The Austronesian Homeland: A Linguistic Perspective

ROBERT BLUST

INTRODUCTION

THE DISCOVERY OF the Indo-European language family has been called one of the great intellectual achievements of the nineteenth century. In important respects it contributed to the Romantic movement in literature, and probably played no small part in the rise of European nationalism. One can readily appreciate the intellectual excitement that must have prevailed among those scholars—primarily in Denmark and Germany—whose researches established the common origin of the Germanic languages with the languages of classical antiquity and with the distant and then only recently-discovered Sanskrit of India. For in founding the discipline of historical linguistics, they opened a new window on an unexpected past.

The reconstruction of Proto-Indo-European led naturally to the question "Where was this prehistoric language, whose modern descendants reach from Iceland to India, spoken?" In attempting to locate the "Urheimat" or homeland of the Indo-European-speaking peoples, linguists came to acknowledge three relevant approaches: (1) linguistic palaeontology (the "Wörter und Sachen technique"), (2) migration theory, and (3) toponymy (the study of place names).

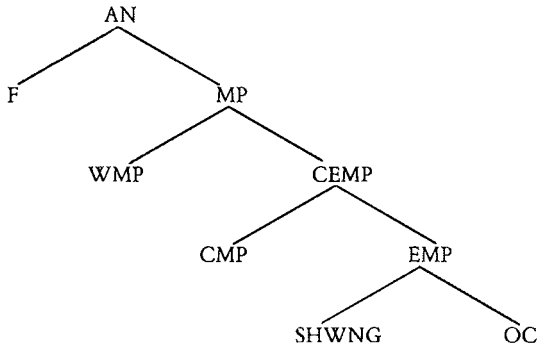
This paper examines the uses of linguistic palaeontology and migration theory in relation to the question of the Austronesian homeland, drawing on cultural vocabulary only where this can provide information about the natural environment.[1] Its conclusions, like those of all homeland theories, depend crucially on the subgrouping of the languages compared. Table 1 presents a subgrouping of the Austronesian (AN) languages which I have proposed and attempted to justify elsewhere.[2]

LINGUISTIC PALAEONTOLOGY

The expression "linguistic palaeontology" is perhaps best known through the celebrated lectures of Saussure (1966), who credits Pictet with creating the approach, if not

Robert Blust is Professor of Linguistics at the University of Hawaii at Manoa, Honolulu, Hawaii.

TABLE 1. HIGHER-LEVEL STRUCTURE OF THE AN LANGUAGE FAMILY (after Dahl 1976; Blust 1977, 1978, 1982, 1983–1984)



F:	Formosan (one or more primary subgroups in Taiwan)
MP:	Malayo-Polynesian (all AN languages outside Taiwan)
WMP:	Western Malayo-Polynesian (the languages of the Philippines and western Indonesia, including Palauan and Chamorro of western Micronesia, the Chamic languages of mainland Southeast Asia, and Malagasy)
CEMP:	Central-Eastern Malayo-Polynesian (all other MP languages)
CMP:	Central Malayo-Polynesian (the AN languages of the Lesser Sundas and of the southern and central Moluccas beginning with Bimanese of eastern Sumbawa and including the Aru Islands in the east and the Sula Archipelago in the northwest)
EMP:	Eastern Malayo-Polynesian (the languages of the SHWNG and OC groups)
SHWNG:	South Halmahera-West New Guinea (the AN languages of Halmahera and of the north coast of New Guinea as far as the Mamberamo River, together with the Raja Ampat Islands)
OC:	Oceanic (the AN languages of Melanesia, Micronesia and Polynesia except as stated elsewhere)
NB:	The languages of the Bomberai Peninsula (Irian) are insufficiently known to permit secure subgrouping, but preliminary data suggest that some of these (including at least Sekar and Onin) are CMP, while others (as Irahutu) may constitute a primary branch of the SHWNG group. The position of Yapese (western Micronesia) within MP is unclear.

the term. The approach to culture history that was pioneered by Pictet is also known by a German name: the "Wörter und Sachen technique." As its name implies, the Wörter und Sachen technique makes use of reconstructed vocabulary (words) as a basis for inferences about the kinds of referents (things) that were known to speakers of a reconstructed language. Reconstructed vocabulary can be used both as a basis for cultural inferences (for example, Benveniste 1973) and as a basis for inferences about the natural environment. Both Thieme (1958) and Friedrich (1970), for example, use tree names (in conjunction with subgrouping) to establish that Proto-Indo-European probably was spoken in extreme eastern Europe. In addition, Thieme argues that an etymology for "salmon" points specifically to a location along the rivers that flow into the North Sea. Similarly, I have noted (Blust 1982) that the distribution of cognate terms for placental mammals in AN languages (in conjunction with subgrouping) cannot easily be reconciled with a hypothesis that the AN homeland was east of the Wallace Line.

The point of these examples is that a well-motivated subgrouping theory together with a collection of cognate sets relating to flora and fauna can be made to yield inferences of a sometimes surprisingly specific character in relation to the homeland question.

MIGRATION THEORY

Linguistic migration theory was first formulated in general terms by Sapir (1968), who noted that the area of highest linguistic diversity (but not necessarily the area with the greatest number of languages) is the most likely center of dispersal of a language family or linguistic subgroup. Dyen (1956) attempted to formalize and to some extent quantify Sapir's views. A parallel principle used to determine the most likely centers of dispersal of cultivated plants was developed by the botanist Vavilov (1926).

All of these approaches are characterized by a belief that areas of greater genetic diversity are more likely to represent older centers of dispersal than areas of lesser genetic diversity. In each case the argument rests on parsimony: a theory that requires fewer and/or shorter movements is preferable to one that requires more and/or longer movements to account for the same distribution. As with linguistic palaeontology, any given application of migration theory to linguistics is thus crucially dependent on the subgrouping of the languages compared.

THE AUSTRALIAN HOMELAND

The discovery of the AN language family or, more accurately, the recognition that certain widely separated AN languages must have a common origin, actually preceded the discovery of the Indo-European language family.[3] For several reasons, however, this realization had little impact on the thinking of European scholars. Even if other things had been equal, one reason would have been sufficient in itself to account for the enormous difference in the European reaction to the two cases: the discovery of Indo-European was more than an intellectual activity—it was a process of linguistic and cultural *self-discovery* (hence its contribution to the Romantic movement). By contrast, the establishment of the AN language family by Europeans was more dispassionate, temporally more gradual, conceptually less intense.

Notwithstanding these important differences the problem of determining the AN homeland is no less challenging or intellectually exciting than that of determining the Indo-European homeland. In each case, unlike the situation in many language families, we are faced with a substantial collection of languages (over 50 IE, more than 800 AN) distributed over a large portion of the earth's surface comprising distinct climatic and biogeographic zones. Historical linguistic theory tells us that these distributions are the product of dispersal over many centuries from a geographically more restricted center of origin. Any well-supported theory concerning the location of such centers will have culture-historical implications of potential interest, not only to the linguist but to the archaeologist and physical anthropologist as well.

A number of views concerning the location of the AN homeland have been advanced since Marsden suggested in 1834 that the inhabitants of the Pacific islands came from "Tartary." Theories proposed up to the mid-1960s are conveniently summarized and critically evaluated by Anceaux (1965).

The resources of linguistic palaeontology were first applied to the problem of the AN homeland by the Dutch linguist Hendrik Kern (1889), whose study has become a classic.[4] After briefly discussing the concept of linguistic palaeontology, Kern noted that the most useful evidence for his purpose would consist of terms for flora and fauna—particularly those with a distinctive geographical distribution.

Kern began his examination of the evidence by pointing out the wide distribution of related words for 'sugarcane' in AN languages, including such forms as Malay *tebu,* Fijian *dovu.* This distribution of cognate forms—together with a geographically similar distribution of cognate terms for coconut, banana, bamboo, rattan, cucumber, pandanus, yam, stinging nettle, and taro—led to his first conclusion: the AN homeland was between the Tropics, or at the very least not far outside them.

A consideration of cognate terms for 'rice'—which even in the 1880s was known to be of Asian origin—led to Kern's second conclusion: the AN homeland was on or near the Asian mainland.[5]

Kern's third conclusion, that the AN homeland was near the sea, was supported by cognate sets that refer to marine fauna and to the art of navigation.

To further narrow the geographical possibilities Kern considered terms that relate to land forms and spatial orientation. He noted that in Malay, spoken in the Malay Peninsula and in eastern Sumatra on both sides of the Straits of Malacca, and in Achinese of northern Sumatra, *selatan* 'south' can be analyzed etymologically as "straits region" *(selat + an),* thus suggesting a point of origin for both languages further to the north. The widespread opposition of directional terms meaning 'toward the sea' and 'toward the interior' in AN languages was further taken as an indication that the homeland was on a large island or mainland. These observations, together with apparent AN loanwords in a number of the non-AN languages of Mainland Southeast Asia, and the presence of AN languages in Viet Nam and Cambodia (the Chamic languages) led to his final conclusion: the AN homeland "was probably located in Champa, Cochin-China, Cambodia and the neighboring regions along the sea."

Only two significant counter-proposals concerning the AN homeland have been advanced since Kern. In the first of these Haudricourt (1954) suggested a homeland in coastal areas of the Asian mainland, but further to the north than Kern had supposed, between the islands of Hainan and Taiwan. His conclusion was based on two considerations: (1) the likely subgrouping of the AN languages themselves (a problem area that was appreciated by Kern, who remarked on it, but which remained *terra incognita* in his day), and (2) a possible external relationship or substratum connection with various languages of southern China, northern Viet Nam and Hainan Island, first noted by Benedict (1942). A radically different view was defended by Dyen (1965), who proposed, on the basis of a lexicostatistical classification of the AN languages, that the primary center of dispersal was in the region of New Guinea and the Bismarck Archipelago.

Next, I will advocate a position very similar to that of Haudricourt, introducing a considerable number of new comparisons bearing on the homeland issue. For expository reasons it will be helpful to use Kern's proposals as a starting point and to react to them in the light of the linguistic evidence now available.

Kern concluded first that the AN homeland was within the tropics, or at the very least not far outside them. This still allows considerable latitude (literally: it is some 47 degrees from the Tropic of Cancer to the Tropic of Capricorn). In the end he opted for a location at about 11–12 degrees N. latitude.

Both the subgrouping of the AN languages and the pattern of lexical reconstructions that follow from it now strongly support Haudricourt in pointing to a homeland that was at best only marginally tropical. As seen in Table 1 at least one, and perhaps as many as three primary AN subgroups (Atayalic, Tsouic, Paiwanic), are located on the island of Taiwan. Individually or collectively these groups are coordinate with the rest of the AN languages as a whole, thus suggesting a primary center of dispersal on or near Taiwan.

	PF	PMP	ENGLISH
a)	CebuS	tebuh	sugarcane
b)	saleŋ	saleŋ	pine
c)	belbel	punti	banana
d)	dakeS	—	camphor laurel
e)	—	kuluR	breadfruit
f)	X	X	*Lactuca indica*

Fig. 1. Distribution types of cognate sets across major AN subgroups. (PF = Proto-Formosan, PMP = Proto-Malayo-Polynesian)

The evidence of linguistic palaeontology, summarized in Tables 2–4 (Appendix) supports this conclusion. Kern assembled a number of etymologies relating to flora and fauna which appeared to indicate a tropical homeland. The distribution of these terms across major AN subgroups, however, suggests that his inference properly applies not to the *Austronesian* homeland (c. 4500 B.C.), but rather to the Malayo-Polynesian homeland (c. 3500 B.C.). A somewhat more detailed examination of this material will be worthwhile, but before proceeding several methodological issues should be addressed. Four issues are of particular importance:

1. ***Independence.*** The linguistic determination of prehistoric centers of population dispersal is often said to depend on *two* factors: subgrouping and lexical reconstruction. There is, however, reason to doubt that these lines of evidence are independent. If, for example, the AN language family were assumed to divide into two primary branches, Oceanic and non-Oceanic, it would follow from lexical reconstruction that Proto-Austronesian speakers were familiar with marsupial mammals and crocodiles, but with no placental mammals apart from the domesticated pig (see Table 3). If, on the other hand, the AN family is assumed to divide into two primary branches, Formosan and non-Formosan, it follows from the distribution of cognate sets that Proto-Austronesian speakers probably were familiar with the dog, monkey, pangolin, wild and domesticated pig, and some kind of buffalo- or antelope-like ruminant, but not with any marsupial mammal, nor with the crocodile. The results of linguistic palaeontology, then, are to a very large extent a consequence of the subgrouping theory adopted. For this reason, it is important that the subgrouping be determined primarily by other kinds of evidence.[6]

2. ***Negative evidence.*** Linguistic palaeontology is dependent on lexical reconstruction, and lexical reconstruction is based on positive evidence. We can therefore be reasonably certain that speakers of a protolanguage were familiar with distinctive features of the natural environment if associated linguistic reconstructions are available. But where reconstructions are not available we cannot be certain that speakers were NOT familiar with a given referent.[7] Figure 1 presents the major possibilities, with actual examples relating to the domain of flora.

In conjunction with Table 2, Figure 1 can be read as follows:

a) Reflexes of PAN *CebuS 'sugarcane' are found in all major AN subgroups (F, WMP, CMP, SHWNG, OC). Even if our subgrouping theory were revised drastically it would be necessary to acknowledge that Proto-Austronesian speakers were familiar with sugarcane. As this comparison vividly illustrates, however, the security of a linguistic reconstruction as measured by its independence from a particular subgrouping theory is bought at the cost of its diagnostic value in relation to the homeland question.

b) Reflexes of PAN *saleŋ 'pine' are widely attested in Formosa and the northern and central Philippines. Two explanations suggest themselves: (1) PAN speakers were familiar

with the pine and lost the associated word when they moved into tropical lowland areas that lacked the tree, (2) *saleŋ was innovated when AN speakers left the tropics and encountered the pine. Advocates of the latter view might attempt to account for the reported distribution through either subgrouping (F + Philippines) or borrowing.

c) Reflexes of *belbel 'banana' are common in Taiwan, and reflexes of *punti are common in the rest of the AN world. Like 'sugarcane', the *referent* is thus found in languages belonging to all major AN subgroups, but unlike 'sugarcane' cognate sets are limited to Formosan or to Malayo-Polynesian languages. Through application of the comparative method we cannot assign a reconstruction for 'banana' to Proto-Austronesian, but neither can we completely rule out the possibility that Proto-Austronesian speakers had a word for 'banana' (see note 7). The major alternatives are: (1) PAN had *belbel, with *punti being a lexical innovation in PMP, (2) PAN had *punti, with *belbel being a lexical innovation in Proto-Formosan, 3) PAN had some other word for 'banana', with *belbel and *punti *both* being lexical innovations, (4) PAN lacked a word for 'banana'.

d) Reflexes of *dakeS 'camphor laurel' are found in a number of Formosan languages, but both the term and the plant appear to be absent outside Taiwan. The explanations proposed in connection with *saleŋ would seem to apply here, except that the innovation (if that is what it is) is confined to Formosan languages.

e) Reflexes of *kuluR 'breadfruit' are widespread outside Taiwan, but neither the word nor the plant have been reported on the latter island. Again, two explanations appear possible: (1) PAN speakers had *kuluR, but the plant and its name were lost in Taiwan, (2) PAN speakers lacked the breadfruit, the plant being acquired only in PMP times.

f) The *Lactuca indica,* a useful plant with a wide geographical distribution, is found both in Taiwan and elsewhere in the AN world. It is possible that PAN speakers had a word for this plant which was replaced in most daughter languages, but to date no positive evidence for reconstructing it has come to light.

Examples a–f illustrate the disconformity between historical inferences that are justified through the comparative method and historical inferences that are logically possible but are not supported by positive evidence. There are at least three reasons why a form that was found in PAN might not be inferable. First, to be assigned to a PAN etymon a cognate set must be distributed over both Formosan and Malayo-Polynesian languages. Only 21 of the approximately 825 AN languages are Formosan, and seven of these are extinct. If by chance a lexical item which was present in PAN failed to survive in any Formosan witness, its PAN status would be irretrievable. Second, only three of the Formosan languages (Atayal, Paiwan, Amis) are represented by dictionaries of any scope; in some cases the evidence for a PAN etymon may exist, but is yet to be made generally available through publication. Third, it is possible that PAN was spoken in Taiwan and that names for the zelkova tree *(Zelkova formosana),* crape myrtle *(Lagerstroemia subcostata),* maple *(Liquidambar formosana),* black alder *(Alnus formosana),* mulberry *(Morus acidosa),* miscanthus, and the like—which are prominent in Formosan languages (cf. Tsuchida 1977) but are less important or even absent in other parts of the AN world—were lost when AN speakers left Taiwan on their epic odyssey through six millennia and over half the tropical surface of the earth.

3. ***Environmental change.*** The environment of a human group can change in either of two ways: by climatic alteration or by migration. The former tends to produce gradual change, the latter abrupt change. Only the latter type of environmental change was con-

sidered above, but given the likelihood that PAN was spoken as early as 4500 B.C. in the neighborhood of Taiwan, there is a possibility that the AN homeland was climatically and biotically somewhat different than the same area today.

Taiwan straddles the tropical and subtropical zones. Its climate is moderated by the warm waters of the Kuroshio (Japan) Current. The summer is long (about 200 days from April–November), with a mean monthly temperature above 68 degrees F (20 degrees C), and a high of 86 degrees F (30 degrees C) from June–September. In the colder months the mean monthly temperature is about 59 degrees F (15 degrees C). Temperatures fall with increasing altitude, and in winter the central mountains (which rise to 13,113 feet, with many peaks above 10,000 feet) are snowcapped.

According to Bellwood (1979:54), during the last glaciation mean temperatures in Southeast Asia may have dropped as much as eight degrees C, even in tropical lowlands. However, the period of the AN expansions seems clearly to have been postglacial. Wang (1984:177) maintains that Early Holocene temperatures in east Asia (8000–6000/5500 B.C.) were 5–6 degrees C cooler than today, and that Middle Holocene temperatures (6000/5500–1000 B.C.) were warmer by 2 to 4 degrees C. A cooling tendency set in during the Late Holocene (the past 3000 years).

It is difficult to assess the possible impact of such climatic variations on floral and faunal assemblages without more exact information. Chung, Huang, and Stamps (1973) suggest on the basis of a pollen analysis that the climate of at least some parts of Taiwan has changed little in the past 4000 years (that is, it has remained subtropical), and during the previous 6000 years it could be characterized as "warm temperate." Based on this observation it would appear that plant assemblages in Taiwan have changed relatively little in the past 6000–7000 years. Merrill (1954:236) points out that the Taiwanese flora "is essentially continental (Asiatic) as compared with the insular (Malaysian) floras of the great islands to the south and southwest." A fairly sharp dividing line separates Taiwan from the northernmost Philippine islands, and even from Botel Tobago Island, some 40 miles distant. To the extent that these differences were already present in the period of the earliest Austronesian migrations they would serve to account for a number of the observations made above (for example, absence of PAN terms for breadfruit, coconut, hibiscus, sago, etc.).

4. ***Borrowing.*** Borrowing as an explanation for lexical distributions that cross major subgroup boundaries was suggested briefly under "negative evidence." Distributions that are due to borrowing have several general characteristics in common. First, they tend to be geographically compact and continuous rather than dispersed and/or discontinuous. Second, they tend to affect certain semantic domains more than others (for example, items of material culture more than body-part terms). Third, irregularities in the phonological correspondences are likely (though not inevitable), especially in more recent loans.

The Formosan languages are geographically concentrated, in marked contrast to the Malayo-Polynesian languages. If more than one primary AN subgroup is represented in Taiwan it should be possible to base some PAN reconstructions on Formosan cognates only. However, given their geographical contiguity it is likely even in the absence of positive evidence that the Formosan languages have borrowed extensively from one another over the millennia of their coexistence on Taiwan. To avoid a possible confusion of directly inherited cognates with widely diffused innovations, no PAN reconstruction is based on Formosan witnesses alone. As a consequence of this procedure all PAN reconstructions in this paper imply a PMP continuation.

We are now ready to look at the evidence of linguistic palaeontology. Kern's first and principal argument for an AN homeland within the tropics was based on widespread cognate sets for flora. Although his use of comparative linguistic data was advanced for his day, considerably more and better descriptive material is now available.[8] Table 2 schematizes the distribution of known cognate sets which refer to plants in terms of reconstructions that have been proposed on the PAN, PMP, PCEMP, PEMP, and POC levels.

The subset of lexical reconstructions in Table 2 that can be attributed to PAN include three types of bamboo, the areca palm and nut (inferred solely from PAN *apuR 'betel chew'), two types of fern, the sword grass *Imperata cylindrica* (often associated with slash-and-burn agriculture), three words for millet, the stinging nettle, pandanus, derris root (used to stupefy fish), rattan, five words connected with rice, sugarcane, the "elephant ear" taro (*Alocasia* spp.), and several plants *(Solanum nigrum, Smilax, Urena lobata)* and trees (*Cordia* spp., *Pinus* spp.). Not attributed to PAN are the banana, breadfruit, coconut, cucumber, ginger, hibiscus, indigo, jackfruit, kapok, mango, mangrove, cordyline, ramie, sago, *Colocasia* taro, many tree names, turmeric and yam. In effect, then, Table 2 acts as a filter to distinguish that part of the widely distributed AN vocabulary which can securely be attributed to PAN from that part which *may* be due to later innovation.

Much the same can be said of Table 3, where the distribution of cognate sets suggests the presence of deer, the dog, some type of dove, freshwater eels, at least one type of monkey, the pangolin, wild and domesticated pigs, and some kind of large ruminant which may have been distinct from a deer. Not attributable to PAN are words for civet cat, chicken, crocodile, flying fox, and python, for which widespread cognate sets are found outside Taiwan, as well as names of other animals for which widespread cognate sets are not available outside Taiwan (e.g., bear and leopard, found both in Taiwan and in Borneo).

Table 4 indicates a tectonically unstable area (*linuR) with hills or mountains, lakes, possibly some influence from the monsoons (though this appears marginal), a distinct cold season and periodic typhoons. Not reconstructible are words for 'cave' and 'river', though the latter may have been identical to 'water' or 'flowing water'.

Several items in Table 4 merit further comment. In many WMP languages a reflex of *bukid means 'hill, mountain' (e.g., Bikol *búkid* 'mountain, hill', Javanese *wukir* 'mountain'). In other languages the cognate term refers distinctively to hills [e.g., Malay *bukit* 'hill, that is, an elevation from about 100 to about 1500 feet, higher elevations being "mountains" *(gunong)*', Palauan *bukl* 'hill; mound of earth' (compare, *róis* 'mountain')]. It appears likely, then, that *bukid meant 'hill' in PWMP. A PAN reconstruction at the present time depends entirely on the inclusion of Paiwan *vukid* 'continuous forest', *vuki-vukid* 'woods on foothills (said to be inhabited by ancestors and evil spirits)'. It is noteworthy that a reference to vegetation also occurs in Itbayaten *vuchid* 'cogon grass: *Imperata cylindrica*' and perhaps Tagalog *búkid* 'country outside of towns and cities; farm, field under cultivation'. To a people whose preferred habitat was the relatively open, unforested regions along the sea, the notions 'forest' and 'hill country' would have been to a large extent inseparable.

The terms *SabaRat and *timuR clearly referred to the monsoons in PMP, though the isolated Formosan reflexes (Amis *savalat* 'west wild', *ka-timul* 'south') make it difficult to associate the PAN forms with these meanings.[9]

Of particular interest is PAN *qamiS(-an) 'north wind, cold weather'. Reflexes of this form appear in all major Formosan subgroups, and in a number of Philippine languages

(cf. Tsuchida 1976:160ff for a sample distribution). In the more northerly languages or those spoken in mountainous areas of Taiwan the word often means 'winter' (Atayal *qemis-an,* Saisiyat *amis-an,* Kanakanabu *ʔamisan,* Saaroa *ʔamisan-a* 'winter'). In areas where seasonal variations of temperature are less extreme it generally means 'north' or 'north wind' (Paiwan *qamis,* Amis *ka-qamis,* Ilokano *amian* 'north wind', Ilokano *amian-an* 'north', Cebuano *amihan* 'wind from the north', *amihan-an* 'north, northern', Maranao *amian* 'north wind', *amian-an* 'north', Proto-Minahasan (Sneddon 1978) **amian* 'north, north wind'. The southernmost reflexes of *qamiS(-an) found to date are from the Minahasa area of northern Sulawesi; together these cognate forms clearly support a PAN reconstruction which referred to a north wind that brought colder weather. They are thus incompatible with the hypothesis of an equatorial or near-equatorial homeland.

Reflexes of *baRiuS appear in southern Taiwan, the Philippines, the Marianas, and parts of Borneo in the meaning 'typhoon'. Together these indicate a PAN homeland within the typhoon belt of the northern hemisphere.

The material of Tables 2–4, then, supports Kern's second conclusion (that the AN homeland was near the Asian mainland), but it leads us to question his first conclusion (that the AN homeland was within the tropics). Table 5 (Appendix) relates to marine life or to a marine environment, and so supports Kern's third major inference: that the AN homeland was near the sea. Although only a few of the reconstructions in Table 5 can be assigned to PAN, it appears likely that PAN speakers, like PMP speakers, lived near the sea. This inference is supported by an Amis reflex of *kaNasay 'mullet', by Paiwan reflexes of *kagaŋ/kaRaŋ 'land crab', *qiSu 'shark', and *Nabek 'surf, breakers', by Siraya *pagig* 'stingray' and, less convincingly, by reflexes of *qenay 'sand' in various Rukai dialects.

Beginning in the seventeenth century an accelerating influx of Chinese immigrants into the western and northern plains led to the sinicization of several Taiwan aboriginal groups; others retreated into less accessible mountain areas. As a result much of the vocabulary that had once related to the sea would have been lost or would have undergone radical semantic changes in the languages formerly spoken in what are now regions of heavy Chinese settlement. Only four of the 14 surviving ethnolinguistic groups native to Taiwan are in contact with the sea, and of these one group (Kuvalan) is heavily sinicized. To a very great extent, then, the evidence that PAN speakers were in contact with the sea depends crucially on a handful of reflexes in three surviving languages (Paiwan, Puyuma, Amis) and in two others (Favorlang, Siraya) that are known only from seventeenth century Dutch records.

One final group of observations in connection with the AN homeland is noteworthy. Blust (1976) refers to widespread evidence for the outrigger canoe complex in the AN world. However, apart from PAN *qabaŋ 'canoe, boat', a word which may have referred to dugout craft for use in coastal waters, all of these terms are PMP. Ferrell (1969:52ff) notes possible evidence of the former use of outrigger canoes among the Kuvalan, but all modern Formosan aboriginal groups use large bamboo sailing rafts for fishing at sea. It is therefore unclear whether the outrigger canoe complex (which would imply a marine environment) was already present in PAN times or was a PMP innovation.

In conclusion, the subgrouping of the AN languages and the use of lexical reconstructions that follow from it support Kern's claims that the AN homeland was near the Asian mainland and the sea, but fail to support his claim that this area was in the region of mod-

ern Viet Nam and Kampuchea, or even that it was tropical. Rather, the AN migrations appear to have begun in a tectonically unstable region with distinct seasonal temperature variations within the Pacific typhoon belt which stretches from the eastern Carolines westward and northward through the Philippines, Taiwan, the Ryukyus, and southern Japan. The flora and fauna of the area included various types of bamboos, ferns, nettle, pandanus, *Cordia, Derris, Alocasia* taro, rattan, and such trees as the pine (in mountain areas?), and areca palm (in lowland areas?), as well as deer, freshwater eels, wild pigs, monkeys, scaly anteaters, and some kind of large ruminant which may have been distinct from a deer. A term for *Imperata cylindrica* suggests deforestation due to swidden agriculture. Cultivated plants and domesticated animals included several types of millet, rice, sugarcane, the dog, and the pig. The overall picture that emerges is thus consistent with a homeland in Taiwan or the adjacent mainland of China, but is not consistent with such areas as Viet Nam–Kampuchea, New Guinea, the Bismarck Archipelago, or Indonesia (which lies outside the typhoon belt).

SECONDARY CENTERS OF DISPERSAL

We might end our discussion here, but the overall subgrouping of the AN languages permits several statements, however tentative, about the postdispersal history of the AN speaking peoples. The mode of presentation that I will adopt might best be characterized, like that of Heine-Geldern (1932) or Solheim (1975), as one of "informed speculation." In the interest of historical continuity this outline will follow the order of splits implied by Table 1, with a rough indication of the likely chronology, beginning with the AN homeland itself.

1. c. 4500 B.C.
   AN → F + MP

If this view of the first split within AN is correct, it implies that the AN homeland was near Taiwan. The center of primary dispersal need not have been Taiwan itself, though this possibility obviously becomes stronger if the Formosan languages are considered to comprise more than one primary subgroup of the AN family.

No AN languages are presently found in China, except the late Chamic intrusion into Hainan Island (Benedict 1941).[10] But if Taiwan was not the AN homeland it must have been settled from the adjacent coast of China during the initial dispersal of AN speakers. As already seen, there are several indications that the AN homeland was in a subtropical or warm temperate region, straddling or perhaps slightly north of the Tropic of Cancer. Acceptance of the various, seemingly conflicting views that have been proposed regarding the external relationships of the AN languages (Austric, Austro-Thai, Austro-Japanese, etc.) will have consequences for the likely center of dispersal of pre-AN speakers, and this in turn inevitably will influence our ideas about the AN homeland. Only with the Austro-Thai hypothesis would the pre-AN and AN homelands remain approximately the same (coastal area of southern China from perhaps Swatow to Foochow). Until better evidence of genetic relationship is available I remain agnostic concerning *all* claims of broader genetic groupings that include Austronesian.

In short, comparative linguistic evidence suggests that the first major geographical region to be settled by AN speakers was Taiwan, either as the AN homeland, or as the

earliest colony after the initial AN dispersal. If the so-called "Changpinian industry" of eastern Taiwan goes back to the late Pleistocene (Chang 1969), Taiwan must have been populated for several millennia before PAN came into being. It is unclear whether these "preceramic" (aceramic?) material data indicate a distinct population (Negrito?), or an ancestor of PAN, though the latter interpretation appears unlikely given the suggested *terminus ante quem* of 3000 B.C. In any case no trace of an earlier language has survived. Similarly, it is possible that some or all of the Ryukyu Islands were settled by AN speakers from Taiwan, but if so the subsequent settlement of this area by Japanese speakers left little evidence of linguistic substratum.

2. c. 3500 B.C.
MP → WMP + CEMP

Given the geographical distribution of MP languages the historical implication of split (2) would appear to be that PMP was spoken somewhere in central Indonesia—perhaps in Sulawesi. But it is unlikely that the Philippines would have been missed in a move from either Taiwan or mainland China to central Indonesia. Given the early settlement of Taiwan by AN speakers the second major geographical region to be settled almost certainly was the Philippines. If the AN homeland was on Taiwan the settlement of the Philippines presumably was effected by direct voyaging across the Bashi Channel and Luzon Strait, probably via Botel Tobago and the Batanes and Babuyan Islands. If the AN homeland was on the Swatow-Amoy-Foochow coast the Philippines could have been settled by direct voyaging from South China. However, I consider this less likely than a move from Taiwan to Luzon. As seen above, terms relating to the canoe complex are reconstructible for PMP, but not for PAN. This may be an accidental by-product of the recent historical condition of most Formosan aboriginal groups, which have either been driven from the fertile coastal plain into the mountain fastnesses or culturally and linguistically absorbed by the Chinese-speaking majority. But it is equally compatible with the view that the outrigger canoe complex developed *after* the breakup of PAN into Formosan and MP descendants. Ferrell (1969) suggests that some Paiwanic groups might have had the outrigger canoe until early historical times, and that Kuvalan of the northeastern coast may have had a cognate of Proto-Oceanic *waŋkaq 'canoe', but the evidence is shaky. The Formosa Strait could easily have been crossed on rafts of the type currently used, and according to Ferrell (1969:53), "These rafts are apparently of quite ancient date in Taiwan."

If the Philippines was the second major geographical area to have been settled by AN speakers it is remarkable how shallow the separation times between Philippine languages generally appear to be. It is true that considerable differences are found between Yami-Itbayaten-Ivatan in the extreme north and Tagabili, Bilaan, or Samal in the extreme south of the Philippine archipelago, but as Zorc (1986) has argued, all of the languages of the Philippines except the Samalan group appear to belong (together with Sangiric, Minahasan, and Gorontalic of northern Sulawesi) in a single subgroup of WMP. This argument suggests (1) exceptional conservativeness in Philippine languages, (2) language levelling—that is, the spread of one prehistoric language at the expense of many or even all others in the area—after the dispersal of PMP, or (3) both. Linguistic levelling could, of course, have been a recurrent phenomenon. In fact, it appears that a major levelling occurred in the central Philippines within the past 1500 years, as the implied time-depths of language splits from the Tagalog-speaking area of Manila Bay through southern Luzon

and the Bisayas are far less than we would expect for an area that almost certainly has been settled by AN speakers for upwards of 5000 years. Similar levellings are known from Borneo, the most recent being the Iban expansion into the Second Division of Sarawak during the past century and a half, causing the extinction of some indigenous groups (Seru) and the decimation of others (Ukit, Bekatan).

When AN speakers entered the Philippines in the fourth millennium B.C. they encountered an autochthonous population that consisted at least in part of hunting and gathering Negrito bands. Fox (1953) has raised the possibility that a second type of pre-AN population represented by the modern Dumagats may once have been more important in the Philippines. PMP speakers evidently designated all such groups as *qaRta 'outsiders, alien people' (Blust 1972).

The subsequent history of Formosan speakers is less dramatic than that of MP speakers, a history of the "stay-at-homes" in contrast to the wanderers. During the 6000 or more years that AN speakers have been on Taiwan their preferred habitat probably was coastal until the major Chinese invasions of the second half of the seventeenth century. Contact between adjacent groups undoubtedly has been considerable, and has tended to level the differences that developed from millennia-long differentiation in situ, thus making it more difficult to determine whether the Formosan languages constitute a single subgroup of AN or several primary branches.

3. c. 3000 B.C.
   WMP →?
   CEMP → CMP + EMP

On leaving the southern Philippines the flow of AN speakers apparently split, one branch which gave rise to the WMP languages entering Borneo, the other which gave rise to CEMP languages entering the northern Moluccas, with perhaps some settlement in Sulawesi. PWMP presumably was spoken in the northern part of the present WMP territory—possibly in the southern Philippines. PCEMP apparently was spoken in the northern Moluccas, near the boundary between CMP and EMP languages.

The internal subgrouping of the WMP languages is unclear. Just as the Formosan languages may constitute more than one primary AN subgroup, the WMP languages may constitute more than one primary MP subgroup. Chamorro and Palauan appear to continue early offshoots of PWMP, although their distinctiveness is possibly due in part to the high degree of isolation they must have experienced after settling the Marianas and Palau, in contrast to other WMP languages, which continued to influence one another long after their separation from a common ancestor. If Spoehr's (1957) radiocarbon date of 1527 B.C. from Saipan is accurate the ancestors of the Chamorros probably had reached the Marianas by 2000 B.C. Judging impressionistically from the degree of linguistic differentiation, Palauan probably has been separated from other AN languages (including Chamorro) for 4000 years, and it too may be the result of a settlement from the southern Philippines or northern Sulawesi by the beginning of the second millennium B.C.

By perhaps 2000 B.C. a language that I have elsewhere (Blust 1974) called "Proto-Northwest Borneo" apparently was spoken in coastal areas of western Sabah. Several centuries later this language split into two groups. One remained in northern Borneo, giving rise to the modern indigenous languages of Sabah, while speakers of the other migrated to the area about the mouth of the Baram River, eventually giving rise to the modern Bintulu, Lower Baram, Kenyah, and Kelabit languages, perhaps among others.

At a considerably later date—probably in the third or fourth century B.C.—an extensive population movement from Southwest Borneo led to the settlement of eastern Sumatra, the Malay Peninsula, and large parts of the coast of Mainland Southeast Asia, perhaps as far north as the Gulf of Tonkin. This dialect complex eventually differentiated into a northern group, which gave rise to the modern Chamic languages and Achinese, and a southern group, which gave rise to the "Malayic Complex" (Malay, in all its dialect forms, Minangkabau-Kerinci, Iban, and other "Malayic Dayak" languages of Borneo, Madurese, Sundanese, and perhaps Lampung). The exact relationship of this group of languages, spoken by what in the earlier literature were sometimes called "Deutero Malays," to the non-Malay languages of Sumatra (Gayo, Simalur, Sichule, the Batak group, Rejang, Mentawai, Enggano), spoken by so-called Proto Malays, is unclear—though several of the latter appear to form a subgroup (Nothofer 1986). Moken, of the Mergui Archipelago, may belong to a larger Sumatran group; its antiquity in the area is unknown, though it—like the other non-Malayic languages of Sumatra—probably antedates the arrival of the Malays in the Sumatran region.

The closest external ties of the Malayo-Chamic languages appear to be with the Java-Bali-Sasak group, and then perhaps with the Barito languages of Southeast Kalimantan (including Malagasy). This tie suggests an earlier protolanguage spoken in Southeast Kalimantan perhaps in the period 1000–1500 B.C., which first split into the ancestor of the Barito group and a language ancestral to Malayo-Chamic and Java-Bali-Sasak. The latter in turn split up by perhaps 800–1000 B.C.

Still later, by the fourth or fifth century A.D., Madagascar almost certainly was settled from Southeast Borneo, as Dahl (1951) has shown. Java, Bali, Lombok, and western Sumbawa either were settled relatively late (within the past 2500 years) from Borneo or Sumatra, or there has been extensive levelling of earlier linguistic differences in these areas (a distinct possibility, given the known political history of Java).

PCEMP, which probably was spoken in the northern Moluccas, apparently split into two daughter languages, one of which (PCMP) moved south to the neighborhood of Seram. Speakers of the other daughter language (PEMP) evidently moved eastward to the area of Cenderawasih Bay in Irian.

4. c. 2000 B.C.
   CMP →
   c. 2500 B.C.
   EMP → SHWNG + OC

The subgrouping of CMP languages as a whole remains unclear, although Collins (1983) has made important clarifications in the area of the central Moluccas. There are some indications that speakers of PCMP moved rapidly through the southern Moluccas and Lesser Sundas, eventually meeting up on Sumbawa with descendants of PWMP that were working their way east through the Lesser Sundas. Sumbawanese, spoken in the western half of Sumbawa, is the easternmost extension of WMP in the Lesser Sundas and Bimanese, spoken in the eastern half of Sumbawa, is the westernmost extension of CMP. After earlier splitting at some point considerably further to the north, two primary branches of MP thus appear to have circled back to meet at a much later date (500 B.C.?) on the island of Sumbawa.

In most parts of the Moluccas and Lesser Sundas a pre-AN population was either lacking or was numerically too insignificant to have a noticeable effect on physical type or

language. This was, however, not uniformly the case. The Aru Islands, in particular, appear to have had a substantial pre-AN population that strongly affected the physical type of the incoming AN speakers and at the same time accelerated the rate of replacement of basic vocabulary. Some CMP speakers evidently also moved into the Bomberai Peninsula of Irian, where they encountered speakers of SHWNG languages.

PEMP, which probably was spoken in the area of Cenderawasih Bay, Irian, split into two daughters, Proto-South Halmahera–West New Guinea and Proto-Oceanic. Speakers of PSHWNG remained in the same general area, gradually spreading westward into southern Halmahera. This migration appears to have taken place within the past 1000–1200 years, as linguistic differences in South Halmahera and the Raja Ampat Islands are on the whole quite minor. Since AN speakers would almost certainly have reached Halmahera before moving into the Pacific, it is likely that the present linguistic situation in South Halmahera and adjacent areas is a product of extensive language levelling, perhaps following a major depopulation. Another depopulation appears to have occurred on the island of Obi in the north-central Moluccas, the vacuum now being filled by immigrants from Sulawesi and other parts of eastern Indonesia. On the whole the SHWNG languages show an accelerated rate of basic vocabulary replacement, possibly due to fairly strong and steady contact influence from speakers of Papuan languages.

Proto-Oceanic probably was spoken in the Bismarck Archipelago, or on the north coast of New Guinea facing the Bismarck Sea. It appears likely that the Admiralty Islands were among the earliest areas settled after the breakup of the POC speech community. Proto-Oceanic retained a large proportion of the basic vocabulary inherited from PMP (perhaps 70%), and its speakers evidently intermarried little with Papuans. After the dispersal of POC there probably was a rapid expansion eastward into island Melanesia, as suggested by Pawley (1981:280ff). By 2000 B.C. a single dialect chain may have extended from the north coast of New Guinea to the central Solomons, and perhaps beyond. A few chance voyages may have led to the settlement of New Caledonia and some neighboring areas by this early date.

In western and perhaps parts of central Melanesia AN speakers evidently were in contact with a previously established population. During the early period relations probably were hostile; the evidence of physical type in Micronesia and Polynesia at least suggests that intermarriage between AN and non-AN speakers was not common until after the migrations to previously uninhabited areas in the eastern Pacific. Following these migrations intermarriage with and linguistic absorption of non-AN speakers gradually altered the physical type of AN speakers in most of Melanesia. A parallel process later occurred with many of the speakers of Polynesian Outlier languages in Melanesia.

By perhaps 1800 B.C. AN speakers from the Southeast Solomons entered Vanuatu; by perhaps 1200 B.C. other AN speakers from the Southeast Solomons settled Kiribati (the Gilberts). From Kiribati the rest of Micronesia, apart from Palau, Yap, and the Marianas, was settled by a gradual south-to-north and east-to-west movement into the Marshalls, Kusaie, Ponape, and the Carolines (probably in that order), as far as the island of Mapia, some 200 miles north of the Vogelkop Peninsula of Irian. More recently Numfor speakers from the area of Cenderawasih Bay have settled Mapia. After earlier splitting in the area of Cenderawasih Bay, then, two primary branches of EMP appear to have met again largely as a result of the recurving migratory route of Micronesian speakers.

The linguistic history of Yap is in many respects obscure. The one fact that seems reasonably clear is that Yapese has borrowed liberally from both Palauan and Woleaian.

Palauan loans appear on the whole to be older, and include a number of basic vocabulary items. Nonborrowed vocabulary is often unfamiliar, and it is difficult to determine whether Yapese is an Oceanic language. A few scattered indications suggest a distant subgrouping connection with the (OC) languages of the Admiralty Islands.

Around 1500 B.C. Fiji appears to have been settled from central Vanuatu. If Fiji was not settled more than once by closely related AN speakers of rather different phenotype, the founding population must have included two distinct and non-intermarrying physical types. From eastern Fiji Tonga was settled around 1200 B.C., and Samoa was perhaps settled from Tonga around 1000 B.C. A period of separation, linguistic innovation, and divergence followed, with later migrations from Samoa to various parts of Melanesia and Micronesia, and to eastern Polynesia during the centuries immediately before and after the beginning of the Christian era.

The AN migrations in the eastern Pacific came to an end only because there were no more suitable islands to settle. East of Hawaii the Pacific stretches unbroken for some 2500 miles before the nearest landfall. Some 300 miles east of Easter Island lies the tiny island of Sala-y-Gomez; the islet of San Felix and the small Juan Fernandez Islands lie another 1500 miles to the east. It is possible that these islands were visited by Polynesian speakers, though no permanent settlements were established. Polynesian speakers may also have reached the coast of South America; it has long been noted that the precontact presence of the sweet potato in Polynesia suggests such a scenario.

Other details of Polynesian culture history have been discussed extensively in the published literature, and need not concern us further here. The one essential conclusion that the linguistic evidence forces upon us again and again is that the dramatic migrations of the Polynesians were but the endpoint of a millennia-long journey that began far to the west and north. The peoples of eastern Polynesia (including the Maori) are the ultimate wanderers, just as the Atayal and similar peoples of Taiwan are the ultimate stay-at-homes, never having ventured beyond the ancestral hearth fires into the tropical world beyond.

ACKNOWLEDGMENT

After the manuscript of this paper had been accepted for publication I received a number of valuable comments on it from U. Sirk of the Institute of Oriental Studies, Moscow. Under the circumstances only a few of these suggestions could be incorporated in the final version.

NOTES

1. Because it would require an independent study of considerable scope, a toponymic approach to the homeland question is not included. For a broad outline of Austronesian culture history based on the study of cultural vocabulary see Blust (1976).
2. Modifications of this schema have been proposed by Harvey (1982) and by Reid (1982). For reasons that I cannot enter into here I find these proposals (for all their merits) unconvincing.
3. The relationship of Malagasy to Malay was recognized as early as 1603 by de Houtman (the captain of the first Dutch ship to sail to Indonesia via the Cape of Good Hope and Madagascar), and the connection of Malay and various languages of the Pacific was noted by Reland in 1708, well before the famous address by Sir William Jones initiated a research program for Indo-European linguistics in 1786.
4. Although I have drawn mainly from the original, I have also found the English translation of McFarland and Tsuchida (1976) useful, and in some cases (as in the statement of Kern's final conclusion) have adapted their wording to my purposes.
5. Kern actually says "in *eastern Asia*" (italicized in the original). As he makes clear in subsequent remarks, however, the area that he intended would today be described as "Southeast Asia."

6. Under certain exceptional conditions linguistic palaeontology can provide a *secondary* means for the determination of subgroups (cf. Blust 1982, where the distribution of cognate terms for marsupial mammals provides evidence for Central-Eastern Malayo-Polynesian).
7. Pawley and Green (1971, postulate 9) appear to advocate the use of negative evidence for drawing conclusions about the probable absence of a referent, though they qualify their position in various ways. While I acknowledge that all scientific inferences are probabilistic, whether based on positive or on negative evidence, the latter are far more difficult to control and should be regarded as carrying an inherently greater potential for error.
8. Kern cited, for example, 17 widely distributed cognate sets for flora and seven (actually five) relating to the sea, where we cite 91 and 54 respectively.
9. All geographical and meteorological sources that I have been able to consult give the predominant wind directions of the Southeast Asian monsoons as NE and SW. The reflexes in a number of widely separated AN languages, however, suggest NW and SE: Tagalog *habágat* 'west or SW wind; monsoon', Cebuano *habágat* 'strong wind that hits Cebu from the southwest, common from June to September', *tímug* 'wind that hits Cebu from the east', Palauan *ngebard* 'west, west wind', *dímes* 'south, south wind', Malay *barat* 'west; etymologically associated with *ribut*, that is, strong westerly winds', *timur* 'east', Yamdena *barat* 'west, westerly, west monsoon', *timur* 'east, easterly, east monsoon; the Tanimbarese from the east coast of Yamdena', Buli *pāt* 'west, west wind', Numfor *barek* 'west', *wam-barek* 'west wind, west monsoon', Motu *lahara* 'NW wind and season'.
10. Solheim (1975:153) maintains that "The earliest historically known people in eastern and coastal South China were Austronesian-speaking peoples," a statement which would imply to most readers that some kind of linguistic documentation of the ancient "Yüeh" peoples of south China is available. As authority for this statement he cites the third, revised edition of Wolfram Eberhard's *A history of China*. However, in both the English translation and the original German edition (1948:19, 22) Eberhard simply speculates that "the Yao culture, an early Austronesian culture" mixed with early Thai culture, giving rise to the Yüeh culture of the Chinese records. This Yüeh culture, termed "a second Proto-Austronesian culture" *(eine zweite proto-austronesische Kultur)* spread over most of Indonesia, and is identified with Heine-Geldern's "quadrangular axe culture." No justification of any kind is given for the linguistic classification of the Yao and Yüeh peoples.

APPENDIX

The following (doubtless incomplete) tables give the distribution of lexical reconstructions which refer to diagnostic features of the natural environment in PAN, PMP, PCEMP, PEMP, and POC. Space limitations prevent the inclusion of supporting evidence, but much of this can be found in Dempwolff (1938) and Blust (1970, 1973, 1980, and 1983–1984). In all tables 1 = PAN, 2 = PMP, 3 = PCEMP, 4 = PEMP and 5 = POC; X = the listed etymon (*qauR, *betuŋ, and so forth) can be attributed to the protolanguage in question; (X) = attribution possible, but with probable semantic difference; X? = attribution possible, but with phonological problems; LI = lexical innovation (an etymon is known, but it is not the one listed); X/Y = X and Y reconstructible as "disjuncts" (cf. Blust 1980); X,Y = X and Y reconstructible as doublets. A number of the reconstructions for flora owe a great deal to the important work of J. A. J. Verheijen (1967, 1982, and personal communication).

TABLE 2. LEXICAL RECONSTRUCTIONS FOR FLORA (INCLUDING DOMESTICATES)

NO.	GLOSS	1	2	3	4	5	FORM
	bamboos (1–8):						
1.	*Bambusa* spp.	X	X	X	X	X	qauR
2.	*Dendrocalamus* spp.		X	X			betuŋ
3.	*Dendrocalamus* spp. (?)		X	X	X	X	bituŋ/pituŋ
4.	*Bambusa* spp.	X	X				buluq
5.	*Bambusa spinosa*	X	X				kawayan
6.	*Bambusa vulgaris (?)*		X	X	X	X	periŋ
7.	*Schizostachyum* spp. (?)		X	X			tamiaŋ
8.	unidentified spp.		X	X			teriŋ
9.	banana		X	X	X	X	punti
10.	betel chew	X	X			LI	apuR
11.	betel nut		X	X	X	X	buaq
12.	breadfruit		X	X	X	X	kuluR
13.	breadfruit sp.		X	X			teRep
14.	coconut (palm/fruit)		X	X	X	X	niuR
15.	coconut husk		X	X	X	X	bunut
16.	cucumber, melon		X	X	X	X	timun
17.	*Asplenium nidus* (?)	X	X	(X)			lukuC
18.	*Lygodium circinnatum*		X	X			ni(n)tuq
19.	*Athyrium esculentum*	(X)	X	X			pah(e)ku
20.	tree fern, *Cyathea* spp.		X	X			puni
21.	ginger		X	X	X	X	laqia
22.	grass, *Imperata cylindrica*	X	X				Riaq
23.	hibiscus		X	X	X	X	baRu
24.	indigo		X	X			taRum
25.	jackfruit		X	X			naŋka
26.	kapok, *Ceiba pentandra*		X	X			kabu
27.	mango, *Mangifera indica*		X	X			paSuq
28.	mango sp.		X	X	X	X	wai
29.	mangrove		X	X	X	X	teŋeR
30.	millet sp.	X	X				baCaj
31.	millet, foxtail millet (?)	X	X	X			beteŋ
32.	millet sp.	X	X				zawa
33.	nettle, *Laportea* spp.	X	X	X	X	X	(za)lateŋ
34.	pandanus	X	X	X	X	X	panDan, paŋeDan
35.	*Solanum nigrum*	X	X	X			ameCi
36.	*Smilax* spp.	X	X				baNaR, banaw
37.	*Leea* spp.		X	X			mali
38.	*Urena lobata*	X	X	X			puluC
39.	*Cordyline, Dracaena* spp.		X	X	X	X	siRi
40.	*Flagellaria indica*		X	X			SuaR
41.	*Millingtonia hortensis*		X	X			taŋga
42.	*Derris* spp.	X	X	X	X	X	tuba
43.	ramie, *Boehmeria nivea*		X	X			rami
44.	rattan (generic?)	X	X	X	X		quay
45.	rattan sp.		X	X			naŋa
46.	rice husk		X	X			qapa
47.	rice husk	X	X				qeCa
48.	rice (husked)	X	X	(X)			beRas
49.	riceplant, rice in the field	X	X	X	X		pajay
50.	rice (cooked)	X	X				Semay
51.	rice stubble	X	X				ZaRami

NO.	GLOSS	1	2	3	4	5	FORM
52.	sago (palm)		X	X	X	X	Rambia, Rumbia
53.	sago branches		X	X			kumbal
54.	sugarcane	X	X	X	X	X	CebuS
55.	taro, *Alocasia* spp.	X	X	X	X	X	biRaq
56.	taro, *Colocasia* spp.		X	X	X	X	tales
57.	palm, *Caryota* spp.	(X)	X				qanibuŋ
58.	*Grewia* spp.		X	X			qanilaw
59.	*Pisonia umbellifera*		X	X	(X)	(X)	qanuliŋ
60.	*Cordia* spp.	X	X				qaNuNaŋ
61.	*Ficus* spp.		X	X			qaRa?
62.	*Casuarina equisetifolia*		X	X	X	X	aRuSu
63.	*Nauclea orientalis*		X	X			baŋkal
64.	*Pterospermum diversifolium*		X	X			bayuR
65.	*Calophyllum* spp.		X	X	X	X	bitaquR
66.	*Barringtonia* spp.		X	X	X	X	butun
67.	*Dracontomelum* spp.	(X)	X	X			daqu
68.	*Heritiera littoralis*		X	X			dunuŋ
69.	*Erythrina indica*		X	X	X	X	DapDap
70.	*Erythrina indica*		X	X			DeDap
71.	*Intsia bijuga*		X	X	X	X	qipil
72.	*Murraya* spp.		X	X			kamuniŋ
73.	*Cananga odorata*		X	X			kanaŋa
74.	*Cordia* spp.		X	X	X	X	kanawa
75.	*Antiaris toxicaria* (?)		X	X	X	X	laji
76.	*Citrus* spp.		X	X			muntay
77.	*Pterocarpus indica*		X	X	X	X	nara/naRa
78.	*Ficus benjamina*		X	X	X	X	nunuk
79.	*Palaquium* spp.		X	X	X	X	ñatuq
80.	*Morinda citrifolia*		X	X	X	X	ñeñu (?)
81.	*Vitex pubescens*		X	X	X?		pa(m)pa
82.	*Pinus* spp.	X	X				saleŋ
83.	*Caesalpinia sappan*		X	X			sepaŋ
84.	*Gnetum gnemon*		X	X			suka
85.	*Terminalia catappa*		X	X	X	X	talisay
86.	*Melochia umbellata*		X	X	X	X	tenu
87.	*Intsia bijuga* (?)		X	X	X	X	teRas
88.	*Dolichandrone spathacea*		X	X	(X)	(X)	tui?
89.	turmeric, *Curcuma* spp.		X	X			kunij
90.	to winnow	X	X	X	X		taSep
91.	yam		X	X	X	X	qubi

TABLE 3. LEXICAL RECONSTRUCTIONS THAT REFER TO OR IMPLY NONMARINE FAUNA

NO.	GLOSS	1	2	3	4	5	FORM
1.	bandicoot			X	X	X	mansar
2.	civet cat		X	X			musaŋ
3.	cock		X	X			laluŋ
4.	crocodile		X	X	X	X	buqaya
5.	cuscus			X	X	X	kandoRa
6.	deer		X	X			Rusa
7.	deer, bark of	X	X				dekiŋ
8.	dog	X	X	X			asu
9.	domesticated animal	(X)	X				qayam
10.	dove sp.	X	X	X	X	X	baluj
11.	dove sp.		X	X			-muken
12.	dove sp.		X	X	X	X	punay
13.	eel, freshwater	X	X	X	X	X	tuNa
14.	flying fox		X	X	X	LI	paniki
15.	horn	X	X				uReŋ
16.	hornbill	(X)	X	X		LI	kalaw
17.	monkey sp.	X	X				luCuŋ
18.	monkey, scream of		X	X			keraq
19.	monkey, scream of	X	X				keriq
20.	monkey, scream of		X	X			kusik
21.	pangolin	X	X				qaRem
22.	pig, wild	X	X	X			babuy
23.	pig, domesticated	X	X	X	X	X	beRek
24.	python		X	X			sawa
25.	ruminant sp.	X	X				(qa)Nuaŋ

TABLE 4. LEXICAL RECONSTRUCTIONS THAT REFER TO WINDS, DIRECTIONS, TOPOGRAPHY, AND THE LIKE

NO.	GLOSS	1	2	3	4	5	FORM
1.	cave		X	X	X	X	liaŋ
2.	earthquake	X	X				linuR
3.	hill	(X)	X				bukid
4.	lake	X	X	X	X	X	danaw
5.	mountain peak		X	X		LI	bunduk
6.	mountain range		X	X			qilih
7.	monsoon, northwest	(X)	X	X	X	(X)	SabaRat
8.	monsoon, southeast	(X)	X	X	X	(X)	timuR
9.	north wind, cold weather	X	X				qamiS(-an)
10.	toward the interior	X	X	X	X	X	daya
11.	toward the sea	X	X	X	X	(X)	lahud
12.	typhoon	X	X				baRiuS

TABLE 5. LEXICAL RECONSTRUCTIONS THAT REFER TO OR IMPLY THE SEA

NO.	GLOSS	1	2	3	4	5	FORM
1.	channel		X	X	X	X	sawaq
2.	clam, *Tridacna* spp.		X	X	X	X	kima
3.	coral sp.		X	X	X	X	buŋa
4.	crab, coconut		X	X	X	X	qayuyu
5.	crab, hermit		X	X	X	X	qumaŋ
6.	crab, mangrove		X	X	X	X	qali-maŋaw/maŋu
7.	crab, sand		X	X	X	X	kaRuki
8.	crab sp.	X	X				kagaŋ/kaRaŋ
9.	cuttlefish, squid		X	X	X?	X?	nus
10.	dugong		X	X	X	X	duyuŋ
11.	fathom		X	X	X	X	depa
12.	barracuda		X	X	X	X	qalu
13.	dolphinfish		X	X	X	X	lajih
14.	goatfish		X	X	X	X	tiqaw
15.	grouper (?)		X	X	X	X	keRteŋ
16.	mackerel, Spanish		X	X	X	X	taŋiRi
17.	milkfish		X	X	X	X	qawa?
18.	mullet (young)		X	X	X	X	qaRuas
19.	mullet (adult)	X	X	X	X	X	kaNasay
20.	perch, sea (?)		X	X	X	X	kurapu
21.	pilotfish, remora		X	X	X	(X)	gemi/kemi
22.	porcupinefish		X	X	X	X	taRutum/taRutuŋ
23.	sailfish		X	X	X	X	saku layaR
24.	scad, big-eyed		X	X	X	X	qatulay
25.	shark (generic?)	X	X	X		LI	qiSu
26.	squirrelfish		X	X	X	X	taRaqan
27.	stingray	X	X	X	X	X	paRiS
28.	stonefish		X	X	X	X	nepuq
29.	trevally		X	X	X	X	bilu
30.	tuna, bonito		X	X	X	X	qatun
31.	unicornfish		X	X	X	X	qumay
32.	wrasse		X	X	X	X	mamin
33.	island		X	X	X	X	nusa
34.	lagoon, harbor		X	X	X	X	namaw
35.	octopus		X	X	X	X	kuRita
36.	octopus arms		X	X	X	X	gaway
37.	oyster		X	X	X	X	tiRem
38.	sand	X	X	X	X	X	qenay
39.	sea, littoral		X	X	X	X	daRat
40.	sea, saltwater		X	X	X	X	tasik
41.	seaweed sp.		X	X			lamu/lamut
42.	seaweed sp.		X	X	X	X	limut
43.	seaweed sp.		X	X	X	X	lumut
44.	shell, cateye		X	X	X	X	qaliliŋ
45.	shell, conch		X	X	X	X	tambuRi
46.	shell, cowrie		X	X	X	X	buliq
47.	shrimp, lobster		X	X	X	X	quDaŋ
48.	snail or barnacle spp.		X	X	X	X	sisi/sisiq
49.	starfish		X	X	X	X	saŋasaŋa
50.	surf, breakers	X	X	X	X	X	Nabek
51.	tide, low; exposed reef		X	X	X	X	ma-qati
52.	tide, high		X	X	X	X	Ruab
53.	turtle, green		X	X	X	X	peñu
54.	turtle, hawksbill			X	X	X	keRaŋ

REFERENCES

ANCEAUX, J. C.
1965 Linguistic theories about the Austronesian homeland. *BTLV* 121:417–432.

BELLWOOD, PETER
1979 *Man's Conquest of the Pacific.* New York: Oxford University Press.

BENEDICT, PAUL K.
1941 A Cham colony on the island of Hainan. *Harvard Journal of Asiatic Studies* 6:129–134.
1942 Thai, Kadai, and Indonesian: a new alignment in South-eastern Asia. *AA* 44:576–601.

BENVENISTE, ÉMILE
1973 *Indo-European Language and Society.* Trans. from the French by Elizabeth Palmer. Miami Linguis-
(1969) tic Series No. 12. University of Miami Press.

BLUST, ROBERT
1970 Proto-Austronesian addenda. *Oceanic Linguistics* 9:104–162.
1972 Note on PAN *qa(R) (CtT)a 'outsiders, alien people'. *Oceanic Linguistics* 11:166–171.
1973 Additions to "Proto-Austronesian addenda" and "Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages—II." *Working Papers in Linguistics* 5 (3): 33–61. Honolulu: Department of Linguistics, University of Hawaii.
1974 The Proto-North Sarawak vowel deletion hypothesis. Unpublished Ph.D. dissertation, Department of Linguistics, University of Hawaii.
1976 Austronesian culture history: some linguistic inferences and their relations to the archaeological record. *WA* 8:19–43.
1977 The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report. *Working Papers in Linguistics* 9 (2): 1–15. Honolulu: Department of Linguistics, University of Hawaii.
1978 Eastern Malayo-Polynesian: a subgrouping argument. *Pacific Linguistics* C61 (Fascicle 1):181–234. Canberra: Australian National University.
1980 Austronesian etymologies. *Oceanic Linguistics* 19:1–181.
1982 The linguistic value of the Wallace Line. *BTLV* 138:231–250.
1983– More on the position of the languages of eastern Indonesia. *Oceanic Linguistics* 22:1–28.
1984

CHANG, K. C.
1969 Review article, on Changpinian: a newly discovered preceramic culture from the agglomerate caves on the east coast of Taiwan, by Wen-hsun Sung. *AP* 12:133–136.

CHUNG, THEIN-FOOK, TSENG-CHIENG HUANG, AND RICHARD B. STAMPS
1973 Palaeoecological study of Taiwan (3)—the P'u-Li Basin. *Taiwania: Scientific Reports of the Botany Department of the National Taiwan University* 18 (2): 179–191.

COLLINS, JAMES T.
1983 *The Historical Relationships of the Languages of Central Maluku, Indonesia.* Pacific Linguistics D47. Canberra: Australian National University.

DAHL, OTTO CHR.
1951 *Malgache et Maanjan.* Oslo: Egede-instituttet.
1976 *Proto-Austronesian,* 2nd, rev. ed. Scandinavian Institute of Asian Studies Monograph Series No.
(1973) 15. London: Curzon Press.

DEMPWOLFF, OTTO
1938 *Vergleichende Lautlehre des austronesischen Wortschatzes,* vol. 3: Austronesisches Wörterverzeichnis. Berlin: Dietrich Reimer.

DYEN, ISIDORE
1956 Language distribution and migration theory. *Language* 32:611–626.
1965 *A Lexicostatistical Classification of the Austronesian Languages.* International Journal of American Linguistics Memoir 19.

EBERHARD, WOLFRAM
1948 *China's Geschichte.* Bern: A. Francke.

FERRELL, RALEIGH
1969 *Taiwan Aboriginal Groups: Problems in Cultural and Linguistic Classification.* Academia Sinica Monograph No. 17. Taipei: Institute of Ethnology.

FOX, ROBERT B.
1953 The Pinatubo Negritos: their useful plants and material culture. *PJS* 81:173–414.

FRIEDRICH, PAUL
1970 *Proto-Indo-European Trees.* Chicago: University of Chicago Press.

HARVEY, MARK
1982 Subgroups in Austronesian, in *Papers from the Third International Conference on Austronesian Linguistics,* vol. 2: *Tracking the Travellers:* 47–99, eds. Amran Halim, Lois Carrington, and S. A. Wurm. Pacific Linguistics C75. Canberra: Australian National University.

HAUDRICOURT, A. G.
1954 Les origines asiatiques des langues malayo-polynésiennes. *JSO* 10:180–183.

HEINE-GELDERN, ROBERT VON
1932 Urheimat und früheste Wanderungen der Austronesier. *Anthropos* 27:543–619.

KERN, HENDRIK
1889 Taalkundige gegevens ter bepaling van het stamland der Maleisch-Polynesische volken. *Verslagen en Mededeelingen der Koninklijke Akademie van Wetenschappen, afdeeling Letterkunde.* 3de reeks 6:270–287. (Reprinted in *Verspreide Geschriften* 6:105–120, The Hague, 1917.)

MCFARLAND, CURTIS D., AND SHIGERU TSUCHIDA
1976 Linguistic evidence for the determination of the original homeland of the Malayo-Polynesian peoples. Trans. of Hendrik Kern, Taalkundige gegevens ter bepaling van het stamland der Maleisch-Polynesische volken. *Oceanic Studies: Linguistics, Anthropology & Sociology* 1:60–81.

MERRILL, ELMER D.
1954 *Plant Life of the Pacific World.* New York: Macmillan.

NOTHOFER, BERND
1986 The Barrier Island languages in the Austronesian language family, in *FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics:* 87–109, eds. Paul Geraghty, Lois Carrington, and S. A. Wurm. Pacific Linguistics C94. Canberra: Australian National University.

PAWLEY, ANDREW
1981 Melanesian diversity and Polynesian homogeneity: a unified explanation for language, in *Studies in Pacific Languages and Cultures in Honour of Bruce Biggs:* 269–309, eds. Jim Hollyman and Andrew Pawley. Auckland: Linguistic Society of New Zealand.

PAWLEY, ANDREW, AND KAYE GREEN
1971 Lexical evidence for the Proto-Polynesian homeland. *Te Reo* 14:1–35.

REID, LAWRENCE A.
1982 The demise of Proto-Philippines, in *Papers from the Third International Conference on Austronesian Linguistics,* vol. 2: *Tracking the Travellers:* 201–216, eds. Amran Halim, Lois Carrington, and S. A. Wurm. Pacific Linguistics C75. Canberra: Australian National University.

SAPIR, EDWARD
1968 (1916) Time perspective in aboriginal American culture: a study in method, in *Selected Writings of Edward Sapir in Language, Culture and Personality:* 389–467, ed. David G. Mandelbaum. Berkeley: University of California Press.

SAUSSURE, FERDINAND DE
1966 (1915) *Course in General Linguistics,* ed. Charles Bally and Albert Sechehaye in collaboration with Albert Riedlinger. Trans. from the French by Wade Baskin. New York: McGraw-Hill.

SNEDDON, J. N.
1978 *Proto-Minhasan: Phonology, Morphology and Wordlist.* Pacific Linguistics B54. Canberra: Australian National University.

SOLHEIM II, WILHELM G.
1975 Reflections on the new data of Southeast Asian prehistory. *AP* 18:146–160.

SPOEHR, ALEXANDER
1957 *Marianas Prehistory.* Chicago: Fieldiana.

THIEME, PAUL
1958 The Indo-European language, in *Human Communication: Language and its Psychobiological Bases* (readings from *Scientific American*): 39–47, ed. William S-Y. Wang. San Francisco: W. H. Freeman and Company.

TSUCHIDA, SHIGERU
1976 *Reconstruction of Proto-Tsouic Phonology.* Monograph No. 5. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
1977 Some plant names in Formosan languages. *Computational Analyses of Asian and African Languages* 7:70–119.

VAVILOV, N.
1926 Studies on the origin of cultivated plants. *Bulletin of Applied Botany and Plant Breeding* 26:1–248 (1–138 in Russian, 139–248 in English).

VERHEIJEN, J. A. J.
1967 *Kamus Manggarai 1: Manggarai-Indonesia.* The Hague: Nijhoff.
1982 *Dictionary of Manggarai Plant Names.* Pacific Linguistics D43. Canberra: Australian National University.

WANG, PINXIAN
1984 Progress in late Cenozoic palaeoclimatology of China: a brief review, in *The Evolution of East Asian Environment,* vol. 1: *Geology and Palaeoclimatology:* 165–187, ed. Robert Orr Whyte. Hong Kong: University of Hong Kong, Centre of Asian Studies.

ZORC, R. DAVID
1986 The genetic relationships of Philippine languages, in *FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics:*147–173, eds. Paul Geraghty, Lois Carrington, and S. A. Wurm. Pacific Linguistics C94. Canberra: Australian National University.