



R. Blust

The linguistic value of the Wallace line

In: Bijdragen tot de Taal-, Land- en Volkenkunde 138 (1982), no: 2/3, Leiden, 231-250

**This PDF-file was downloaded from <http://www.kitlv-journals.nl>**

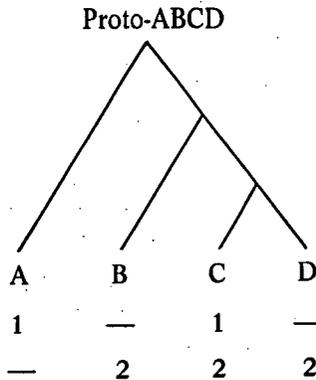
ROBERT BLUST

## THE LINGUISTIC VALUE OF THE WALLACE LINE

It is well known that reconstructed vocabulary can be of value in determining the center of dispersal ("homeland") of a language family.<sup>1</sup> It is also commonly recognized that the level of assignment (chronological value) of reconstructed vocabulary is dependent on the subgrouping of the languages compared. In particular, a linguistic feature can safely be attributed to the common ancestor of a collection of languages only if it is attested in at least two primary branches of that collection, and is not likely to result from borrowing or convergence. We might thus say that the chronological value of a lexical reconstruction is a consequence of our subgrouping assumptions (Fig. 1):

FIGURE 1

*Distribution of non-borrowed cognate sets in linguistic subgroups*



A, B, C, D = languages, or (if non-terminal) linguistic subgroups;  
1, 2 = cognate sets; hyphen = absence of a cognate term

Conclusion: 1 is assignable to Proto-ABCD, but 2 only to Proto-BCD  
(middle non-terminal node)

A subgrouping can, of course, also be a consequence of lexical reconstruction, as when we find evidence of exclusively shared lexical innovations. But where lexical reconstructions are used to establish a bipartite subgrouping on the highest level (as A vs. BCD in Fig. 1) the method can be regarded, with some justification, as circular. Although generally overlooked, it is noteworthy that a lower-level subgroup (as CD in Fig. 1) can under favorable circumstances, and without danger of circularity, be determined by lexical reconstruction on a higher level. In the following remarks I hope to show that, given an independently motivated subgrouping at the highest level, 1) the distribution of cognate terms for placental mammals in Austronesian (AN) languages supports a hypothesis that the AN homeland was west of the Wallace Line, and 2) the distribution of cognate terms for marsupial mammals supports a hypothesis that the languages of eastern Indonesia subgroup with those of Oceania. In so doing I will defend the view that non-linguistic facts such as major zoogeographical boundaries can have important consequences for historical linguistics.

1. The subgrouping of the AN languages on the highest levels is controversial. Traditionally (prior to Dempwolff 1934-8) the subject was avoided by implicit appeal to a four-way geographical classification into Indonesian, Melanesian, Micronesian (sometimes included with Melanesian) and Polynesian languages. In contrast to most earlier writers Dempwolff concluded that the languages of the last three geographical divisions (exclusive of Palauan and Chamorro of western Micronesia) belong to a single subgroup. He did not, however, commit himself to a view regarding the internal affiliations of the languages of Indonesia, the Philippines and Madagascar, and he did not consider the languages of Formosa.

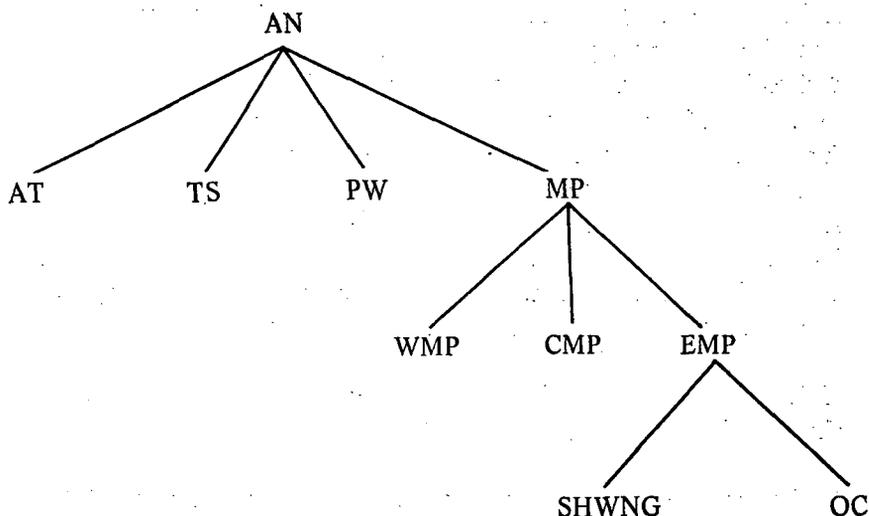
Subsequent scholars have supported Dempwolff overwhelmingly in recognizing a large eastern Austronesian subgroup, now generally called 'Oceanic'. They nonetheless remain divided with respect to the classification of the other members of this great language family.

Perhaps the most common recurrent view concerning the internal relations of the non-Oceanic AN languages is one which segregates some or all of the languages of Formosa from those of other areas at the highest level of distinction. Haudricourt (1965) divides the AN languages into three primary subgroups: 1) Western including the languages of the Philippines, Indonesia, mainland Southeast Asia and Madagascar, 2) Northern (= Formosan) and 3) Eastern (= Oceanic).

Dahl (1976) accepts a grouping into 1) Formosan and 2) the rest. Elsewhere (Blust 1977) I have proposed a division into four primary sub-groups, as shown in Figure 2:

FIGURE 2

*A subgrouping of the AN languages (after Blust 1977) <sup>2</sup>*



- AT: Atayalic (Formosa)  
 TS: Tsouic (Formosa)  
 PW: Paiwanic (Formosa)  
 MP: Malayo-Polynesian (all AN languages outside Formosa)  
 WMP: Western Malayo-Polynesian (the MP languages of the Philippines and western Indonesia, including Palauan, Chamorro, Chamic and Malagasy)  
 CMP: Central Malayo-Polynesian (the MP languages of the Lesser Sundas east of the Bima-Sumba group, and of the southern and central Moluccas)  
 EMP: Eastern Malayo-Polynesian (the MP languages of Halmahera and the Pacific region)  
 SHWNG: South Halmahera-West New Guinea (the MP languages of Halmahera, and of the adjacent north coast of New Guinea as far as the Mamberamo river)  
 OC: Oceanic (the MP languages of Melanesia, Micronesia and Polynesia except as stated elsewhere)

NOTE: CMP and EMP may share a common node below MP. The position of Yapese within MP is unclear.

Sirk (1978) has also suggested that the AN languages form two primary subgroups: 1) Formosan (or possibly just Atayalic) and 2) the rest. Most recently Harvey (1979) has argued that the first split in Austronesian is between certain of the languages of Formosa (including Atayalic, Tsouic and some Paiwanic languages) and the rest.

One scholar who not only rejects this view, but who has taken issue with the Oceanic hypothesis is Isidore Dyen. In a classification of the AN languages based on his interpretation of lexicostatistical percentages for 245 languages, Dyen (1965) maintained that the AN family divides into 40 primary subgroups, of which all but three are confined to western Melanesia and adjacent areas. Following a principle first formulated in linguistics by Edward Sapir, he concluded that the AN homeland probably was in the area of New Guinea and the Bismarck Archipelago. No other Austronesian linguist has publicly accepted Dyen's results. Moreover, Dyen himself (1978) now adopts the radically different hypothesis that the AN languages divide into two primary subgroups: 1) Oceanic and 2) the rest. As an advocate of the Oceanic hypothesis I have previously opposed the major conclusions of Dyen (1965);<sup>3</sup> it is Dyen's more recent views that I wish to challenge here.

2. Because the essential details have been published elsewhere (Blust 1977) I will not attempt to justify my subgrouping of the AN languages on the highest level in this paper beyond a few summary remarks. The Proto-Austronesian pronouns fall into at least two partially distinct sets, which for convenience we can call long form (= actor/patient) and short form (= agent/possessor) pronouns. Short form pronominal roots consist of the last vowel plus any preceding consonant of the corresponding long form pronoun (e.g. \*aku:ku '1st sg.'). Internal Formosan comparison indicates an ancestral system in which this formal correspondence is regular, but in all major geographical regions and island groups outside Formosa we find that the short form pronoun corresponding to \*kamu '2nd pl.' typically, though not exclusively, is a *singular* pronoun. Considerations of symmetry and simplicity favor our treating this usage as an innovation. Given the unnaturalness of a plural-to-singular shift which affects the short form pronoun but not the corresponding long form, it is, moreover, simplest to attribute the singular use of \*-mu to *one* change rather than to a collection of convergent innovations. The change \*-mu '2nd pl.' > -mu '2nd sg.' is therefore taken as evidence for a non-Formosan (Malayo-Polynesian)

subgroup of Austronesian languages. Corroboration of this hypothesis is found in the disappearance of Proto-Austronesian preconsonantal and final \*S, and in the merger of a resultant final shwa with \*a in all Malayo-Polynesian languages: PAN \*kuSkus > PMP \*kuku 'nail (of finger, toe)', PAN \*tuqaS > PMP \*tuqa 'old', PAN \*CumeS > PMP \*tuma 'clothes louse', etc.<sup>4</sup>

It is possible, of course, that other types of evidence will turn out to support a subgrouping hypothesis different from that defended here. But in any such alternative the foregoing observations must be explained in some other way. Although the data taken to support the Formosan/Malayo-Polynesian split are not extensive, they include apparent innovations both in grammar and in phonology, and thus present a serious challenge to any competing view.

3. Given a division of the AN languages into Formosan and Malayo-Polynesian subgroups, certain Proto-Austronesian lexical reconstructions prove to be of diagnostic value in connection with the homeland question.

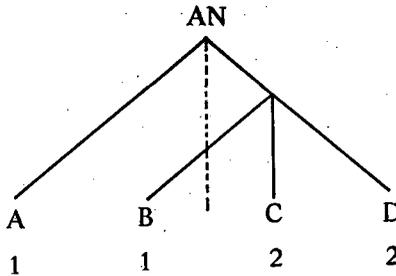
In the mid 19th century, during a prolonged collecting expedition to the Malay archipelago, the celebrated English naturalist A. R. Wallace discovered that the floral and faunal assemblages of the Greater Sunda islands closely resemble those of the Asian mainland, whereas those of the islands further to the east resemble those of Australia. The biota of Sulawesi (Celebes) partake to some extent of both regions, but are in other respects *sui generis*. In honor of his discovery, the division between these major zoological regions came to be called the 'Wallace Line'. What concerns us here is the indigenous distribution of mammals in relation to the Wallace Line. It is generally recognized today (Darlington 1957) that apart from the rat, which may have found transport aboard drifting logs, no terrestrial placental mammal established itself in areas east of the Moluccas before the arrival of man. Conversely, no marsupial mammal of any kind is native to areas west of the Wallace Line, apart from Sulawesi.<sup>5</sup> Since speakers of AN languages are found on both sides of this important natural boundary it stands to reason that the distribution of cognate terms for placental and marsupial mammals in these languages may shed light on the history of AN population movements.

The key to the linguistic value of the Wallace Line is the lack of correspondence between linguistic subgroups and biotic zones in the Austronesian world. If the Formosan/Malayo-Polynesian linguistic di-

vision corresponded to the placental/marsupial zoogeographical division there would be little hope of determining from the distribution of cognate terms for mammals whether speakers of Proto-Austronesian resided west or east of the Wallace Line, since neither placental nor marsupial terms could then safely be attributed to Proto-Austronesian (as each would be confined to a single primary subgroup). For the linguist it is a fortunate accident of history not only that the principal zoogeographical boundaries fail to coincide with the major AN subgroups in island Southeast Asia, but that cognate terms for placental mammals are found both in Formosan and in Malayo-Polynesian languages, and that cognate terms for marsupial mammals are found both in Central Malayo-Polynesian and in Eastern Malayo-Polynesian languages (Fig. 3):

FIGURE 3

*Distribution of cognate sets relating to mammals in major AN subgroups*



A = Formosan, B = Western Malayo-Polynesian, C = Central Malayo-Polynesian, D = Eastern Malayo-Polynesian  
 1 = cognate sets for placental mammals, 2 = cognate sets for marsupial mammals  
 dotted line = Wallace Line

4. The comparisons relating to placental mammals are: <sup>6</sup>

1) \*qaRem 'scaly anteater, pangolin'

FORMOSA

Atayal (Squliq)

qom

Sediq

?áruŋ

Kanakanabu	kani-arúm-ai
Saaroa	?arəmə
Tsou	hi-arm-uz-a
Pazeh	?arúm
Bunun	halum
Puyuma	?arəm
Paiwan	qam
<b>BORNEO</b>	
Batu Belah	am
Kiput	arəm
Long Jegan	akəm
Baluy Kayan	həm
Ma'anyan	ayəm
Katingan	ahəm
2) *(qa)Nuaŋ 'ruminant sp.'	
<b>FORMOSA</b>	
Pazeh	núaŋ 'buffalo'
Thao	qnú:wan 'buffalo'
Rukai (Tona)	nwáŋə 'cow'
Bunun (Takitodoh)	qanváŋ 'deer', qanvaŋ bonól 'buffalo'
<b>PHILIPPINES</b>	
Ilokano	nuáŋ 'carabao'
Hanunóo	?anwáŋ 'tamarau, timarau, a more or less straight-horned wild buffalo peculiar to the island of Mindoro: Bubalus mindorensis Heude'
Aklanon	ánwaŋ 'carabao'
<b>SULAWESI</b>	
Tae'	nuaŋ 'bush antelope: Anoa depressicornis' (Dutch <i>gemsbuffel</i> )
Buginese	anuaŋ idem
Makasarese	anuaŋ 'wild mountain buffalo: Anoa depressicornis' (Dutch <i>gemsbuffel</i> )
<b>MALAYA/SUMATRA</b>	
Malay	b-enuaŋ 'big, heavily built; but only in the expressions <i>kerbau b.</i> , <i>rusa b.</i> , ... <i>kerbau b.</i> 'the larger of two Sumatran varieties of buffalo', <i>rusa b.</i> 'a big-bodied <i>Cervus equinus</i> (deer)'

## 3) \*luCuŋ 'monkey sp.'

## FORMOSA

Kavalan	Ro:tóŋ
Saisiyat (Taai)	Loson
Thao	rúthun <sup>7</sup>
Bunun (Takitodoh)	hotóŋ
Ami	rotóŋ
Puyuma	Lutuŋ

## PHILIPPINES

Subanun	gutun
Sarangani Manobo	loton

## MALAYA/SUMATRA

Malay	loton 'black or grey long-tailed monkey: Semnopithecus spp.'
-------	---

## JAVA

Javanese	lutun
Sundanese	lutun

## BALI

Balinese	lutun
----------	-------

## 4) \*(q)uReŋ 'horn'

## FORMOSA

Sediq	?uruŋ
Pazeh	?uxúŋ
Yami	?urúŋ

## BORNEO

Long Terawan	ugun
Long Anap	uəŋ
Uma Juman	huəŋ (met.) 'horn, antler'

## 5) \*salajeŋ 'male, of ruminants'

## FORMOSA

Proto-Rukai	*salaonəŋ 'stag, male deer'
-------------	-----------------------------

## PHILIPPINES

Western Bukidnon	
Manobo	selazeŋ 'male deer'
Maranao	saladeŋ 'deer'

## MALAYA/SUMATRA

Malay	seladaŋ 'wild ox: <i>Bos gaurus</i> ; tapir (in Sumatra)'
-------	--

In evaluating the evidence for a linguistic reconstruction four considerations are particularly important: 1) the number of *languages* in which an apparent cognate is found; as this increases, the probability that the comparison is due to chance decreases, 2) the number of *primary subgroups* in which a cognate set is distributed; this determines the chronological value of the reconstruction, 3) the regularity of phonological development; irregularities may suggest borrowing, 4) the spatial character of the distribution; a discontinuous distribution is less likely than a continuous distribution to be a product of diffusion.

Because it is widely attested both in Formosa and in Borneo, comparison 1) cannot reasonably be attributed to chance. Within historical times the scaly anteater has been recorded in the Philippines only from the island of Palawan, and from the smaller neighboring islands of the Kalamian and Cuyo groups. Known terms for the scaly anteater in these areas and in Sabah (north Borneo) do not reflect \*qaRem.<sup>8</sup> Reflexes of this root thus display a discontinuous distribution, with a consequently low probability that they are due to borrowing. This conclusion is further supported by the nearly problem-free character of the phonological correspondences.

Comparison 2) is widely attested in Formosa and the Philippines as far south as northern Panay in the meaning 'carabao, domestic water buffalo'. Occasionally (as in Hanunóo) the word refers to a wild species. It is unrecorded in the southern Philippines, Borneo and northern Sulawesi, but appears in central and southern Sulawesi as the name of the peculiarly Sulawesi *Anoa depressicornis*, an indigenous dwarf buffalo.<sup>9</sup> As with cp. 1) the distribution is discontinuous, and thus not likely to be a product of borrowing. Phonological correspondences are regular, except that some languages in each primary subgroup indicate a disyllable (\*Nuɑŋ), whereas others indicate a trisyllable (\*qaNuɑŋ).

Comparison 3) was first recognized in principle by Kern (1889), who noted that semantically similar cognates of Malay *lotɔŋ* 'black or grey long-tailed monkey: *Semnopithecus* spp.' are found in Javanese, Sundanese, Balinese and in Formosa (*rutɔŋ*). He argued further that Makasarese, Buginese *lotɔŋ* 'black' is a cognate with narrowed meaning. Kern knew of no cognates in the Philippines, and he did not specify the Formosan language from which *rutɔŋ* was drawn.<sup>10</sup> Although somewhat less well attested outside Formosa, this comparison, like cps. 1) and 2), exemplifies a discontinuous distribution, and thus is not likely to be due to diffusion. Except for Subanun *gutɔŋ* (expected Sindangan

\*\*dlutuŋ, Siocon \*\*glutuŋ) correspondences appear to be regular.

Comparison 4) is less well-attested than the preceding. As with cps. 1) - 3), however, the known distribution is discontinuous, and phonological correspondences appear to be regular (last-syllable \*e normally > Long Terawan /ə/, but sometimes > /u/, as in \*?enem > *num* 'six'). Despite its superficial plausibility, the comparison Tsou *suŋu*, Javanese *suŋu* 'horn', Toba Batak *suŋu* 'rhinoceros horn' apparently is due to chance (cf. Tsuchida 1976: 215, Proto-South Formosan \*səquŋ).

Comparison 5) is problematic. Malay *seladan* and related Philippine forms point to Proto-Western Malayo-Polynesian \*sala(dDj)eŋ (Blust 1970). Thus far suspected cognates in Formosa are limited to Rukai, and these can be compared to non-Formosan forms only on the assumption 1) that PAN \*d, \*D or \*j disappeared in Proto-Rukai, 2) that PAN \*e sometimes became PR \*o and 3) that the form in question contains a suffix \*-an. Li (1977: 35) indicates that PAN \*d and \*D merged as PR \*D, and that PAN \*j became PR \*g, except after \*i, where it disappeared (\*pija > \*pia). However, he cites only one example of the latter change, and I find no statement on the development of secondary postvocalic shwa. This comparison may be a product of chance, but such near-agreement in the phonological correspondences for a relatively long word, and the specific semantic identity of the Western Bukidnon Manobo and Rukai terms raises the possibility that cognation ultimately will be established.

Whatever decision is reached with regard to comparison 5), cps. 1) - 4) support an inference that clearly is of major interest to students of Austronesian culture history. Since the scaly anteater, all forms of ruminant and all forms of monkey are placental mammals, and since no marsupial has horns, it follows that speakers of a language ancestral to at least the languages of Formosa, the Philippines and western Indonesia (including Sulawesi) had terms for placental mammals. On the subgrouping assumptions defended in sect. 2 this language would be Proto-Austronesian, a probable dialect-complex which must have begun to differentiate by about 4,000 B.C. (Pawley and Green 1973; Blust 1976; Bellwood 1979). If Proto-Austronesian speakers had terms for placental mammals presumably they lived in an area in which such animals were found, and hence were located west of the Wallace Line, or on Sulawesi. But on present knowledge this area could not have been Sulawesi, which has no extant pangolin.<sup>11</sup>

5. The comparisons relating to marsupial mammals are:<sup>6</sup>

6) \*kandoRa 'cuscus'

MOLUCCAS

Watubela	kadola
Kesui (Keldor village)	udora
Misool (Fofanlap village)	do:?
Buli	do 'small marsupial'

MELANESIA

Manam	?odora
Sori	ohay (met.)
Penchal	kotay (met.)
Lenkau	ŋohay (met.)
Nauna	kocay (met.)
Mussau	aroa
Nggela	kandora

7) \*mansar 'bandicoot, marsupial rat'

LESSER SUNDAS

Leti-Moa	mada, made
Damar	madar

MOLUCCAS

Yamdena	mande 'cuscus'
Ngaibor (Aru)	medar
Ujir (Aru)	meday
Kei	medar 'cuscus'
Elat (Kei)	mender
Amblau	mate
Asilulu (Ambon)	marel
Elpaputi (Ceram)	makele

MELANESIA

Motu	mada
Takia	madal
Wogeo	m <sup>w</sup> aja 'cuscus'
Duke of York	man
Likum	muc-ay
Lou	m <sup>w</sup> as
Nauna	m <sup>w</sup> ac

Many of the above languages are as yet poorly described, and the cognation of the terms in question remains to be established. For cp. 6)

Watubela regularly reflects \*k as /k/ and \*R as /l/ (\*kuCu > *kutu* 'head louse', \*Rumaq > *lumaq* 'house'); \*nd is not well attested, but there is no known counterevidence to the assumption that \*nd > Wtb /d/. Kesui (reportedly a dialect of Watubela) shows unexplained /u/ for \*ka-.

On the assumption that \*nd > /d/, the expected reflex of \*kandoRa both in Buli and in Misool would be \*\*ado (Blust 1978b). However, at least one other Buli word has lost the sequence \*ka-: \*kanaway > *mani naw* 'sea gull'.

Manam ?odora (anticipated \*\*?adora) possibly involves assimilation.

All languages of the Admiralty islands (Sori, Penchal, Lenkau, Nauna) have metathesized the vowels of \*kandoRa, but generally exhibit regular reflexes of \*k (\*kuDen > Sori *uh*, Penchal, Nauna *kul*, Lenkau *kuh*, 'clay cooking pot'), \*nd (\*panDan > Penchal, Nauna *pac*, Lenkau *pah* 'pandanus sp.') and \*R (\*(d)aRaq > Sori *hay*, Nauna *cay* 'blood', \*paRa > Penchal, Lenkau *pay* 'firewood rack above the hearth'). The initial consonant of Lenkau *ɲohay* is unexplained. Proto-Admiralty \*nd > Penchal /t/ also appears in \*nduRi > *tui* 'bone' (for a discussion of this root cf. Blust 1978a: 111-3).

Regular Mussau changes are \*k > Ø (\*kaSiw > *ai* 'tree'), \*nd > /r/ (\*panDan > *arana* 'pandanus sp.') and \*R > Ø (\*kuRita > *uita* 'octopus').

Comparison 7) is somewhat more problematic. Leti-Moa *mada* 'bandicoot', Damar *madar* 'marsupial sp.' can be compared directly with Motu *mada*, Takia *madal* and similar forms in other Oceanic languages, but the Leti-Moa variant *made* exhibits a different last-syllable vowel. Various languages of eastern Indonesia agree with Leti-Moa in suggesting an original vowel sequence \*a . . . e in this root, but others point instead to a sequence \*e . . . a. The reason for these discrepancies is unknown.

In many of the languages of eastern Indonesia \*nd and \*ns are not distinguished. That Leti-Moa *mada*, *made* can reflect a prototype with \*ns, however, seems clear from PAN \*isa (> *insa*) Leti, Moa *ida* 'one'. Details for Damar are unclear, but Stresemann (1927: 38-43, 45) reconstructed Proto-Ambon \*maɣer 'Beutelmarder' with an intervocalic consonant that is compatible with earlier \*nd or \*ns.

Among Oceanic witnesses Motu and Takia reflect earlier \*ns as /d/ (\*nsai > Motu *dai-ka* 'who?', \*nsapa > Motu *daha-ka* 'what?', \*ɲunso > Motu *udu* 'mouth, nose, beak', Takia *ɲudu* 'nose, snout', \*nsalan

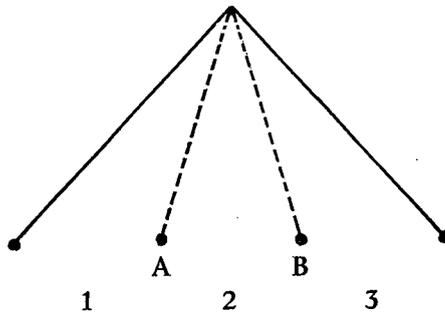
> Motu *dala*, Takia *dal* 'road'). It is not known whether Duke of York *man* is a regular development, but reflexes in the Admiralties (Likum, Lou, Penchal, Nauna) and in Wogeo indicate Proto-Oceanic \*m<sup>w</sup>ansad.<sup>12</sup>

As can be seen, although comparisons 6) and 7) are assigned to etyma \*kandoRa, \*mansar, the language to which these etyma are assigned has not been stated. Either etyma 6) and 7) are retentions from Proto-Austronesian, or they are innovations: there is no other possibility. Is there a principled basis for deciding between these alternatives?

As already observed, the Wallace Line is absolute with regard to the western limit of marsupials, but not with regard to the eastern limit of placental mammals; what it marks, then, is not the boundary between two mutually exclusive zoological regions, but the boundary between the Asiatic zoological region and a transitional zone leading to the Australian zoological region. The eastern boundary of this transitional zone is marked by the Lydekker Line, which separates the mixed fauna of the Moluccas from the (almost) exclusively marsupial terrestrial fauna of New Guinea and areas east (Fig. 4):

FIGURE 4

*Tripartition of island Southeast Asia by the Wallace and Lydekker Lines*



A = Wallace Line, B = Lydekker Line

1 = Asiatic region, 2 = transitional zone, 3 = Australian region

Given a subgrouping which separates the Formosan languages from all other AN languages there seems no alternative to the conclusion that

Proto-Austronesian speakers had terms for the scaly anteater, a species of ruminant, a species of monkey, horn and possibly 'male, of ruminants'. This conclusion excludes any area east of the Lydekker Line as a possible AN homeland. But, as Wallace first noted, the placental mammals of the Lesser Sundas and Moluccas have the character (relatively superficial genetic differentiation from their relatives in the Greater Sundas or Sulawesi) of an adventive fauna, of which at least the monkey, deer and civet cat probably were introduced during the AN occupation of island Southeast Asia. This leaves Sulawesi as the only possible homeland east of the Wallace Line. As seen earlier, the absence of a recorded Sulawesi pangolin is difficult to reconcile with this hypothesis. Finally, the hypothesis of a Sulawesi homeland is also difficult to square with the relatively shallow differences observed among the languages of this island as compared with, say, Formosa.

Since Proto-Austronesian speakers evidently resided in the Asiatic faunal region then, we are forced to regard etyma 6) and 7) as innovations. If the languages of eastern Indonesia in fact subgroup with the languages of western Indonesia as Dyen (1978) maintains, speakers of CMP and Oceanic languages would have entered the Australian zoological region not as a linguistically united group, but through historically independent population movements.<sup>13</sup> Under such circumstances, however, terms for marsupials would have been innovated independently and hence would not be cognate.

There are two types of conditions under which the preceding statement might prove false: 1) if etyma 6) and 7) are substratum terms, borrowed from non-Austronesian languages already spoken east of the Wallace Line at the time of the AN penetration of eastern Indonesia and Oceania, or 2) if etyma 6) and 7) involve inherited AN morphemes which have undergone parallel semantic shifts in CMP and Oceanic languages. Neither of these possibilities seems to me to be strongly supported by present evidence.

Hypothesis 1) assumes without justification that cognate terms for 'cuscus' and 'bandicoot' were found in non-Austronesian languages that probably were widely separated either in space or in time or both, and that at least the former term underwent independent yet identical phonological adaptations in Proto-Central Malayo-Polynesian and in Proto-Oceanic. The argument that a word would undergo identical phonological adaptations if borrowed independently by speakers of distinct (but related) languages not only implies that the phonological systems of the borrowing languages are very similar, but suggests that

the phonological shape of the borrowed word is virtually identical in each of the two (or more) lending languages. Chances of such a situation occurring are extremely small.

Collins (p.c.) feels that reflexes of \*kandoRa may be cognate with Popalia (Tukangbesi islands) *kadola* 'chicken', Malay *burong kedera* 'Pallas' mynah: *Agrospar sturnius*', *b. kedera laut* 'the curlew sandpiper: *Tringa ferruginea*' and similar forms in western Indonesia. He argues that Geser (Southeastern Ceram) *manuk* 'bird', *manuk fanakit* 'cuscus', together with semantic parallels in other languages, provides evidence that the arboreal cuscus is a classificatory avian in the folk taxonomies of the central Moluccas. Given such a conception a semantic shift from 'bird sp.' to 'cuscus' would not be improbable. Even if this interpretation is adopted, however, it need not imply that the marsupial reflexes of \*kandoRa are the result of parallel semantic innovations in eastern Indonesia and Melanesia.<sup>14</sup>

Although the bandicoot is absent, two species of phalanger reportedly are found on Sulawesi. Only one term appears to be reconstructible for both of these: \*kusay (Sangir *kusai*, *kuse* 'phalanger ursinus', Tontemboan *kuse* 'cuscus', Bolaang Mongondow *kutoi* 'a marsupial: *Phalangista ursina*', Tae' *kuse* 'marsupial sp.', Mandar *use* 'phalanger sp.').<sup>15</sup> This situation is entirely compatible with the hypothesis that \*kandoRa and \*mansar were innovated in a speech community ancestral to many of the languages of eastern Indonesia and of Oceania, but not to the languages of Sulawesi. Indeed, without such a hypothesis the distribution of cognate sets relating to marsupial mammals in AN languages is inexplicable.

The subgrouping problem treated here is complex in that a solution to it requires the acceptance of an antecedent subgrouping condition (the Formosan/Malayo-Polynesian dichotomy) which at first sight might seem remote from the facts to be explained. But it clearly follows that if cps. 1) - 5) are assigned to Proto-Austronesian cps. 6) and 7) must be innovations. Since no basis for a convergent development is apparent it is simplest to attribute these innovations to a single speech community (Proto-Central-Eastern Malayo-Polynesian) that was immediately ancestral to the languages assigned by Blust (1977) to the Central Malayo-Polynesian and Eastern Malayo-Polynesian groups. With these data we are given not only the first strong confirmation of the Central-Eastern Malayo-Polynesian hypothesis, but also a clear reminder of the potential importance of non-linguistic facts to the enterprise of linguistic subgrouping.

## ACKNOWLEDGEMENT AND SOURCES

Portions of this paper would have been all but impossible to write were it not for the generosity of Jim Collins in supplying me with data from his voluminous fieldnotes on the languages of eastern Indonesia. For his assistance I am deeply grateful.

Material has been drawn from standard published sources for the following languages: Aklanon, Balinese, Bolaang Mongondow, Buginese, Buli, Hanunóo, Hawaiian, Ilokano, Javanese, Kapingamarangi, Lau, Leti-Moa, Malay, Maranao, Mokilese, Mota, Motu, Nggela, Rennellese, Samoan, Sangirese, Sundanese, Tae', Toba Batak, Western Bukidnon Manobo, Woleai and Yamdena. Other sources are: Blust (n.d.): Batu Belah, Kiput, Lenkau, Likum, Long Anap, Long Jegan, Long Terawan, Lou, Manam, Mussau, Nauna, Penchal, Sori, Takia, Uma Juman, Wogeo; Collins (n.d.): Amblau, Asilulu, Damar, Elat, Elpaputi, Kesui, Misool, Ngaibor, Popalia, Ujir, Watubela; Ferrell (1969): Ami, Atayal (Squiliq), Bunun, Paiwan, Pazeh, Sediq, Thao, Yami; Hudson (1967): Katingan, Ma'anyan; Li (1977): Rukai (Tona), Proto-Rukai, (1978): Saisiyat; Lanyon-Orgill (1962): Duke of York; Mills (to appear): Mandar; Reid (1971): Sarangani Manobo, Subanun (Sindangan and Siocon dialects); Rousseau (1974): Baluy Kayan; Sneddon (1978): Tonsawang, Tonsea; Ting (1978): Puyuma; Tsuchida (1971): Bunun (Takitodoh), Kavalan, (1976): Kanakanabu, Saaroa, Tsou. Makasarese material is drawn from the dictionary of Cense (1979).

*University of Leiden*

## NOTES

- <sup>1</sup> For a classic illustration of the technique in the Malayo-Polynesian, or as it is now more generally known Austronesian family of languages, cf. Kern (1889).
- <sup>2</sup> Dyen (1963) and Tsuchida (1976) argue that the Formosan languages form a subgroup. Given the proximity of these languages to one another and the likelihood of extensive borrowing between them over a period of millennia, I find their evidence unconvincing (Blust to appear a.). For the same reason, however, I find it convenient in reconstruction to treat the Formosan languages as a single witness. In addition, I now regard CMP as including the languages of the Bima-Sumba group.
- <sup>3</sup> Blust 1972a, 1972b, 1973, 1978a, 1978b.
- <sup>4</sup> Formosan reflexes of PAN \*kuSkus and \*tuqaS are well known (Dyen 1965). Likely Formosan cognates of Dempwolff's \*tuma 'clothes louse' are Saisiyat *somaeh*, Kavalan *tu:məs* 'body louse' and Ami *toməs* 'flea'. A second probable example of this change is seen in Tonsea *dəse?ə*, Tonsawang *ləse?ə* (next Paiwan *liseqes*) < \*liseqeS 'nit, louse egg'.
- <sup>5</sup> As originally formulated the Wallace Line corresponds closely to the submerged boundary of the Sunda shelf, thus dividing Bali from Lombok in the

Lesser Sundas and Borneo from Sulawesi further north. Wallace nonetheless recognized (1962: 160, 300-01) that although this boundary is absolute with regard to the western limit of marsupials, it is not absolute with regard to the eastern limit of placentals. The special position of Sulawesi has already been mentioned, but placental mammals — including species of monkey, civet cat, deer, wild pig and shrew — also appear in progressively diminishing variety east of Bali in the Lesser Sundas, and in the Moluccas (hence the subsequent Lydekker Line). Because most of these species are very closely related to species in the Greater Sundas, Wallace concluded that their occurrence in the Australian faunal region is due to relatively recent human intervention.

- <sup>6</sup> (Sect. 4 and 5.) Unless noted otherwise the meaning of a reflex is that attributed to the etymon. Forms recorded by Collins in Damar, Ngaibor, Ujir, Elat, Amblau, Asilulu and Elpaputi were noted simply as 'marsupial sp.'
- <sup>7</sup> Typographical limitations require the use of a digraph for Ferrell's *theta*.
- <sup>8</sup> Formosan reflexes of \*qaRem designate *Manis pentadactyla*; the extant pangolin of western Indonesia (and the extinct Philippine type) is a distinct species, *Manis javanica*. I am indebted to R. Ziegler, curator of mammalogy, Bernice P. Bishop Museum in Honolulu, for information on the distribution of the scaly anteater in Southeast Asia, and for supplying local names recorded for this animal during the last century in the western Philippines.
- <sup>9</sup> Groves (1969) classifies the Anoa as a subgenus of the genus *Bubalus* (which includes the Indian Buffalo, and the Tamarao of Mindoro). He notes further (1) that "Two distinct types of Anoa are found all over Celebes, a large one with white legs, long tail and rugged horns, and a small one with legs mainly the same colour as the body, short tail and conical horns. There is some indication that the small Anoa inhabits mountainous areas. In the absence of any evidence of intergradation they must be classed as two distinct species . . . *Bubalus (Anoa) depressicornis* and *B. (A.) quarlesi*." The frequent lay description of the Anoa as an antelope is perhaps owing to the fact that in the larger species (*depressicornis*) the horns are not bowed, as in the Indian Buffalo (carabao), but project straight backward, as in the South African oryx (Dutch *gemsbok*).
- <sup>10</sup> Evidently Siraya: cf. van der Vlis (1842) *roulong* (/rutoŋ/) 'monkey'.
- <sup>11</sup> Proponents of the view that the AN homeland was east of the Wallace Line might object that the terms in question actually referred to marsupials, but have undergone parallel semantic shifts in the languages which retain them. Thus \*qaRem might have referred to the echidna (or spiny anteater), a monotreme comprising distinct genera in New Guinea (*Zaglossus*) and Australia (*Tachyglossus*). However, no reflex of \*qaRem has yet been found east of the Wallace Line, and the chances that cps. 2) — 5) originally referred to marsupials seem slight.
- <sup>12</sup> Cf. Blust 1978a: 160. The appearance of labiovelar consonants in some Oceanic languages corresponding to simple labials in non-Oceanic AN languages is a longstanding problem (Goodenough 1962; Haudricourt 1962).
- <sup>13</sup> The proposed Central Malayo-Polynesian subgroup has thus far been assumed without qualitative demonstration. Dyen (1978) makes a similar assumption, presumably on the basis of his 1965 lexicostatistical results.
- <sup>14</sup> Collins' argument in fact carries no such implication. For the record it should be noted that although Malay *kedera* corresponds regularly with reflexes of \*kandoRa, Popalia appears to reflect \*R as /h/: \*(d)jaRaq > *raha* 'blood', \*SaReZan > *hosa* 'ladder'. Moreover, in many Oceanic languages reflexes of \*manuk mean 'bird (generic)', but the same root or a derivative

also refers to animals other than birds, seemingly with no implication of a narrow classificatory equivalence: Motu *manu* 'bird', *manumanu* 'beetle, insect; kind of stinging fish', Lau *manu* 'bird, any creature that flies (bee, beetle, etc.)', Mota *manu* 'bird, flying creature, beetle, bat; the bird or beetle which attaches itself to a tree or plant is the *manu* of it', Woleai *mal* 'animal; bird; animate object', Mokilese *mahn* 'animal', *mahn in lohjapw* 'kind of insect', *mahnsang* 'bird', Rennellese *manu* 'fauna except human beings and fish, including birds and flying insects, hairy animals, reptiles except turtles', Samoan *manu* 1) bird, 2) animal, cattle, 3) cow, cattle, horse, Hawaiian *manu* 'bird; any winged creature', fig. 'person'.

- <sup>15</sup> Also cf. Mills (to appear), Proto-South Sulawesi or Proto-Toraja \*kuse 'a marsupial animal, phalanger spp.'. Despite its resemblance this term is distinct from phonetically similar forms in the Admiralty islands (cp. 6), as a proto-type \*kusay would yield Proto-Oceanic \*kuse, Sori \*\*uh, Penchal \*\*kut, etc.

## REFERENCES

Bellwood, Peter

- 1979 Plants, climate and people; the early horticultural prehistory of Australasia. Ms.

Blust, Robert A.

- 1970 'Proto-Austronesian addenda', *Oceanic Linguistics* 9: 104-62.  
 1972a 'Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages: a preliminary list', *Working Papers in Linguistics* 4.1: 1-43, Honolulu, Department of Linguistics, University of Hawaii.  
 1972b 'Additions to "Proto-Austronesian addenda" and "Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages"', *WPL* 4.8: 1-17, Honolulu, Department of Linguistics, University of Hawaii.  
 1973 'Additions to "Proto-Austronesian addenda" and "Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages"' — II, *WPL* 5.3: 33-61, Honolulu, Department of Linguistics, University of Hawaii.  
 1976 'Austronesian culture history: some linguistic inferences and their relations to the archaeological record', *World Archaeology* 8.1: 19-43.  
 1977 'The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report', *WPL* 9.2: 1-15, Honolulu, Department of Linguistics, University of Hawaii.  
 1978a 'The Proto-Oceanic palatals', *Memoir* No. 43, The Polynesian Society, Wellington.  
 1978b 'Eastern Malayo-Polynesian: a subgrouping argument', Proceedings, Second International Conference on Austronesian Linguistics (SICAL), Fascicle 1: 181-234, *Pacific Linguistics* C61, Canberra, Australian National University.  
 to appear a. Review of Tsuchida, *Reconstruction of Proto-Tsouic phonology, Language* 1981.  
 n.d. Fieldnotes on languages of central and western Borneo, and western Melanesia.  
 —, ed.  
 to appear b. *Historical linguistics in Indonesia*, NUSA (Linguistic Studies in Indonesian and Languages in Indonesia), Jakarta.

Cense, A. A.

- 1979 Makassaars-Nederlands woordenboek, KITLV, The Hague, Nijhoff.

- Collins, J.  
n.d. Fieldnotes on languages of eastern Indonesia.
- Dahl, Otto Christian  
1976 *Proto-Austronesian*, 2nd. ed., Scandinavian Institute of Asian Studies Monograph Series, No. 15, Curzon Press.
- Darlington, Philip J., Jr.  
1957 *Zoogeography: the geographical distribution of animals*, New York, Wiley and Sons.
- Dempwolff, Otto  
1934-8 'Vergleichende Lautlehre des austronesischen Wortschatzes', *Zeitschrift für Eingeborenen-Sprachen*, Supplements 15 (1934), 17 (1937), 19 (1938), Berlin, Reimer.
- Dyen, Isidore  
1963 'The position of the Malayopolynesian languages of Formosa', *Asian Perspectives* 7.1-2: 261-71.  
1965 *A lexicostatistical classification of the Austronesian Languages*, International Journal of American Linguistics Memoir No. 19, Supplement to Vol. 31, No. 1.  
1978 'The position of the languages of eastern Indonesia', Proceedings, Second International Conference on Austronesian Linguistics (SICAL), Fascicle 1: 235-54, *Pacific Linguistics* C61, Canberra, Australian National University.
- Ferrell, Raleigh  
1969 *Taiwan aboriginal groups: problems in cultural and linguistic classification*, Institute of Ethnology Academia Sinica Monograph No. 17, Taipei.
- Goodenough, Ward H.  
1962 'Comment on Capell, "Oceanic linguistics today"', *Current Anthropology* 3: 406-8.
- Groves, Colin P.  
1969 'Systematics of the Anoa (Mammalia, Bovidae)', *Beaufortia* (series of miscellaneous publications, zoological museum of the University of Amsterdam), Vol. 17, No. 223: 1-12.
- Harvey, Mark  
1979 Subgroups in Austronesian. Ms.
- Haudricourt, André G.  
1962 'Comment on Capell, "Oceanic linguistics today"', *Current Anthropology* 3: 410.  
1965 'Problems of Austronesian comparative philology', *Lingua* 14: 315-29.
- Hudson, Alfred B.  
1967 *The Barito isolects of Borneo*, Department of Asian Studies, Southeast Asia Program, Data Paper No. 68. Ithaca, Cornell University.
- Kern, Hendrik  
1889 'Taalkundige gegevens ter bepaling van het stamland der Maleisch-Polynesische volken', *Verslagen en Mededeelingen der Koninklijke Akademie van Wetenschappen, Afd. Letterkunde*, 3de reeks, dl. 6: 270-87 (reprinted in *Verspreide Geschriften* 6: 105-20, The Hague, 1917).
- Lanyon-Orgill, Peter A.  
1962 *A dictionary of the Raluana language*, Victoria, B.C.

Li, Paul Jen-kuei.

1977 'The internal relationships of Rukai', *Bulletin of the Institute of History and Philology, Academia Sinica* 48.1: 1-92.

1978 'A comparative vocabulary of Saisiyat dialects', *Bulletin of the Institute of History and Philology, Academia Sinica* 49.2: 133-99.

Mills, Roger F.

to appear 'Additional addenda', in: Blust (to appear b.).

Pawley, Andrew and Roger Green

1973 'Dating the dispersal of the Oceanic languages', Papers of the First International Conference on Comparative Austronesian Linguistics (FICCAL), *Oceanic Linguistics* 12: 1-67.

Reid, Lawrence A., ed.

1971 *Philippine minor languages: word lists and phonologies*, Oceanic Linguistics Special Publication No. 8, Honolulu, University of Hawaii Press.

Rousseau, Jérôme

1974 'A vocabulary of Baluy Kayan', *Sarawak Museum Journal* 22.43 (NS): 93-152.

Sirk, Ü.

1978 'Problems of high-level subgrouping in Austronesian', Proceedings, Second International Conference on Austronesian Linguistics (SICAL), Fascicle 1: 255-73, *Pacific Linguistics* C61, Canberra, Australian National University.

Sneddon, J. N.

1978 'Proto-Minahasan: phonology, morphology and wordlist', *Pacific Linguistics* B54, Canberra, Australian National University.

Stresemann, Erwin

1927 'Die Lauterscheinungen in den ambonischen Sprachen', *Zeitschrift für Eingeborenen-Sprachen*, Supplement 10, Berlin, Reimer.

Ting, Pang-hsin

1978 'Reconstruction of Proto-Puyuma phonology', *Bulletin of the Institute of History and Philology, Academia Sinica* 49.3: 321-92.

Tsuchida, Shigeru

1971 List of words of Formosan languages. Ms.

1976 *Reconstruction of Proto-Tsouic phonology*, Studies of Languages and Cultures of Asia and Africa Monograph Series, No. 5, Tokyo.

Vlis, C. J. v. d.

1842 'Formosaansche woorden-lijst, volgens een Utrechtsch Handschrift', *Verhandelingen van het Bataviaasch Genootschap* 18: 437-88.

Wallace, Alfred Russel

1962 [1869] *The Malay Archipelago*, New York, Dover.