

The Linguistic Macrohistory of the Philippines: Some Speculations¹

Robert Blust
University of Hawai'i

1. History and Macrohistory in Language Change

Historical linguistics is normally concerned with details of form-meaning pairing in particular languages as these relate to issues of reconstruction, change, subgrouping, borrowing and the like. Its focus is thus the history of individual language communities, a topic that can conveniently be called 'linguistic microhistory'. But there are larger issues of linguistic prehistory which can be approached entirely without reference to such details. The geographical distribution of language families and major subgroups, for example, provides information about the relative age of language groups in a given region, centers of dispersal ('homelands'), and prehistoric episodes of language extinction. Following Blust (1992) historical inferences founded upon these types of observations can conveniently be grouped together under the general rubric of 'linguistic macrohistory'.²

Blust (1992) adopted a distributional approach in outlining the linguistic macrohistory of mainland Southeast Asia, where five distinct language families are historically attested: 1. Austroasiatic (AA), 2. Sino-Tibetan (ST), 3. Austronesian (AN), 4. Tai-Kadai (TK), and 5. Hmong-Mien (HM). Documentary evidence tells us that Hmong-Mien (in earlier literature 'Miao-Yao') languages are newcomers which arrived from Guizhou in southern China during the past four or five hundred years, and that Thai and its closest relatives (Shan, Lao, etc.) moved south from Yunnan no earlier than the thirteenth century (Çoedes 1968:189ff). The historical evidence for the arrival of Austronesian speakers in mainland Southeast Asia is more sketchy, but the Indianized state of Champa (under the name 'Lin-yi') was recognized in Chinese dynastic records as early as 192, with oblique references to the same ethnic group dating back to 137 AD

¹ In the summer of 1966 an eager young undergraduate who had just been released from the U.S. Army and was relishing his newfound freedom enrolled in an introductory linguistics course at the University of Hawai'i taught by a brand-new Ph.D. This proved to be the start of a long and productive relationship. I want to thank Laurie for giving me my foundation in linguistics, and for his continuing valuable feedback as a colleague over these four decades, a period of time that has passed so quickly that it is difficult for either of us (young-at-heart as we both are) to believe or accept.

² Nichols (1992), who speaks of 'macroareas' and 'macrogender', but not of 'macrohistory', illustrates a typologically-oriented approach to issues of the kind addressed here. The variables with which she works are language families, typological features, geographical distributions and the like. At this level of abstraction linguistic forms — the basic material of the standard comparativist — hardly enter the picture at all. An earlier and rather different approach to issues of linguistic macrohistory which also is concerned primarily with explanations for the geographical distribution of diversity is seen in Sapir (1968 [1916]).

(Coedes 1968[1964]:43). Moreover, the evidence for a Malayo-Chamic subgroup of Austronesian languages allows us to infer that Chamic speakers probably reached mainland Southeast Asia through a major population movement out of southwest Borneo in the period 2,000–2,300 BP, with subsequent migrations to Hainan island and northern Sumatra following the fall of the northern Cham capital of Indrapura to the Vietnamese in 982 AD (Blust 1992, Thurgood 1999).

The relative antiquity of the two remaining language families in mainland Southeast Asia lies beyond the reach of historical documentation. Here methods appropriate to linguistic macrohistory take precedence. The following observations are relevant:

1. AA languages have a wide east-west extent from eastern India to the South China sea, the interior of the Malay peninsula, and the Nicobar islands.
2. The distribution of these languages is patchy, with many AA languages separated from their relatives by Tai-Kadai or ST speakers.
3. The westernmost (Munda) languages are only distantly related to those further east (Mon-Khmer).
4. Apart from a few islands of Mon (AA) within the territory of the Karen, an intrusive wedge of Shan (TK) speakers in upper Burma, and Garo (ST), which lies just east of the great bend of the Brahmaputra river in a region dominated by Assamese (Indo-Aryan) speakers, the ST languages of Southeast Asia occupy a more-or-less continuous geographical block, reaching from the Chittagong, Chin and Naga hills of eastern India, through Burma and southward to the upper quadrant of the Malay peninsula. Their east-west extent is far more restricted than that of the AA languages.
5. Many ST languages are spoken in the Himalayas and throughout China, while few AA languages reach as far north as southern China.

Together these observations support an inference that AA languages were in mainland Southeast Asia earliest, and that ST (or, more particularly, Tibeto-Burman) languages arrived in this area as a result of later migrations southward from the eastern Himalayas, presumably following the main river valleys. In so doing they split the AA territory into discontinuous eastern and western segments. Major population movements such as the splitting migration implied by the present distribution of TB languages in relation to AA languages would almost certainly have had dislocating consequences for some smaller or weaker groups. Benedict (1990) cites the puzzling case of Lai, a small isolated AA language currently spoken in Guangxi, but apparently having an ancestry further west in Guizhou and southern Yunnan. Apart from the expected Chinese loans, this language shows numerous TB loanwords. Some of these are quite basic ('sky/rain', 'ox/cattle', 'fowl', 'head', 'needle', 'run'), and most appear to have been borrowed very early, rather than more recently from the Lolo-Burmese languages with which the Lai are presently in contact. Although Benedict does not draw this conclusion, these observations are consistent with a scenario in which TB languages, expanding into Southeast Asia from the eastern Himalayas via the Irrawaddy drainage, displaced some AA groups southward into the Malay peninsula and the Nicobar islands, and scattered others northward and eastward as far as Guangxi at a fairly early time. The Lai appear to be the descendants of just such a fugitive group.

2. Resetting The Clock: Language Extinction and Its Distributional Consequences

Language change appears to be a more or less continuous process. As a result of change and separation language communities diverge over time until eventually the very fact of relationship can no longer be established with certainty. In general, greater separation time translates into greater divergence, although there is evidence that rates of change in basic vocabulary may vary significantly.

Diamond (1992) has coined the apt expression 'resetting the clock' for episodes of linguistic extinction which initiate a new cycle of linguistic differentiation with shallower time depth than would have been the case had extinction never occurred. A well-known example of this process is seen in the Italian peninsula, where about 2,500 years ago Etruscan, Latin, Faliscan, Sabellian, Oscan, Umbrian, and perhaps other languages were spoken within a fairly confined region of what is now central Italy. As a result of the political and military success of the Roman Empire the use of Latin expanded at the expense of the other languages of this area and beyond (Gaul, Iberia). Today, none of the five (or more) languages which were contemporaries of Latin circa 500 BC has left any descendants: apart from recent immigrant communities and some overlapping of German or South Slavic dialects in northern border areas, only Italian dialects or closely related languages are spoken in the Italian peninsula. In effect, then, the spread of Latin reset the linguistic clock in Italy, since without this event we would expect the present linguistic diversity of this part of Europe to be much greater than it is, given its known diversity some 2,500 years ago.

The development of Latin to the Romance languages is sometimes cited as a precious control on linguistic reconstruction, since it provides us with the rare opportunity of comparing a language inferred by use of the Comparative Method with historical records of the same language (recognizing, of course, that the two may represent different social registers). Less often appreciated is the fact that the linguistic macrohistory of the Italian peninsula also provides us with an invaluable model of how language distributions may be skewed by episodes of extinction — a model which can be applied to other areas of the world where historical documentation is far more scanty, or is lacking entirely.

As noted in Blust (2000b), the distribution of Austronesian languages on Sumatra shows a major discrepancy between expectation and observation. Voorhoeve (1955) recognized 25 languages for the island, which fall into eleven conservatively-defined microgroups: 1. Acehnese, 2. Gayō, 3. Batak, 4. Malay and Minangkabau, 5. Rejang, 6. Lampung, 7. Islands east of Sumatra, 8. Simalur, 9. Nias-Sichule, 10. Mentawai, and 11. Enggano. Of these groups 1 (Acehnese) and 7 (Orang Laut, Orang Lom, Lonchong, Belitung Malay) form a larger subgroup with Malay and Minangkabau. The greatest linguistic diversity in Sumatra is thus found in the mountainous interior (Gayō, Batak, Rejang) and in the chain of Barrier Islands (Simalur, Nias, Mentawai, Enggano) which lies west of the Sumatran mainland. By contrast, large tracts of southern and eastern Sumatra are peopled by speakers of various forms of Malay or closely related languages (Minangkabau, Kerinci, Kubu, and the various 'Middle Malay' dialects such as Besemah and Seraway).

Most scholars today believe that Proto-Austronesian was spoken in Taiwan. If so, the general direction of the Austronesian expansion must have been southward into the Philippines, and then westward into Borneo, mainland Southeast Asia, Sumatra, and eventually Madagascar on the one hand, and eastward into Sulawesi, the Moluccas and the Pacific, on the other. Given this larger trajectory it is hardly likely that Sumatra was settled from the west, and it is clearly impossible that interior or

highland portions of the island could have been settled before coastal or lowland regions. The pattern of linguistic diversity in Sumatra is thus contrary to expectation: if this area had been settled from Borneo, Java or the Malay peninsula, we would expect greater diversity in the lowlands and coastal areas of the south and east than in the Barrier islands and the interior highlands of the main island. As in the case of the Italian peninsula, where we have documentary controls, such discrepancies between expectation and observation can provide important clues to linguistic macrohistory. In the present case what the geographical distribution of diversity in Sumatra appears to reflect is not primary settlement history at all, but rather residual conservativeness. Stated differently, the most diverse areas are not the areas which have been settled longest, but rather those which have been least affected by secondary leveling: the diversity which should have been produced by earliest settlement in the coastal and lowland areas of southern Sumatra apparently has been eliminated by language extinction. Since virtually all lowland communities in the southern half of Sumatra north of the Lampung districts speak dialects of Malay or closely related languages, the most reasonable inference is that the linguistic clock in this area was reset by the expansion of Malayic speakers, who now occupy territories which were once occupied by speakers of other, unknown Austronesian languages. Similar inferences have been made for the history of Chinese. Since both archaeology and the earliest documentary attestation of the Chinese writing system point to an origin in the Yellow River valley, we would expect northern China to show the deepest splits within Sinitic, but in fact the greatest diversity is found south of the Yangzi River, in areas that were not settled by Han Chinese until about 2,000 years ago (Norman 1988:183ff). Again, the most plausible explanation for this discrepancy is linguistic leveling — in this case the expansion of an early form of Mandarin at the expense of other Sinitic languages which presumably were once spoken in northern China.

A hypothesis of linguistic leveling in Sumatra is also consistent with observations from neighboring areas. Nearly a century ago Skeat and Blagden (1906) observed that many of the Aslian (AA) languages of the Malay peninsula contain apparent Austronesian loanwords which cannot be attributed to borrowing from Malay. In some cases plausible cognates of these presumptive loanwords can be found in languages of Borneo. These observations suggested to Skeat and Blagden that the Malays may have been preceded by other Austronesian speakers in the Malay peninsula. Given the pattern of language distribution in Sumatra this would not be surprising: if an earlier linguistic diversity had been leveled by the expansion of Malayic speakers into southeast Sumatra, there is no obvious reason why the Malay peninsula should not show a parallel history. Similar examples of expansion and extinction can be cited from Borneo, albeit on a smaller scale, as with the expansion of the Kayan and later the Iban into the basin of the Rejang river of Sarawak during the nineteenth century, leading to the extinction of the Seru Dayaks and the decimation of other groups, such as the Ukit, and Bekatan.

The purpose of these prefatory remarks has been to sketch the kinds of assumptions and methods of inference that will be used in the remainder of the paper. The problem which I wish to address using these methods is the linguistic macrohistory of the Philippines.

3. The Evidence for Proto-Philippines

Since the pioneering studies of Blake (1906), Scheerer (1908), and Conant (1910), impressionistic claims have been made for a Philippine subgroup of Austronesian

languages.³ Thomas and Healey (1962) were among the first scholars who tried to support this impression with evidence (in the form of lexicostatistical percentages). Although they left open the possibility that some of the languages of Borneo, Sulawesi, western Micronesia, or Taiwan might be members of their ‘Philippine Superstock’, and although the Sama-Bajaw languages are inexplicably omitted from their classification, the domain of comparison was essentially defined by geography: for them, the term ‘Philippine languages’ was understood to mean ‘all and only the languages of the Philippine archipelago’.

A Philippine subgroup which includes Yami, spoken on Botel Tobago, or Orchid island off the southeast coast of Taiwan, the Sangiric, Minahasan and Gorontalic languages of northern Sulawesi, and all languages of the Philippines apart from the Samalan (Sama-Bajaw) group, was first proposed by Zorc (1986). Blust (1991) suggested that this collection of more than 180 languages be divided into at least nine microgroups, as shown in Table 1:⁴

Table 1. Philippine Microgroups (after Blust 1991)

1. Bashiiic
2. Cordilleran
3. Central Luzon
4. Inati
5. Kalamian
6. GCP (Central Philippines, South Mangyan, Palawanic, Manobo, Danaw, Subanun, Gorontalic)
7. Bilic
8. Sangiric
9. Minahasan

Zorc presented 98 proposed lexical innovations in support of Proto-Philippines. External cognates have since been found for 12 of these (Appendix 1). However, more lexical evidence supporting Proto-Philippines is now available. Appendix 2 lists 241 lexical, morphological and semantic innovations not in Zorc (1986) which appear to be confined to members of the proposed Philippine group. Together, Zorc’s list and the material in Appendix 2 thus come to an impressive 327 items. It should be noted, moreover that the material in Appendix 2 is drawn from Blust (1999), a resource which is only about 25% complete. Maximum search efficiency required that reconstructions that begin with a vowel, *h, *q or *S be done at the same time. As a result, certain sections of

³ Earlier writers, such as Kern (1917 [1882]) do not appear to use the terms ‘Philippine group’, or ‘Philippine language’, but refer to the languages of the Philippines simply as ‘Malayo-Polynesian’.

⁴ Dyen (1965) recognized a lexicostatistically-defined subgroup which he called the ‘Philippine Hesion’. However, his classification includes the Dusunic and Murutic languages of Sabah, and at least one Sama-Bajaw language (Yakan) within the Philippine group, and leaves Ilongot unclassified. Problems of either overinclusion, or underinclusion (or both) are also found in the classifications of Charles (1974) and Walton (1979). As argued in Blust (1998), despite their general typological similarity to Philippine languages and the presence of a number of Greater Central Philippine loanwords in their vocabularies, the languages of Sabah do not appear to belong to this group.

the dictionary (*a, *b, *e, *h, *i, *q, *S, *u and *w) have been worked quite thoroughly, while others have hardly been touched. The coverage is therefore very uneven, and many innovations that are confined to Philippine languages do not show up in the list given here. If the material analyzed so far is representative, however, the completed dictionary should contain around $241 \times 4 + 86$, or about 1,050 ‘Philippine-only’ cognate sets.

The sheer quantity of this material probably is sufficient to dispel doubts as to the reality of a Philippine subgroup, but questions might be raised about its quality. In assessing the value of proposed lexical innovations the notion of ‘replacement innovation’ is critical. In ordinary innovations the semantic category associated with a novel form is not clearly associated with any reconstructed form of greater antiquity; rather, novel forms are associated with novel meanings. Thus, PPH *biglâq ‘suddenly’ appears to be a lexical innovation, but since no higher-level reconstruction is available for this meaning the matter is unclear. In replacement innovations, on the other hand, a form-meaning pairing which can be assigned to a given proto-language is replaced by a different but equivalent form-meaning pairing in a language descended from it, as with PMP *ñeRab or *niRab ‘yesterday’ (with reflexes extending from Sabah to Fiji), but PPH *ka-Rabiqi ‘yesterday’. In these cases novel forms are paired with familiar meanings, and the etymon confined to the lower-level genetic grouping is a replacement innovation. Other examples include the paired semantic shift of PMP *Rumaq ‘domicile, residential unit’ to the meaning ‘sheath of a knife’, and of PMP *balay ‘meeting hall, guest house’ to ‘domicile’. These changes occur with the distribution shown in Table 2, where + marks evidence for the innovation, 0 marks a non-cognate morpheme (or a loanword), and ? marks lack of evidence. There are no cases (-) in which *balay and *Rumaq are reflected with their PMP meanings; numbers of Philippine microgroups follow Table 1:

Table 2. Distribution of the changes PMP *balay ‘meeting hall, guest house’ > ‘house’ and *Rumaq ‘domicile’ > ‘scabbard’

Microgroup	1	2	3	4	5	6	7	8	9
*balay	+	+	+	0	+	+	+	+	+
*Rumaq	?	0	0	?	0	+	+	+	+

These innovations are not evenly attested over all proposed Philippine microgroups, but both are widely distributed: the first is found in all microgroups except Inati, for which we have only one rather limited source (Pennoyer 1986/87). The second is found in Greater Central Philippines, Bilic, Sangiric and Minahasan languages, but has not been identified further north (Botolan Sambal *guma?* ‘sheath’ shows phonological changes which mark it as a likely Tagalog loanword). Although many languages outside the Philippines also reflect *balay in the meaning ‘house’, there are few parallels to the semantic change seen in *Rumaq, and the combination of both changes is highly distinctive for members of the Philippine group of languages.⁵

⁵ As noted by Blust (1989a) in some South Sulawesi languages reflexes of *banua ‘inhabited territory’ acquired the meaning ‘house’, and then through metaphor ‘sheath’. A similar repetition of semantic history is seen in Bontok *baley* ‘home, shelter; spiderweb; placenta; sheath of a bolo.’ These examples show the same conceptual processes at work, but with different morphological material.

Finally, a striking morphological innovation is seen in the historically double layer of suffixation found in PMP *qamih-an > Casiguran Dumagat *amiyan* 'northeast monsoon': *k-amiyan-an* 'northeast', Ilokano *amián* 'north wind': *amian-an* 'north', Kapampangan *amian* 'winter, cold season': *amian-an* 'north', Bikol *amíhan* 'northeast wind': *amihán-an* 'direction of the northeast wind', Aklanon *amíhan* 'northeast wind': *amihán-an* 'north', Hiligaynon *amính-an* (met.) 'north', Cebuano *amíhan* 'north wind': *amihán-an* 'north, northern', and Maranao *amian* 'north wind': *amian-an* 'north'. Here languages representing the Cordilleran, Central Luzon and Central Philippines groups reflect a historically suffixed disyllable (PMP *qamih-an) as a reanalyzed trisyllable (*qamihan), with resuffixation of the reanalyzed base (*qamihan-an). Although this change is not found in the Sangiric and Minahasan languages (which reflect *qamihan, but not *qamihan-an) its distribution is most simply explained by the reconstruction of PPH *qamihan 'north wind', *qamihan-an 'north' (= location/direction of the north wind).⁶

Together with the 327 cognate sets in Zorc (1986) and Appendix 2, these replacement innovations leave little doubt that there is a Philippine subgroup of Austronesian languages. However, while there may be a certain satisfaction in establishing sizeable subgroups of some time-depth in a given language family, the recognition of a Philippine subgroup raises conceptual problems which have hardly been acknowledged. If Taiwan was settled before any other area historically occupied by speakers of Austronesian languages, and the direction of movement out of Taiwan was southward into insular Southeast Asia and the Pacific, the Philippines would be the second longest-settled area in the Austronesian world. If so, why *should* there be a Proto-Philippines? Given the archaeological evidence for the arrival of Neolithic cultures in the Cagayan Valley by at least 4,500 BP, we would expect the languages of the Philippines to belong to several primary branches of Malayo-Polynesian, just as the Formosan languages belong to several primary branches of the Austronesian family as a whole. The fact that we do not find such an order of linguistic diversity, but rather a single linguistic subgroup plus an intrusive population of sea nomads and their settled relatives, provides an important clue to the linguistic macrohistory of the Philippines. Since the second major linguistic extinction in the Philippines has already been documented in some detail and so provides a model for the first extinction, it will be best to begin with it.

4. The Second Extinction: The Expansion of Proto-Greater Central Philippines

As noted in Blust (1991), the Philippine archipelago can be divided into three regions of roughly equal size, northern, central, and southern, which differ markedly in their patterns of linguistic diversity. The northern region contains three flourishing microgroups: Bashiic, Cordilleran, and Central Luzon. The southern region (which extends into northern Sulawesi) contains a number of others (Central Philippines, Manobo, Danaw, Subanun) that are subsumed under Greater Central Philippines, along with Bilic, Sangiric, Minahasan, and the non-Philippine languages of the Samalan group. Both the northern and southern regions of the Philippines are thus areas of

⁶ Since neither *qamihan nor *qamihan-an is attested in any Cordilleran language except Ilokano it might be argued that *qamihan-an is a Greater Central Philippines innovation which has spread to a few other languages in more accessible (lowland) areas. However, it is difficult to see why words that mean 'north' or 'north wind' would be borrowed by more northerly languages from those spoken further to the south.

substantial linguistic diversity. By contrast, in southern Luzon, the Bisayas, Palawan, and northern Mindanao Central Philippine languages predominate to the almost total exclusion of others. It is true that three primary branches of the Philippine subgroup are represented by languages spoken in the central region, but Microgroup 4 contains a single language, Inati, spoken by a nomadic population of some 900–1,000 persons (Pennoyer 1986/87), and Microgroup 5 contains just three languages, Kalamian Tagbanwa, Central/Aborlan Tagbanwa, and Agutaynen, spoken by a total of perhaps 25,000 persons (Grimes 2000). In sharp contrast with the linguistic situation in the northern or southern regions, almost the entire central region of the Bisayas and southern Luzon constitutes an extended dialect network with roughly 45 million first-language speakers of Tagalog, Bikol and intergrading varieties of Bisayan. The distribution of phylogenetic complexity in the central Philippines can be compared to that in southeast Sumatra: in both cases areas that have been settled by Austronesian speakers long enough for us to expect considerable linguistic diversity show a surprisingly high degree of homogeneity. There is only one way to interpret this contradiction, namely, that in addition to the usual processes of language split and diversification over time the linguistic history of the central Philippines included a major episode of linguistic expansion/extinction.

Blust (1991) gave the name ‘Proto-Greater Central Philippines’ to the hypothetical language which brought about this linguistic leveling in the Bisayas and southern Luzon.

Given the relatively close genetic relationship of all GCP languages it was inferred that the extinction event occurred no earlier than about 500 BC. To recapitulate the argument, the principal pieces of evidence for the GCP expansion are:

1. the unexpectedly low level of linguistic diversity in southern Luzon, the Bisayas and northeast Mindanao, given the higher levels of diversity in other parts of the Philippines or northern Sulawesi, and the known radiocarbon dates for the introduction of Neolithic cultures into the central Philippines by at least 4,000 BP (Bellwood 1997:219ff).
2. the evidence for a discontinuously distributed linguistic subgroup, linking the Gorontalic languages of northern Sulawesi with the languages of the central Philippines and much of Mindanao, but skipping the intervening Bilic languages of southern Mindanao and the Sangiric and Minahasan languages of northern Sulawesi.
3. the presence of what Conant (1910) called ‘the stereotyped *g*’ in Philippine languages, referring to sporadic instances of $*R > g$ in languages which normally reflect PPH $*R$ as some other phoneme. This development suggests that contact with GCP languages, in which $*R > g$ was regular, reached as far north as northwest Luzon, and as far south as central Sulawesi, although some instances of this development in Ilokano and Central Cordilleran languages could be due to borrowing from Northern Cordilleran languages in which $*R > g$ is regular (Reid 1973), and some instances in central Sulawesi could be due to borrowing from Tomini languages in which the same change has occurred (Himmelman 2001).
4. a few examples of $*R > r$ for expected *g* in Molbog of Balabac island (a Palawan language), and in Bonggi of Banggi island between Sabah and Palawan (a Sabahan language). These instances of what might be called ‘stereotyped *r*’ suggest the former presence of a substrate language or languages in this area. Since $*R$ is reflected as a liquid (*l* or *r*) in both Kalamian Tagbanwa to the north of Palawan, and in the Bilic languages of

southern Mindanao, either the Kalamian microgroup, the Bilic microgroup (or both) may have extended to Balabac and Banggi islands before the GCP expansion.

5. The First Extinction: The Expansion of Proto-Philippines

The Greater Central Philippines hypothesis shows that several superficially disparate observations regarding modern Philippine languages are the expected consequences of a single prehistoric event, which may have taken some generations to complete. At the same time it provides a cautionary reminder that the history of related languages is not always a uniform process of differentiation and divergence, but may be punctuated by important episodes of extinction. The linguistic leveling in the central Philippines which apparently resulted from the GCP expansion is hardly unique. Something rather similar happened in the early recorded history of the Italian peninsula, and must have taken place in southeast Sumatra, and perhaps in the Malay peninsula as a result of the Malayo-Chamic expansion out of southwest Borneo (Blust 1992). As it happens, all of these events evidently took place within the past 2,000–2,500 years. But if linguistic leveling has occurred within the past two and one half millennia in such places as the Italian peninsula, the central Philippines, southeast Sumatra and the Malay peninsula, there is no reason why similar expansion/extinction events could not have taken place even earlier. As noted already, *a priori* logic would suggest that a number of primary branches of Malayo-Polynesian should be found in the Philippines. However, this is not the case: there is persuasive evidence that all languages of the Philippines apart from Sama-Bajaw belong to a single subgroup. If the GCP hypothesis is valid as an explanation for the greater-than-expected linguistic homogeneity of the central Philippines, then, the same type of argument should be valid as an explanation for the greater-than-expected linguistic homogeneity of the Philippines as a whole.

The time-depth of Proto-Philippines is contentious, but clearly must exceed that of Proto-Greater Central Philippines by some centuries. Greater Central Philippines languages are, at least impressionistically, somewhat more divergent than groups like Romance or Polynesian, which have a time-depth of roughly two millennia, and it therefore seems unlikely that Proto-Greater Central Philippines was spoken less than 2,300–2,500 years BP. Since the earliest radiocarbon dates that are generally accepted for a Neolithic presence in the Philippines cluster around 4,500 BP we are able to bracket the break-up of Proto-Philippines between roughly 2,500 and 4,500 BP. The earlier of these two dates is associated with the initial Austronesian settlement of the Philippines and so cannot be associated with Proto-Philippines. If we place the break-up of Proto-Philippines roughly midway between these two chronological points of reference, then, we arrive at a provisional date of about 3,500 BP. The assumption of a time-depth of about 3,500 years for the separation of the Philippine languages seems reasonable in the light of both radiocarbon chronologies and degrees of linguistic difference relative to other parts of the Austronesian world. But it clearly implies that at least a millennium of previous linguistic differentiation in the Philippines was wiped clean by the expansion of a single prehistoric language early in the Neolithic settlement history of the archipelago. The how and why of such an expansion probably will never be known. However, much the same is true of historically more recent expansions in the Austronesian world which produced linguistic leveling on a smaller scale. Sutlive (1978:20ff) considers both demographic and cultural explanations for the dramatic nineteenth-century Iban migrations, which led to the eventual extinction of the Seru

Dayaks (already weakened through depredations by the Kayan) and the decimation of other groups in southern Sarawak and neighboring parts of Kalimantan. Whether the need for new lands was driven by population growth, or whether population growth was a consequence of the acquisition of new lands can be argued tediously to no conclusion. In any case a significant territorial expansion accompanied by linguistic leveling on small scale did take place, and is well-documented for this group. Although we lack documentary evidence for the early history of the Philippines, the clear discrepancy between radiocarbon chronologies for the beginning of the Neolithic and the less-than-expected degree of linguistic differentiation points unambiguously to a similar episode of territorial expansion and linguistic leveling. The principal difference is one of scale: whereas the Iban migration out of the upper Kapuas basin covered an area of roughly 325 miles from south to north in the span of three or four generations and led to the complete elimination of only one known language, the Proto-Philippine territorial expansion that is hypothesized here apparently covered a greater territory and led to more widespread linguistic leveling.

As noted in Blust (2000a:104ff), the hypothesis that Proto-Philippines expanded around 3,500 BP at the expense of other descendants of Proto-Malayo-Polynesian which were present in the Philippine archipelago at that time is supported by more than one line of evidence. First, there is the evidence considered here, which suggests that at some point after the initial Austronesian colonization of the Philippine islands the linguistic clock was 'reset', and divergence began anew from a single founding community. Second, there is evidence in the form of language displacement. Historically documented cases of language expansion such as that of Latin are accompanied by language extinction, and this is clearly one major consequence of such historical events. However, given the right circumstances language expansion can also lead to the displacement of dominated groups, either by driving them into refugia where the competition for land is less intense, as with the Ainu in Japan, or by forcing them to take flight to some fairly distant location, as with the Austroasiatic-speaking Lai of southern China. There is compelling evidence that pre-Chamorro speakers reached the Marianas islands around 3,500 BP, directly from some part of the Philippines north of Mindanao. Yet there is no evidence that Chamorro subgroups with the languages of the Philippines. The conjunction of these observations presents us with a potentially fruitful contradiction: if pre-Chamorro speakers left the Philippines around 3,500 BP but Chamorro is not descended from Proto-Philippines, the Philippine islands at 3,500 BP must have been home to various descendants of Proto-Malayo-Polynesian, of which only Proto-Philippines survived *in situ*. At the time of the Proto-Philippines expansion Austronesian languages presumably would have been thinly distributed throughout the Philippine islands, and it is difficult to conceive of a demographic or economic advantage by which one group would have succeeded in expanding at the expense of all others. Nonetheless, this inference follows from the evidence for Proto-Philippines, and what it suggests is that Austronesian languages in the Philippines circa 3,500 BP were confined to a fairly narrow range of environments, including only the coastal zones of the larger islands.

In this connection it might be asked how the pre-Neolithic Negrito populations were affected by the expansion of Proto-Philippines. The answer is 'probably little at all.' Reid (1987, 1989, 1991, 1994a) has done pioneering work in unraveling the linguistic history of Philippine Negritos. What is clear is that this population once spoke languages unrelated to those of the incoming Austronesians, but eventually adopted Austronesian languages as a result of contact, and that this happened throughout the Philippine archipelago. When this contact began on a meaningful scale is difficult to determine (and probably varied from area-to-area), but the expansion of Proto-

Philippines would have been essentially a competition between agricultural populations for the same territory, and it is doubtful that the foraging Negrito populations of remote mountain areas would have been affected by it. Since all documented Negrito groups in the Philippines today speak Philippine languages, however, it appears that the 'early switch' in language affiliation by these groups postdated the expansion of Proto-Philippines around 3,500 BP. If this were not the case, we would expect some Negrito languages in the Philippines to be Austronesian, but not descendants of Proto-Philippines.

6. The Position of the Samalan Languages

In general, the geographical distribution of Philippine languages corresponds to the boundaries of the Philippine archipelago. There are, however, three exceptions. First, Yami, a Bashiic language, is spoken on Botel Tobago island within the political boundaries of Taiwan. Second, the Sangiric, Minahasan, and Gorontalic languages of northern Sulawesi subgroup with languages in the Philippines, not with other languages in Sulawesi. Finally, one set of languages spoken within the Philippines is not a member of the Philippine group. This is the collection of diverse dialects or closely related languages known as Sama (an endonym), Bajaw (an exonym, apparently from Buginese), Sama-Bajaw, or Samalan (linguistic creations which I will use interchangeably).

McFarland (1980:106) was willing to say only that 'The relationship of the Sama languages to the other languages of the Philippines is not clear.' Zorc (1986:156), however, firmly excludes them from the Philippine group: 'The number of exclusively shared lexical innovations that I have gathered thus far suggests that the languages of the Philippine archipelago (exclusive of the Sama-Bajaw group) form a single AN subgroup.' The exclusion of the Samalan languages is perhaps the most surprising exception to the general correlation of linguistic subgrouping boundaries with major geographical features in the Philippines. Yami is simply an extension of the Bashiic languages along an island chain which happens to cross a modern political boundary, and the Sangiric, Minahasan, and Gorontalic languages represent expansions of Philippine languages slightly beyond the Philippine archipelago into neighboring parts of Indonesia. But the Sama-Bajaw languages present a problem of a different order: they do not belong to the Philippine group, and to date no one has been able to establish their linguistic position. What, then, is the origin of this seemingly extraneous group of languages?

The first two peculiarities of the Samalan languages which are likely to raise questions about their inclusion in a Philippine group have nothing to do with language. Unlike most Austronesian speakers in the Philippines and elsewhere, many speakers of Samalan languages live on boats rather than on the land, and they are consequently known in the popular literature as 'sea nomads' or 'sea gypsies'. This is not to say that there are no sedentary speakers of Samalan languages. Such groups as the Jama Mapun of Cagayan de Sulu and the Yakan of Basilan island are traditional agriculturalists, but many other Samalan speakers are boat nomads, and this clearly is connected with the second non-linguistic peculiarity of Samalan languages: their wide geographical range. In mapping the distribution of Samalan languages Pallesen (1985:2) shows groups in the Philippines from Capul island, near the southeastern tip of Luzon (Abaknon), to the southernmost part of the Sulu archipelago (Southern Sama), in western and eastern Sabah, various parts of northern and eastern Sulawesi, the island of Kayoa off the west coast of Halmahera, and on Roti and the western end of Timor in the Lesser Sunda chain. In addition, Pallesen (1985:44) speculates 'It is possible that an SB language is still spoken in the Anambas or Natuna Islands, northeast of Singapore, and that some of the Orang Laut languages spoken

in the Java Sea (Riouw, Lingga, Bangka and Billiton (Belitung) island groups) will prove to be closely affiliated with PSB, enabling the reconstruction of a protolanguage for a greater time depth. Several communities described as Bajaws have been reported in coastal Pattani Province, on the east coast of South Thailand (Kenneth Smith, pers. comm., 1976), but no linguistic data are yet available on this group.⁷

These remarks almost certainly go too far. Wherever linguistic data are available for 'Orang Laut' populations in the South China Sea or Strait of Malacca, they show that the languages used by these peoples are either dialects of Malay, or Moken (in the Mergui archipelago), and are therefore linguistically very distinct from Samalan languages.⁷ The Sama-Bajaw communities of the Lesser Sunda islands are known to be historically late arrivals, first appearing in the area around the beginning of the eighteenth century, at which time they were closely associated with the Buginese and Makasarese of south Sulawesi as suppliers of sea cucumbers for the Chinese market (Fox 1977). Excluding these late expansions, then, most Sama-Bajaw communities are confined to the southern Philippines, northern Borneo, and northern Sulawesi, where they are regarded by the sedentary populations as outsiders who subsist by fishing, trade, and in some cases piracy.

When the mobile boat-dwelling orientation of the Sama-Bajaw peoples began is an open question. According to Pallesen (1985:245ff), who is concerned primarily with contact-induced linguistic convergence between Tausug and one subset of Samalan languages, Proto-Samalan probably was spoken in the southern Zamboanga-Basilan area of southern Mindanao about 800 CE, and all Samalan languages except Abaknon, Yakan, and Sama Batuan 'were a single subgroup prior to 1,000 A.D., with the division into distinct subdialects becoming clear about the time of the TSG-SLU contact' (contact between Tausug, and the core group of sea-based Samalan languages in the southern Philippines). This suggests that Proto-Samalan speakers were mobile maritime traders and fishermen who arrived in the southern Philippines some 1,200 years ago, and began to interact with the local populations in a kind of economic symbiosis, which has continued for at least some groups up to the present.

With regard to the origin of the Sama-Bajaw, Pallesen suggests (1985:245) 'an Indonesian origin rather than any close relationship to the Central Philippine languages with which many SB daughter languages are currently in geographical proximity.' Collins, Collins, and Hashim (2001), in the Introduction to their *Mapun-English dictionary*, venture a little further: 'The Mapun language is the language of the *Jama Mapun*, or 'People of Mapun' whose origins are said to have been somewhere on the island of Borneo.' No linguistic evidence is presented to support this statement, which presumably reflects the oral traditions of the people themselves. Casiño (1976:8) notes that Antonio Pigafetta, the Italian chronicler of the Magellan expedition which was in the Philippines in 1520 remarked of the Jama Mapun that they 'are Moros and were banished from an island called Burne.' For at least some Sama-Bajaw peoples, then, the tradition of a Bornean origin goes back some five centuries or more.

If Proto-Samalan was spoken in the Zamboanga-Basilan area of southern Mindanao about 800 CE, as Pallesen proposes, the overwhelming probability is that the Sama-Bajaw reached the Philippines from either Borneo or Sulawesi. An examination of

⁷ For the language situation among the sea nomads of the Riau-Lingga archipelago and the islands of Bangka and Billiton cf. Kähler (1946-49, 1960). For the Urak Lawoi' of peninsular Thailand cf. Hogan (1988), and for the Moken cf. Larish (1999). Dahl (1991:92ff) also assumes incorrectly that the languages of all sea nomads in insular Southeast Asia form a subgroup, apparently in part because they are widely known in the area by various forms of the name 'Bajau', given to them by the Buginese.

both phonological and lexical evidence strongly favors Borneo, and so is consistent with the oral traditions of at least some Samalan speakers. Oral traditions, however, must be treated with care. All attested Sama-Bajaw peoples are Moslems, and Islam was introduced into the southern Philippines by Brunei Malays no more than 150–200 years before Pigafetta's observations were made in 1520. For this reason it is conceivable that Sama-Bajaw traditions of a Bornean origin postdate the arrival of Islam, and reflect the importance of Brunei as a formative element in the syncretic Samalan cultures we now find rather than an orally preserved record of a pre-Philippine 'homeland'.

To pursue the question of Sama-Bajaw origins further we need to consider linguistic evidence. Until recently this was difficult, as little lexical information was publicly available for any Samalan language. That situation has now changed dramatically with the appearance of two splendid dictionaries, one for Mapun (Collins, Collins, and Hashim 2001), and the other for Yakan (Behrens 2002). I will use these sources below, together with the more limited data in Reid (1971), which comes from the barrios of Sisangat and Siganggang, Siasi municipality, a dialect region that Pallesen (1985:19) assigns to Central Sulu (hereafter: CS). Unless otherwise indicated, material from Barito languages is taken from Hudson (1967).

The historical phonology of the Samalan languages is explored in greater depth in Blust (To appear), and the present discussion will be limited to a summary of the major claims of that paper and some of the evidence supporting them. Without entering into details here, the first thing to note is that the vocabulary of Samalan languages can be divided into three principal strata: 1. Philippine loanwords, 2. Malay loanwords, 3. native material. As demonstrated by Pallesen (1985), most Philippine loanwords in the Zamboanga-Sulu region come from Tausug. Abaknon, located between northern Samar and Masbate presumably has been more strongly influenced by Waray or Masbatenyo, although little material is available to determine the facts. Pallesen (1985:6) notes in passing that 'there has been prolonged contact between SB languages and Malay, early phases of which probably predate Arab and Chinese contacts, and perhaps predate the arrival of speakers of SB languages in the Philippines.' Given his focus on Tausug-Samalan contact and convergence, Pallesen did not follow up these preliminary remarks, but his first impressions were certainly correct: if anything, borrowing from Malay in the Samalan languages of the Zamboanga-Sulu region has been even heavier than borrowing from Tausug, and there are intriguing indications that some Malay loans entered the Samalan languages before the arrival of Islam. There is some phonological evidence which supports the view that the Samalan languages have a Bornean origin, but it cannot completely rule out other areas in western Indonesia as possible places of origin. For this reason I will confine myself here to lexical evidence, with comments on the phonology of individual forms as the need arises.

The following inferences are advanced:

1. the Samalan languages originated in Borneo,
2. the specific 'homeland' of the Sama-Bajaw peoples is the area which today forms the basin of the Barito river and its tributaries — the same area from which the Malagasy derive,
3. the Samalan languages apparently belong to Hudson's (1967) 'Barito Family', but do not subgroup closely with any of the modern Barito languages, including Malagasy,
4. the Samalan languages contain probable loanwords from early Southeast Barito languages,

5. some loanwords suggest that early Samalan speakers were located further inland than early Southeast Barito speakers,
6. loanwords in both Samalan and Southeast Barito languages show that Sriwijayan Malays ancestral to the modern Banjarese community were trading preserved fish to the native peoples before AD 800, probably in return for jungle products such as dammar, rattan, wax and the like,
7. through these trade contacts some speakers of Barito languages were drawn into a large-scale Malay trade network centered on spices which connected eastern and western Indonesia with the outside world.

Inference 1: The Samalan languages originated in Borneo

As noted earlier, this claim evidently is part of the oral tradition of at least some Sama-Bajaw groups in the Philippines, in particular the Jama Mapun. However, concrete linguistic evidence for a Bornean origin had never previously been presented. Lexical items that appear to be shared exclusively by one or more Samalan languages with languages throughout much of Borneo include the following:

1. *payaw:⁸ Yakan *payew*, Tindal Dusun, Tarakan, Bolongan *payaw*, Paitan Murut *payow*, Kelabit, Kayan *payo*, Bintulu *pazaw*, Mukah Melanau, Lahanan, Iban *payaw* ‘sambhur deer’.
2. *pesi: Yakan *pessi*, Long Anap Kenyah *pesi*, Uma Juman Kayan *pesi?*, Mukah Melanau *pesey*, Katingan *pesi?*, Dohoi *posi?* ‘fishhook’.
3. PAN *sapaw ‘field hut’ > Mapun *sapaw*, Yakan *sapew*, Kelabit *apo*, Bintulu, Ngaju Dayak *sapaw* ‘roof’, Highland Kenyah, Mukah Melanau *sapaw* ‘roof; thatch’, replacing reflexes of PMP *qatep ‘roof, thatch’.
4. *iban: Mapun *ivan* ‘son-in-law, daughter-in-law’, Kadazan Dusun *ivan*, Kelabit *iban*, Kiput *ibin* ‘parent-in-law’, Long Jeeh Kenyah *iban* ‘component of affinal terms for +1, -1 and Ego’s generation’, Bintulu *ivan* ‘parent-in-law’, *anak ivan* ‘son-in-law, daughter-in-law’, Ma’anyan *ivan* ‘brother-in-law, sister-in-law’.

The first three cognate sets are clear replacement innovations for PMP *uRsa ‘sambhur deer’, *kawil ‘fishhook’, and *qatep ‘roof, thatch’ respectively, and point unambiguously to a Sama-Bajaw derivation from Borneo. The fourth may also be a replacement innovation, but this is dependent upon its gloss (cf. PMP *hipaR ‘sibling-in-law’, and *ma-tuqah ‘parent-in-law’, but PMP terms for ‘child-in-law,’ or ‘in-law’ are unknown). The wide and distinctive distribution of these forms in Borneo indicates a Bornean origin for the Sama-Bajaw, but it does not pinpoint a particular area.

⁸ Beech (1908:71) notes that Malay *air payau* means ‘brackish, of water’, and speculates that the name of the sambhur deer in Borneo may derive from its habit of obtaining salt from brackish water: ‘Curiously enough the Tidongs always say that deer are passionately fond of the brackish water that is often to be found in the inland swamps of Borneo, and such places are the favourite hunting grounds for natives who wish to obtain these animals. Perhaps *air payau* = ‘water such as the deer loves.’ This is an interesting observation, but it suggests no reason why the sambhur would be singled out from other deer, and provides no evidence that *payaw is a word of any antiquity in the meaning ‘brackish.’

Inference 2: The Samalan languages originated in southeast Kalimantan

Additional lexical comparisons help to narrow the possibilities for a Samalan homeland region in Borneo. Surprisingly, these point to a connection between Sama-Bajaw and the languages of southeast Kalimantan that Hudson (1967) called the 'Barito Family'.⁹ Most of these innovations single out the branch that Hudson called 'Southeast Barito', which includes Ma'anyan and Malagasy, but not Ngaju Dayak. Although other kinds of evidence for a Samalan-Southeast Barito connection are discussed in Blust (To appear), the following seven linguistic comparisons plus a few others that will be included in the discussion of borrowing should be sufficient to establish the case:

5. PMP *ma-quḍip replaced by *belum 'living, alive': Mapun *allum/llum* 'alive, living', Yakan *ellum* 'living, alive (of animate beings)', Ba'amang, Katingan *belum*, Taboyan, Lawangan *bolum*, Paku *wolum*, Ma'anyan, Samihim, Dusun Witu, Dusun Malang *welum*, Malagasy *velona* 'living, alive'. This innovation is very distinctive, as it is known only from languages of the Barito group (Ray 1913), Malagasy, and Sama-Bajaw.
6. PMP *kamu/kamuyu replaced by *kaʔam '2pl': Mapun *kaam* 'You. Second person plural emphatic actor focus. Also second person plural non-focus patient, goal, beneficiary in the core of the clause', Yakan *ka'am* 'You. Second person plural absolutive and oblique', CS *ka'am* '2sg', Tunjung *kaam*, Lawangan *kam* '2pl'. Again, a striking replacement innovation. In this case the form is shared with a Northeast Barito language (Lawangan), and with Tunjung, a language which Hudson (1967) regarded as forming a primary branch of the Barito Family. If forms such as Uma Juman Kayan, Katingan, Murung, Tamuan *ikam* '2pl; ye' are admitted under an analysis *i-kam* the collection of languages reflecting this innovation will be larger. However, a disyllabic base without *i-* has been reported so far only for Sama-Bajaw and Barito languages.
7. PMP *ma-piaq replaced by *ma-halap 'good': Mapun *ma-hap* 'good; pretty, beautiful', Yakan *halap/hap* 'good (quality of things and moral character); nice, pretty, beautiful', CS *haap/halap* 'good', Kapuas (Hudson 1967) *bahalap* 'good', Ngaju Dayak (Hardeland 1859) *ba-halap* 'good, honest, prosperous; favorable, friendly', *ha-halap*, *hala-halap* 'suitable, good'. Possibly a chance resemblance, as Hudson (1967:89) gives this form only for Kapuas (of the other two 'Ngaju Dayak' groups Ba'amang has the Malay loanword *baik*, and Katingan the unrelated form *lamus*).
8. PMP *tuduq 'drip, leak (as roof)' replaced by *petak 'a drop; to drip': Yakan *pettak* 'a drop', *mag-pettak* 'to drip', Ma'anyan (Ray 1913:122) *petak* 'drip'.
9. PMP *sempit replaced by *kiput 'narrow': Mapun *kiput* 'narrow (limited in space or size)', Yakan *kiput* 'narrow, small (of openings or long objects,

⁹ The Barito Family has three primary branches: 1) Barito-Mahakam, with one language (Tunjung), 2) West Barito, with two subdivisions, Northwest Barito (Dohoi, Murung-1, Murung-2, Siang), and Southwest Barito (Ba'amang, Kapuas, Katingan), 3) East Barito with three subdivisions, Northeast Barito (Taboyan, Lawangan), Central East Barito (Dusun Deyah), and Southeast Barito (Dusun Malang, Dusun Witu, Paku, Ma'anyan, Samihim). Traditionally, Northwest and Southwest Barito have been called 'Ot Danum', and 'Ngaju Dayak' respectively. As shown by Dahl (1951, 1977) East Barito also includes Malagasy.

e.g. road)', CS *kiput* 'narrow (of space)', Samihim *kiput* 'narrow'. This comparison is somewhat shaky, since a known Bornean reflex of **kiput* 'narrow' is found only in Samihim.

10. PMP **ini* 'this (1st person deictic)' replaced by **itu* 'that (2nd person deictic)': Mapun *itu* 'a discourse particle that is used for emphasis when referring back to a specific topic that has already been discussed, or a topic that has already been in the mind of the speaker. The topic is very close to the speaker', Yakan *itu*/'*itu*'-en 'this here (close to speaker, may be far from hearer)', CS *itu* 'this', Paku, Dusun Deyah, Ba'amang *itu* 'this'.
11. PMP **ina* 'that' replaced by **idu* 'that (3rd person deictic)': Yakan *lu* 'there, that (the person or object spoken to or about is somewhat removed from the speaker)', CS *ilu* 'there (2nd person)', Taboyan, Dusun Witu *iro*, Ma'anyan, Paku *iru* 'that'. Comparisons 10) and 11) must be considered together. The PMP demonstrative system included at least **ini*, **itu*, and **ina*, all of which incorporated the generic locative marker **i* 'in, at, on', and so might be written **i-ni*, etc. Each of these terms probably was associated with a degree of personal reference (first, second, and third person respectively), but there is evidence for additional terms such as **idi*, which cannot easily be glossed within the parameters of the semantic system assumed here. What is significant about comparisons 10) and 11) is that there was a double innovation in the deictic system of both Samalan and Southeast Barito languages: the second person deictic **itu* moved into first person semantic space, and the third person deictic **ina* was replaced by an apparent lexical innovation, **idu*. The first of these innovations is found in a number of Philippine languages but the second is not, and the combination of the two must be given some importance as evidence of exclusively shared history.

Inference 3: The Samalan languages form a primary branch of the Barito Family

Most of the innovations shared exclusively by Sama-Bajaw and Barito languages favor the Southeast Barito group. This includes comparisons 8), 9), 11), and as will be explained below, comparisons 3) and 13). Yet there is little evidence for including the Samalan languages even as a primary branch of the East Barito group. Cognate percentages between Samalan and East Barito languages do not appear to be high, and the historical phonology of the two groups diverges in a number of respects. This contrasts sharply with the similarities linking Malagasy and Barito languages, which can be derived in a fairly straightforward manner from a reconstructed Proto-East Barito.

So, where does this leave us? As will be seen, some of the innovations shared only by Samalan and East Barito languages almost certainly are loanwords. However, others probably are not. In particular, it is unlikely that a language would borrow personal or demonstrative pronouns from another language which has the same type and level of culture. Minimally, then, we can assume that reflexes of **ka'am* '2pl' and **idu* 'that (3rd person)' are directly inherited. Reflexes of the former are found in Samalan languages, Tunjung (Barito-Mahakam), and Lawangan (Northeast Barito). Reflexes of the latter are found in Samalan, Northeast Barito (Taboyan), and Southeast Barito languages (the rest). Given the fragmentary nature of the evidence and the lack of any clear directionality in subgrouping, the default hypothesis would appear to be that the Samalan languages form a primary branch of the Barito Family, which can now be reformulated as in Table 3:

Table 3. The ‘Extended Barito Family’ of southeast Borneo

1. Barito-Mahakam: Tunjung
2. West Barito: Ot Danum and Ngaju Dayak
3. East Barito: Lawangan, Ma’anyan, Malagasy, etc.
4. Samalan: languages of the Sama-Bajaw

Inference 4: Proto-Sama-Bajaw borrowed from early East Barito languages

Because the historical phonology of Samalan languages is discussed in detail elsewhere (Blust to appear), irregularities in phonological development have been largely ignored. However, one form which clearly involves an irregularity is Mapun *iwan* ‘son-in-law, daughter-in-law’. In all Samalan languages for which published data is available PMP *b is reflected as *b*: *balabaw > Mapun, CS *babaw* ‘rat’, *batu > Mapun, Yakan *batu* ‘stone’, CS *batu* ‘coral rock’, *bulan > Mapun *buwan*, Yakan, CS *bulan* ‘moon’, *qabu > Mapun, Yakan, CS *abu* ‘ash’, *labuq > Mapun *labu*, Yakan *labo*, CS *labu* (error for *labu*?) ‘fall, drop’. The only known exceptions to this statement are reflexes of Proto-Barito (PB) *belum ‘living, alive’, *iban ‘affinal relative’, PMP *babuy (> Mapun, Yakan, CS *bawi* ‘pig’), and one or two other forms. The phonological development *belum > Mapun *allum/llum*, Yakan *ellum* ‘living, alive’ is irregular, but not unparalleled (cf. *besuR > Mapun *asso*, *sso*, Yakan *esso* ‘full from eating, satiated’). It nonetheless remains puzzling, since a directly inherited reflex of *belum should have *b-*, and a loanword from an early East Barito language apparently should have *w-*.¹⁰ Irregularities in reflexes of *iban ‘in-law’ are more tractable. Since all known Barito reflexes of *iban ‘affinal relative’ contain *-w-*, the irregular change in Mapun *iwan* finds a natural explanation as a borrowing from an ancestral Barito source. A similar explanation may apply to Mapun, Yakan, CS *bawi* ‘pig’ in light of Kapuas *bawuy*, Taboyan *bawi* ‘pig’, although this is less certain.

Inference 5: Proto-Sama-Bajaw was inland from early Southeast Barito languages

Another comparison supports the inference that Proto-Sama-Bajaw speakers borrowed from early Southeast Barito languages, and also sheds light on the relative geographical positions of the two language communities. Since PMP *s is normally reflected as a sibilant in Samalan languages (*hasaq > Mapun, Yakan *asa* ‘whet, sharpen’, *asiq > Mapun, Yakan *ase* ‘pity; love; favor; mercy; kindness’, *isi > Mapun, Yakan *isi* ‘flesh of any human or animal’, *kasaw ‘rafter’ > Mapun *kasaw* ‘purlin’, Yakan *kasew* ‘supports in a roof for tying the shingles on’), the change *s > *h* in *tasik > Yakan, CS *tahik* ‘sea, saltwater’ must be regarded as irregular. In Ma’anyan, on the other hand, PMP *s became *s* or *h* (Dahl 1951:69), and the same appears to be true of all other languages in the Southeast Barito group (Hudson 1967). Without a larger context reflexes of PMP *tasik ‘sea, saltwater’ in Sama-Bajaw would simply appear to show an unexplained change. Given Ma’anyan, Paku, Dusun Witu, Dusun Malang *tahik* ‘sea’, however, Proto-Sama-Bajaw *tahik has the expected shape for a loanword from an early

¹⁰ This depends on the reflex of *w- in directly inherited forms, which are few in number, and often difficult to distinguish from Philippine loanwords. *walu > Mapun, Yakan, CS *walu* ‘eight’ shows *w- > *w-*, but the unexplained glottal stop raises questions about its history.

Southeast Barito source. This comparison serves to narrow the range of potential donors, as the medial consonant of **tasik* is preserved as a sibilant in other Barito languages, including Malagasy (where it means ‘lake’, and may be a loan from Malay *tasék*, which shows the same semantic change). The adoption of a word for ‘sea, saltwater’ from an early Southeast Barito source can signify only one thing: speakers of Proto-Sama-Bajaw must have been located further inland than the languages from which they borrowed this word.

This inference raises an important point. Ras (1968:194) has drawn attention to the fact that the geomorphology of southeast Borneo has changed significantly over the past millennium. The present basin of the Barito river and its tributaries for nearly 200 miles into the interior (almost as far as Muara Tewe in what is now the middle course of the Barito) has formed from silts deposited at very rapid rate to form coastal mud flats which over time became low-lying swampy interior forest. Prior to AD 800, when Pallesen believes the Sama-Bajaw arrived in the southern Philippines, the present basin of the Barito river was a vast bay at least 80 miles wide at the mouth, bounded on the east by a 200-mile long headland or peninsula formed by the Meratus mountains, and on the west by the foothills of the Schwaner range. To say that the ancestral Sama-Bajaw lived inland from speakers of early Southeast Barito languages therefore does not imply that they were located in the remote interior of Borneo. What is now the territory of the East Barito languages between the Barito river and the western flank of the Meratus range was then located on the eastern shore of this bay. Central-East and Northeast Barito speakers probably inhabited territory around the head of the bay and northward into the interior, and West Barito speakers were at some remove along the western shore. The pattern of Barito loanwords in Samalan languages suggests that the ancestral Sama-Bajaw were in close contact with early Southeast Barito languages, but not with West Barito languages, and that they were away from the sea in relation to Southeast Barito speakers. By a process of elimination this would place them either on the western flank of the Meratus mountains facing the bay (with Southeast Barito speakers between), or further north in the broken hill country near the terminus of the Meratus range, around modern Tanjung.

It may seem strange that a people who have been known to history as ‘sea nomads’ have borrowed the word for ‘sea’ from the language of an interior riverine people (Hudson 1972). But Southeast Barito speakers live on alluvial lands that have only formed within the past millennium, and these lands almost certainly were at the eastern shore of the great bay of the Barito estuary at the time the Sama-Bajaw departed. In relative terms, then, early Southeast Barito speakers were more familiar with the sea than the early Sama-Bajaw. More enigmatically, after they reached the southern Philippines Samalan speakers borrowed the word for ‘sea’ again, this time from Malay, as seen in Mapun *dilawut* ‘sea, ocean’, Yakan *dila’ut/diya’ut* ‘the sea as spoken about from a distance; also when at the shore, further out is *diya’ut*’, CS *kaut/kalaut* ‘sea’ (from Malay *laut*, with the ‘adhesive locative’ noted in Blust 1989b). It is noteworthy that no reported Samalan language has borrowed the Central Philippine form *dagat* ‘sea’. This difference in the semantic fields from which loanwords are drawn undoubtedly reflects significant differences in attitudes toward the speakers of the lending languages. To the Sama-Bajaw the Tausug were landlubbers who gained their familiarity with the sea through contact with the sea nomads. On the other hand, the Malays, as will be seen below, probably were viewed as masters of maritime commerce, and in some sense models for the Sama-Bajaw themselves.

Inference 6: Sriwijayan Malays were trading preserved fish with early Barito speakers

Two other linguistic comparisons are important for different reasons, but both serve to demonstrate that Sriwijayan Malays ancestral to the modern Banjarese must have been trading with speakers of early Southeast Barito languages before the ancestral Sama-Bajaw reached the southern Philippines.¹¹ Each of these comparisons requires a somewhat lengthy discussion. The first is a word for ‘animal’. No term with this meaning can be reconstructed for early Austronesian proto-languages, and the word for ‘animal’ in the modern languages is thus either innovated or borrowed (Blust 2002).

12. Mapun *sattuwa* ‘any kind of creature, beast or animal (domesticated or not)’, Yakan *sattuwa/settuwa* ‘creature (if the name is unknown, esp. of big ones)’, Taboyan, Lawangan *setua*’, Ma’anyan, Paku, Dusun Witu *satua*’, Banjarese *satua* ‘animal’.

For Mapun Collins, Collins, and Hashim (2001) also give *binatang* ‘(Mal.). An animal. Often used when distinguishing between animals that are permissible (*halal*) to eat and those that are forbidden (*haram*),’ and *hayop* ‘animal (usually domesticated, as water buffalo, cattle, chickens, goats, dogs, etc.).’ For Yakan Behrens (2002) gives similar terms, including *binatang* ‘animal; esp. meat (of animals, as carabao, cow, goat, or chicken)’, *hayep* ‘animal (esp. large ones, domestic and wild)’, and *da’at* ‘a destructive animal (like wild pig, rat, monkey, rice bird)’. What lends special interest to comparison 12) is that it derives from Sanskrit *sattva*- ‘creature, animal.’ Gonda (1973) mentions this word in connection with six languages: Javanese *sato* ‘animal’, Old Javanese *sato/sattuwa* ‘living being; creature; animal’, Balinese *sato* ‘animate being, animal, esp. wild beast’, Sasak *sato/sesato* ‘creature; animal; character’, Malay *setua* ‘wild beast; monster’, Toba Batak *santua* ([*sattuwa*]) ‘mouse, rat’. The Balinese and Sasak forms are transparent loans from Javanese, which acquired them through direct contact with Sanskrit. The remaining words raise an interesting point. It is clear that all Sanskrit loanwords in languages of Borneo and the Philippines were acquired through the medium of Malay, as there is no historical or archaeological evidence of direct Indianization of any of these areas, apart from fragmentary inscriptions suggestive of casual contact or abortive colonization. It is also clear that Malay has simplified consonant clusters, including geminate consonants, which it once had in Sanskrit loanwords. Some of these survived long enough to be transmitted to Philippine languages which preserved them after they were lost in the lending language, as with Malay *muka*, but Tagalog *mukhá?* ‘face’, Skt. *mukha* ‘face’. Taboyan, Lawangan *setua*’,

¹¹ The modern city of Banjarmasin is near the mouth of the Barito river, and forms the nucleus of the language community using Banjarese, usually described as a dialect of Malay (Ras 1968:8ff), but lexically no closer to standard Malay than is Iban, which is usually regarded as a different language (Blust 1988). No one knows with certainty how long Banjarese Malay has been used in Borneo. The one certainty is that, unlike the Malayic Dayak languages spoken over much of south and southwest Borneo it is not native to the area. Ras (1968:8) suggests that ‘The Malay linguistic community in South-East Borneo may in fact have had its beginning in a small Malay colony founded in the time of the empire of Sri-Wijaya.’ Its *raison d’être* was trade ‘for Borneo’s traditional forest produce of wood, rattan, dammar, wax, *kaju putih* and ... gold’ (Ras 1968:189). Although the early Banjarese/Sriwijayan Malays appear to have had their most intensive linguistic contacts with speakers of Southeast Barito languages, in more recent centuries (following the merger of *e and *a) the locus of primary contact has shifted to Southwest Barito (Dempwolff 1937, Dyen 1956).

Ma'anyan, Paku, Dusun Witu *satua* 'animal' must, then, have been acquired from Malay, and the same is true of the related words in Samalan languages. But *setua* is an archaic word in Malay, used more in literary than in colloquial contexts, and this apparently has been the case for some centuries. The ordinary spoken word for 'animal' in Malay is *binatang*, and this word has been borrowed, probably from Brunei Malay, into a number of the coastal languages of Sarawak and Sabah (Bintulu, Narum *binatang*, Mukah Melanau *benaténg*, Limbang Bisaya *benatang* 'animal'), into languages of the southern Philippines (Maranao, Tausug *binatang* 'animal', Tiruray *binatang* 'all four-legged, warm-blooded animals, including monkeys', Western Bukidnon Manobo *binatang* 'to have no shame, as an animal has no shame'), and into a few Barito languages (Tunjung, Ba'amang *binatang* 'animal').

Most language communities which have borrowed Malay *binatang* are Islamic, and although this word is not an Arabic loan, the gloss for Mapun *binatang* makes it very clear that this word is associated with Islamic dietary regulations. It is thus a reasonable inference that *binatang* was the common colloquial term for 'animal' from the time that Brunei Malays began to actively spread Islam in coastal regions of Borneo and the southern Philippines, a process that is usually assumed to have begun between 1,350 and 1,400 CE (Saleeby 1963:42ff). Given the association of *binatang* 'animal' with the post-Islamic Malay world, it seems likely that Malay *setua* was last used as a colloquial term during the pre-Islamic period. It follows that the words for 'animal' which are traceable to Sanskrit *sattva*- in both Barito and Samalan languages almost certainly were borrowed from the Sriwijayan Malays who settled in southeast Borneo and gave rise to modern Banjarese. This inference is further supported by the observation that the geminate consonant of the Sanskrit word has been lost in both standard Malay *setua* and in Banjarese *satua*, but is preserved in the other languages.¹²

This brings us to the second comparison which can be used as evidence that Sriwijayan Malays ancestral to the modern Banjarese must have been trading with speakers of early Southeast Barito languages before the ancestral Sama-Bajaw reached the southern Philippines. Comparison 13), which requires the most extensive discussion, shows that PMP *hikan was replaced by *kenah 'fish':

13. Yakan *kenna* 'fish (generic)', Ma'anyan, Samihin, Paku, Dusun Witu, Dusun Malang *kenah* 'fish', Merina Malagasy (Richardson 1885:249) *hena* 'beef; flesh meat',¹³ Tañala Malagasy (Beaujard 1998:317) *hèna* 'fish; meat'.

This is a particularly striking innovation which is both complicated and enriched by two considerations: 1) the shared innovation is semantic, not lexical, and 2) the agreement between Yakan and Southeast Barito languages clearly is a product of borrowing. First, *kenas can be attributed to Proto-Malayo-Polynesian in the meaning 'preserved meat or fish' (Blust 1999). The following glosses illustrate: (WMP) Kadazan Dusun *kanas* 'food taken with rice dishes (fish, meat, vegetables, etc.)', Iban *kenas* 'fish (*ikan*) preserved by steeping or parboiling to remove bones, salting and smoke-curing in bamboo', Malay *kenas* 'shellfish, salted and mixed with rice, sago and other ingredients for preservation as food', Old Javanese *kenas* 'small game, in particular deer', Sangir

¹² Banjarese shows some Javanese linguistic influence, but this cannot account for the occurrence of *sattuwa* in Samalan languages for two reasons: 1. Javanese contact with Banjarmasin began long after the Sama-Bajaw had left southeast Borneo, and 2. as a result of regular 'guna assimilation' in the history of Javanese, by late Old Javanese times the form of the word had already become *sato*.

¹³ Also used in names of fish, as *hena lahy*, or *hena laza*, both given as 'k.o. fish.'

kina? ‘fish (first element in many fish names); flesh in general; prey, victim’, (CMP) Manggarai *kenas* ‘roasted meat which is preserved in a bamboo container’. Reflexes of **kenas* have undergone a semantic change to ‘fish (generic)’ in five cases which appear at first to be historically independent: 1) Yakan, and presumably other Samalan languages for which published data is not available, 2) all languages in the Southeast Barito group, 3) Soboyo (Sula archipelago, central Moluccas) *kena*, 4) Bolongan (northeast Borneo) *kenas*, and 5) Proto-Sangiric **kinas*.

Because it is somewhat different from the other cases, we can consider Proto-Sangiric **kinas* first. Sneddon (1984:87) gives Sangir *kina?* ‘fish; meat (general term)’, but also Sangir, Sangil *kinas-ey* ‘be eaten (of meat)’. Sangiric reflexes of **kenas* thus appear to have a wider signification than ‘fish’. In Yakan, Soboyo, and so far as we can tell from the limited lexical source materials in Southeast Barito languages, the related word is the generic term for ‘fish’, and nothing more. These semantic innovations initially appear to provide evidence for a parallel change with no historical connection, but closer attention to detail shows that this probably is not true. Yakan should reflect **kenas* with a final sibilant, as in **Ratus* > *hatus* ‘hundred’, **nipis* > *nipis* ‘thin, as paper’, **tanjis* > *tangis* ‘cry, weep’, or **beRas* > *buwas* ‘rice (pounded and/or cooked)’. As already seen, in Ma’anyan and other Southeast Barito languages, on the other hand, PMP **s* became *s* or *h*, and most commonly *-h* in final position. Malay loanwords that end in *-h* are borrowed without the final fricative in Samalan languages (**lebiq* > Malay *lebih*, Mapun, Yakan *labi* ‘excess’, **susaq* > Malay *susah*, Mapun *susa*, Yakan *suse* ‘sad, worried’), and *-h* from **R* in native words disappeared after lowering a preceding high vowel. It may be assumed, then, that loanwords with *-h* from any source would be borrowed without the final fricative. If so, the absence of *-s* in Yakan *kenna* ‘fish’ has a simple explanation: this word is a regular continuation of *kenah*, which was borrowed from an early Southeast Barito source. Intriguingly, Soboyo *kena* ‘fish’ is also irregular (expected ***kona*), but in its vocalism, not the loss of **-s*. Since both this phonological irregularity and the shared semantic change can be accounted for by borrowing from a Sama-Bajaw source, it seems likely that Soboyo *kena* is a loan from one of the Sama-Bajaw communities of the northern Moluccas. The initial impression that reflexes of **kenas* show five independent changes from ‘preserved fish’ to ‘fish’ is thus reduced as three independent changes, of which one appears to cover a broader range.

But why would a semantic change from ‘preserved fish’ to ‘fish (generic)’ take place in the first place? What this comparison suggests is that some form of preserved fish was an important trade item at the time of the Sama-Bajaw migrations to the Philippines. While the Sama-Bajaw may have played a part in conducting this trade they could hardly have initiated it, since they themselves evidently borrowed the word *kenah* from an early Southeast Barito source. What, then, might have motivated a semantic change of this kind in the Southeast Barito languages themselves? If preserved fish came to be more important in the daily diet than fish freshly caught, the word for ‘preserved fish’ might well come to replace the inherited reflex of PMP **hikan* ‘fish’. To gain perspective on how this might happen we need to widen our view. Several of the West Barito languages also have an innovative word for ‘fish’, seen in Kapuas, Katingan *lauk*, Ba’amang *lau?*. This word is cognate with, and apparently borrowed from Malay *lauk* ‘solid food (fish or flesh) to be eaten with rice’; etym. of fish-food ... and often used in that sense, esp. in Minangkabau; *lauk pada* ‘small fish pickled in brine’; *lauk pauk* ‘curried fish and meat of all sorts’; cry of the fish-sellers (Minangkabau) PMP **lahuk* ‘to mix (types of food)’. Revealingly, Mapun and Central Sulu Sama have another apparently borrowed word for ‘fish’: Mapun *daying* ‘fish (generic)’, CS *daing* ‘fish’. These words exhibit phonological irregularities which indicate borrowing, almost certainly

from Malay *daéng/dahing* ‘slicing into thin strips and drying in the sun. Esp. of fish preserved (*ikan daéng*) after being cut in two along the line of the vertebrae.’ With this set of etymologies, then, we see a word that means ‘preserved fish’ in the lending language borrowed as the generic term for ‘fish’ in the receiving language. Given this context it would not be surprising if *kenah* in Southeast Barito languages were also an early borrowing of Malay *kenas*.¹⁴

Inference 7: Trade drew both the Malagasy and the Sama-Bajaw out of southeast Borneo

The linguistic comparisons presented here force us to conclude that the Sama-Bajaw originated in what is now the basin of the Barito river. Along with some speakers of early Southeast Barito languages, they were drawn into trade with Sriwijayan Malays, and eventually became sea nomads centered in the southern Philippines. The parallels with the Malagasy migration from the Barito basin to Madagascar can hardly escape notice, nor can these parallels simply be an accident. Adelaar (1989) has pointed out the crucial role of Sriwijayan Malays in making the Malagasy migration possible, since the ancestral Malagasy would have been riverine people, while the Malays of Sriwijaya traded widely throughout insular Southeast Asia, and were skilled sailors. One important piece of evidence that Adelaar uses to illustrate the dependence of the early Malagasy on Malay seafarers is the terminological system for wind directions, critical to sailing, which he argues were borrowed from Sriwijayan Malay into Malagasy. Virtually the same eight-point system has been borrowed from Malay into Samalan languages (Collins, Collins and Hashim 2001:87). If we are to draw inferences from the linguistic data, then, it would appear that both the ancestral Malagasy and the ancestral Sama-Bajaw were schooled in sailing by Sriwijayan Malays. In fact, it may not be too extreme to suggest that were it not for Sriwijayan Malay contact with southeast Borneo neither the Malagasy nor the Sama-Bajaw would have come into existence.

From this point we have little linguistic data to guide us, but it does seem possible to venture a few speculations that are not entirely groundless. First, why would Sriwijayan Malays have been interested in southeast Borneo rather than in other parts of Borneo that are closer to southern Sumatra? The answer appears to be topography. As noted above, in the late first millennium A.D. what is now the Barito river basin was an enormous bay, sheltered on the southeast by the Meratus mountains, which formed a 200-mile long peninsula, and on the west by the peaks and foothills of the Schwaner range. According to Hall (1985:78) ‘The state of Sriwijaya dominated maritime commerce passing through Southeast Asia between A.D. 670 and 1025.’ Commerce in an insular environment obviously depends on shipping, and the success of shipping depends on good anchorage, an association which is amply demonstrated through the repeated rise of commercial centers large and small in locations with sheltered waters. Examples in insular Southeast Asia include Manila Bay, where Fukienese merchants had established important trade connections with the local Filipinos before the advent of the Manila Galleon, Brunei Bay, which fostered the Brunei sultanate, and Ambon harbor, the most important collection point for the spice trade in eastern Indonesia. In enlarging their trade contacts Malays, who were excellent sailors, would certainly have sought that part of the Borneo coast which provided the best anchorage and shelter

¹⁴ Modern Banjarese has merged PMP *e and *a, but this is a change which presumably took place after a number of Sriwijayan Malay loanwords had already entered early Southeast Barito languages.

from the annual monsoons. The unnamed bay in the Barito basin provided an ideal location for vessels to load jungle products from the Bornean hinterland, as the waters were shallow, and were protected on both sides (but especially on the east) by mountain barriers. In time some of these Sumatran Malays established a permanent colony near the mouth of the Barito river, the ancestor of the modern Banjarese. It is a matter of speculation when this trade began and when a Malay-speaking colony was established. However, if Pallesen is correct in placing the arrival of the Sama-Bajaw in the Zamboanga-Sulu area by about AD 800, it would be reasonable to place the onset of Sriwijayan contact with southeast Borneo in the period 670–800 AD. Dahl (1991:49ff) argues that the Kota Kapur stone inscription from the island of Bangka, which carries the date 686 AD contains one line of Old Ma'anyan text, suggesting that Sriwijayan Malays were already in contact with southeast Borneo at that time.

Although Sriwijayan trade may have been multifaceted, one key element for many centuries clearly was the shipment of cloves and nutmeg from the tiny islands of eastern Indonesia where they were produced, to the vicinity of the Malay-controlled Strait of Malacca, where they could be exported to external markets. So far as I know, the connection has never been explicitly drawn, but the distribution of Malay dialects outside the Malay 'homeland' of southwest Borneo, the Malay peninsula and southern Sumatra, has left a linguistic trail showing the routes by which spices were shipped from Moluccan producers to Sumatran Malay distributors (always under the control of Malay speakers). The northern route is marked by Ambonese Malay, Manado Malay, and Brunei Malay, the southern route by Kupang Malay, Larantuka Malay, and Jakarta Malay. Why did the ancestral Sama-Bajaw move north, while the Malagasy moved west? In many ways the motivation for the better-known Malagasy migration is more difficult to fathom than that for the movement of the Sama-Bajaw. The early Sama-Bajaw may well have been drawn into the southern Philippines by the Malay-dominated spice trade. A location in the southern Zamboanga-Sulu area would have placed them almost midway between Manado and Brunei, and the heavy influx of Malay loanwords into Samalan languages attests to their continuing intense contact with Malay speakers over a period of several centuries. For reasons that remain obscure, the Malagasy left this lucrative intra-archipelagic network to seek their fortune elsewhere, but the Sama-Bajaw may have continued to participate in it until Malay control was lost to the incoming European powers in the latter part of the sixteenth century.

7. The Proto-Philippines Homeland

The last topic I would like to touch on, however briefly, is the location of Proto-Philippines. Radiocarbon dates associated with plain or red slipped pottery and with small, square, double-walled houses which contain interior hearths, show that Neolithic cultures were established in the Cagayan Valley of northern Luzon by 4,800 B.P. (Bellwood 1997:220, Spriggs 1989:593).¹⁵ As noted already, however, these remains mark the earliest known presence of Neolithic cultures in the Philippines, and it is unlikely that Proto-Philippines began to differentiate until perhaps a millennium later. Sometime around 3,500 B.P., then, Proto-Philippines speakers began a territorial expansion that effectively eliminated other early Austronesian languages in the Philippines either through absorption, or through flight. A priori this scenario seems

¹⁵ According to Peter Bellwood (pers. comm., December 17, 2002) these dates are not to be trusted, and the oldest reliable dates from the Cagayan valley probably are no earlier than 4,000 B.P.

implausible, as it is difficult to imagine how one Austronesian-speaking community would be able to secure such a decisive advantage over others at a similar stage of culture. Latin replaced all other ancient languages of the Italian peninsula, but this was through its role as the language of a vast military, political, economic, and cultural empire. Clearly, whatever conditions were conducive to language expansion and extinction in the early Philippines must have differed substantially from those which favored the spread of Latin.

What would the population distribution have been like in the prehistoric Philippines circa 3,500 B.P.? First, speakers of Proto-Malayo-Polynesian had a rich vocabulary relating to the sea and its economic products. This marine orientation was fundamental to early Austronesian societies, and is reflected in all parts of the Austronesian world by evidence that coastal zones were preferred over inland areas until some time after the initial period of colonization. We are probably safe in assuming, then, that 3,500 years ago the mountainous interior regions of the Philippines were the exclusive preserve of the pre-Neolithic Negrito populations, and that wherever possible Austronesian-speaking groups continued to maintain contact with the sea and its rich inventory of economically useful products. At the same time, these Austronesian speakers were growing rice and millet, as shown by widespread cognate sets relating to grain agriculture in general, and to rice in particular (Blust 1995). By 3,500 B.P. the Austronesian languages of the Philippines had been diverging from one another for a millennium or more, and the situation we might expect during this period is a collection of several dozen closely related but distinct languages, spread around the coastal zones of the major islands.

Bellwood (1995) has argued that the Austronesian expansion was driven by rice agriculture. The cultivation of grain crops encouraged territorial expansion for two reasons. First, in the beginning at least, it almost certainly was based on swiddening, and so required frequent shifting of the area put under cultivation. Second, rice agriculture permits greater population densities than are found in foraging or horticultural societies, even with swidden agriculture. With wet rice soils are not exhausted so quickly, but population densities increase much more rapidly, putting pressure on local resources and encouraging migration. Historically, the 'rice granary' of the Philippines has been the volcanically-enriched Central Plain stretching from Lingayen Gulf in the north to Manila Bay in the south, and from the Zambales mountains in the west to the southern Cordillera in the east. Once early Austronesian speakers reached this area the opportunities for population expansion would have increased dramatically. Linguistic evidence does not permit us to determine with any degree of certainty the antiquity of wet rice agriculture in the Philippines (Reid 1994b), but either through swiddening or wetfield rice cultivation the language group that arrived in this area first probably would have had opportunities for rapid population growth that greatly exceeded those of other areas. While the territorial expansion of Latin was due to political and military factors, then, the territorial expansion of Proto-Philippines may have been due to a kind of geographical determinism: linguistic groups that reached an advantageous location first were most likely to expand at the expense of others.

Wernstedt and Spencer (1967:368ff) nonetheless make it clear that the agricultural productivity of the Central Plain is a relatively recent phenomenon. In 1571, when the Spaniard Martin de Goiti quickly pacified central Luzon, the interior regions of the Central Plain in the present provinces of Tarlac and Nueva Ecija were heavily forested, and almost unpopulated. However, in Lower Pampanga Province, within about 10 miles of Manila Bay the population density increased dramatically, and a similar pattern on a lesser scale was observed within about 10 miles of Lingayen Gulf.

The economy of early Philippine societies thus depended on striking a balance between exploiting the resources of the sea and those of the land: rice farmers needed good rice lands, but never far from the coast. Since this presumably was true throughout the archipelago, the richer rice lands surrounding Manila Bay would have provided a relative abundance of food, and consequent opportunities for rapid population growth in relation to other areas. Considering all of the relevant facts, then, it seems likely that the Proto-Philippines expansion began from the region of central Luzon surrounding Manila Bay. Moreover, the attested distribution of primary subgroups of Philippine languages is consistent with this view: Bashiic, Cordilleran, and most of Central Luzon to the north, Inati and Kalamian slightly to the south, and GCP, Bilic, Sangiric and Minahasan further to the south, with a possibility that Sangiric and Minahasan share a common node under Proto-Philippines.

Many questions naturally remain unanswered. Does the extreme southern region of the Central Plain of Luzon just to the north of Manila Bay offer sufficiently great agricultural advantages in itself to support the inference I make, or is it necessary to assume without the benefit of any known supporting evidence that wet rice agriculture was first practiced in this area, and that this supplied the needed demographic advantage for expansion? Why did the Proto-Philippines expansion stop at the geographical boundaries of the Philippines, while the GCP expansion leaped beyond to northern Sulawesi, and produced a major loan influence in Sabah? How could all earlier Austronesian languages in the Philippines, including those that must have begun to diverge very early in the Cagayan valley, have been replaced by a single expanding population speaking Proto-Philippines? These and other questions undoubtedly will continue to nag us, but we cannot ignore the observations which led us to ask them in the first place. Much of what happened in the past is lost to us forever, but more than is commonly appreciated can be retrieved by resourceful use of the available evidence. In historical linguistics this need not be limited to the traditional use of the Comparative Method, but as I hope to have shown here, much can also be gained from a consideration of the large-scale distributional observations that form the basis for inferences relating to 'macrohistory.'

Appendix 1¹⁶

Proto-Philippine lexical innovations proposed in Zorc (1986) for which external evidence has since been found.

1. PPH *qinit ‘heat of the sun’. BUT: Totoli *init* ‘hot’, Pamona, Wolio *ini* ‘sweat’.
2. PPH *lújan ‘ride; load’. BUT: Mandar *ruray*, Makasarese *luray* ‘load, cargo’, Bugotu *lu-luja* ‘cargo’, Nggela *lunda* ‘load a canoe or ship with cargo; embark passengers; cargo’, Arosi *ruta* ‘load a canoe; carry (s.t.) to canoe and stow’.
3. PPH *púnas ‘to wipe’. BUT: Iban *punas* ‘barren, childless, with no direct heirs; died out, having no survivors; wipe out, destroy’, Chamorro *funas* ‘eradicate, erase, wipe out, put an end to’. All known Philippine reflexes of *punas mean ‘to wipe away (dirt, crumbs), take a sponge bath’ and the like. Given the semantic agreement of the Iban and Chamorro forms it appears that PPH *púnas ‘to wipe’ may contain a semantic innovation.
4. PPH *sujud ‘fine-toothed comb’. BUT: Bintulu *surud*, Nias *suxu* ‘comb’.
5. PPH *a-núh ‘what?’. BUT: Murik *ano?* ‘which, which one?; thing mentioned’, Bintulu *anew ineh* ‘what?’, Moken *ano-ŋ* ‘what?’
6. PPH *ʔudu ‘medicine’. BUT: Rungus Dusun *uru* ‘medicine’, Minangkabau *uduah* ‘incantation; k.o. basket with cooked rice, sugar and grated coconut, set in the field during rice-planting to guard against calamity’, Pamona *uru* ‘taboo sign or charm used to protect fruit trees, temporary shelters, etc. from thieves’. Since Rungus Dusun *uru* may be a GCP loanword, it might be argued that PPH *ʔudu ‘medicine’ shows a semantic innovation. However, Sangir *uro* ‘charm, spell (against thieves of fruit in plantation)’, Tontemboan *uru?* ‘charm, spell concealed on a fruit tree to prevent the theft of fruit’, Bolaang Mongondow *uyu?* ‘charm used to safeguard one’s property, but also used to inflict harm on others’ suggests that the change of meaning to ‘medicine’ postdated the break-up of PPH.
7. PPH *ʔuŋá? ‘child’. BUT: Lauje, Tialo, Dondo *uŋa* ‘child’. Possibly a GCP loanword.
8. PPH *buqél ‘leg joint (knee, ankle, etc.)’. BUT: Buli *pu-puo* ‘knee’.
9. PPH *keRáj ‘scab’. BUT: Kavalan *keran* ‘scab’. These forms are etymologically discrepant with respect to the final nasal.
10. PPH *ma- ‘one unit’ (10, 100, 1,000). BUT: *ma-Ratus > Long Semadoh *me-ratu* ‘100’, *me-ribu* ‘1,000’, Bintulu *matus* (Ray 1913), Miri *ma-rataw* ‘100’, *ma-ribuh* ‘1,000’, Kiput *ma-lataw* ‘100’, *ma-libo* ‘1,000’
11. PPH *n-atáy ‘dead’. BUT: Kelabit *n-ate* ‘was killed’.
12. PPH *tulqid ‘straight’. BUT: Sebop, Long Anap Kenyah *telit* ‘straight’.

¹⁶ David Zorc offered useful comments on an earlier version of this appendix. I am much indebted to him, and am immensely gladdened by his return to the field of Austronesian linguistics.

Appendix 2

Additional innovations which can be attributed to Proto-Philippines (subgroup distribution is indicated by numerical code, where 1 = Bashiic, 2 = Cordilleran, 3 = Central Luzon, 4 = Inati, 5 = Kalamian, 6 = Greater Central Philippines, 7 = Bilic, 8 = Sangiric, 9 = Minahasan; for supporting evidence cf. Blust 1999).

1.	PPH *aba (PMP *baba) ‘carry pick-a-back’	(26)
2.	PPH *abag ‘join forces, cooperate in working’	(26)
3.	PPH *abal-abal ‘beetle species’	(27)
4.	PPH *abat ‘spirit that causes sickness’	(26)
5.	PPH *abij ‘to copulate’	(26)
6.	PPH *abijay ‘to sling over the shoulder’	(26)
7.	PPH *abuR ‘to chase, drive away’	(26)
8.	PPH *alinaw (PMP *qaninu) ‘shadow’	(26)
9.	PPH *ma-anay ‘termite-infested’	(36)
10.	PPH *aŋal ‘groan, cry of pain’	(26)
11.	PPH *paŋ-ampu ‘descendants’	(26)
12.	PPH *ampu-ampu ‘forefathers, ancestors’	(26)
13.	PPH *apúy ‘to break out in boils’	(26)
14.	PPH *ka-asu ‘canine companion’	(29)
15.	PPH *asúd ‘to pound rice in tandem’	(26)
16.	PPH *atá ‘expression of displeasure or surprise’	(26)
17.	PPH *atáb ‘flood tide’	(236)
18.	PPH *antábay ‘to stay with, accompany’	(26)
19.	PPH *atáj ‘a sacrifice to the spirits’	(26)
20.	PPH *átaj-átaj ‘poisonous arthropod’	(26)
21.	PPH *ántij ‘to hear at a distance’	(26)
22.	PPH *ayáw ‘to depart, separate from’	(26)
23.	PPH *bábad ‘to soak’	(26)
24.	PPH *balalaŋ ‘sandpiper, snipe’	(28)
25.	PPH *balaqi ‘co-parent-in-law’	(367)
26.	PPH *bala(R)baR ‘crosswise, athwart’	(267)
27.	PPH *balát ‘sea cucumber’	(26)
28.	PPH *balatuŋ ‘mung bean’	(26)
29.	PPH *baláw ‘negative marker’	(26)
30.	PPH *balaw ‘small shrimp sp.’	(26)
31.	PPH *baláyaŋ ‘banana sp.’	(26)
32.	PPH *báli ‘join, participate in, accompany’	(269)
33.	PPH *balíkes ‘encircle, wrap around’	(26)

34.	PPH *baliketád ‘reverse, turn around’	(26)
35.	PPH *balikid ‘reverse, turn over or around’	(26)
36.	PPH *balíkis ‘tie around; belt’	(26)
37.	PPH *balikutkút ‘to bend, curl up’	(26)
38.	PPH *balilit ‘edible snail sp.’	(26)
39.	PPH *balisúsu ‘kingfisher’	(26)
40.	PPH *balítíq ‘banyan, strangler fig’	(1236)
41.	PPH *balñáw ‘to rinse, rinse off’	(26)
42.	PPH *balúlaŋ ‘large open-work basket’	(26)
43.	PPH *balútu ‘dugout canoe’	(26)
44.	PPH *banabá ‘a tree: Lagerstroemia speciosa’	(236)
45.	PPH *banÉR ‘to bruise, raise welts’	(26)
46.	PPH *banhéd ‘numb, as a limb’	(26)
47.	PPH *baniákaw ‘tree sp.’	(26)
48.	PPH *banug/banuR ‘hawk, eagle’	(26)
49.	PPH *banuít/benuit ‘fishhook’	(26)
50.	PPH *banútan ‘tree sp.’	(26)
51.	PPH *baraŋgay (dbl. *baraŋay) ‘communal boat’	(26)
52.	PPH *baranbán ‘small marine fish’	(26)
53.	PPH *barat ‘tree with fragrant flowers’	(26)
54.	PPH *baráyeŋ ‘tree sp.’	(26)
55.	PPH *baRat ‘to meet’	(16)
56.	PPH *behñat ‘to stretch’	(26)
57.	PPH *beláy ‘tired, weary; grow tired’	(26)
58.	PPH *beŋát ‘to state, express in words’	(26)
59.	PPH *beŋet ‘moustache, whiskers’	(67)
60.	PPH *beŋqáw ‘chasm, precipice’	(26)
61.	PPH *beŋuit ‘fishhook’	(26)
62.	PPH *beRnát ‘relapse’	(26)
63.	PPH *berták ‘to collide, of hard objects’	(26)
64.	PPH *beskáj ‘to open, of a flower in bloom’	(26)
65.	PPH *maR-besuR ‘satiated’	(26)
66.	PPH *beték ‘bundle of rice stalks’	(26)
67.	PPH *pa-betu ‘make an exploding sound’	(67)
68.	PPH *biaR ‘satiated’	(56)
69.	PPH *biglák ‘suddenly’	(236)
70.	PPH *bijaw ‘winnowing basket’	(26)
71.	PPH *biŋkit ‘joined along the length’	(567)
72.	PPH *biklaj ‘to spread out, unfurl’	(169)
73.	PPH *biklát ‘scar’	(26)

74.	PPH *ma-bílanj ‘countable’	(36)
75.	PPH *bílu ‘blackened’	(26)
76.	PPH *binuŋa ‘a tree: <i>Macaranga tanarius</i> ’	(68)
77.	PPH *bisíbis ‘sprinkle water on something’	(26)
78.	PPH *bítay ‘to hang’	(267)
79.	PPH *bitek ‘intestinal worm’	(679)
80.	PPH *bítin ‘to hang, suspect’	(23678)
81.	PPH *bítu ‘hole, cavern, pitfall trap’	(126)
82.	PPH *biu ‘tree sp.’	(26)
83.	PPH *buág/buál ‘uproot a tree’	(267)
84.	PPH *buás ‘tomorrow’	(26)
85.	PPH *maka-buta ‘become blind’	(68)
86.	PPH *búgaq ‘pumice’	(236)
87.	PPH *bugrís ‘diarrhoea’	(26)
88.	PPH *bugtúnj ‘alone, single’	(26)
89.	PPH *bujas ‘to pluck, as fruit’	(12)
90.	PPH *bujiq ‘fish eggs, roe’	(267)
91.	PPH *buklad ‘unfold, open up, blossom’	(16)
92.	PPH *buknúl ‘knot, lump’	(26)
93.	PPH *buktút ‘bulge, as hunchback’	(236)
94.	PPH *bulág (dbl. PMP *bulaR) ‘ocular cataract’	(2378)
95.	PPH *bulalákaw ‘spoiled coconut’	(26)
96.	PPH *bulalákaw ‘meteor, shooting star’	(2368)
97.	PPH *bulilit ‘dwarfish, small (of people)’	(26)
98.	PPH *bulínaw ‘anchovy: <i>Stolephorus</i> spp.’	(26)
99.	PPH *bulud ‘to borrow, lend’	(12)
100.	PPH *bunal ‘to beat up, bruise someone’	(68)
101.	PPH *búnuj ‘edible mushroom sp.’	(26)
102.	PPH *buŋáŋaq ‘to open the mouth wide’	(68)
103.	PPH *buŋaw ‘scrotum, testicles’	(26)
104.	PPH *buŋtút ‘stench, bad odor’	(12)
105.	PPH *buqetís ‘pregnant; pregnant woman’	(26)
106.	PPH *búquŋ ‘broken, shattered’	(26)
107.	PPH *busiq/busiqsiq ‘split, rip open’	(26)
108.	PPH *busuáŋ ‘strong free flow of water’	(126)
109.	PPH *butakál ‘male pig, boar’	(26)
110.	PPH *butí ‘swine disease; pockmarks’	(268)
111.	PPH *butík ‘spotted, dappled, speckled’	(236)
112.	PPH *butikíq ‘house lizard, gecko’	(26)
113.	PPH *bútil ‘satiated’	(26)

114. PPH *butíq (doublet *beRtiq) ‘popped rice’ (26)
115. PPH *butúy ‘swollen, of the flesh’ (26)
116. PPH *buybuy ‘silk cotton tree: *Ceiba pentandra*’ (26)
117. PPH *buyug ‘bumblebee’ (26)
118. PPH *buyúk ‘rotten, stinking’ (26)
119. PPH *dayaw ‘fame, glory; praise’ (268)
120. PPH *dalág ‘freshwater mudfish’ (126)
121. PPH *dúdun ‘grasshopper, locust’ (236)
122. PPH *edeg ‘back of humans or animals’ (26)
123. PPH *enép ‘conceal one’s feelings’ (29)
124. PPH *eɲus ‘to sniffle, pant’ (69)
125. PPH *epit ‘fibers at base of coconut frond’ (26)
126. PPH *maka-esa ‘once’ (69)
127. PPH *eták ‘bush knife, machete’ (26)
128. PPH *hábas ‘tumor in the mouth of an animal’ (26)
129. PPH *hajek-an ‘to kiss’ (126)
130. PPH *hambúg ‘proud, boastful’ (26)
131. PPH *haɲút ‘to chew or gnaw on’ (26)
132. PPH *hampak ‘to slap, to smack’ (26)
133. PPH *haprus/hapRus ‘to rub, massage’ (26)
134. PPH *hápun ‘to roost, of fowls; time of roosting’ (26)
135. PPH *háyp ‘animal’ (236)
136. PPH *hebás ‘to evaporate, dry up’ (26)
137. PPH *hendék ‘to moan, to groan’ (26)
138. PPH *hediq (dbl. PMP *hadiq) ‘no, not’ (26)
139. PPH *helék ‘to sleep’ (256)
140. PPH *henaq ‘to think, consider; thought, idea’ (68)
141. PPH *hideRáq ‘to lie down’ (267)
142. PPH *higpít ‘to pinch or squeeze’ (26)
143. PPH *hikñat ‘to stretch’ (26)
144. PPH *hiláw ‘raw, unripe’ (236)
145. PPH *híli ‘village, town’ (12)
146. PPH *hilít ‘outskirts, edge of settlement’ (26)
147. PPH *hílut ‘to massage, set bones’ (236)
148. PPH *himatún ‘to notice, observe’ (26)
149. PPH *hínam ‘to crave, desire intensely’ (267)
150. PPH *híɲak ‘sound of rushing air or water’ (26)
151. PPH *hípuq ‘to feel, touch’ (68)
152. PPH *hírig ‘to lean, incline’ (236)
153. PPH *hírud ‘to scrape or rub off’ (26)

154. PPH *hiRét 'to tighten; constriction' (26)
155. PPH *hium 'to close the lips; smile' (26)
156. PPH *hubád 'to untie, unravel' (26)
157. PPH *i-huluR 'to lower, drop' (26)
158. PPH *huñat 'to stretch, straighten out' (26)
159. PPH *huyáp 'to count, enumerate' (26)
160. PPH *iba 'bad, evil' (267)
161. PPH *íbun 'bird sp.' (26)
162. PPH *ígat 'eel' (26)
163. PPH *inguk 'deep throaty sound' (26)
164. PPH *ikamen 'woven mat' (26)
165. PPH *íliw 'homesick; yearn for something' (26)
166. PPH *imún 'jealous, envious' (26)
167. PPH *ímut 'stingy, selfish' (236)
168. PPH *inqiñ 'shrill sound' (26)
169. PPH *ipa- 'verb prefix: 3rd p. causative-command' (26)
170. PPH *ipil-ípil 'a shrub: *Leucaena glauca*' (267)
171. PPH *ípun 'rice variety' (27)
172. PPH *irij 'similar; to imitate' (267)
173. PPH *ísip 'thinking, thought, opinion' (236)
174. PPH *isul 'to retreat, move backward' (26)
175. PPH *íwas 'to avoid, evade (as people)' (26)
176. PPH *íwik 'to squeal' (26)
177. PPH *íwit 'rear part' (26)
178. PPH *pala pala 'temporary shed' (26)
179. PPH *palakaq 'frog' (16)
180. PPH *p-al-ikpik 'fin of a fish' (26)
181. PPH *pális 'animal sacrifice' (26)
182. PPH *palít 'to change, exchange' (26)
183. PPH *panúqus 'stench, sour smell' (26)
184. PPH *paqlín 'visible defect of the eyes' (26)
185. PPH *paRbu 'one of the four principal rafters' (26)
186. PPH *p-aR-ukpuk 'sound of bubbling, knocking' (26)
187. PPH *p-aR-utput 'gaseous defecation' (26)
188. PPH *pulpul 'blunt, lacking a point' (26)
189. PPH *alik-qabuk 'dust' (26)
190. PPH *qabut 'to overtake, catch up with' (256)
191. PPH *qambijay 'carry slung over the shoulder' (26)
192. PPH *qamih-an-an 'north; rainy season' (26)
193. PPH *pañ-qanup 'hunting dog' (26)

194. PPH *qaŋesú 'stench of urine' (26)
195. PPH *qaŋetéj 'stench of burning hair' (26)
196. PPH *qasawa-en 'to marry, take a spouse' (26)
197. PPH *qelad 'sheet; flat, wide object' (56)
198. PPH *qémel 'squeeze into a ball' (56)
199. PPH *qeŋqéŋ 'whining sound' (26)
200. PPH *qepés 'deflated, shrunken' (256)
201. PPH *maR-qeti 'to lower, of water level' (26)
202. PPH *qimún 'sexual jealousy' (256)
203. PPH *qínut 'use sparingly' (26)
204. PPH *maŋ-quay 'to gather rattan' (26)
205. PPH *quhay 'rice panicle' (56)
206. PPH *quhut 'sheaf of rice grains' (56)
207. PPH *quRis 'unusually white; albino' (568)
208. PPH *ipa-qútaŋ 'to lend, give a loan to someone' (26)
209. PPH *quyag 'living, alive' (567)
210. PPH *sigpít 'clamp, clasp; tongs' (26)
211. PPH *simsím 'to taste, sip something' (26)
212. PPH *si-nuh 'who?' (1236)
213. PPH *siqit 'fish bone, spine' (26)
214. PPH *tampipiq 'telescoping basket' (26)
215. PPH *túkud 'to sound, fathom; guess' (26)
216. PPH *i-uliq 'to take something home' (16)
217. PPH *úlit 'to repeat' (26)
218. PPH *ulitau 'bachelor, young unmarried man' (26)
219. PPH *unah-an 'first' (26)
220. PPH *uŋaq 'child' (26)
221. PPH *úŋur 'to moan, growl' (26)
222. PPH *uRtuh 'zenith; noon' (67)
223. PPH *wákat 'to scatter, strew about' (26)
224. PPH *waksí 'to shake, flick off' (26)
225. PPH *waliŋ-waliŋ 'orchid sp.' (16)
226. PPH *walwál/walwár 'work from side to side' (26)
227. PPH *waŋáwaŋ/waŋwaŋ 'wide open space' (26)
228. PPH *warák 'to scatter, strew' (26)
229. PPH *wáras 'to distribute, deal out' (26)
230. PPH *wasay-wasay 'insect sp.' (26)
231. PPH *wasíwas 'wave back and forth (hand, flag)' (26)
232. PPH *waswás 'rinse clothes by shaking in water' (26)
233. PPH *waswás 'tear apart, undo something made' (26)

- | | | |
|------|--|------|
| 234. | PPH *watnág ‘to scatter, disperse’ | (26) |
| 235. | PPH *wáwaq ‘mouth of a river’ | (26) |
| 236. | PPH *waywáy ‘to dangle, hang down loosely’ | (26) |
| 237. | PPH *wigwíg ‘to shake something’ | (26) |
| 238. | PPH *wíqwiq ‘to slit open’ | (26) |
| 239. | PPH *witwít ‘to chirp’ | (26) |
| 240. | PPH *witwít ‘to shake, wave (hand)’ | (26) |
| 241. | PPH *wiwí ‘shrill whistle or chirrup’ | (12) |

References

- Adelaar, K. Alexander. 1989. Malay influence on Malagasy: Linguistic and culture-historical implications. *Oceanic Linguistics* 28:1–46.
- Beaujard, Philippe. 1998. *Dictionnaire Malgache-Français: dialecte Tañala, Sud-Est de Madagascar*. Paris: L'Harmattan.
- Beech, Mervyn W.H. 1908. *The Tidong dialects of Borneo*. Oxford: The Clarendon Press.
- Behrens, Dietlinde. 2002. *Yakan-English dictionary*. Linguistic Society of the Philippines Special Monograph Issue No. 40, Vol. 2. Manila: Linguistic Society of the Philippines.
- Bellwood, Peter. 1995. Austronesian prehistory in Southeast Asia: Homeland, expansion and transformation. *The Austronesians: Historical and comparative perspectives*, ed. by Peter Bellwood, James J. Fox, and Darrell Tryon, 96–111. Canberra: Department of Anthropology, The Australian National University.
- Bellwood, Peter. 1997. *Prehistory of the Indo-Malaysian archipelago*. Revised ed. Honolulu: University of Hawai'i Press.
- Benedict, Paul K. 1990. How to tell Lai: An exercise in classification. *Linguistics of the Tibeto-Burman Area* 13(2):1–26.
- Blake, Frank Ringgold. 1906. Contributions to comparative Philippine grammar (Part One). *Journal of the American Oriental Society* 27:317–396.
- Blust, Robert. 1988. Malay historical linguistics: A progress report. In *Rekonstruksi dan cabang-cabang Bahasa Melayu Induk*, ed. by Mohd. Thani Ahmad and Zaini Mohamed Zain, 1–33. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Blust, Robert. 1989a. A note on semantic cycles in historical change. *Diachronica* 6:297–300.
- Blust, Robert. 1989b. The adhesive locative in Austronesian languages. *Oceanic Linguistics* 28:197–203.
- Blust, Robert. 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30:73–129.
- Blust, Robert. 1992. The Austronesian settlement of mainland Southeast Asia. *Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society*, ed. by Karen L. Adams and Thomas John Hudak, 25–83.
- Blust, Robert. 1995. The prehistory of the Austronesian-speaking peoples: A view from language. *Journal of World Prehistory* 9:453–510.
- Blust, Robert. 1998. The position of the languages of Sabah. *Pagtanáw: Essays on language in honor of Teodoro A. Llamzon*, ed. by Ma. Lourdes S. Bautista, 29–52. Manila: Linguistic Society of the Philippines.
- Blust, Robert. 1999. Austronesian Comparative Dictionary (ACD). Available online by request.
- Blust, Robert. 2000a. Chamorro historical phonology. *Oceanic Linguistics* 39:83–122.

- Blust, Robert. 2000b. Whence the Malays? Paper presented at a Colloquium on Borneo as the homeland of Malay, Institut Alam dan Tamadun Melayu (ATMA), Universiti Kebangsaan Malaysia, March, 2000. Ms., 33pp.
- Blust, Robert. 2002. The history of faunal terms in Austronesian languages. *Oceanic Linguistics* 41:89–139.
- Blust, Robert. To appear. The linguistic position of Sama-Bajaw. Ms.
- Casiño, Eric. 1976. *The Jama Mapun: A changing Samal society in the southern Philippines*. Quezon City: Ateneo de Manila University Press.
- Charles, Mathew. 1974. Problems in the reconstruction of Proto-Philippine phonology and the subgrouping of the Philippine languages. *Oceanic Linguistics* 13:457–509.
- Coedes, Georges. 1968 [1964]. *The Indianized states of Southeast Asia*, ed. by Walter F. Vella, and trans. by Susan Brown Cowing. Honolulu: The University Press of Hawai'i.
- Collins, Millard A., Virginia R. Collins, and Sulfilix A. Hashim. 2001. *Mapun-English dictionary*. Manila: Summer Institute of Linguistics — Philippines, Inc.
- Conant, Carlos Everett. 1910. The RGH law in Philippine languages. *Journal of the American Oriental Society* 30(1):70–85.
- Dahl, Otto Chr. 1951. *Malgache et Maanjan: une comparaison linguistique*. Oslo: Egede-Instituttet.
- Dahl, Otto Chr. 1977. La subdivision de la famille Barito et la place du malgache. *Acta Orientalia* 38:77–134.
- Dahl, Otto Chr. 1991. *Migration from Kalimantan to Madagascar*. Oslo: The Institute for Comparative Research in Human Culture, Norwegian University Press.
- Dempwolff, Otto. 1937. *Vergleichende Lautlehre des austronesischen Wortschatzes, vol. 2: Deduktive Anwendung des Urindonesischen auf austronesische Einzelsprachen*. Berlin: Dietrich Reimer.
- Diamond, Jared. 1992. *The third chimpanzee: The evolution and future of the human animal*. New York: Harper Collins.
- Dyen, Isidore. 1956. The Ngaju-Dayak 'Old speech stratum'. *Language* 32:83–87.
- Dyen, Isidore. 1965. *A lexicostatistical classification of the Austronesian languages*. Indiana University Publications in Anthropology and Linguistics. Memoir 19 of the International Journal of American Linguistics. Baltimore: Waverly Press.
- Fox, James J. 1977. Notes on the southern voyages and settlements of the Sama-Bajau. *Bijdragen tot de Taal-, Land- en Volkenkunde* 133:459–465.
- Gonda, J. 1973. *Sanskrit in Indonesia*. 2nd ed. New Delhi: The International Academy of Indian Culture.
- Grimes, Barbara F., ed., 2000. *Ethnologue: Languages of the world*. 14th ed. Dallas, Texas: SIL International.
- Hall, Kenneth R. 1985. *Maritime trade and state development in early Southeast Asia*. Honolulu: University of Hawai'i Press.

- Harceland, August. 1859. *Dajacksch-Deutsches Wörterbuch*. Amsterdam: Frederick Muller.
- Himmelmann, Nikolaus P. 2001. *Sourcebook on Tomini-Tolitoli languages*. Pacific Linguistics 511. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- Hogan, David W. 1988. *Urak Lawoi': Basic structures and a dictionary*. Pacific Linguistics C-109. Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.
- Hudson, Alfred B. 1967. *The Barito isolects of Borneo: A classification based on comparative reconstruction and lexicostatistics*. Southeast Asia Program, Department of Asian Studies, Cornell University, Data Paper No. 68. Ithaca: Cornell University.
- Hudson, Alfred B. 1972. *Padju Epat: The Ma'anyan of Indonesian Borneo*. New York: Holt, Rinehart and Winston, Inc.
- Kähler, Hans. 1946-49. Ethnographische und linguistische Studien von den Orang laut auf der Inseln Rangsang an der Ostküste von Sumatra. *Anthropos* 41:1-31; 44:757-785.
- Kähler, Hans. 1960. *Ethnographische und linguistische Studien über die Orang darat, Orang akit, Orang laut und Orang utan im Riau-Archipel und auf den Inseln an der Ostküste von Sumatra*. Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg, Band 2. Berlin: Dietrich Reimer.
- Kern, Hendrik. 1917 [1882]. Over de taal der Philippijnsche Negrito's. Reprinted in *Verspreide Geschriften* 10:289-306. The Hague.
- Larish, Michael David. 1999. *The position of Moken and Moklen within the Austronesian language family*. Ph.D. dissertation, University of Hawai'i.
- McFarland, Curtis D. 1980. *A linguistic atlas of the Philippines*. Study of Languages and Cultures of Asia and Africa Monograph Series, No. 15. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: The University of Chicago Press.
- Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.
- Pallesen, A. Kemp. 1985. *Culture contact and language convergence*. Manila: Linguistic Society of the Philippines.
- Penoyer, F. Douglas. 1986/87. Inati: The hidden Negrito language of Panay, Philippines. *The Philippine Journal of Linguistics* 18/19:1-36.
- Ras, J.J. 1968. *Hikajat Bandjar: A study in Malay historiography*. KITLV, Bibliotheca Indonesica. The Hague: Martinus Nijhoff.
- Ray, Sidney H. 1913. The languages of Borneo. *The Sarawak Museum Journal* 1(4):1-196.
- Reid, Lawrence A., ed. 1971. *Philippine minor languages: Word lists and phonologies*. Oceanic Linguistics, Special Publication, No. 8. Honolulu: University of Hawai'i Press.

- Reid, Lawrence A. 1973. Kankanay and the problem *R and *l reflexes. *Parangal kay Cecilio Lopez: Essays in honor of Cecilio Lopez on his seventy-fifth birthday*, ed. by Andrew Gonzalez, F.S.C., 51–63. Philippine Journal of Linguistics Special Monograph Issue No. 4. Quezon City: Linguistic Society of the Philippines.
- Reid, Lawrence A. 1987. The early switch hypothesis: Linguistic evidence for contact between Negritos and Austronesians. *Man and Culture in Oceania* 3 (Special Issue):41–59.
- Reid, Lawrence A. 1989. Arta, another Philippine Negrito language. *Oceanic Linguistics* 28:47–74.
- Reid, Lawrence A. 1991. The Alta languages of the Philippines. *Papers from the Fifth International Conference on Austronesian Linguistics, VICAL 2, Western Austronesian and contact languages*, ed. by Ray Harlow, 265–297. Auckland: Linguistic Society of New Zealand.
- Reid, Lawrence A. 1994a. Possible Non-Austronesian lexical elements in Philippine Negrito languages. *Oceanic Linguistics* 33:37–72.
- Reid, Lawrence A. 1994b. Terms for rice agriculture and terrace building in some Cordilleran languages of the Philippines. *Austronesian terminologies: Continuity and change*, ed. by Andrew K. Pawley and Malcolm D. Ross, 363–388. Pacific Linguistics C–127. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- Richardson, J. 1885. *A new Malagasy-English dictionary*. Antananarivo: The London Missionary Society.
- Saleeby, Najeeb M. 1963 [1908]. *The history of Sulu*. Manila: Filipiniana Book Guild, Inc.
- Sapir, Edward. 1968 [1916]. Time perspective in aboriginal American culture: A study in method. Reprinted in *Selected writings of Edward Sapir in language, culture, personality*, ed. by David G. Mandelbaum, 389–462. Berkeley and Los Angeles: University of California Press.
- Scheerer, Otto. 1908. *The Batan dialect as a member of the Philippine group of languages*. Bureau of Science, Division of Ethnology Publications, Vol. 5, Part 1. Manila: Bureau of Printing.
- Skeat, Walter William, and Charles Otto Blagden. 1906. 2 vols. *Pagan races of the Malay peninsula*. London: Macmillan and Co.
- Sneddon, James N. 1984. *Proto-Sangiric and the Sangiric languages*. Pacific Linguistics B–91. Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.
- Spriggs, Matthew. 1989. The dating of the island Southeast Asian Neolithic. *Antiquity* 63:587–612.
- Sutlive, Vinson H. 1978. *The Iban of Sarawak*. Arlington Heights, Illinois: AHM Publishing Corporation.
- Thomas, David, and Alan Healey. 1962. Some Philippine language subgroupings: A lexicostatistical study. *Anthropological Linguistics* 4(9):21–33.

- Thurgood, Graham. 1999. *From ancient Cham to modern dialects: Two thousand years of language contact and change*. Oceanic Linguistics Special Publication No. 28. Honolulu: University of Hawai'i Press.
- Voorhoeve, P. 1955. *Critical survey of studies on the languages of Sumatra*. KITLV Bibliographical Series 1. The Hague: Nijhoff.
- Walton, Charles. 1979. A Philippine language tree. *Anthropological Linguistics* 21:70–98.
- Wernstedt, Frederick L., and J.E. Spencer. 1967. *The Philippine island world: A physical, cultural, and regional geography*. Berkeley and Los Angeles: University of California Press.
- Zorc, R. David. 1986. The genetic relationships of Philippine languages. *FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics*, ed. by Paul Geraghty, Lois Carrington, and S.A. Wurm, 147–173. Pacific Linguistics C-94. Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.