R. Blust

The reconstruction of proto-Malayo-Javanic: an appreciation

In: *Bijdragen tot de Taal-, Land- en Volkenkunde* 137 (1981), no: 4, Leiden, 456-469

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

ROBERT BLUST

THE RECONSTRUCTION OF PROTO-MALAYO-JAVANIC: AN APPRECIATION

In 1965 Isidore Dyen published a lexicostatistical classification of some 245 Austronesian languages — the most ambitious undertaking of its kind ever attempted. The purpose of this classification was to specify, in accordance with an implicit family tree model of linguistic differentiation, the hierarchy of pairings between all languages and language groups in the Austronesian family. By definition the structure of the resulting dendrogram was taken to mirror the history of language splits from the break-up of the most remotely reconstructible ancestor (Proto-Austronesian) to the attested languages. The problem confronted was thus the traditional problem of subgrouping in comparative linguistics. The method employed, however, was one only recently developed by Swadesh (1950, 1955), and this study inevitably became an important test not only of the correctness of received views on Austronesian subgrouping, but of the reliability of lexicostatistics itself.

Dyen's results were strikingly at variance with the traditional view that Austronesian speakers originated on the Asian mainland and gradually expanded through island southeast Asia into the farthest reaches of the Pacific. To account for the reported distribution of lexicostatistical percentages — where more than 30 *prima facie* primary branches of the Austronesian family were represented in Melanesia — Dyen concluded instead that the Austronesian expansion had begun in the area of New Guinea and the Bismarck archipelago. Some Austronesian speakers remained in this area over the succeeding millennia, he maintained, while one descendant group (Proto-Malayo-Polynesian) set out on great voyages of discovery both to east and to west.

Almost immediately (Grace 1966) it was recognized that Dyen's conclusions not only clashed with what had long been supposed on the basis of physical and cultural evidence, but also clashed with linguistic evidence of a more traditional and better-established kind. The result has been that over the past decade and a half researchers in the field

of Oceanic linguistics almost without exception have rejected Dyen's interpretation in favor of the view that all of the languages east of 136° E longitude fall into a single subgroup of Austronesian, now generally called 'Oceanic'.

While disagreement with higher levels of Dyen's classification was sounded quickly, and continues to be sounded (Blust 1978), many of the lower-level subgroups have remained unchallenged. In view of the number of Austronesian languages in relation to the number of scholars engaged in research on them this neglect is understandable — it will simply take time to critically evaluate each of the hundreds of nodes in Dyen's lexicostatistically-generated family tree.

One of Dyen's lower-level subgroups which should repay prompt critical evaluation because it includes such relatively well-known languages as Malay and Javanese, is the 'Javo-Sumatra Hesion'. In a revised version of a PhD dissertation completed at Yale University under Dyen in 1973, Nothofer (1975) has rechristened the ancestor of this putative subgroup 'Proto-Malayo-Javanic', and attempted to reconstruct its system of phonological contrasts. According to its author (V) the work in question (hereafter RPMJ) "concerns the reconstruction of the phonemes of Proto-Malayo-Javanic, the last proto-language which Sundanese, Javanese, Malay, and Madurese directly continue".

Part 1 presents an internal subgrouping of the four languages under investigation, together with sketches of their synchronic phonologies (including both Old and Modern Javanese). Part 2 is concerned with the reconstruction of the PMJ phonological system. Among his major conclusions Nothofer maintains 1) that the long-puzzling distinction of a mid-central and a high-central vowel in Sundanese, both deriving from Dempwolff's *ə, can be explained through borrowing (of words containing ə) from Javanese, and 2) that the PMJ phoneme inventory included several segments not previously reconstructed for proto-languages of a higher order.

Judged without reference to its subgrouping assumptions, RPMJ can be considered a competently executed, generally successful study. The descriptive phonologies of attested languages, though brief, are for the most part well thought out and clearly illustrated. In a few cases questions arise, as with the claim (24, 127) that the semivowel in e.g. Madurese /buwá/ 'fruit' is phonemic, whereas the similar segment in the traditional Sundanese orthography is predictable (thus Sund. /buah/ 'fruit'). Perhaps the most questionable of Nothofer's synchronic phonological interpretations involves the functional relationship of the

Javanese phones [ʔ] (written q) — which in non-interjections occurs only in final position after vowels other than /ə/ — and [k] — which occurs elsewhere. Given the fact that these phonetic elements are in complementary distribution, and were accordingly united as /k/ in Nothofer's primary source (Pigeaud 1938), it is difficult to see the justification of his claim (15) that they should be distinguished in order "to bring the material into a phonemic writing".¹

The discussion of methodological principles governing the reconstruction (36-9) provides a clear set of guidelines for decisions later reached in the treatment of the comparative material. Unlike some scholars concerned with the reconstruction of lower-order proto-languages in the Austronesian family — and much to his credit — Nothofer sees the need to reconstruct simultaneously 'from the bottom up' (using two or more internal witnesses) and 'from the top down' (using one internal witness and at least one external witness). Known historical conditions, and linguistic criteria useful for determining the likelihood and probable direction of proto-historic or prehistoric borrowing, are precisely stated and applied as constraints on the reconstruction of PMJ morphemes.

One would not expect to find significant oversights in a work that displays such an evident concern for method. Unfortunately, however, these occur. Nothofer commences his investigation with the acknowledgement that its scope (in terms of languages included) was defined lexicostatistically by Dyen (1965: 26). In view of the circumstances under which the original manuscript was written it is perhaps understandable that Nothofer would have been reluctant to challenge this basic assumption. Moreover, by the time the manuscript was revised for publication the commitment to a subgrouping theory had so determined its content that any change in basic assumptions would have necessitated considerable rewriting. Various observations nonetheless suggest that the subgrouping theory which defines the scope of Nothofer's inquiry and hence the form of his reconstruction, could profitably be reconsidered.

Under the heading 'Javo-Sumatra Hesion' Dyen included three major groups: 1) the Malayic Hesion, 2) Sundanese and 3) Javanese; 1) was further divided into 1.1 The Malayan Subfamily (Malay, Minangkabau, Kerinci), 1.2 Madurese, 1.3 Achinese and 1.4 The Lampungic Subfamily (Lampung, Kroë). The first thing likely to strike one familiar with the wider linguistic situation in western Indonesia and mainland southeast Asia is that Dyen's Javo-Sumatra Hesion appears to contain

both too little and too much. Some omissions (as of Iban) in this classification may be due to the sheer obviousness of the subgrouping relationship of the omitted language to one or more of those included, but without explicit indications it is often difficult to know whether a language is not mentioned because it is implicitly included, or because it is excluded. The objection that a linguistic reconstruction does not take account of material in all languages which it is implicitly posited to explain can, of course, be a trivial one. Dempwolff reconstructed the sound system of Proto-Austronesian on the basis of correspondences holding between only 11 of the many hundreds of attested languages and, with the exception of certain features which did not become fully apparent until the Austronesian languages of Formosa were belatedly incorporated into the comparative picture, he appears to have been remarkably successful. But the 11 languages which figured in Dempwolff's reconstruction of Proto-Austronesian were carefully selected with a view to economy of exposition from among many more that had previously been investigated. This seems not to have been the case with Nothofer's choice of languages used in the reconstruction of Proto-Malayo-Javanic.

In addition to a number of apparently unconditioned phonemic splits, Nothofer recognizes five phonemic mergers in the transition from Proto-Austronesian to Proto-Malayo-Javanic. In the first of these it is maintained (82) that "PAN *w* and PAN *b* between **a*'s merge in PMJ". But Iban (Scott 1956), a language whose subgrouping relation to Malay is evident on inspection (Cense and Uhlenbeck 1958) clearly preserves this contrast: 1) **aba* > *aba*: **cabanj* > *cabanj* 'branch of a tree', **kaban* > *kaban* 'group, company, shoal, herd', **laban* > *laban* 'opposed to, against', **taban* 'what is captured, booty' > *taban* 'carry off, elope, carry (with it)', **tabaR* > *tabar* 'tasteless, insipid, unsalted, unsweetened'; 2) **awa* > *awa*: **awaj* 'space between earth and sky, atmosphere' > *awan* 'air', **zawa* > *jawa* 'millet', **ñawa* 'soul, breath' > *ñawa* 'life, existence', **sawa* > *sawa* 'python', **tawaR* > *tawar* 'cure, antidote', etc. In the second phonemic coalescence described for PMJ (104) "PAN *ɖ* and PAN *r* in preconsonantal position and final position and PAN *R*₃ and *R*₄ merge in PMJ *r*".² Maloh of southwest Borneo — a language which Hudson (1970) places in an immediate subgrouping relationship with Malay and Iban — seems, however, to retain the **D/r* distinction in at least **SanteD* > *antat* 'accompany, escort', but **tutur* > *tutul* 'tell, relate'. Although Hudson's argument for including Maloh in a Malayic group is assumed without justification

(Hudson presents evidence only for distinguishing Maloh and its close relatives from 'Malayic Dayak' and 'Ibanic' languages), thoroughness demands at least some discussion of the claim and its possible relevance to the reconstruction.³

If Nothofer's reconstruction is distorted by the non-inclusion of diagnostic witnesses for certain PAN distinctions which reportedly were lost in Proto-Malayo-Javanic, there is reason to believe that it is further distorted by the inclusion of at least one language which in important respects appears to be extraneous to the comparison. For the six possible pairings of Sundanese, Javanese, Malay and Madurese Nothofer (4) reports the following cognate percentages, corrected from earlier figures by the elimination of determinable⁴ intra-Malayo-Javanic loans: SUND-JAV: 33, SUND-MAL: 37, SUND-MAD: 36, JAV-MAL: 33, JAV-MAD: 37, MAL-MAD: 47. As can be seen, the figure for Malay-Madurese exceeds all others by a full ten percent, and the lowest percentages are shared by Javanese both with Sundanese and with Malay. The difficulty with taking any of these percentages at face value is that many loanwords leave no tell-tale irregularities in phonological development, and hence are undetectable on formal grounds. In some cases a single sound change provides the touchstone for distinguishing direct from indirect inheritance in a significant number of morphemes. This is the case, for example, with the change of PAN *e (apparently a mid-central vowel) to a high-central vowel in Sundanese: for a number of items on the Swadesh 200-word list Sundanese shares a *prima facie* lexical innovation exclusively with Javanese, but nine of these items contain a mid-central vowel in both languages, thus pointing to borrowing from Javanese as the most likely explanation for the agreement. In other cases, however, a sound change can have the opposite effect. Thus, if two languages A and B each independently undergo a phonological change which affects a large number of lexical items that undergo no other alteration, a criterion which may be important for distinguishing loans between either of these languages and a third language C can become inoperative for A-B comparisons, thereby inflating the percentages for A-B comparisons relative to others (A-C, B-C, etc.). This may well be the case in Javanese-Malay, Javanese-Madurese and Malay-Madurese comparisons, as all three languages have lowered original high vowels (though not necessarily with phonological restructuring) in certain closed final syllables. Whether Javanese, Malay *jantong* 'heart', for example, is a lexical innovation shared by these languages with a few close relatives, or is product of borrowing is not

determinable from the evidence of phonological irregularities, whereas a determination would be possible if either language did not regularly lower *u in the stated environment. In view of the probability that many loans from Javanese or Malay remain unrecognized, then, it is remarkable that Javanese should share such relatively low percentages with languages that have long been in intimate contact with it.

The foregoing point is worthy of at least some elaboration. As Nothofer points out (36-9), both Malay and Javanese have been the official languages of Indianized states which for centuries exercised political control over populations that spoke some other vernacular. An important result of the periodic alternation of Malay-speaking centers of influence in Sumatra or Malaya with Javanese-speaking centers of influence in Java has been that Malay and Javanese have borrowed from one another over a time-span of more than a millennium. Sundanese, by contrast, appears to have been effectively isolated from Malay (though not from Javanese) influences until the nineteenth century, and Madurese, while it may have borrowed heavily both from Javanese and from Malay, probably contributed little in return. As Madurese scores high with all other languages in Nothofer's sample it may well have had a lower-than-average lexical replacement rate, thereby scoring higher with Sundanese than would normally be expected despite the absence or low incidence of borrowing between these languages. Apart from this complication the Sundanese-Malay (37) and Sundanese-Madurese (36) percentages presumably are little disturbed by secondary contact (except for recent, detectible, borrowing from bahasa Indonesia in the former case). For all other language pairs, however, the undistorted cognate percentages undoubtedly are lower than the percentages cited, although the degree of distortion in particular cases cannot presently be determined. In disagreement with the internal subgrouping assumptions which Nothofer adopts, there thus appears to be some lexicostatistical support for a grouping that includes Sundanese and Malay but not Javanese.

Much more serious for the Malayo-Javanic hypothesis is the existence of qualitative evidence which suggests that not only Sundanese, but also certain languages of south-central Sumatra, Iban, Maloh and other languages of southwest Borneo, Achinese and apparently the Chamic languages of mainland southeast Asia are all genetically closer to Malay than are Javanese, Madurese or Lampung. A consideration of the PAN numerals 7-9 and their replacements in these languages suffices to illustrate:

TABLE 1

Replacements of the PAN numerals 7-9 in languages of western Indonesia and mainland southeast Asia ⁵

PAN	pitu '7'	walu '8'	siwa '9'
1) Malay	tujoh	delapan	sembilan
Minangkabau	tujueh	salapan	sambilan
Kerinci	tujeuh	salapan	sambileŋ
Middle-Malay ⁶	tujoh	delapan	sembilan
2) Selako	tujuh	lapan	sembilan
Iban	tujoh	lapan	semilan
3) Sundanese	tujuh	dalapan	salapan
Maloh	tuju	lapan	samilan
Rejang	tujua	delapan	semilan
4) Achinese	tujōh	lapan	(sikureuëŋ)
Cham	tijuh	dalapan, salapan	samilan
Jarai	təjuh	səpan	dua rəpan

The data in Table 1 serve to demonstrate that Malay has innovated roots for the PAN numerals 7-9, and that similar innovations appear in various other languages of western Indonesia and the Asian mainland. It is noteworthy that neither Javanese (*pitu*, *wolu*, *saŋa*), Madurese (*pittu*, *ballu*, *saŋa*) nor Lampung (*pitu*, *walu*, *siwa*) shows evidence of similar innovations. One might argue that numerals are readily borrowed, particularly where trade plays an important role in the relations between linguistic communities, but if this is the explanation for these striking agreements it is surprising that Javanese and Madurese — both of which were subject to strong and protracted Malay influence in a mercantile context — did not also borrow the Malay numerals.⁷ Moreover, a hypothesis of borrowing fails to account for the fact that 1) the Malay innovations in the numerals 'one' (*satu*) and 'three' (*tiga*) do not also appear in Sundanese, Maloh, Rejang, Achinese or the Chamic languages, 2) Malay has *sembilan* 'nine', but the cognate forms in Iban, Maloh, Rejang and Cham have a simple medial nasal, and 3) Jarai is and probably has been for well over a thousand years a language of the Vietnamese highlands, and hence has been relatively inaccessible to maritime contacts.

Perhaps the most important phonological consequence of discarding Javanese from a comparison involving Malay and its closest relatives is the elimination of a portion of the evidence for the PMJ *b/B distinction. Dempwolff (1934-8) reconstructed the PAN source of PMJ *b/B as *b, but noted that the Javanese reflexes are unpredictably *b* or *w*, with occasional *b/w* doublets. In essential agreement with the early claim of Kiliaan (1897), Nothofer (123-45) points out that Madurese also exhibits double reflexes of Dempwolff's *b, and that these enter into regular correspondence classes with Javanese *b*, *w* from this source. Since there is purely internal support for a PMJ *b/B distinction (neutralized in final position), the exclusion of Javanese alone thus does not remove the problem of multiple reflexes of Dempwolff's *b to some other comparative context. Yet, as already seen, despite its high lexicostatistical percentage with Malay, Madurese — like Javanese — evidently did not participate in the innovations which affected the numeral system of Malay and other languages. We cannot, therefore, dismiss the possibility that the lexicostatistical evidence for a Malay-Madurese subgroup is due to the combination of an unusually high retention rate in Madurese with heavy and prolonged borrowing from Malay. But if so Madurese can also be excluded from the comparison, and with it all remaining evidence of a *b/B distinction and of geminate consonants (posited only to account for Madurese reflexes) in a language immediately ancestral to Malay and its closest relatives.

The tentative subgroup that results (Table 1) is substantially different from Dyen's 'Javo-Sumatra Heslon', of which Nothofer's 'Malayo-Javanic' is purportedly an essential representative. Using Malay as a point of reference, this group — to which the land Dayak languages of southwest Borneo evidently must be added under 3) ⁸ — can conveniently be called 'Malayic'. A tentative internal subgrouping of the Malayic languages is possible on the basis of certain noteworthy phonological, semantic and lexical innovations in basic vocabulary. Thus, with the exception of group 4) all languages in Table 1 show metathesis of the vowels in reflexes of PAN *ma-quDip 'living, alive'. In some cases there are variants with and without metathesis, as where Hudson (1970) cites such Land Dayak forms as Bekati' *idup*, Semandang *midup*, whereas Ray (1913) recorded Quop, Sentah, Beta *udip*, Sadong *mudip*. Since Sundanese also exhibits reflexes with and without metathesis (*hirup* 'living, alive', *hurip* 'revive; flourish, thrive (as vegetation)') it is perhaps simplest to assume that *hidup was innovated in the immediate ancestor of 1) - 3) (Table 1), but coexisted with *hudip

until the unmetathesized variant disappeared in the immediate ancestor of 1).⁹

The evidence for the position of Selako, Iban and Sundanese is contradictory. On the one hand Selako and Iban have cognates of Malay *satu*, *tiga*, replacing **esa/isa* 'one', **telu* 'three'. Sundanese, on the other hand — with *hiji* 'one' (apparently an irregularly altered borrowing of Javanese *siji*), *tilu* 'three' — has a cognate of Malay *anjin* 'dog', whereas Selako, Iban and similar languages of southwest Borneo retain a reflex of PAN **Wasu*. What emerges clearly from these observations is that the closest relatives of standard Malay which might conceivably be treated as distinct languages are Minangkabau, Kerinci and Middle-Malay (Běšmah, Sěraway). Under the next node in a family-tree diagram certain criteria support the inclusion of Selako and Iban apart from Sundanese, but others support the inclusion of Sundanese apart from Selako and Iban. This conflict of evidence suggests that the differentiation of Malayic languages has not always followed the abrupt cleavages implied by the family-tree model. As Malay appears to share a great deal more cognate vocabulary (both basic and non-basic) with Selako and Iban than with Sundanese, it seems likely that these languages once belonged to a larger dialect chain in which Malay intervened between Sundanese and the Bornean communities with which it was generally in closer contact.

Apart from distortions that were almost certainly introduced by the exclusion of diagnostic witnesses and the inclusion of extraneous languages, other distortions in Nothofer's reconstruction may have resulted from insufficient attention to available material for certain languages explicitly included in Dyen's 'Javo-Sumatra Hesion'. I have already mentioned the probability that *-D became *t*, but that **r* disappeared in Achinese. A subtler problem involves the relative chronology of changes. Like other Austronesianists before him, Nothofer views the phonological history of a language as an unordered inventory of reflex-statements, a conception of historical phonology dominated by what might be called 'the fiction of simultaneity'. But, ideas to the contrary notwithstanding, even the most meticulously detailed descriptions of reflexes made without attention to ordering relations are apt to miss important features of historical development.

From a specialist's point of view, one of the most interesting results of Nothofer's reconstruction of PMJ lexical items is the discovery that some attested roots with a simple intervocalic consonant in Malay and many other languages of western Indonesia derive from inter-

mediate prototypes which contained a previously unsuspected heterorganic consonant cluster, preserved in Sundanese (e.g. *ruksak > Sund. *ruksak*, Jav. *rusak*, Mal. *rosak* 'spoiling, ruining'). Nothofer maintains (46) that "In PMJ dissyllabic bases there is evidence for only one cluster which is not a nasal cluster or a double consonant, namely the cluster *-ks-".¹⁰ Comparison of Malay with Philippine languages in fact initially suggests that Proto-Malayic (PM) contained other clusters of this type. Thus Malay *berkas* 'colligation, binding together (with rope or rattan); a bale so bound' next to Tagalog *bigkis* 'abdominal band, girdle; a bundle, packed by tying together' < *beRkes (Dempwolff 1924-5) would appear to point to an early cluster *-Rk- which was retained in Proto-Malayic. More detailed analysis, however, raises the question whether some heterorganic clusters might not have developed in the separate histories of Malay and Sundanese.

It is well known that the heterorganic consonant clusters of Philippine languages typically result from the syncopation of a shwa (PAN *e) in the environment VC — CV. Given the general constraint against heterorganic consonant clusters in Dempwolff's (1934-8) reconstruction of Proto-Austronesian, consonant sequences like that in Sundanese *ruksak* presumably have a similar history. On this assumption explicit derivations of various Malay forms which appear to have once had a medial cluster will work only if SYNCOPE is ordered after the change of Proto-Malayic prepenultimate *a to Malay shwa. Consider the comparisons 1) Cebuano Bisayan *dagpak* 'slap hard enough to make a noise', Malay *depak* 'a smack', 2) CB *lagtub* 'inflamed, blistered', MI. *letup* 'blister on skin', 3) CB *saylag* 'roast something in a pan with little or no oil', MI. *selar* 'branding' (= Minangkabau *salar*, *saylar* 'broiling, cooking at an open fire'), 4) CB *taklap* 'spread, lay something flat over something so as to cover it', MI. *tekap* 'covering with a flat surface'. In each of these comparisons Malay exhibits penultimate *e* (= shwa) corresponding to *a* in Cebuano Bisayan. Since this correspondence can be regarded as regular only in prepenultimate position, it appears simplest to assume that at the time prepenultimate vocalic oppositions were neutralized as shwa in Malay (1) - (4) were trisyllabic:

(1) PM *dagepak

1	degepak	prepenultimate neutralization
2	degpak	syncope
3	depak	cluster reduction

Malay *depak*

Note that a general neutralization of vocalic oppositions before heterorganic consonant clusters apparently did not occur, as the first vowel in Malay reduplicated monosyllables remained unaffected (e.g. *kaskas > *ML. kakas* 'scratch, scrape'). But if 1 must be ordered before 2 to produce the observed results in (1) - (4), and if PREPENULTIMATE NEUTRALIZATION postdated the break-up of Proto-Malayic (since it did not occur in Sundanese), it follows that the attested heterorganic clusters of Sundanese and the (largely) inferential heterorganic clusters of early Malay are parallel developments in the separate histories of these languages.

Surprisingly enough, other derivations will work only if it is assumed that SYNCOPE preceded PREPENULTIMATE NEUTRALIZATION in Malay, as in 5) *Paiwan qajelits* 'be scorched (rice or millet)', *Isneg aylit* 'to smell of burning hair', but Malay (*h*)*ajit* (with /a/) 'foul-smelling', 6) *CB aslum*, *Kelabit laam*, but Malay (*m*)-*asam* (with /a/) 'sour', etc. These differences suggest that PREPENULTIMATE NEUTRALIZATION may have spread gradually through the lexicon of early Malay, but where the two changes came into competition for the same lexical items (Wang 1969) PREPENULTIMATE NEUTRALIZATION was arrested in mid-course by the emergence of SYNCOPE.

Although many more details could be discussed, the foregoing remarks should provide an adequate idea of the nature and qualities of this book. There are a number of (mostly minor) typographical errors, and some conspicuous gaps in the references; on the whole, however, the work was carefully proofread, and gives a thorough treatment of background literature. In summary, *RPMJ* is an important contribution to the comparative study of the languages considered. It presents a mass of comparative material much of which had not previously been subjected to systematic analysis, and in this respect alone it is a contribution of enduring value. But in Nothofer's choice to uncritically adopt a given comparative framework a golden opportunity to conduct an important qualitative test of a lexicostatistically-motivated subgrouping hypothesis obviously was missed.

University of Leiden

NOTES

- ¹ Not only does this analysis contravene Pigeaud's orthography, but it also conflicts with Uhlenbeck's (1949: 41f) carefully argued conclusion that [k] and [ʔ] in Javanese are variants of a single phoneme. Although Nothofer includes the latter work among his references, no mention is made of Uhlenbeck's conclusion in the appropriate section.

- ² I follow Dyen (1971 and earlier works) in writing *D and *r.
- ³ Whatever decision ultimately is reached regarding the position of Maloh (see King 1976, and discussion later in this review), Achinese appears also to distinguish *-D from *r: *SanteD > *antat*, (*eu*)*ntat* 'accompany, escort, convey', *Satur > *atō* 'order, regulations', *tutur 'communication' > *tutō* 'speak, talk'. D. J. Prentice (p.c.) informs me that some Malay dialects also reflect *SanteD with final /t/. He regards this fact as vitiating the value of *antat*-type reflexes for the purpose at hand, since such forms could reflect a variant with *-t, and thus shed no light on the development of *-D. However, Achinese also shows *D > /t/ in *likuD > *likōt* 'back'. The loss of *-D in *bayaD > Ach. *bayeuē* 'pay' may be due to borrowing from Malay.
- ⁴ I.e. demonstrable comparative irregularities which become intelligible on an assumption of borrowing.
- ⁵ Malay *tujoh* and related forms generally are thought to derive from the word for 'index finger', the second finger of the second hand in finger-counting. The innovations for 'eight' and 'nine' appear to have originated as subtractives, from the roots *alap and *ambil 'fetch, take' (8 = 2 taken away, 9 = 1 taken away). It is likely that *dalapan varied freely with, or was a secondary development from *dua lapan (cf. Maloh Kalis *dua lapan* 'eight', Jarai *dua rapan* 'nine'), and that *salapan, *sambilan/samilan were synonymous. The function of the original morphology presumably was lost at an early date, as reflexes of *salapan appear with the meaning 'eight' in Minangkabau, Kerinci and Cham, and the forms originally meaning 'eight' and 'nine' have been interchanged in Jarai.
- ⁶ Material is cited from the Bēsēmah dialect.
- ⁷ Trade contacts almost certainly lie behind the otherwise baffling appearance of Makasarese *salapaŋ* 'nine', since no other non-borrowed lexical innovation is known to be shared exclusively by Makasarese or other languages of South Sulawesi with the languages of Table 1. The most likely source language for Makasarese *salapaŋ* is Malay, but contemporary standard Malay lacks a cognate form. That some Malay dialects once had a reflex of *salapan, however, seems probable from material recently collected for the language of Bacan by Dik Teljeur (ms). Along with many other items indicating that the lexicon of Bacan derives at least in part from a Malay dialect (e.g. *kepala* 'head', *rambut* 'hair', *hiduŋ* 'nose', *gigi* 'tooth', *perempuan* 'woman') we find the numerals *dua lapan* 'eight', *selapan* 'nine' (but *pitu* 'seven'). The appearance of these numerals and of Makasarese *salapaŋ* can be explained plausibly on the assumption that Proto-Malayic *salapan and *sambilan/samilan persisted as synonymous competing forms in early Malay, one or the other eventually replacing its competitor in any given dialect.
- ⁸ All Land Dayak languages, together with various other languages of central, western and northern Borneo, have cognates of Malay *tujoh* 'seven'. In most Land Dayak languages, however, the numerals 'eight' and 'nine' are represented by lexical innovations of limited geographical distribution (e.g. Sadong *mahi* 'eight', *prie* 'nine'). To account for other observations to be discussed below these forms are assumed to be secondary replacements of the Proto-Malayic numerals (*walu > *dua lapan/ dalapan > *mahi*, etc.).
- ⁹ Balinese *hidup* 'living, alive' is taken to be a Malay loan, as *q- normally disappears.
- ¹⁰ In general usage among Austronesianists the expression 'nasal cluster' refers to a homorganically prenasalized obstruent, or to the rarer cluster -ŋs-. Nothofer uses it to refer to any cluster of consonants of which the first is a nasal.

REFERENCES

- Aymonier, E. and A. Cabaton
1906 *Dictionnaire čam-français*, Paris, E. Leroux.
- Blust, R.
1978 'Eastern Malayo-Polynesian: a subgrouping argument', Second International Conference on Austronesian Linguistics: Proceedings, Fascicle 1, Western Austronesian (= *Pacific Linguistics* C61): 181-234, Canberra, Australian National University.
- Cense, A. A. and E. M. Uhlenbeck
1958 *Critical survey of studies on the languages of Borneo*, KITLV, Bibliographical Series, no. 2, The Hague, Nijhoff.
- Coolsma, S.
1913 *Soendaneesch-Hollandsch woordenboek*, 2nd. ed., Leiden, A. W. Sijthoff.
- Dempwolff, O.
1924-5 'Die L-, R- und D-Laute in austronesischen Sprachen', *Zeitschrift für Eingeborenen-Sprachen* 15: 19-50, 116-38, 223-38, 273-319.
1934-8 *Vergleichende Lautlehre des austronesischen Wortschatzes*, ZfES, Supplements 15 (1934), 17 (1937), 19 (1938), Berlin, Dietrich Reimer.
- Djajadiningrat, H.
1934 *Atjèhsch-Nederlandsch woordenboek met Nederlandsch-Atjèhsch register*, Batavia.
- Dyen, I.
1965 'A lexicostatistical classification of the Austronesian languages', *International Journal of American Linguistics*, Memoir 19 (vol. 31, no. 1).
1971 'The Austronesian languages and Proto-Austronesian', in: Sebeok 1971: 5-54.
- Grace, G. W.
1966 'Austronesian lexicostatistical classification: a review article', *Oceanic Linguistics* 5: 13-31.
- Helfrich, O. L.
1904 'Bijdragen tot de kennis van het Midden Maleisch (Běšěmahsch en Sěrawajsch dialect)', *VBG* 53: 1-284.
- Hudson, A. B.
1970 'A note on Selako: Malayic Dayak and Land Dayak languages in western Borneo', *Sarawak Museum Journal* 18: 301-18.
- Kiliaan, H. N.
1897 *Madoereesche spraakkunst*, 2 vols., Batavia.
- King, V. T.
1976 'The Maloh language: a vocabulary and summary of the literature', *Sarawak Museum Journal* 24: 137-71.
- Lafont, P-B.
1968 *Lexique Jarai*, Paris, Adrien-Maisonneuve.
- Panganiban, J. V.
1966 *Talahuluganang Pilipino-Ingles*, Manila, Government Printing Office.
- Pigeaud, Th.
1938 *Javaans-Nederlands Handwoordenboek*, Groningen, J. B. Wolters.

- Ray, S. H.
1913 'The languages of Borneo', *Sarawak Museum Journal* 1.4: 1-196.
- Scott, N. C.
1956 *A dictionary of Sea Dayak*, London, University of London, School of Oriental and African Studies.
- Sebeok, T. A. (ed.)
1971 *Linguistics in Oceania*, Current trends in linguistics, vol. 8, The Hague, Mouton.
- Swadesh, M.
1950 'Salish internal relationships', *International Journal of American Linguistics* 16: 157-67.
1955 'Towards greater accuracy in lexicostatistic dating', *International Journal of American Linguistics* 21: 121-37.
- Teljeur, Dik
ms. Short wordlists from South Halmahera, Ternate, Tidore, Makian, Kayoa and Bacan.
- Toorn, J. L. v. d.
1891 *Minangkabausch-Maleisch-Nederlandsch woordenboek*, 's-Gravenhage.
- Uhlenbeck, E. M.
1949 *De structuur van het Javaanse morpheem*, Bandung, A.C. Nix and Co.
- Usman, A. H.
ms. Swadesh 200-word list for Kerinci.
- Walker, D. F.
1976 *A grammar of the Lampung language: the Pesisir dialect of Wai Lima*, Jakarta, Badan Penyelenggara Seri NUSA.
- Wang, W. S-Y.
1969 'Competing changes as a cause of residue', *Language* 45: 9-25.
- Wilkinson, R. J.
1959 *A Malay-English dictionary* (romanised), London, Macmillan.
- Wink, P.
1926 'Maleisch-Redjangsche woordenlijst, Lais', *VBG* 66.2: 111-24.
- Wolff, J. U.
1972 *A dictionary of Cebuano Visayan*, Manila, Linguistic Society of the Philippines.