʔow	-wiyaw	iwow	-roŋow	suliw	-lamow
PDUS	*rugi	*arigi	*liʔəw	*wiyaw	*iwəw	*rəŋəw	*s(ui)liw	*laməw
RUNG	rugi	origi	liʔow	-vizaw	—	—	—	—
DTDL	rugi	torigi	liyow	-ayaw	—	-roŋow	—	-lamow
DKRG	rugi	torigi	liyow	-wiyaw	—	-roŋow	—	-lamow
DTOB	rugi	torigi	liyow	-wizaw	—	-roŋow	—	-lamow
DTGS	rugi	—	—	—	—	-roŋow	—	—
DTLT	rugi	torigi	—	-wijaw	—	-roŋow	—	-lamow
DTMB	rugi	torigi	liyow	-wiyaw	—	-roŋow	—	-lamow
MKOK	rugi	torigi	fiyow	-vizaw	tivow	-roŋow	—	-famow
KUJAU	rugi	torigi	fiyow	—	—	-rəŋəw	sufiw	-famow
SKNG	rugi	torigi	liyow	-iyaw	—	-roŋow	—	—
KDZPE	lugi	—	—	-vizaw	—	-oŋow	hisiw	—
DPPr	yugi	oigi	hiʔow	-vizaw	iwow	-yoŋow	sihiw	-hamow
KLIAS	ugi	—	liyow	-wizow	iβow	-roŋow	siliw	-lamow
DMPS	rugi	—	liyow	-iyaw	iwow	-roŋow	—	—
PGMUR	*duhi	*arihi	*liʔog	*[b]iyag	*iwog	*-roŋog	*sulig	*lamog
PAPAR	duhi	arihi	iʔog	biyag	iwog	-roŋog	suwig	-famog
TAT	duwi	rigi (<GDUS)	liʔog	biyag	iwog	-roŋog	sulig	-lamog
GANa	duwi	—	liyog	---	iwog	-əŋəg	sulig	-laməg
MNAB	duwi	—	liyog	biyag	iwog	—	sulig	-lamog
MBOK	duwi	arii	liyog	-ayag	iwog	—	—	-lamog
MTIM	ruwi	—	liyoy	-ayag	tiwoy	-roŋoy	suliy	-lamoh
MPAL	ruwi	alii	liyoy	-ayah	tiwoh	—	sulih	-lamoh
MTAG	ruwi	—	liyoh, liyog	-ayah	-tiwoh, -tiwog	-roŋoh, -riŋog	sulih	-lamoh
KOLOD	duwi	—	liyog	-ayag	—	-roŋog	—	-lamog
MKAL	duy	—	liyoy	-uyay	-iwoy	—	—	-lamoy
MSEL*	duwi	—	—	-ayah	tiwoh	-riŋoh	—	—
MSEM†	duwi	—	—	ayag	—	-kiŋog	—	—
TING	duwi	—	liyog	-ayag	—	kinog	sulig	-lamog
ABSK	liduy	—	liyog	-ayag	—	diŋog	—	lamog
ABTU	duwi	arii	liyok	-ayak	—	ŋinok	—	—
BLSU	duwi	—	liyog	-uyag	iwog	diŋog	—	lamog
TIDUNG	duwi, udui	—	liyog	-uyag	-iwog ~ -uyog	diŋog	sulig	lamog

* Data for these two languages from SIL wordlists provided by Dave Moody.

Table 14b. *R medial / (aiu)_(aou)

	'short (time)'	'add'	'blood'	'slurp'	'vein'	'order'	'look up'
PSWSAB	*duRay	*duRaŋ	*daRa?	*iRup	*uRat	*suRu?	*tiŋaRa?
PBISLO	*ruhay	*ruaŋ	*raha?	*ihup	*uhat	*su ju?	*tiŋaha?
DBRU	—	—	raa?	iyup	uwat	suu?	—
BISLI	—	—	raa?	—	uwat	suu?	—
BISSA	—	ruwaŋ	ra?	—	urat (< MLY)	—	tiŋaa?
LOTUD	oruhay	---	raha?	ihup	uhat	—	tiŋaha?
PDUS	*ruhay	*ruhaŋ	*raha?	*ihup	*uhat	*suhu?	*tiŋaha?
RUNG	—	—	raha?	—	uhat	suhu?	tiŋaha?
DTDL	oruhay	—	raha?	—	tuhat	—	tiŋaha?
DKRG	oru?ay	—	ra?a?	—	tuwat	—	tiŋa?a?
DTOB	oruway	—	raa?	—	tuwat	—	tiŋaa?
DTLT	oruway	ruwaŋ	raa?	—	tuwat	suu?	tiŋaa?
DTMB*	oruway	ruwaŋ	raa?	—	tuwat	suu?	tiŋaa?
MKOK	oruway	ruwaŋ	raa?	—	tuwat	suu?	tiŋaa?
KUJAU	oruhay	ruhaŋ	raha?	—	tuhat	suhu?	tiŋaha?
SKNG	—	ruwaŋ	raa?	iyup	tuwat	suu?	tiŋaa?
KDZPE	—	—	zaa?	—	—	—	—
DPPR	—	—	haa?	ihup	uhat	—	tiŋaha?
KLIAS	oduvay	—	raa?	—	uwat	suu?	—
DMP5	—	ruwaŋ	raa?	—	uwat	suu?	—
PGMUR	*ru jay	*du jaŋ	*daha?	*ihup	*uhat	—	*tiŋaha?
PAPAR	—	ruwaŋ (<GDUS)*	daha?	ihup	uhat	—	tiŋaha?
TAT	—	—	daa?	—	urat (< MLY)	—	tiŋaa?
GANAN	—	—	daa?	—	uwat	—	tiŋaa?
MNAB	ruwoy	duwaŋ	daa?	—	uwat	—	tiŋaa?
MBOK	—	—	daa?	—	uwat	—	—
MTIM	—	ruwaŋ	raraa?	—	uwat	—	tiŋaa?
MPAL	—	duwaŋ	daa?	—	uwat	—	tiŋaa?
MTAG	—	ruwaŋ	---	—	uwat	—	—
KOLOD	—	—	daa?	—	uwat	—	—
MKAL	—	ruwaŋ	rara?, raa?	—	uwat	—	—
MSEL†	—	—	raa?	—	—	—	—
MSEM†	—	—	daa?	—	uwat	—	—
TING	—	duwaŋ	daa?	—	urat (< MLY)	—	tiŋaa?
ABSK	—	duwaŋ	daa?	—	uwat	—	—
ABTU	—	—	daa?	—	uwat	—	—
BLSU	—	duwaŋ	dada?	—	uwat	—	—
TIDUNG	—	duwaŋ	dada?	—	awat	—	tiŋaa?

* SIL wordlists for Dusun Tambunan communities show *R > /h/ in these forms; my informants reflected *R > *h > Ø; it is unclear if this is a difference in village of origin, or if a further shift has occurred in the past 30 years.

† Initial /r/ in Papar *ruwang* suggests that this is a borrowing from a Greater Dusunic language.

Dusunic and Murutic languages. The fortition of *R > g before *i in Dusunic is in some sense unsurprising, as /i/ has been observed to have a fortifying effect on a preceding consonant: for example, in certain Philippine languages in which *l generally became /y/, and /ɣ/ zero or an interdental lateral, the presence of an adjacent /i/ or /y/ blocks the shift.¹⁹

It is noteworthy that, in addition to the aforementioned environmentally conditioned reflexes of *R, there are a number of other forms where *R is reflected as /g/ in Dusunic lan-

guages in environments where another reflex would be expected. This parallels the situation with *b and *d, which in certain languages usually lenited to *w and *r, respectively, but for which /b/ and /d/ reflexes are also often found. Therefore, it seems likely that *b, *d, and *R generally split in Proto-Dusunic, with each having both stop and lenited reflexes. Note that Blust (2010) also found many exceptions to the expected reflexes of *b, *d, and *R, and while it is unclear how to explain them (although borrowing is one possibility), they have no bearing on the current discussion, which is based on the environmentally conditioned reflexes that help define the various Southwest Sabah subgroups.

2.3 DUSUNIC *ma- > *a-/*ə-. As illustrated by the examples in table 15, the initial consonant of the adjectival prefix *ma-, and of the homophonous accidental/abilitative prefix, was lost when prefixed to roots beginning with any consonant except *ʔ in Proto-Dusunic, Proto-Paitanic, and all Bisaya-Lotud languages except Sabah Bisaya. The distribution of this innovation suggests that it was a dialectal feature in Proto-Greater Dusunic, and that Sabah Bisaya was ultimately the only language in this subgroup that retained the full *ma- prefix in all environments. Note that the initial *m- of this prefix was later dropped altogether in most Dusunic languages except Dumpas and Southern Kadazan, which along with Lotud and the Paitanic languages retain the complementary distribution of *a- on consonant-initial roots and *ma- on vowel-initial roots.

It is interesting to note that a similar, but not identical, shift of *ma- to *a- is also found in some Murutic languages, but curiously only in the central and southern Murutic languages²⁰ that have had little or no contact with Dusunic languages: Paluan, Tagol, Kalabakan, Kolod, Sembakung, Selungai, and the various varieties of Tidung. The shift in these Murutic languages is considered to be independent of the shift in the Greater Dusunic languages, especially since there are two key differences in its realization: (1) while only the adjectival prefix *ma- and the homophonous abilitative/accidental prefix are affected in the

TABLE 15. *ma- > *a-/*ə- IN DUSUNIC

PSWSAB	PDUS	PPAIT	PBISLO	PMUR	PIDAAAN
*ma-asin ‘salty’	*mə-əsin	*ma-asin	*mə-ʔəsin	*ma-asin	—
*ma-itəm ‘black’	*mə-itəm	*mə-itəm	*mə-ʔitəm	*ma-itom	—
*ma-əmis ‘sweet’	*mə-əmis	*mə-əmis	*ma[ta]ʔəmis	*ma-amis	*ammis
*ma-a[n]səm ‘sour’	*mə-ənsəm	*mə-əsəm	*mə-ʔənsəm	*mo-o[n]som	*as[s]om
*ma-paʔit ‘bitter’	*ə-pəʔit	*ə-pəit	*mə-pəʔit	*ma-paʔit	*a-pait
*ma-raat ‘bad’	*a-raat	—	*ma-ra[]at	*ma-raat	*arat
*ma-ləmiʔ ‘soft’	*ə-ləmiʔ	*ə-ləmiʔ	*mə-ləmiʔ	*ma-lamiʔ	—
*ma-luwaR ‘loose’	*ə-luwaw	*ə-luwaw	*mə-luwaw	*ma-luwag	*a-luwag
*ma-ratuʔ ‘fall’	*a-ratuʔ	*a-ratuʔ	*ma-ratuʔ	*ma-ratuʔ	(aratuʔ)

* Thanks to an anonymous reviewer for bringing this Begak form from Goudswaard (2005:479) to my attention.

19. Blust (pers. comm., November 14, 2011) also notes that the strengthening of *l > d / _i is found in Malagasy, Maanyan, and several other Southwest Barito languages. However, it is not always the case that *i strengthens a preceding consonant, as *R became zero before *i in Umiray Dumaget (Himes 2002), and languages like Tagalog, Tausug, and Southern Binukidnon are among the rare Central Philippine languages in which *l > zero even adjacent to /i/.

20. Note that Abai Sembuak, Abai Tubu, and Bulusu have lost this prefix altogether, which mirrors the trend in non-Philippine-type languages found to the south.

Greater Dusunic languages, the innovation in the selected Murutic languages affects all prefixes of the shape *ma[C]-; and (2) the innovation in the selected Murutic languages affects the *ma[C]- prefixes on all roots whether consonant-initial or vowel-initial.

Note that the initial *n- of the past accidental/abilitative prefix *na- is universally reflected as /n/, even in languages where the initial *m of its nonpast counterpart is lost.

2.4 SPLIT OF PMP *b > DUSUNIC *b, *w. In initial position, *b sporadically lenited to *w in Proto-Dusunic, Proto-Paitanic, and Proto-Bisaya-Lotud, as illustrated in table 16.²¹ Note that this did not occur in all forms, as reflexes of words like *bu[ʔ]ayə ‘crocodile’, *buyuʔ ‘betel leaf’, and *bəŋəl ‘deaf’ all reflect *b as /b/.

2.5 SPLIT OF PMP *d > DUSUNIC *d, *r. Just as *b- sporadically lenited to *w- in Proto-Dusunic, *d- likewise sporadically lenited to *r-, as illustrated in table 17.

2.6 DUSUNIC AND MURUTIC REFLEXES OF *ə. PMP/PSWSAB *ə was retained as *ə in Proto-Dusunic, and is retained as /ə/ in some inland Dusunic languages such as Kujau and Minokok. Elsewhere, PDUS *ə is reflected as /o/. The exception to this is that if the vowel of the word-final syllable in PSWSAB was *a, then *ə > *a in PDUS, which is reflected as /a/ in the Dusunic languages.

TABLE 16. REFLEXES OF *b IN NORTHERN BORNEO

PMP or PSWSAB	PDUS	PBISLO	PPAIT	PGMUR	PIDAAAN
*bətiqis ‘calf’	*wətis	—	*wətis	*batis	*b(əi)tis
*baŋkiŋ ‘bedbug’	*wəŋkiŋ	*bəŋkiŋ	*bəŋkiŋ	*baŋkiŋ	*baŋkiŋ
*bəRas ‘uncooked rice’	*wagas	*wagas	*wəgas	*bagas	*bəgkas
*bulan ‘moon’	*wulan	*(bw)ulan	*ulan	*bulan	*bulan
*buluq ‘bamboo type’	*[w]uluʔ	*buluʔ	*buluʔ	*buluʔ	—
*busak ‘flower’	*[w]usak	*usak	*(buŋa)	*busak	*(tasak)
*buaq ‘fruit’	*uwaʔ	*[b]uwaʔ	*buwaʔ	*buwaʔ	*buwaʔ
*baRəqaŋ ‘molar’	*wiyaŋ	*bagaŋ (not L)	*bagaŋ *waŋaŋ	*bagaŋ	*bagaŋ
*baliw ‘move, change’	*waliw	*[w]aliw	(undaliw)	*baluy	(Begak <i>boluy</i> *)
*balay ‘house’	*walay	*walay	*waləy	*baloy	*balay
*batu ‘stone’	*watu	*(bw)atu	*watu	*batu	*batu
*bəRək ‘pig’	*wəgək	*bəgək	*wəgək	*bo(gr)ok	—
*bəRəqat ‘heavy’	*wagat	*wagat	*wəgat	*bagat	*bəgkat
*buhək ‘hair’	*əbuk	*əbuk	*əbuk	*abuk	*əbpuk
*bənəŋ ‘frog’	*bənəŋ	*bənəŋ (L)	—	*bonəŋ	—
*bawaŋ ‘river’	*bawaŋ	*bawaŋ	(bawaŋ)	*bawaŋ	—
*ma-buway ‘long (time)’	*ə-buway	*ma-buway	—	*ma-buwoy	*buway
*səbu ‘urine, urinate’	*səbu	*səbu	*səbu	*sabu	*(sidu)
*təbuh ‘sugarcane’	*təbu	*təbu	*təbu	*təbu	*təbpu
*[R]ibaŋ ‘left’	*gibaŋ	*[k]ibaŋ	*gibaŋ	*[g]ibaŋ	*gibaŋ
*dəbas ‘face, forehead’	*rabas	*rabas	(rabas)	*dabas	—

* Thanks to an anonymous reviewer for bringing this form, cited in Goudswaard (2005:454), to my attention.

21. Note that Prentice (1974) was the first to discuss this, although as Blust (2009:568–69) notes, his claims about the implications of this split are problematic.

TABLE 17. REFLEXES OF *d IN MURUTIC AND DUSUNIC

PMP or PSWSAB	PDUS	PBISLO	PPAIT	PGMUR	PIDAAAN
*daRaQ ‘blood’	*rahaʔ	*rahaʔ	*raaʔ	*dahaʔ	*ddaʔ
*daʔan ‘branch’	*raʔan	*raʔan	*raan	*daʔan	*daan
*daʔun ‘leaf’	*rəʔun	*rəʔun	*rəun	*daʔun	*daun
*dapuRan ‘stove’	*rapuhan	*rəpuhan	(dəmpuran)	*dapu[jaŋ	(dəpuRan)
*duRaŋ ‘add’	*ruhaŋ	*ru[jaŋ	*ruwaŋ	*du[jaŋ	—
*duRi ‘thorn’	*rugi	*rugi	*duwi	*duhi	*duwi
*dabas ‘face’	*rabas	*rabas	*rabas	*dabas	—
*dasam ‘rain’	*rasam	*rasam	—	*dasam	—
*daŋaw ‘hand span’	*raŋaw	*raŋaw	(raŋan)	*daŋow	(raŋŋan)
*dəŋəR ‘hear’	*rəŋəw	*rəŋəw	*rəŋəw	*rəŋəg	*kiŋog
*duRay ‘short (time)’	*ruhay	*ruhay (L)	*ruwəy	*ru[joy	—

In Proto-Murutic, on the other hand, PSWSAB *ə became *o word-finally and *a elsewhere, with the exception that if the vowel of the word-final syllable was *o, then the *ə of the preceding penultimate syllable became *o.

Although similar on the surface to vowel harmony, Blust notes that this type of shift in the languages of Sabah is “sequence-sensitive,” not a “harmonic pattern . . . targeting *oCa* but not *aCo*” in Dusunic languages (2009:250), and the same could also be said for Murutic, where the sequence *-aCo-* is targeted but not *-oCa-*.

2.7 PMP *a > DUSUNIC *ə. PMP/PSWSAB *a is generally reflected as *a in the various modern Southwest Sabah languages, except under three conditions.

First, in the pronouns *aku, *kai, *kau, and *[ki]ta-kau, PSWSAB *a became *ə in Proto-Dusunic and Proto-Bisaya-Lotud, but not in Proto-Paitanic.

Second, in Proto-Dusunic, the *a of the prefixes *ma-, *maki-, *maN-, and *maG- became *ə, except where the vowel of the following syllable was /a/, in which case it is reflected as /a/.²² Conversely, in Proto-Murutic, the *a of these prefixes became *a, except where the vowel of the following syllable was /o/, in which case it became /o/. Tables 18 and 19 illustrate the reflexes of PSWSAB *a in the various branches of the Southwest Sabah subgroup. Note that segments in parentheses in the pronouns in table 18 are present in the Topicalized Nominative forms but not in the Short Nominative forms. Where two separate pronoun forms are given, the first is the Topicalized Nominative form, while the second is the Short Nominative form.

Finally, as illustrated in table 19, *a became *ə in closed penults and in all prepenultimate syllables in Proto-Dusunic, Proto-Bisaya-Lotud, and Proto-Paitanic, unless the vowel of the following syllable was *a.

2.8 DUSUNIC *nu ‘2SG.GEN’. In spite of not being an innovation, it is noteworthy that the distribution of second person singular genitive pronouns *mu vs. *nu almost perfectly matches the innovation-defined subgroupings of Southwest Sabah languages. In northern Borneo, *nu is found in Bonggi, in Dusun Puawang (a Paitanic language in spite of its name), and in all Dusunic languages except Dumpas. The form *mu, on the other

22. It is unclear whether *a never shifted to *ə in this environment, or if there was a universal shift of *a > *ə in these prefixes, which was later “erased” by a subsequent shift of *ə > /a/ when the vowel of the following syllable was *a, as described in 2.6.

hand, is found in all other languages in northern Borneo, except for the Idaanic languages (Idaan, Begak, Sungai Seguliud, and Subpan) which have lost the PMP genitive pronouns altogether. Dumpas may have adopted the form *mu* under influence from multiple languages with which it is or has been in contact, such as Sungai Beluran and other Paitanic languages, as well as Tidung and Tausug, all of which reflect **mu* instead of **nu*. Likewise, Dusun Puawang may have borrowed *mu* from neighboring Dusunic languages such as Rungus and Sonsogon. Considering this, it is noteworthy that **nu* can be reconstructed

TABLE 18. REALIZATIONS OF **a* IN VARIOUS PRONOUNS IN DUSUNIC AND MURUTIC LANGUAGES

PMP	PDUS * <i>a</i> > / <i>o</i> /				PMUR * <i>a</i> > / <i>o</i> /
	1SG.NOM	1EXCL.NOM	1INCL.NOM	2PL.NOM	2SG.NOM
	* aku	* kami	* kita+kamu	* kamu	* ika[h]u
PSWSAB	* aku	* ə-kai	* [ki]ta+kau	* kau	* (iə)ka[w]
PBisLO	* əku	* (iə)kəy	* [i]təkəw	* (iə)kəw	* ikaw
DBRU	kuij?	(jami? < OBL)	(jati? < OBL)	(muyun < OBL)	ikow
BISLI	aku	(jami? < OBL)	(jati? < OBL)	(muyun < OBL)	ikaw
BISSA	oku	okoy	tokow	okow	ikow
LOTUD	oku	ikoy	itokow	ikow	ikaw
PDUS	* əku	* i-(iə)kəy	* [i]-təkə[w]	* i-ka(wy)u, *kəw	* -ka[w]
DMEM	zaʔu, oku	ziʔoy, ikoy	zitokow	ziyazu, kow	ziyaw
DKMS	joʔu, oku	jiʔoy	tokow	jiyoju, kow	jiʔaw
DPPR	zow, oku	ziʔoy, koy	(zi)tokow	ziyozu, kow	ziʔaw
KDZOK	zoʔu, oku	ziʔoy, okoy	tokow	zozu, kow	ziyaw
KDZPE	(i)zou	(i)zikoy	(izo)tokow	iziyozu, kow	iziyaw
MKAK, MKOK	(y)oku	ikoy	tokow	ikovu, kow	ika
SONS	(y)oku	(y)okoy	toko	ikow, kow	ika
DTGS	(y)oku	()koy	itoko	ikow, kow	ikaw
DKRG	(y)oku	(y)okoy	(i)tokow	ikoo, kow	ikaw
RUNG	(y)oku	(y)okoy	(i)tokow	(i)kow	ikaw
DTMB	(iy)oku	koy	tokow	(i)kow, ikowu	ika
KUJAU	iyoku	(i)koy, yokoy	(i)tokow	(i)yoju, kow	ikaw
KLIAS, DBFT	oku	ikoy	(i)tokow	ikovu, kow	ikaw
DMPS	oku	ikoy	toko	ikow, kow	ikaw
DTLT	oku	okoy	tokow	ikou, kou	ika
DTD	oku	{yahay}	toko	yokoyu, kou	iya
DTOB	(iy)oku	(i)koy	tokow	ikoyu, kow	ikaw
PGMUR	* aku	* akay	* [i]takaw	* aka[w]	* oko[w]
PAPAR	aku	akay	{kitaw}	akaw	okow
TAT, GANA, MNAB, MTIM, TING, MSEM, BLSU	aku	akay	(i)takaw	akaw	okow
MBOK	aku	akay	(i)takaw	aka	oko
MPAL	aku	akay	(i)takaw	akaw, ka	oko
MTAG	au	akay	(i)taka(w)	akaw, ka	oko
MDAL	aku	akay	(i)takaw	akaw	oko
KOLOD	au	akay	itaka	akaw	oko
MKAL	aku	akay	takaw	kaw	—
ABSK, MSEL	au	akay	taka	akaw	oko
ABTU	ou ~ oo	axay	taxa	oxow	oxo
TIDUNG	aku	(a)kay	taka	kaw	—

TABLE 19. *a- > ə- IN THE BRANCHES OF SOUTHWEST SABAH

	PSWSAB	PDUS	PBISLO	PPAIT	PMUR
PENULT	*[ta]ʔəmis ‘sweet’	*əmis	*ma[ta]ʔəmis	*əmis	*amis
	*paʔit ‘bitter’	*pəʔit	*pəʔit	*pəit	*paʔit
	*daʔun ‘leaf’	*rəʔun	*rəʔun	*rəun	*daʔun
	*asin ‘salty’	*əsin	*əsin	*asin	*asin
	*paʔə ‘thigh’	*pəʔə	*pəʔə	*paa	*paʔa
PREPENULT	*adiRi ‘ladder’	*ərigi	*ərigi	*əndii	*arihi
	*taliŋa ‘ear’	*təliŋə	*taliŋə	*təliŋə	*taliŋə
	*aninipət ‘firefly’	*əninipət	*ə(nd)i(nd)ipət	*əninipət	*a(nd)i(nd)ipət
	*dapuR-an ‘stove’	*rəpuhan	*rəpuhan	(dəmpuran)	*dapu[ʔan
_CC	*sandiR ‘lean on’	*səndiw	*səndiw	*səndiw	*sandig
	*kəndiw ‘eagle’	*kəndiw	*kandiu	*kəŋ[d]iw	*kanduy
	*alsəm ‘sour’	*ənsəm	*ənsəm	*əsəm	*o[n]som
	*bəŋkiŋ ‘bedbug’	*wəŋkiŋ	*bəŋkiŋ	*bəŋkiŋ	*bəŋkiŋ
	*ənduʔ ‘woman’	*ənduʔ	—	—	*anduʔ
	*ə[n]tut ‘flatulence’	*əntut	*əntut	*ə[n]tut	*antut

only for Proto-Dusunic, but not for any other intermediate subgroup within Southwest Sabah, while *mu can be reconstructed for Proto-Greater Murutic, Proto-Paitanic, and Proto-Bisaya-Lotud, but not for Proto-Dusunic.

3. MURUTIC INNOVATIONS. Having discussed the innovations that define the Dusunic subgroup, we will now discuss those that define Murutic. Note that the Murutic reflexes of *R, *ə, and *a have been discussed in 2.2, 2.6, and 2.7, respectively.

3.1 PSWSAB *aw/*ay > MURUTIC *ow/*oy. In the Greater Murutic languages, earlier *aw and *ay shifted to *ow and *oy, respectively, as illustrated in table 20. Note that this most likely occurred as centralization of the *a of *aw and *ay to *ə, and then the shift of the *ə of *əw and *əy to /o/, yielding /ow/ and /oy/, respectively. Note that this shift also occurred in the Paitanic languages, and is a common shift, also found in some members of the Manobo, Subanen, Palawan, and Mongondow-Gorontalo subgroups.

3.2 PSWSAB *iw > MURUTIC *uy. Although there are only a few examples in the data, it appears that in all Murutic languages, including Papar (but not Tatana), earlier

TABLE 20. *aw/*ay > *ow/*oy IN MURUTIC

PSWSAB	PDUS	PBISLO	PPAIT	PMUR
*daŋaw ‘handspan’	*raŋaw	*raŋaw	(*raŋan)	*daŋow
*adaw ‘day/sun’	*adaw	*adaw	—	*odow
*(əi)-ka[w] ‘you (2SG.NOM)’	*i-ikaw	*ik(aə)w	*(əi)kaw	*oko[w]
*ə-kau ‘you (2PL.NOM)’	*i-kə(wy)u	*(iə)kəw	*kau	*aka[w]
*takaw ‘steal’	*takaw	*takaw	*takəw	*takow
*uway ‘rattan’	*uway	*uway	—	*owoy
*buway ‘long (time)’	*buway	*buway	—	*buwoy
*kusay ‘man’	*kusay	—	*kusəy	*kusoy
*balay ‘house’	*walay	*walay	*waləy	*baloy
*mamatay ‘kill’	*mamatay	*mamatay	*məmatəy	*mamatoy

*iw merged with *uy, as illustrated in table 21. As a result, while Dusunic and Paitanic have both /iw/ and /uy/, Murutic only has /uy/. Note that, apparently independently, Limbang Bisaya has also undergone this shift: for example, PSWSAB *suliR ‘floor’ > Limbang Bisaya *sihuy*, Brunei Dusun *siliw*.

3.3 PSWSAB ADJECTIVAL *g- > MURUTIC Ø-. Another innovation found in all of the core Murutic languages and in Papar (but not in Tatana) is the deletion of the root-initial *g- of adjectives when prefixed with *ma-, as illustrated in table 22. Prentice (1974) also notes that the same phenomenon affects *b when prefixed with *ma-, but examples of this are lacking in my data set.

3.4 PSWSAB NONFINAL *ə > MURUTIC *a. As noted in 2.6, PSWSAB *ə became PMUR *a in nonfinal syllables unless the vowel of the final syllable was *o (that is, *ə in PSWSAB), as illustrated in table 23.

4. DISCUSSION. We can now discuss the classifications of the Southwest Sabah languages in light of the above evidence. Table 24 presents an overview of the major innovations found in the Greater Dusunic and Greater Murutic subgroups as discussed in sections 2 and 3. Note that innovations of more limited distribution are not included, as these do not affect the subgrouping argument at the higher levels under discussion here.

It is clear from table 24 that nine innovations define the Dusunic subgroup (four of which are shared with Paitanic, and six or possibly seven with Bisaya-Lotud languages), while seven define the Murutic subgroup (two of which are not shared with Tatana, and one of which is not shared with Papar). Thus, the core Murutic languages, along with Tatana and Papar, can be grouped together as three coordinate branches of the Greater Murutic subgroup, as illustrated in figure 2.

Likewise, the Dusunic, Paitanic, and Bisaya-Lotud languages can be grouped together as three coordinate branches of a Greater Dusunic subgroup, as illustrated in figure 3, based on the fact that Dusunic and Bisaya-Lotud share at least six innovations,

TABLE 21. *iw > *uy IN MURUTIC

	PSWSAB	PMUR	PDUS	PBISLO	PPAIT	TATANA	PAPAR
*iw	*kəgiw ‘orangutan’	*kaguy	*kəgiw	—	*kəgiw	kagiw	—
	*kəndiw ‘eagle’	*kanduy	*kəndiw	*kəndiu kanuy (BISLI)	*kən[d]iw	kandiw	kanduy
	*baliw ‘move s.t.’	*baluy	*waliw	aluy (BISLI)	—	baliw	baluy
*uy	*hapuy ‘fire’	*apuy	*apuy	*apuy	*apuy	apuy	apuy
	*ləbuy ‘float’	*labuy	—	lobuy (BISSA)	—	labuy	la?buy

TABLE 22. *-g- > Ø IN MURUTIC ADJECTIVES

PGMUR	PMUR	PAPAR	TATANA
*ma-gayo ‘big’	*ma-ayo	mayo	magayo
*ma-galud ‘far’	*ma-alud	malud	magalud
*ma-ga?ad ‘near’	*ma-a?ad	ma?ad	magaad
*ma-gawad ‘long’	*ma-awad	mawad	magawad

TABLE 23. REFLEXES OF *ə IN SOUTHWEST SABAH

PSWSAB	PGMUR	PDUS	PPAIT	PBISLO	PIDAAN
*ə > a / _C(aiu)					
*kau '2PL.NOM'	*aka[w]	*i-kə(wy)u	*kau	*(iə)kəw	—
*(ə)kai '1EXCL.NOM'	*akay	*i-(iə)kəy	*kai	*(iə)kəy	*k(əu)mmi
*ətud 'knee'	*atud	*ətud	*ətud	*ətud	—
*əbuk 'hair'	*abuk	*əbuk	*əbuk	*əbuk	*əbpuk
*[]əRis 'sand'	*agis	*əgis	*əgis	*əgis	*bəris
*ə[n]tut 'flatulence'	*antut	*əntut	*ə[n]tut	*əntut	*[ot]tut
*səbu 'urinate'	*sabu	*səbu	*səbu	*səbu	*sidu
*təbu 'sugarcane'	*tabu	*təbu	*təbu	*təbu	*təbpu
*pənu? 'full'	*panu?	*pənu?	*pənu?	*pənu?	—
*bəli 'buy'	*bali	*bəli	*bəli	*bəli	—
*bəti 'calf of leg'	*bati	*wəti	*wəti	—	*b(əi)ti
*kəgut 'burnt rice'	*kagut	*kəgut	—	*kəgut	—
*bəgas 'uncooked rice'	*bagas	*wagas	*wagas	*wagas	*bəgkas
*ə > o / _Co					
*ədaw 'day, sun'	*odow	*ədaw	—	*ədaw	*mato'dtaw
*(iə)ka[w] '2SG.NOM'	*oko[w]	*i-ikaw	*(əi)kaw	*ik(əə)w	*ik(əə)w
*pədəs 'spicy'	*podos	*pədəs	*pədəs	*pədəs	*podos
*tələn 'swallow'	*tolon	*tələn	*tələn	*tələn	*tollon

Dusunic and Paitanic share four innovations, and three innovations are shared by all three of these subgroups.²³

The Greater Dusunic and Greater Murutic subgroups appear to form two branches of the Southwest Sabah subgroup, based on the five weak phonological innovations listed in 1.2.

This new subgrouping differs from that of King and King (1984) and/or Lewis (2009) regarding the position of several languages, including Tatana, Papar, Lotud, Gana, Dumpas, and Murut Serudung. Furthermore, the more geographically comprehensive approach to the current study allows the following languages to be incorporated into the Southwest Sabah subgroup: (1) Abai Sembuk and Abai Tubu, in the northern part of the Indonesian province of Kalimantan Timur; (2) Bulusu, also in northern Kalimantan Timur; (3) Limbang Bisaya, in the Limbang district of northern Sarawak, Malaysia; and (4) the three dialects of Brunei Dusun (upland, lowland, and coastal).

4.1 TATANA AND PAPAR. In King and King (1984), Tatana is grouped with Sabah Bisaya in a Southern Dusun branch of the Dusunic subgroup, and Papar is included as an unclassified branch of the Dusunic subgroup. That these two languages would be mistaken for Dusunic languages based on lexicostatistics and intelligibility testing is understandable, considering their location and their sociolinguistic situation. Both Tatana and Papar occupy small patches of territory toward the northern tip of Kuala Penyu town, and are proximate to and in frequent contact with speakers of Sabah Bisaya

23. Blust (pers. comm., September 6, 2012) points out that the innovations mentioned in this paper suggest that Dusunic and Bisaya-Lotud share a common node apart from Paitanic, but in the absence of a larger set of data representing the Paitanic languages, the current author considers it safer to reserve judgment on the internal structure of the proposed Greater Dusunic subgroup until such time that a more extensive study can be done.

**TABLE 24. MAJOR INNOVATIONS IN SOUTHWEST SABAH
(with key to shared innovations)**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
BisLO																	
BISAYA SABAH	-	+	+	-	+	+	+	+	-	-	-	(+)	-	-	-	-	+
BISAYA LIMBANG	-	+	+	(+)	+	+	+	?	-	-	-	-	-	-	-	-	+
DUSUN BRUNEI	-	+	+	(+)	+	+	+	?	-	-	-	(+)	-	-	-	-	+
LOTUD	-	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-
DUSUNIC																	
RUNG	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
DTD	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
DKRG	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	+
DTOB	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	+
DTGS	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	+
DTLT	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	+
MKOK	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	+
DMEM, DKMS	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
KUJAU	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
KDZPE	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	+
DPPR	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
KLIAS	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	+
DMPS	+	+	+	+	+	+	+	+	(-)	-	-	-	-	-	-	-	+
PAITANIC																	
	-	+	-	+	+	-	-	-	-	+	-	-	-	-	-	-	+
GR. MURUTIC																	
PAPAR	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-
TAT	-	-	-	-	-	-	-	-	-	+	+	+	-	+	+	+	+
GANA	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
MNAB	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
MBOK	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
MTim	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
MPAL	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
MTAG	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
KOLOD	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
MKAL	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
MSEL	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
MSEM	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
ABAI (BOTH)	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
BLSU	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
TIDUNG (ALL)	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+

Key to table 24:

- | | |
|---|-----------------------------------|
| 1) three replacement pronoun innovations (cf.2.1) | 10) *R > h~Ø / _i |
| 2) *R > w / _# | 11) *-R > *-g |
| 3) *R > g / _i | 12) *aw, *ay > *ow, *oy |
| 4) *ma- > a- on consonant-initial roots | 13) *iw > *uy |
| 5) *b/*w split | 14) *g- > Ø after Adjectival *ma- |
| 6) *d/*r split | 15) *a > *a / _C(aiu) |
| 7) *a > a / _Ca | 16) PMUR *-in 'Location Focus' |
| 8) *a > *a in four pronoun forms | 17) PDUS *h > Ø or PMUR *h > Ø |
| 9) *nu '2SG.GEN' | |

FIGURE 2. THE PRIMARY BRANCHES OF GREATER MURUTIC

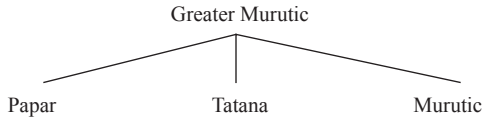
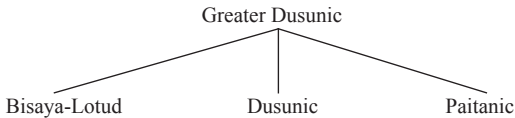


FIGURE 3. THE PRIMARY BRANCHES OF GREATER DUSUNIC



and various Dusunic groups. In fact, it is highly unlikely that there are currently any remaining adult speakers of Papar who have not acquired an understanding of Tatana and Sabah Bisaya, and it is likewise unlikely that there are many adult speakers of Tatana who have not acquired at least a passive understanding of Sabah Bisaya. Because of this, borrowing would be expected to have an effect on the lexicostatistical scores, and contact would likewise have an effect on the intelligibility tests. As can be observed from table 24, however, the phonological and morphological innovations clearly indicate that Tatana belongs with Murutic and not with Dusunic, and likewise indicate that Sabah Bisaya cannot be considered a member of the Murutic subgroup.

4.2 GANA. The Gana language is listed in *Ethnologue* (Lewis 2009) as belonging to the Dusunic subgroup, even though Smith (1984) classified it as Murutic. As can be observed from table 24, Gana clearly shares the innovations of the Murutic subgroup but not any of those that are characteristic of the Dusunic languages. Lexical similarity between Gana and the Dusunic languages is likely explainable by the fact that Gana and Kujau, a Dusunic language, usually occupy either the same community or neighboring communities throughout Keningau town.

4.3 LOTUD. Under the present analysis, Lotud belongs to a Bisaya-Lotud subgroup along with Sabah Bisaya, Limbang Bisaya, and Brunei Dusun. While the Lotud themselves have no oral history of a connection to the Sabah Bisaya, or of any significant role in the area, they consider themselves distinct from neighboring Dusun groups, and their geographical location seems to support this: the Lotud occupy the areas around the town center of Tuaran, just north of Kota Kinabalu, the present capital of Sabah. Their domain does not extend to the coastal areas, however, as the arrival of more recent immigrant groups such as the Sama (called *Bajau* in Sabah), Iranun, and Tausug (called *Suluk* in Sabah) has prompted formerly coastal populations to move further inland in many parts of Sabah.

That lexicostatistics indicated a particularly close relationship between Lotud and the Dusunic languages is explainable by the fact that Lotud is surrounded by, and in constant contact with, Dusun dialects (Tindal, Kiulu, Bukid) that surround its relatively small territory. As such, it would be expected that its neighbors would have had a considerable amount of lexical and phonological influence on Lotud. However, its pronouns are virtually identical to those of Sabah Bisaya, and even more conservative in some areas (for

example, the retention of a distinct 1st person exclusive genitive form, *nyami*, where Sabah Bisaya has generalized the oblique form *jami* to the genitive), and like other languages in the Bisaya-Lotud group, it shares none of the Dusunic pronominal innovations. It is likewise the most phonologically conservative language in this group, retaining both /h/ (< PSWSAB *h and *R in certain environments) and intervocalic /ʔ/, as well as the genitive common noun case marker *nu, which has been lost in the other Bisaya-Lotud languages and in Dusunic. As can be observed from table 24, Lotud also shares seven innovations with other Bisaya-Lotud languages, and shares none of the diagnostic Murutic innovations.

4.4 DUMPAS. King (1984:235) claims that “it is apparent that the Dumpas language has a closer relationship with the Paitanic languages ... than ... with the Dusunic languages,” and that “the historical roots of this language ... are most likely Paitanic.” A paragraph later, King even speculates that “it is possible that further testing would reveal Dumpas to be a dialect of one of the Paitanic languages or a separate language within the Paitanic family.” To the contrary, Dumpas appears to be a close relative of the Dusun language called Sungai Karamuak or Sukang, and local historian Sanen Marshall (pers. comm., July 18, 2011) relates a report by Shim (2007:103–5) that independently confirms this:

During their 600–700 years of settlement in the Kinabatangan, three groups split off from the Sukang tribe. These became the Dumpas, Mangkaak and Gunatong. ... Only the Dumpas migrated down the Kinabatangan and overland to the Labuk. It was because of their close association with the downstream tribes that they were mistaken for *Paitanics* or Orang Sungai.

It used to be common practice amongst the Kadazandusuns for a person susceptible to frequent illness to change his or her name in the hope that the illness would not follow the person with the new name. ... This practice was followed when some Sukang started the Dumpas tribe according to the following legend ...

After they had fled Bukit Linggang, Batulong thought that the tribal name, Sukang, brought only sickness and bad luck so he decided to change the name in the hope that their luck would change. He called the survivors Dumpas and his followers gave him the name Raja Tua Batulong.

During their wanderings, the Dumpas mixed and intermarried with other tribes. As a result, their dialect is now slightly different from the Sukang dialect.

Historical research and oral history, therefore, corroborate the phonological and functor evidence, that Dumpas is a close relative of the language of the Sukang or “Sungai Karamuak” people. Its Paitanic features were probably acquired after moving downriver toward its present location north of Beluran town, where it is surrounded by Paitanic languages such as Sungai Paitan, Tombonuwo, Lingkabau, and Sungai Beluran. Complicating the picture even further are likely contacts with Tidung—speakers of which have been in the Beluran (formerly “Labuk-Sugut”) district since the 1800s—and with Tausug, the influence of which in the area dates back well over 500 years. Again, the pitfalls of King and King’s methodology is apparent: that intelligibility testing demonstrated that the Dumpas could understand Paitanic languages is only natural, since the Dumpas have been surrounded by Paitanic languages for genera-

tions. Likewise, high lexicostatistical scores between Dumpas and Paitanic languages are not surprising, given borrowing between Dumpas and its neighbors; note, however, that King (1984:235) herself points out that “the initial lexicostatistical classification of Smith [(1984)] ... placed Dumpas with the Dusunic languages.”

4.5 THE MURUTIC LANGUAGES OF KALIMANTAN TIMUR, INDONESIA. Other than several varieties of Tidung (besides the two or three found in Sabah), there are a half-dozen Murutic languages that are found primarily or wholly in the Indonesian province of Kalimantan Timur just south of the border with eastern Sabah: Kolod, Murut Sembakung, Tingalan, Abai Sembuak, Abai Tubu, and Bulusu. The first three are located near the border between Kalimantan and Sabah, and were mentioned at least in passing in King and King (1984), but the second three were overlooked. Abai Sembuak, Abai Tubu, and Bulusu represent the southernmost extreme of the Murutic subgroup, but are still very clearly and uncontroversially Murutic languages. They are distributed in areas to the west and south of Malinau town, although like many of the tribes in Kalimantan, they report that they originated further uphill or upriver a generation or so ago, before the Indonesian government persuaded them to move downhill or downriver closer to established towns.

4.6 LIMBANG BISAYA AND BRUNEI DUSUN. Limbang Bisaya and Brunei Dusun represent the closest relatives of Sabah Bisaya and Lotud, rounding out a group of languages spoken by a people known to have had close sociopolitical interaction with the Sultanate of Brunei centuries ago (Okushima 2003:238), and rumored in and around Brunei to have even had a genealogical relationship to the first Sultan of Brunei.²⁴ Most of the remnants of this population are found around Brunei Bay, but their closest relatives, the Lotud, are found in Tuaran town further north. The location of the Lotud near politically important Kota Kinabalu is likely no coincidence, and they quite possibly controlled the coastal areas in the vicinity of Kota Kinabalu and Tuaran until the Iranun arrived from the Philippines three or four centuries ago, pushing the Lotud further inland.

4.7 PAITANIC. The Paitanic languages, while apparently part of the Greater Dusunic subgroup, are problematic because they do not always agree in reflecting certain innovations, or do not reflect them consistently. In many cases, this is quite possibly because the evidence has been obscured by borrowing from neighboring languages, usually whichever member of the core Dusunic subgroup that a particular Paitanic language has been in historic contact with. Further work on the Paitanic subgroup, including reconstruction of its historical phonology, is needed before its position can be determined with a greater degree of confidence.

4.8 THE POSITION OF MURUT SERUDUNG. One minor issue related to the current discussion is the position of a language called Murut Serudung, originally spoken in Serudung Laut (also called “Serudung Lama” [‘Old Serudung’] by speakers in

24. Note that the Sabahan Bisaya are not the only claimants to this relationship, as the Iranun of Sabah and a number of groups in the Philippines also claim a genealogical connection to the Sultan of Brunei. However, the geographical proximity of the Sabahan Bisaya makes their claim seem at least more likely than the claims of other groups.

Tawau) south of Kalabakan near the border with Kalimantan Timur, but also spoken by descendants of a group that migrated to the community of Serudung Baru ('New Serudung') in Tawau town. Evidence from the pronouns and demonstratives, along with suspect reflexes of *R, indicate that this language does not belong to the Murutic subgroup, but subgroups with the Paitanic languages. Table 25 lists the pronouns of Murut Serudung, which can be compared to the Proto-Paitanic pronouns in table 8 and the Proto-Murutic pronouns in table 7. While there are admittedly few significant differences between the pronouns of Proto-Murutic and Proto-Paitanic, it is noteworthy that Murut Serudung reflects the Proto-Paitanic forms *kai '1EXCL.NOM', *sirə '3PL.NOM', and *niyə '3SG.GEN', as well as sharing the formation of the Oblique pronouns with the *sa- ~ *sə-formative plus the Nominative base, a formation that is quite widespread in the Philippines but not found elsewhere in Southwest Sabah outside of the Paitanic subgroup. Elsewhere, in the verb system, Murut Serudung also reflects the Proto-Paitanic reciprocal action prefix *mu-, which is otherwise not found in the Murutic languages.²⁵ The presence of a number of Proto-Paitanic and Proto-Greater Dusunic lexical innovations in Murut Serudung is also noteworthy, especially since no Dusunic or Paitanic language is spoken anywhere near either Serudung Laut or Serudung Baru.

Note that the '2PL.NOM' form *kuwo* is an innovation unique to Murut Serudung, as is the vowel assimilation in the forms *kee* '1EXCL.NOM', *mee* '1EXCL.GEN', *sekee* '1EXCL.OBL', and *seeyo* '3SG.OBL'. Although *kuwo* '2PL.NOM' is an innovation, it still resembles the Proto-Paitanic 2PL.NOM form *kau, which lacks the initial vowel found in synonymous forms in the Dusunic, Bisaya-Lotud, and Greater Murutic languages.

4.9 THE IDAANIC LANGUAGES. The Idaanic languages—including Idaan, Begak, Subpan, Sungai Segulud, and the elusive "Buludupi"—do not generally share in the Southwest Sabah phonological or morphological innovations, and are lexically, phonologically, and grammatically quite distinct. It is likely that, as Blust (2010:46) believes, the Idaanic languages are coordinate with the Southwest Sabah and North Sarawak subgroups in a North Borneo macrogroup. However, I differ from Blust in placing Bonggi with Molbog, based on functor evidence that Blust himself admits is "difficult to evaluate", "unusually challenging", and "would normally be sufficient to support an argument

TABLE 25. MURUT SERUDUNG PRONOUNS

	NOM	GEN	OBL
1SG	aku	ku	saaku
2SG	okow, =ko	mu	sookow
3SG	iyo	nyo	seeyo
1EXCL	kee	mee	sekee
1INCL.DU	toduwo	toduwo	sitoduwo
1INCL.PL	taka	taka	sitaka
2PL	kuwo	muyu	sakuwo
3PL	siro	niro	siro

25. All Dusunic languages except Dumas reflect Proto-Dusunic *mi- 'Actor Focus reciprocal action'. Dumas has replaced earlier *mi- with *mu-*, unsurprising considering the fact that it has been surrounded by Paitanic languages for at least a century.

that Molbog and Bonggi form a node” (2010:66–67), and he predicts that “this case will surely challenge scholars for some time to come” (2010:68).

4.10 TIDUNG.²⁶ While certainly deserving of a careful study of their own, for the purposes of this paper it is sufficient to simply state that the Tidung languages clearly belong to the Murutic subgroup, as they share all of the diagnostic Murutic innovations, and none of the Dusunic innovations. This may come as a surprise to those whose only background in the Tidung languages is Beech (1908), where two alleged varieties of Tidung are presented: the Tarakan dialect of Tidung, and the Bulungan language. Unfortunately, Bulungan does not appear to be closely related to Tidung, to the Murutic subgroup,²⁷ or to Southwest Sabah in general, and Tidung Tarakan turns out to be the least ideal variety of Tidung for comparative purposes, as it is the least conservative Tidung variety, having lost most of the Philippine-type structure still found in the Tidung varieties further north as well as in those upriver in Kalimantan Timur (of which the Bangawong, Sambal/Sombol, Kuala Merotai, Kalabakan, Nunukan, Malinau, and Mansalong varieties are also represented in my data).²⁸ A more complete set of data representing more Tidung varieties reveals much more of the Murutic nature of Tidung than data from only Tidung Tarakan would. In fact, the only thing surprising when traveling from north to south is just how different Tidung Tarakan is from the other varieties of Tidung, still quite clearly belonging to the same subgroup, but having undergone phonological, morphological, and structural shifts that are not found in the Tidung varieties further north and further upriver.

It is worth noting that Prentice (1971:375) likewise places Tidung in the Murutic subgroup, although his data on Tidung languages and other Murutic languages were much more limited than those of the current author.

5. CONCLUSION. This paper has attempted to fill a long-standing void in the use of phonological and morphological innovations to determine the relationships of the Southwest Sabah languages, that is, the languages traditionally assigned to the Dusunic, Murutic, and Paitanic subgroups. At the same time, a more geographically inclusive approach has been taken, surveying not only those languages located in Sabah, but also languages in adjacent parts of northern Sarawak, Brunei, and northern Kalimantan Timur. The result is a more accurate picture of the membership of these subgroups than was previously achieved with the methodologies of lexicostatistics and intelligibility testing, the former often discredited as unreliable, and the latter never shown to prove anything other than that members of two language communities can understand one other to some extent, without adequately considering the reasons why. In fact, lexicostatistics and

26. While previous authors have used the spelling “Tidong,” the proper spelling according to all Tidung groups visiting in Malaysia and in Indonesia is “Tidung.” There is even a minimal pair in all known dialects of Tidung between *Tidung* /tiduŋ/ ‘Tidung’ and *tidong* /tidonŋ/ ‘mountain’.

27. This is based on the analysis of my own data, but Moody (1984:127) similarly notes that “Prentice says this is neither [Tidung] nor Murutic.”

28. The Tidung of Tarakan can easily recite a list of 20 or so varieties of Tidung, although one gets the impression that these are geographical designations, not linguistic distinctions. There has been so much intermingling of various Tidung groups, as Okushima (2003) also reports, that linguistic fieldwork on each group at present would be rather difficult, if not impossible.

intelligibility testing prove to be even less accurate than one's quick impressions after looking at phonological, functor, and lexical data for these languages.

Some languages are prototypically Dusunic or Murutic, and for these it makes little difference which method is used. Not surprisingly, a number of these languages were accurately classified under previous approaches: speakers of Dusun Tambunan and Dusun Tindal can largely understand each other, as can speakers of Murut Paluan and Murut Tagol, as they will tell you. For the languages on the fringes, however, and especially those that have long been in contact with languages from other subgroups, these approaches fare much worse: these languages, including Lotud, Tatana, Papar, Sabah Bisaya, Dumpas, Gana, and Murut Serudung require a much more reliable methodology, such as the use of phonological and morphological innovations. These innovations, as shown in table 24, clearly indicate the position of these languages in a way that lexicostatistics and intelligibility testing never could. It is likewise no less significant that the historical record, including oral histories, corroborates the innovation-based subgrouping, with Shim (2007) noting the historical relationship between the Dumpas and Sukang, and Okushima (2003) noting a historical relationship between the Tidung and Tatana, not to mention the local belief among the Tidung that they are linguistic cousins of the Murutic-speaking peoples of Sabah and the northernmost extremes of Kalimantan Timur.

While it is expected that the subgrouping presented in this paper will hold up to further scrutiny, a number of issues still need to be resolved, not the least of which is the internal structure of each subgroup, and especially the Bisaya-Lotud and Paitanic languages, which appear to have been influenced particularly heavily by members of neighboring subgroups.

REFERENCES

- Beech, Mervyn W. H. 1908. *The Tidong dialects of Borneo*. Oxford: The Clarendon Press.
- Blust, Robert A. 1977. The Proto-Austronesian pronouns and Austronesian subgrouping: A preliminary report. *University of Hawai'i Working Papers in Linguistics* 9(2):1–15.
- . 1981. *Variation in retention rate among Austronesian languages*. Paper presented to the Third International Conference on Austronesian Linguistics, Bali, Indonesia.
- . 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30:73–129.
- . 1992. On speech strata in Tiruray. In *Papers in Austronesian Linguistics 2*, ed. by Malcolm D. Ross, 1–52. Canberra: Pacific Linguistics.
- . 1998. The position of the languages of Sabah. In *Pagtanaw: Essays on language in honor of Teodoro A. Llamzon*, ed. by Ma. Lourdes S. Bautista, 29–52. Manila: Linguistic Society of the Philippines.
- . 2000. Why lexicostatistics doesn't work: The 'universal constant' hypothesis and the Austronesian languages. In *Time depth in 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.
- . 2009. *The Austronesian languages*. Canberra: Pacific Linguistics.

- . 2010. The Greater North Borneo hypothesis. *Oceanic Linguistics* 49:44–118.
- Buck, Gill P. 2009. Tobilung language lessons. Unpublished MS.
- Burton, Scott L. 1996. A case study of lexical borrowing between language families: The East Mindanao and Manobo languages. MA thesis, University of Texas at Arlington.
- Goudswaard, Nelleke. 2005. The Begak (Ida'an) language of Sabah. PhD diss., Free University of Amsterdam. Utrecht: Landelijke Onderzoekschool Taalwetenschap (Netherlands Graduate School of Linguistics).
- Grace, George W. 1964. The linguistic evidence. *Current Anthropology* 5(5):361–68.
- . 1992. How do languages change? (More on “aberrant” languages). *Oceanic Linguistics* 31:115–30.
- Himes, Ronald S. 2002. The relationship of Umiray Dumaget to other Philippine languages. *Oceanic Linguistics* 41:275–94.
- King, Julie K. 1984. The Dumpas language. In *Languages of Sabah: A survey report*, ed. by Julie K. King and John Wayne King, 231–35. Canberra: Pacific Linguistics.
- King, Julie K., and John Wayne King. 1984. *Languages of Sabah: A survey report*. Canberra: Pacific Linguistics.
- Kroeger, Paul R. 2005. Kimaragang. In *The Austronesian languages of Asia and Madagascar*, ed. by Sander Adelaar and Nikolaus Himmelmann, 397–428. London/New York: Routledge.
- Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the world*, 16th ed. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com/>.
- Lobel, Jason William. 2010. Manide: An undescribed Philippine language. *Oceanic Linguistics* 49:480–512.
- McFarland, Curtis D. 1974. The dialects of the Bikol area. PhD diss., Yale University.
- Moody, David C. 1984. The Tidong language. In *Languages of Sabah: A survey report*, ed. by Julie K. King and John Wayne King, 125–37. Canberra: Pacific Linguistics.
- Okushima, Mika. 2003. Ethnic background of the Tidung: Investigation of the extinct rulers of coastal northeast Borneo. *Journal of Sophia Asian Studies* 21:233–60.
- Pallesen, A. Kemp. 1985. *Culture contact and language convergence*. Manila: Linguistic Society of the Philippines.
- Pawley, Andrew. 1999. Chasing rainbows: Implications of the rapid dispersal of Austronesian languages for subgrouping and reconstruction. In *Selected papers from the Eighth International Conference on Austronesian Linguistics*, ed. by Elizabeth Zeitoun and Paul Jen-kuei Li, 95–138. Taipei: Academia Sinica.
- Penroyer, F. Douglas. 1986–87. Inati: The hidden Negrito language of Panay, Philippines. *Philippine Journal of Linguistics* 17(2)–18(1):1–36.
- Prentice, D. J. 1971. *The Murut languages of Sabah*. Canberra: Pacific Linguistics.
- . 1974. Yet another PAN phoneme? *Oceanic Linguistics* 13:33–75.
- Ray, Sidney H. 1913. The languages of Borneo. *The Sarawak Museum Journal* 1(4):1–196.
- Ross, Malcolm D. 1991. How conservative are sedentary languages? Evidence from western Melanesia. In *Currents in Pacific linguistics: Papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, ed. by Robert A. Blust, 433–51. Canberra: Pacific Linguistics.
- . 2005. The Batanic languages in relation to the early history of the Malayo-Polynesian subgroup of Austronesian. *Journal of Austronesian Studies* 1(2):1–24.
- . 2006. Reconstructing the case-marking and personal pronoun systems of Proto Austronesian. In *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*, ed. by Henry Y. Chang, Lillian M. Huang, and Dah-an Ho, 521–63. Taipei: Academia Sinica.
- Rubino, Carl R. G. 2005. Utudnon, an undescribed language of Leyte. In *Current Issues in Philippine linguistics and anthropology: Parangal kay Lawrence A. Reid*, ed. by Hsiu-chuan Liao and Carl R. G. Rubino, 306–37. Manila: Linguistic Society of the Philippines and SIL-Philippines.

- Shim, P. S. 2007. *Inland people of Sabah: Before, during and after Nunuk Ragang*. Kota Kinabalu (Malaysia): Borneo Cultural Heritage Publisher.
- Smith, Kenneth D. 1984. The languages of Sabah: A tentative lexicostatistical classification. In *Languages of Sabah: A survey report*, ed. by Julie K. King and John Wayne King, 1–49. Canberra: Pacific Linguistics.
- Wolff, John U. 1967. History of the dialect of the Camotes Islands, Philippines, and the spread of Cebuano Bisayan. *Oceanic Linguistics* 6:63–79.
- Yamada, Yukihiro, and Shigeru Tsuchida. 1983. *Philippine languages*, 2nd ed. Asian and African Grammatical Manual 15b. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Zorc, R. David. 1974. Internal and external relationships of the Mangyan languages of Mindoro. *Oceanic Linguistics* 13:561–600.
- . 1977. *The Bisayan dialects of the Philippines: Subgrouping and reconstruction*. Canberra: Pacific Linguistics.
- . 1978. Functor analysis: A method of quantifying function words for comparing and classifying languages. *The Fifth LACUS Forum*, 510–21. Columbia, SC: Hornbeam Press.