PROBLEMS OF HIGH-LEVEL SUBGROUPING IN AUSTRONESIAN

Ü. SIRK

This paper aims to discuss some problems arising in connection with a task of enormous complexity - subgrouping the Austronesian (AN) languages on the high level, i.e. clearing up into which groups the AN family divides immediately or near-immediately.

The AN languages of Oceania evidently have some features in common, which at the same time oppose them to the languages of more western regions. This observation has given rise to the idea of bipartite division of the family into Eastern AN, or Oceanic (OC), and Western AN, or Indonesian.[1]

The notion of Eastern AN, or OC, covering Polynesian as well as Melanesian and Micronesian languages (or at least greater parts of these traditional units) is, in fact, rejected by schools of thought which attach peculiar importance to the lexical diversity of the Melanesian languages.

One of these schools of thought is best represented by the 'pidginisation hypothesis' as developed by S.H. Ray and A. Capell. According to this hypothesis, the Melanesian languages have come into being as a result of pidginisation of Indonesian languages in a Papuan linguistic environment, and this situation has led to a heavy influx of Papuan vocabulary into the resulting languages (it is also assumed that such pidgins might undergo new pidginisation processes).

A rather detailed evaluation of the 'pidginisation hypothesis' has been given by G.W. Grace (1968, 1971, and other works). I cannot add much to Grace's analysis, but I would like to emphasise some points.

Modern pidgins (in any case, almost all those which have been better described) are characterised by a fairly high degree of material homogeneity. It seems unlikely therefore that the process of pidginisation of Indonesian languages in ancient Melanesia included, or directly

Sirk, U. "Problems of High-level Subgrouping in Austronesian". In Wurm, S.A. and Carrington, L. editors, *Second International Conference on Austronesian Linguistics: Proceedings.* **C-61:255-273**. Pacific Linguistics, The Australian National University, 1978. DOI:*10.15144/PL-C61.255*
©1978 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

induced, some kind of mass borrowing. More conceivably (in view of the lexical inadequacy of a pidgin in comparison with a native language) many vocabulary items might have been borrowed later, in the creolisation period, from Papuan adstrata which were still living at that time.

In principle, such developments are possible. Nonetheless, however, the 'pidginisation hypothesis' is highly speculative. The early Papuan adstrata - suppose they ever existed - have become extinct (or perhaps unrecognisable as old adstrata). Apparently it is only in extremely rare cases that a sizable amount of lexical material of the Papuan language (now extinct) which has served as adstratum or substratum for a Melanesian language can be reconstructed by comparing this Melanesian language with nearby situated Papuan languages (Cowan 1965).[2]

The result of the language contact has been partly determined in every case by various factors which are no longer reconstructible, namely: the exact situation of the Indonesian colony, relations between the colonists and the aborigines, etc.

Further, the very notion of 'Indonesian languages' is rather vague. As several authors have pointed out, the Melanesian languages, at least in Insular Melanesia, show in certain aspects more uniformity than do the Indonesian languages taken in the traditional sense. The aspects to be considered are grammar (both grammatical structure and morpheme inventory) and the list of 'Austronesian' (in the Dempwolffian sense) lexical roots.[3] If the linguistic map had been formed by contacts of different (albeit related) Indonesian languages with extremely diverse Papuan languages, it would be much more particoloured than the map as actually seen - we must agree with Grace (1962:409-10).

Another school of thought attempts to explain the lexical diversity of the Melanesian languages by their early - in comparison with other AN groups - divergence which resulted in long periods of independent development. I. Dyen's scheme of lexicostatistical classification of the AN languages contains no unit comparable with the traditional 'Melanesian'. The vast majority of the languages which are usually called Melanesian and which Dyen includes in his study are scattered among numerous small (most often one-language) groups; all these groups are co-ordinated with the huge Malayopolynesian group and, like it, labelled as immediate subdivisions of AN. Such small groups are concentrated in the area of New Guinea - Melanesia - Micronesia; only two or three of them are found farther west (on Taiwan and Enggano). The Malayopolynesian group covers, turning to the traditional terminology, most of the included Indonesian languages, all Polynesian and a number of Melanesian languages.

Is the lexicostatistical/glottochronological method a reliable means of classifying languages? This question has been the matter of unceasing controversies since the fifties. Now it can safely be said that the opponents of lexicostatistics have not succeeded in demonstrating its absolute inadequacy. On the other hand, however, the success scored by its advocates is far from being brilliant. The lexicostatistical method manifestly has its value as one of the shortcut ways which supply preliminary indications about the relations between languages. Sometimes, however - for most various reasons - the method under discussion does not justify the hopes set on it.

As applied by Dyen to the AN languages of Oceania, lexicostatistics clearly has laid no firm foundation for subgrouping them on the high level. As for my arguments, I refer to evident discrepancies between the divisions of Dyen's classification and numerous isoglosses which are characteristic of the AN languages of Oceania. Among these isoglosses there are morphemic (both lexical and grammatical), structural and historico-phonetic ones (the first group may be exemplified by *kianjo *'outrigger'*, *moli *'citrus'*, *ma 'future marker'; the second group - by the Oceanic possessive constructions; the third group - by merger of *p and *b, merger of certain palatals and *s, change of *ə to o, etc.). At the same time, such Oceanic isoglosses seem to be territorially closer to one another than to the divisions of Dyen's classification.

We have good reason to interpret a number of morphemic and structural features widely spread in Oceania but not reported from the West as innovations (this problem is discussed, particularly in connection with grammar, by Pawley (1973)).

What has been said above evidently shows that Melanesian linguistic diversity has no bearing on the problem of first-order subgrouping of AN.

Then, there does not remain much evidence in support of the intra-Oceanic localisation of the deepest split, or splits, of the AN family.

Further, at least part of the languages of Eastern Indonesia, most evidently those assigned by Esser (1938) to the South Halmahera-West New Guinea group, show considerable common features with the languages of Oceania. Basing himself upon this observation, Blust (1974, preliminary publication) suggested a larger group, 'Eastern Austronesian', including both the Oceanic and the South Halmahera-West New Guinea languages (in 1977, also in a preliminary publication, Blust calls this group Eastern Malayo-Polynesian, as he does in the paper in this volume).

All the AN languages not belonging to Blust's Eastern Austronesian/ Eastern Malayo-Polynesian are labelled here, quite preliminarily, as 'Western Austronesian'.

The Eastern versus Western distinction, established in such a way, is a useful starting-point for further discussion.

Let us turn to Western Austronesian now.

So far, nobody has advanced weighty qualitative evidence to show that the Western AN languages form one group.[4] Most likely, Western AN is indeed not a group. If so, however, which high-level groups does it cover (either fully or partly)?

I would like to draw your attention to some features which, most probably being of areal nature, are at the same time of interest in respect to the just named complicated problem.

Within the territory of Western AN, we can outline a number of more or less clearly marked areas defined by characteristics (both structural and material) of verb morphology. Among them, those most strictly opposed to one another are undoubtedly the areas that I call here Philippine and South Sunda. The former of the two areas covers, besides the Philippines, also Taiwan[5], the Marianas islands, the Sangir-Talaud islands and the northern parts of Sulawesi and Kalimantan. The South Sunda area includes the languages of Sumatra, (except Acehnese), those of the islands placed west of Sumatra, the Malay dialects of the Malay Peninsula (and nearby islands), the languages of Java, Madura, Bali, Lombok, those of the greater part of Sulawesi (and nearby islands) and some of the languages of Kalimantan (e.g. Kendayan). In Kalimantan, there are also (probably numerous) languages which belong to neither of the two mentioned areas.

Let us examine the verbal morphology characteristic of the Philippine area.

The most striking structural characteristic of this morphology is the so-called focus system: the verb-object relation and the relation of the verb to the subject (i.e. topic of sentence or determinée of attributive construction) are not marked by separate combinable sets of verbal affixes. At the same time, a clear formal distinction is made (by the use of different forms of pronouns syntactically bound with the verb, etc.) between active and passive forms.[6] The passive forms, besides expressing the known (passive) verb-subject relation, also differentiate verb-object relations. What results is the asymmetric shape of the paradigm: one active and several passive forms, the passive forms differing in object meaning.[7]

The most characteristic formants of the voice forms are the affixes assignable to the prototypes *-um-/mu- and *maR- (in active), *-ən (in so-called direct passive), *-an and *-i (in so-called local passive, usually in different modal forms) and *Si- (in so-called instrumental passive). For example, in Tagalog the verb stem -bilí *'buy'* produces

the following voice forms (exemplified by the so-called infinitives; with rough translation):

ACTIVE:	'DIRECT PASSIVE':
b-um-ilʃ	bilh-ʃn
'to buy'	*'to be what (who) is bought'*
	'LOCAL PASSIVE':
	bilh-án
	'to be from whom something (somebody) is bought'
	'INSTRUMENTAL PASSIVE':
	i-bilʃ
	'to be for what something (somebody) is bought (said, e.g., about money)*';* *'to be used to buy something (somebody)';* *'to be for whose use something (somebody) is bought'*

The languages of the Philippine area are further characterised by synthetical tense forms (exceptions, like Tsou and Chamorro, are extremely rare). The past tense marker originating from *-in-/ni- is found almost everywhere.[8]

In the languages of the South Sunda area the verbal affixation systems include two separate sets of affixes: those object-oriented and those subject-oriented. The verb-object and verb-subject relations are expressed separately from one another. As a result, the scheme is symmetric: verbs including an object relation marker have both active and passive forms.[9] Here is an example taken from Malay:

ACTIVE	PASSIVE
meng-ambil	di-ambil
'to take'	
me-nerbang-i	di-terbang-i
'to fly over (something)'; *'to fly (covering a distance)'*	

ACTIVE	PASSIVE
me-nerbang-kan	di-terbang-kan
'to fly away with... (snatching, carrying somebody, or something)'; 'to make to fly'	

Notwithstanding the differences in the structure of paradigm and the functional load of affixes, the two areas under discussion show remarkable similarity in the inventory of verbal affixes participating in expression of verb-object and verb-subject relations. In my opinion, the most important point of difference is that the Philippine area languages do not have verbal suffixes assignable to the prototypes *-kən and *-akən whose reflexes are widely spread in the South Sunda area.[10] In particular, *-kən is the source of Malay and Minangkabau -kan, Karo -kěn, Toba -hon, Sundanese -keun, Kroe and Banggai -kon; most likely the suffix -ka visible in most of the Toraja languages also stems from it.[11] -aka of the greater part of the Toraja languages and Wolio, -ako/-akon- of Mori and -akěn of Javanese and Bawean Madurese constitute evidence permitting the reconstruction *-akən (cf. also the Malay preposition akan).[12]

As far as I can judge the South Sunda area, it is only in the majority of the Toraja languages that both the suffixes under discussion coexist side by side, retaining special meanings. In these languages, the suffix -ka exhibits considerable semantic analogy with the Philippine 'instrumental passive', the most typical shades of meaning being the following: *'for (somebody)'; 'in connection with (somebody, something)'; 'with (somebody, something)', 'using (something) as the means, instrument'*. The meaning expressed by -aka most often seems to be factitive (also called 'weak causative').

The other languages of the South Sunda area have only one suffix, which usually covers so wide range of meanings that it is difficult to refrain from supposing that this suffix may be a result of some merger.[13]

A further contrastive feature of the South Sunda area is the following. In the languages of this area the affixal elements assignable to *-in-/ni- do not express tense meanings (nor do they express meanings like *'accomplished', 'not yet accomplished'*), and the alternation m~n practically does not occur in the prefix-initial position.[14] In addition, synthetical tense (and presumable 'aspect-tense') forms are rare in the South Sunda area, and if found somewhere, very markedly differ in structure from the forms of the Philippine area.

The South Sunda area as a whole (with the exception of a few languages, such as Sundanese, Madurese and Balinese) is characterised by

use of the so-called conjugated forms (Dutch: **vervoegde vormen**) of verb, which can be schematically presented as: 'agens-denoting pronominal element + verb stem'. For instance, in Makassarese: **ku-alle** *'I take, (is) taken by me'*, **nu-alle** *'you take, (is) taken by you'*, and so on. The Philippine area, as a whole, with the exception of Chamorro (and possibly some poorly known languages), lacks this construction.

What about the historical relations between the Philippine and South Sunda systems of verbal morphology?

Abstractly speaking, alternatives must be borne in mind: one of the systems may be a transformation of the other, or both systems may have developed independently from a third one. If the last alternative is preferred, it is reasonable to regard the third, or source system as somewhat less sophisticated, lacking differentiation of verb-object relations by means of verbal affixes (analogies might be found either in Maanyan, or in Acehnese, in the languages of South-Eastern Indonesia, in the Chamic languages, etc.).

The idea of development directed from something simple to something complex is of course correct in principle. But only in principle! It seems that Austronesian verbal morphology passed through the concrete phase of development 'from simple to complex' at an early pre-Austronesian stage, which may even lie beyond the limit of feasible reconstruction.

I would like to suggest for consideration here the two following theses. (None of them is entirely new for Austronesianists because they have already been proposed in one or another form.)

1. The Philippine and South Sunda systems of verbal morphology have not developed independently of each other. The voice forms of the two systems show so much similarity in semantics that the contrary opinion is clearly unlikely. The semantic similarity of the systems affects even small details. For example, both the 'instrumental passive' and South Sunda verbs with suffixes assignable to ***-kən** and ***-akən** express (among other relations) the relations to the beneficiary object and to the object of actions like *'to replace, carry, etc., something in a direction away from the agens or the speaker'* (cf. Wolff 1973:79). The verb-object relations expressed by the South Sunda verbs with suffixes assignable to ***-i** and ***-an** are nearly identical with those expressed by the 'local passive'.

2. No common ancestor of the Philippine and South Sunda systems of verbal morphology is to be sought among the less sophisticated systems which are observable in Maanyan and other languages of the interior of Kalimantan, in Acehnese, in the Chamic languages and in the languages of South-Eastern Indonesia. Particularly, the position of Maanyan seems

to be rather clear in view of the semantically recognisable relics of the 'direct passive' suffix *-ən (e.g. kotaen *'food'*, cf. -kota *'eat'*; kakanen *'dainty dish'* - see Dahl 1951). Probably the 'local passive' *-an has also left traces in Maanyan: its vestige is apparently encountered in the noun ondan *'pillow'*, which word must then be interpreted as *'the place where someone's head is laid'* (NB! It is not **onden); see Aichele 1956:221. In Busang, another Kalimantan language, there is kanĕn *'boiled rice'*; does it also reflect the 'direct passive' *-ən? As regards Acehnese and the Chamic languages, they have been exposed to extremely strong foreign influences and therefore can hardly serve as witnesses of early AN morphology. In connection with the languages of South-Eastern Indonesia (namely with Esser's Bima-Sumba group and the languages of Timor region, Kei and Tanimbar), some investigation is perhaps still needed to determine their validity as witnesses, but what seems to me most probable is that these languages have lost the older AN verbal morphology. It is not to be forgotten that in the area under discussion, too, non-AN influences have been strong (which has more than once been emphasised by A. Capell).

If the two theses presented are sound, the question may be put as follows: which system - that of the Philippine area or that of the South Sunda area - has retained more characteristics of the ancestral morphology?

The scheme of verbal inflection reconstructed by J.U. Wolff (1973) and tentatively ascribed to the proto-AN level resembles quite closely the morphology observable in the Philippine area. No doubt, some caution is needed in accepting Wolff's reconstruction because it is mainly based only on four languages, and three among them (viz. Samar-Leyte, Atayal and Tsou, in contrast to Javanese) belong to the Philippine area. However, Wolff's argument can be considerably strengthened.

Prima facie evidence for the conservative nature of the Philippine area is offered by the Malagasy language. In verbal morphology Malagasy bears much more resemblance to the languages of the Philippine area, than to those of the South Sunda area and the interior of Kalimantan. The 'Philippine' traits of Malagasy are the following ones: syncretism of the categories 'verb-object relation' and 'verb-subject relation', asymmetric configuration of paradigms, use of n- in past tense forms, which n- is in many cases opposed to m- of the present tense. Malagasy has neither suffixes assignable to *-kən or *-akən nor 'conjugated forms'. In other spheres, however - at least in vocabulary - the situation is clearly different. It seems that Malagasy has its closest lexical ties with the South-East Barito group. O.C. Dahl has devoted a monograph to the relations of Malagasy with the Maanyan language,

which was at that time (1951) practically the only known member of the said group. I. Dyen (1971:236-7) gives a fairly long list of lexical correspondences (as it seems, culled out of Dahl's monograph) which are evaluated by him, with caution, as apparently exclusive Malagasy-Maanyan.[15] To all appearances, Malagasy shares a part of phonetic history with the South-East Barito group (Dahl 1976, 1977). There are also material resemblances with other languages of Kalimantan (e.g. the Malagasy passive formant voa-). Still more significantly, we can find widely spread isoglosses which unite Malagasy with numerous languages of the western part of Indonesia, but apparently do not go into the Philippine area very far (seemingly no farther than Northern Kalimantan), namely: Malagasy olona *'man, person'* (such semantics of reflexes of *qulun is characteristic of Kalimantan), vorona *'bird'* (from Dempwolff's *buluŋ), vintana *'fate, destiny'*, kintana *'star'*.[16] After all, Malagasy resembles the languages of Western Indonesia (including the interior of Kalimantan) in respect of the semantics of the prefix (m)aN- (on this prefix see J[onker] 1918).

The described picture is most simply and naturally explained by assuming that in the main features of verbal morphology the proto-language ancestral to Malagasy, to the languages of interior Kalimantan (at least to the South-East Barito group) and to the South Sunda area languages has been similar to the modern languages of the Philippine area (however, the early semantics of (m)aN- remains a problem). This assumed initial status has been most faithfully preserved by Malagasy (although a number of fairly significant changes have taken place in that language, too). In Kalimantan the old morphology has been greatly simplified, and in the South Sunda area it has been rebuilt on a rather different basis.

The hypothesis claiming conservative nature for the Philippine area can be further strengthened by adducing evidence of phenomena which are most naturally explained as relics of Philippine-type morphology. Such phenomena occur in many AN languages. For example, the reflexes of *-ən and *-an recognisable in the so-called passive imperatives of Javanese are, in respect to verb-object relations, quite similar to the Philippine suffixes. Javanese, Toba and some other languages replace -i by -an in forms containing the infix -in- (e.g. Toba: ma-nortor-i *'to dance (upon, about something)'*- passives di-tortor-i, t-in-ortor-an). Concerning Javanese, more details can be found in Wolff 1973. I have already noted apparent relics observable in Maanyan and Busang.

Speaking *a priori*, without any elaborated theoretical basis, a change from an asymmetric scheme of voice forms to some kind of symmetric

scheme seems more probable than a change in the opposite direction. If this is the case, I am bold enough to come to the conclusion that the passage to the South Sunda scheme of voice forms might have had its prerequisite in Philippine-type morphology. Then it would be reasonable to expect that verbal forms separating the verb-object and verb-subject relations would also be discovered in the languages of the Philippine area. Indeed, in Casiguran Dumagat, besides n-i-ginan *'was the thing with which somebody ran away'* (passive), there is nang-i-ginan *'ran away with (something)'* (active) (judging from Headland and Headland 1974:xxxvii).[17]

Although it seems that in the Philippine area some steps were taken in the direction away from the asymmetric scheme of voice forms ('focuses'), nevertheless the **radical change itself** can undoubtedly be considered an innovation of the South Sunda area. The other innovations attributable to the last-mentioned area are: rise of the verbal suffixes *-kən and *-akən; enlargement of the sphere of functioning of -i and -an from passive to both active and passive; loss of the past tense forms marked by the reflexes of *-in-/ni-; in all probability, the 'conjugated forms' may also be referred to.[18]

Of course, these few innovations do not constitute any reliable base for uniting the languages of the South Sunda area in one group - the probability that the innovations have spread non-simultaneously from different areal centres is too high for that. It seems that to postulate a 'South Sunda' language group is out of the question at this moment.

To all appearances, the following is of considerable interest for conceiving the relationships within the AN family. Namely, most of the aforementioned innovations of the South Sunda area seem to be shared by the languages of Oceania. Consequently, supposing that the observed resemblances, in fact, represent South Sunda-Oceanic **common innovations** (not convergences or coincidences), these innovations must be of rather deep antiquity, no less than several millennia. As to the concrete correspondences, at least one of them - South Sunda *-akən ∿ proto-OC *-aki(ni) - can hardly be explained in any other way than by common origin. So we have some clue for introducing *-akən and - through it, but much more hypothetically - also the radical change of the scheme of voice forms into the relative chronology of the AN historic grammar.

In view of the conservative nature of the Philippine area and considerable lexical differences found in some regions (especially in Taiwan; also in Minahasa, etc.) it seems unlikely that all the languages of this area belong to one group. If we also take into account the

seeming South Sunda-Oceanic common innovations, we must logically come to a conclusion rather different from nearly all the schemes of subdivision of the AN family which have been suggested so far.[19] It appears that the deepest split (or possibly a number of first-range splits of unclear internal hierarchy) passes through the languages of the Philippine area, while a group of those languages is (or several groups are) opposed to all other AN languages. In search of possible immediate subgroups of AN, undoubtedly the languages of Taiwan must be considered first (the Atayalic group being a particularly probable candidate), but also various other languages (e.g. those of Minahasa) may be interesting in this respect.

It is not to be forgotten that the processes which have brought the present-day linguistic map of the AN family into existence were of an intricate nature. Areal contacts of most various scope (not only those conventionally named 'local') and of various depth have played an important and sometimes perhaps decisive role in this development. In many cases contacts have obliged ethnic groups to change their linguistic allegiance. Ethnic groups which acted thus were different in their cultural heritage, linguistic prehistory and, consequently, language/speech habits. As a result, old splits may have been strongly if not completely disguised by later contacts. The same value of the supposition just made is therefore open to question.

If we, nevertheless, decide to give preference to the supposition localising the deepest split (or splits) of the AN family in the Philippine area, we must assume that the languages of Oceania and the South Sunda area, together with many other languages, belong to the same first-order subgroup of the family.

However, it is not likely that the languages of Oceania are equally related to all the languages of the South Sunda area. Within the last-mentioned area, the Sulawesi region exhibits a higher degree of similarity with Oceania, than do more western regions (lexical correspondences between the languages of Sulawesi and those of Oceania have been noted more than once; in addition, I point out the object-related enclitic pronominal morphemes). In the light of this, it is reasonable to suppose that the suffix *-akən originates in the region of Sulawesi. Provided the Toraja -ka indeed reflects *-kən, a similar claim can be made regarding *-kən (and then, on account of the probable merger of *-akən and *-kən practically everywhere except the Toraja languages, the supposition made concerning the 'homeland' of *-akən receives additional support).

Loss of the past tense forms with *-in-/ni- can also be hypothetically localised in the Sulawesi region. The main evidence is as follows: on

one hand, the languages of Minahasa differ in using either ni- or -in-/n-; on the other hand, Toraja languages differ in having generalised either m- or n- as the initial sound of verbal suffixes (note that m- may, besides continuing an earlier initial m-, also reflect a segment 'ni- + initial m-'; n- reflects 'm-in-').

In view of the probable location of the dispersion centre of some important South Sunda-Oceanic isoglosses in the Sulawesi region we obviously have the right to hypothesise that this region has been the starting-point of a number of migrations which carried Austronesians both to Western Indonesia and to regions lying in the east (the Moluccas, Oceania, etc.). It also seems natural to think that languages ancestral to those of the present-day Philippine area were located somewhere northward of the said centre: possibly in Northern Sulawesi, and/or in the Philippines, and/or in the northern and eastern parts of Kalimantan.

The linguistic evidence that I can cite in support of this prehistoric 'language map' is, at least so far, rather weak. Perhaps it would not at all have been worth-while to speak about this 'map' if it had not remarkably conformed with the hypothesis of the AN homeland which has been suggested by W.G. Solheim (1976) on the basis of an analysis of non-linguistic data.[20]

N O T E S

1. During the last few years, one more term - 'Hesperonesian' (originally coined by I. Dyen for designating a unit of his lexicostatistical classification) - has sometimes been used in the sense of Western Austronesian or Indonesian.

2. However, it is quite possible that the perspectives of such reconstruction will improve due to the progress in Papuan linguistics.

3. The number of items of this list, which are present in individual languages of Melanesia, varies on a large scale.

4. This problem is discussed by A. Pawley (1973:170). The concept of Western AN assumed here differs somewhat from that used by Pawley (Pawley: AN minus OC; here: AN minus Eastern AN), but this difference can hardly be important in the given context. I would like to add: if somebody shows that absence of labiovelars in a part of the AN languages is due to a loss, it will be possible to regard this loss as an innovation of Western AN. Innovations of such kind are, however, almost useless for language subgrouping.

5. It may be that some of the less-known languages of Taiwan do not belong to this area (cf. P. Jen-kuei Li, in this volume). However, there is no doubt that, according to the criteria used here, at least Paiwan, Saaroa, Kanakanabu, Tsou and Atayal do belong.

6. The terms 'active', 'passive' and 'voice' are used here in their traditional sense. In the given context, the theoretical problem of applicability of these terms to AN languages is not of interest. What is relevant for our discussion is the clear-cut formal opposition between the two sets of forms labelled here as active and passive.

7. True, the full set of passive forms permitted by the affixal inventory is not produced from all verb stems. Nevertheless, the paradigm is asymmetric on the level of invariants (cf. Shkarban 1974: 226,fn.).

8. The initial n- of the nag-type prefixes of numerous Philippine languages is a relic of the infix -in- (Brandes 1903). Taking into account the intermediate stage m-i-a-, which is observable in Magindanao and Maranao (and also some parallel developments in Old Balinese), we can also regard *m-in-a- as the source of the Kapampangan past tense prefix me-.

9. The specific passive formation called by Dutch authors 'verbrede passief' (Kern 1934:76-91) can give rise to asymmetric paradigms in some languages. For example, in Toba (from the verb stem -takko-i):

a. the simple passive: di-takko-i *'to be stolen (said about a great quantity of something)'*;

b. the 'verbrede passief': di-pa-nakko-i *'to suffer losses as a result of a theft'*;

c. the active form: ma-nakko-i.

However, in contrast to the paradigms of the Philippine area languages, such asymmetric paradigms meet the requirements of separate expression of the verb-object and verb-subject relations.

10. Nevertheless, it seems that several groups of languages of the South Sunda area reflect none of these etyma.

11. I am not quite sure regarding the source of the Toraja -ka: this suffix can also be compared with the Iban preposition ka, which expresses comparable meanings in certain constructions.

12. The Javanese suffix -ake may have originated from the combination of *-akən with some enclitic particle, e.g. -e; then *-akən-e > *-ake-ne (through assimilation and metanalysis). The Madurese -aghi may well be due to similar processes. The idea of metanalysis in development of verbal suffixes finds support in the Balinese and Sasak suffix -in, which originates from another etymon, most probably -i. In all likelihood, this -in has arisen from metanalysis of the following sequence: 'verb suffixed by -i + ligative nasal n- + the word linked by it to the verb'. Ligative nasal functions in a great number of AN languages.

For Balinese, compare the allomorphs -e and -ne of an article-like enclitic. At the beginning -in was, of course, a positional allomorph of -i, and was generalised later.

13. A piece of evidence for this supposition is offered by Mori. In this language, the suffix -ako(n-) is linked to the verb root by means of a ligative (in many cases etymological) consonant when the meaning is factitive, and without any ligative consonant when the suffix expresses meanings, similar to those of the 'instrumental passive'. This picture can be explained by supposition that in the second case -ako(n-) has been substituted for a suffix which had begun from a consonant - most probably the old suffix was a reflex of *-kən (expected Mori reflex: -ko).

14. There may be exceptions in poorly described languages (e.g. Muna).

15. This list requires some qualification.

It is natural that many if not most of those lexical items which Dyen labels as apparent exclusive Malagasy-Maanyan correspondences are also shared by other South-East Barito languages/dialects and by Paku, which according to Dahl (1977:132-3) does not belong to that group, but has borrowed very much from it. Judging from Hudson (1967), the Maanyan-Malagasy correspondence represented by Maanyan walenon *'ashes'* has cognates in Dusun Deyah and Samihim. Analogically: Maanyan ŋg-anro *'cook'* - cf. Samihim, Paku; Maanyan mamo *'mosquito'*, lawit *'distant'* - cf. Samihim, Dusun Witu; Maanyan wintan *'hook'* - cf. Dusun Witu, Paku. Of course, such cognates do not invalidate the data presented by Dyen, but it becomes obvious that the question stands about Malagasy-South East Barito rather than about Malagasy-Maanyan unity.

It must be noted that in some cases, what is regarded as (apparent) exclusive Malagasy-Maanyan correspondence is not the etymon, but some peculiarity of development which can hardly be sufficiently unusual to minimise the probability of convergence. Dyen is clearly mistaken at least in two cases:

a. Maanyan ondan and Merina ondana *'pillow'* are due to a development which is also exemplified by Cebu (on this development see: Aichele 1956:221; on the underlying root see Blust 1970: note 162).

b. Maanyan a-limoken *'pigeon'* and its Malagasy cognates (Merina domohina, Old Betsimisaraka demohin *'pigeon'*, Sakalava limohe *'turtle-dove'*) can obviously be reduced to *limbukən (Karo limbukĕn; cf. Malay limbuk). In Maanyan, the change -mb- > -m- may even turn out to be

regular in this position (see examples in Sundermann 1913:204).

It seems that the value of some items as subgrouping indicators is subject to reasonable doubts. Such items are the following:

a. Maanyan, Tandroy mamo *'mosquito'* (from *ñamuk through assimilation?).

b. Maanyan kariwe, Merina hariva *'afternoon, evening'* - the underlying etymon, as it seems, might have contained a metathesised reflex of *Rabi (*'evening'*). Compare also with the South Sulawesi languages (Buginese arawíng, Sadanese karoen, etc., all meaning *'evening'*).

c. Maanyan butoh, Merina boto *'youngster'* - compare with Malay, etc., butuh *'penis'*, and note the similar usage in the South Sulawesi languages, e.g. Sadanese: laso *'penis'*; laso?, aso?, etc. *'terms of address to young boys'*.

If the irrelevant and suspect items are excluded from the list, even then there remain more than 25 comparisons whose evidence for the Malagasy-South-East Barito unity I cannot question at present. However, it is not to be forgotten that the vocabulary of almost all the languages of Kalimantan - with only a few exceptions - is very poorly known.

16. O. Dempwolff (1938:31) assigns vintana to *bintaŋ and kintana to *intaŋ; these etyma are labelled as Nebenformen. About Malagasy kintana see also Dahl 1976:218.

17. The scheme of affixation given in the Headlands' dictionary (p. xxxix) requires a critical approach.

18. These forms/constructions deserve a much more detailed discussion than I am able to go into in this paper.

19. Our scheme has a considerable degree of parallelism with the tentative subgrouping scheme which was recently suggested by R.A. Blust (1977:2, preliminary publication) on the basis of entirely different evidence (personal pronouns).

20. W.G. Solheim hypothesises that Austronesians have evolved in the 'general area' of Eastern Indonesia and Southern Philippines (1976:152).

BIBLIOGRAPHY

AICHELE, W.

1956 'Beiträge zur indonesischen Sprachgeschichte'. *Oriens Extremus* 3:215-33.

BLUST, R.A.

1970 'Proto-Austronesian Addenda'. *OL* 9/2:104-62.

1974 'Eastern Austronesian: a Note'. *WPLUH* 6/4:101-7.

1977 'The Proto-Austronesian Pronouns and Austronesian Subgrouping: A Preliminary Report'. *WPLUH* 9/2:1-15.

1978 'Eastern Malayo-Polynesian: A Subgrouping Argument'. See pp.181-234 in this volume.

BRANDES, J.

1903 Het infix in ... *Album-Kern*, 199-204. Leiden: Brill.

COWAN, H.K.J.

1965 'On Melanesian and the Origin of Austronesian'. *CAnthr* 6/2:217-20.

DAHL, O.C.

1951 *Malgache et maanjan: une comparaison linguistique.* Oslo: Egede-Instituttet.

1973 *Proto-Austronesian.* Scandinavian Institute of Asian Studies Monograph Series No.15. Lund: Studentlitteratur.

1976 'Semantics in Lexicostatistics and the Problem of Borrowing'. *Norsk Tidsskrift for Sprogvidenskap* 30:203-33.

1977 'La subdivision de la famille Barito et la place du malgache'. *Acta Orientalia* 38:77-134.

DEMPWOLFF, O.

1938 *Vergleichende Lautlehre des Austronesischen Wortschatzes.* Band 3: *Austronesisches Wörterverzeichnis.* ZES Beiheft 19.

DYEN, I.

1965 *A Lexicostatistical Classification of the Austronesian Languages.* IUPAL 19. IJAL Memoir 19.

1971 'Malagasy'. CTL 8:211-39.

ESSER, S.J.

1938 'Talen [a linguistic map of the former Dutch East Indies with notes]'. In: *Atlas van Tropisch Nederland.* Amsterdam: Koninklijk Nederlandsch Aardrikskundig Genootschap.

GRACE, G.W.

1962 Comment on: 'Oceanic Linguistics Today', by A. Capell. CAnthr, 3/4:408-10.

1968 'Classification of the Languages of the Pacific'. In: A.P. Vayda, ed. *Peoples and Cultures of the Pacific: an Anthropological Reader,* 63-79. Garden City.

1971 'Languages of the New Hebrides and Solomon Islands'. CTL 8:341-58.

HEADLAND, T.N. and J.D. HEADLAND

1974 *A Dumagat (Casiguran) - English Dictionary.* PL, C-28.

HUDSON, A.B.

1967 *The Barito Isolects of Borneo.* Ithaca, New York: Cornell University.

J[ONKER], J.C.G.
1918 'Indonesische taalstam'. In: *Encyclopaedie van Nederlandsch-Indië* (2nd edition), vol.2:143-7. The Hague: Nijhoff/Leiden: Brill.

KERN, R.A.
1935 'De partikel pa in de Indonesische talen'. *BijdrTLV* 92: 5-121.

LI, P(aul) Jen-kuei
197 'The Case-Marking Systems of the Four Less Known Formosan Languages'. See pp.569-615 in this volume.

PAWLEY, A.K.
1973 'Some Problems in Proto-Oceanic Grammar'. *OL* 12:103-88.

SHKARBAN, L.I.
1974 'Problemy izučenija zalogov v tagal'skom jazyke'. In: *Jazyki Kitaja i Jugo-Vostocnoj Azii: Voprosy Grammaticeskogo Stroja*, 224-44. Moscow: Nauka.

SOLHEIM, W.G.
1976 'Reflections on the New Data of Southeast Asian Prehistory: Austronesian Origin and Consequence'. *Asian Perspectives* 18/2:146-60.

SUNDERMANN, H.
1913 'Der Dialekt der "Olon Maanjan" (Dajak) in Süd-Ost-Borneo'. *BijdrTLV* 67:203-36.

WOLFF, J.U.
1973 'Verbal Inflection in Proto-Austronesian'. *Parangal kay Cecilio Lopez*, 71-91. Quezon City.

Sirk, U. "Problems of High-level Subgrouping in Austronesian". In Wurm, S.A. and Carrington, L. editors, *Second International Conference on Austronesian Linguistics: Proceedings.*
C-61:255-273. Pacific Linguistics, The Australian National University, 1978. DOI:*10.15144/PL-C61.255*
©1978 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.