



Malay Influence on Malagasy: Linguistic and Culture-Historical Implications

K. A. Adelaar

Oceanic Linguistics, Vol. 28, No. 1, A Special Issue on Western Austronesian Languages.
(Summer, 1989), pp. 1-46.

Stable URL:

<http://links.jstor.org/sici?&sici=0029-8115%28198922%2928%3A1%3C1%3AMIOMLA%3E2.0.CO%3B2-X>

Oceanic Linguistics is currently published by University of Hawai'i Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/uhp.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

MALAY INFLUENCE ON MALAGASY: LINGUISTIC AND CULTURE-HISTORICAL IMPLICATIONS

K. A. ADELAAR

RIJKSUNIVERSITEIT TE LEIDEN

INTRODUCTION. In 1951 O. C. Dahl wrote a dissertation on the relationship between Malagasy and Maanyan, a language of the Southeast Barito region in South Kalimantan. Acting on a suggestion from W. Aichele about a possible close genetic relationship between these two languages, he showed convincingly that they were indeed more closely related to each other than either is to any other language.¹ Today few Austronesianists will doubt the outcome of Dahl's dissertation, which supersedes previous theories about the origin of Malagasy. Dahl estimates the migration of the Malagasy to their present homeland to have taken place in the fifth century A.D. at the earliest, because the first traces of Indian influence in Indonesia, which were found in East Kalimantan,² are from this period, and Malagasy has a few Sanskrit loanwords which are almost all found in other Indonesian languages as well. Dahl (1951) leaves open the question whether or not contacts between the Malagasy and Indonesians continued after the migration, and if they did, until what period. He gives a survey of Arabic and Portuguese sources mentioning such contacts, but he cannot determine whether they are decisive evidence or not. He does not consider the possibility of borrowing from other Indonesian languages in Malagasy, although in a later publication he uses borrowing (from Malay and Madurese) as a possible explanation for some irregular sound correspondences in the Malagasy lexicon (Dahl 1977a:95). However, this borrowing is always supposed to have taken place at a period when the ancestors of the Malagasy still lived in South Kalimantan.³

In this article I would like to present evidence for lexical borrowing from Malay and Javanese into Malagasy.⁴ Rather than giving a full list of all possible—or even probable—Malay borrowings, I prefer to present a small number of lexical items that can only be from Malay (§ 1). Once I have established borrowing from Malay and Javanese into Malagasy as a

fact, I shall discuss some of the linguistic and culture-historical aspects of such borrowing.

The linguistic aspects involve the elimination of some sound correspondences between Proto-Austronesian (PAN) or Proto-Malayo-Polynesian (PMP) and Malagasy which appear to apply only to loanwords, and the possible implications of their elimination for the reconstruction of Proto-Austronesian or Proto-Malayo-Polynesian (§ 2). In a number of cases, a Proto-Austronesian or Proto-Malayo-Polynesian protophoneme has two or more Malagasy reflexes (§ 2.1). It appears that some of these reflexes occur only in loanwords. Furthermore, for a number of Proto-Austronesian or Proto-Malayo-Polynesian protophonemes, the evidence is rather weak. The Malagasy lexemes allegedly reflecting these problematic protophonemes appear to be false reflexes or loanwords (§ 2.2).

The culture-historical aspects of this borrowing (§ 3) include a new hypothesis concerning the period of the migration of the Malagasy people (§ 3.1), postmigratory contacts between Madagascar and Southeast Asia (§ 3.2), the nature of the loanwords (§ 3.3), and evidence from the seventh century South Sumatran inscriptions for contacts between the early Malagasy and South Sumatra (§ 3.4).

1. LEXICAL EVIDENCE FOR BORROWING OF MALAY AND JAVANESE INTO MALAGASY. My first concern is to show that there was borrowing from Malay (ML) and Javanese (JV) into Malagasy (MLG), before as well as after the migration of the Malagasy people. The following list consists of MLG lexical items which must be attributed to borrowing because of some unique phonological, semantic, or formal characteristics they show. More structural implications of this borrowing (i.e. the reduction of certain reflexes of protophonemes which appear to occur only in loanwords) are discussed in § 2.

Some of the lexical evidence I present here has already been used for historical-comparative purposes in previous literature (Van der Tuuk 1865, Dempwolff 1934–1938, Dahl 1951). I will not present much new lexical evidence to the reader; rather, I shall attempt to show that some of the existing lexical evidence can be viewed in a different way.

1.1 MLG *taranaka* ‘descendants, posterity, race of the same stem, said of humans, animals and plants; adopted child who is entitled to inherit’. This word never caught the attention of comparativists, but it is clearly a MLG borrowed version of ML *tərnak* (with variant forms *tərənak*, *təranak*, Minangkabau *tarana?*) ‘cattle’. Its original meaning, however, is ‘breed (of any locality) = Min.[Minangkabau] *taranak* or *təranak* native, of the

natives or aborigines of a country (*orang tərnak*); also of live-stock', and it was originally polymorphemic (< ⁺tAr-anak, Zorc 1978:70; see Blust 1982 for apocope of the penultimate vowel). The prefix ⁺tAr- denoted noncontrolledness (Adelaar 1985:182); the base ⁺anak meant 'child of someone', and was also used as a verbal base. The form *taranaka* must be borrowed from ML because MLG does not have a prefix ⁺tar- denoting noncontrolledness. Moreover, the semantics of *taranaka* are very similar to the original meaning of *tərnak*, and there are (as far as I know) no other languages with a correspondence of *taranaka* and *tərnak*.

1.2 MLG *ranto* '1. go trading to far-out places or countries; 2. product of such trading'. Dempwolff made a reconstruction of the phonemes and lexicon of PMP on the basis of material from Tagalog, JV, Toba Batak, ML, Ngaju Dayak, MLG, two Melanesian languages, and three Polynesian languages. He reconstructed PMP *rantaw 'reach (of a river); sail from reach to reach' on the basis of evidence from only three languages, namely, MLG *ranto*, Toba Batak *ratto* 'reach of a river; go wandering' and ML *rantaw* 'reach of a river; go abroad for trading'. Apart from the fact that *rantaw is poorly represented within the group of languages forming the basis of Dempwolff's work (and also outside this group), it is also weak from another point of view. The Malays traditionally travelled along rivers and coasts; they went from reach to reach, and traded along these reaches. It is evident that an original meaning 'reach of a river' extended via a verbal form to 'follow these reaches, trade along these reaches, go trading to foreign places'.⁵ A reverse semantic shift—or any other shift of meaning—is not very likely. *ranto/rantaw/ratto* has correspondences in Sumatran languages (Acehnese, Minangkabau, Gayo (Hazeu 1907), and Lampung). These correspondences have a similar basic meaning, and they have an extended meaning referring to wandering or trading along coasts or in foreign places. It is unlikely that they are inherited: they must have spread from a single Sumatran language (ML or Minangkabau) to other Sumatran languages and to MLG. Leaving one's place of origin, sailing along the rivers, and trading in foreign countries in a certain period of one's life used to be a cultural phenomenon particular to the Minangkabau people and the Sumatran Malays. No correspondences were found outside Sumatra, the Malay Peninsula and Madagascar, except for Iban (Sarawak) *rantaw* '1. reach of river; 2. distinct part of a journey or other undertaking'.

1.3 MLG *rotsaka* 'fall downwards, drop from a high place; slip on a slope'. Dempwolff reconstructed *rusak on the basis of Tagalog *uslak* 'evil person', Toba Batak, JV, ML *rusak* 'destroy', Ngaju Dayak *rusak* 'rotten',

Fijian *rusa* ‘rotten, decomposed, destroyed’. Nothofer reconstructed Proto-Malayo-Javanic *ruksak, with a consonant cluster on account of Sundanese *ruksak*. (This is the only Proto-Malayo-Javanic lexeme with a consonant cluster that is neither caused by doubling of root-morphemes nor made up of a nonnasal consonant preceded by a homorganic nasal; Nothofer 1975: 46). But *rotsaka* and its correspondences can best be explained as originally deriving from Old Javanese *ru:g*, rug ‘to fall in, fall to pieces, collapse; be smashed, scattered, broken up’ and *sa:k* ‘fallen apart, loosened, opened, dispersed, scattered’, which apparently occurred as a compound.⁶ The consonant cluster in this compound is still attested in Sundanese, as it is in MLG, which has an affricate *-ts-*. At the time MLG borrowed the compound form, the cluster was still present in the lending language, which must have been Old Javanese or Malay (which in that case must have acquired the form through borrowing from (Old) Javanese). Fijian *rusa* is not related to this correspondence set, and its phonological similarity is coincidental.

1.4 MLG *tsingy* ‘mountain peak’. Dempwolff reconstructed *tingi ‘high’ on the basis of Toba Batak and ML *tingi* ‘id.’, JV *pratingi* ‘village school’, Ngaju Dayak *tinggi* ‘respectable’, and MLG *tsingy*. Except for ML, each of these languages uses another term for ‘high’ (namely Toba Batak *timbo* (and *ginjaj*), JV (low register, *duwur*, (high register) *ingil*, Ngaju Dayak *ambo*, MLG *avo*, *ambony*). JV, Ngaju Dayak, and MLG narrowed down the meaning of *tinggi/tingy* to a very restricted semantic area. Moreover, correspondences of *tinggi/tingy* are not widespread among other Austronesian languages, which more often have a reflex of PMP *q(a, i) (N)baw for the notion of ‘high’. The only explanation for the semantic diversity of the Toba Batak, Ngaju Dayak, and other correspondences is that they were borrowed from ML, where *tinggi* is the general term for ‘high’.

1.5 ARABIC LOANWORDS INTRODUCED VIA ML: MLG *SOMBILY* AND *SETRA*. Taimoro MLG (on the Malagasy east coast) has *sombily* ‘slaughter in a ritual way’ (cf. also *sombidy* ‘to slaughter’ in de Houtman 1603). This corresponds to ML *səmbəleh* or *səmbəlih* ‘slaughter according to Muslim law’. Dempwolff (1937:45) (from Tendeloo 1901: 160) derived *səmbəleh* from the Arabic formula *b'ismi'lla:hi* ‘in the Name of God’. But Dahl (1951:319) reasoned that, since a correspondence of *səmbəleh* was found in Taimoro MLG, it had to be assigned to PMP, and it apparently pointed to some early Austronesian custom of ritual slaughter. (With this line of reasoning, Dahl (1951) implicitly ruled out the

possibility of postmigratory contacts between Indonesia and Madagascar.) But other correspondences of *səmbəleh* are very restricted. Moreover, they generally have the notion of slaughtering according to Muslim law, and they often show irregularities that must be attributed to borrowing. Examples are Toba Batak *sambol* (with loss of ⁺-i (from */-ih) through back-formation), Iban *səməlih*, Maanyan *samalis* (with -s through hypercorrection), Ngaju Dayak *sambalih* (with irregular penultimate *a* and -*h*). The evidence is strong in support of Dempwolff, although his etymology requires some phonological and semantic explanation. In the context of Arabic pronunciation, *Alla:h* is rather exceptional in that its *a*: is pronounced as a long slightly rounded close low back vowel [ɔ̄:]. But in *b'ismi'lla:hi* the genitive ending -*i* of *ismi* palatalizes the geminated *l*, which in turn produces a long close low front vowel [æ:] in the segment -*lla:h*;- the genitive ending is usually not pronounced at the end of a noun phrase. This yields a pronunciation [besmel:æ:h] (or even [bəsməleh]; cf. Barthélemy 1935), which would easily lead to ML *səmbəleh*. The Arabic short-short-long vowel sequence would be reinterpreted as a trisyllabic word structure with stress on the last syllable, which in ML implies that the preceding syllables contain a shwa. As to the consonants of *səmbəleh*, their positions may be the result of metathesis (caused by the fact that homorganic consonant clusters are indigenous to the word structure of ML, whereas heterorganic clusters are not). It may also be the result of a combination of back-formation and the well-attested phenomenon of epenthesis of a voiced stop⁷ (in this case *b*) between a nasal (*m*) and a following liquid (*l*): the first syllable of [besmel:æ:h] is interpreted as the ML prefix *bə(r)-*, and an epenthetic -*b*- is introduced after -*m*- . This *m*, in combination with the following geminated *l*, may have been perceived as a consonant cluster.

As to the semantics of this etymology, the shift from a formula invoking the name of God to a term for slaughtering animals according to Muslim law becomes intelligible if one considers the Muslim obligation to utter this formula at the moment the animal's throat is being cut. In some Arabic dialects and in other languages spoken by Muslim peoples, the formula *b'ismi'lla:hi* developed into a separate lexical unit referring to the practice of uttering this formula. In Persian (a language from which Malay took many loanwords), the formula became synonymous with killing animals in a ritual fashion (cf. Persian *bismil* 'sacrificed, slaughtered', *bismili* 'for sacrifice, for slaughter', *bismil-gāh* 'slaughter-house, butchery'). The change from *b'ismi'lla:hi* to *səmbəleh* can be accounted for, and it seems evident that Dempwolff was right in deriving the latter from the former. In this light *sombily* must be a ML borrowing. Its antepenultimate *o* is colored by the following labial nasal: this may be the way the MLG perceived the word

from ML speakers. It is also possible that a back vowel already occurred in the form of Malay from which MLG borrowed the term. The possibility of independent borrowing from Arabic into Taimoro can be dismissed as it would presuppose too much coincidence. It is too unlikely that the same complicated phonological and semantic changes could have taken place in two independent cases. On the other hand, the phonology of *sombily* does not give a clue as to the period in which this lexeme was borrowed into MLG. One way of measuring the relative authenticity of a MLG lexeme is the Bantu “substratum,” a rather ill-chosen term for a number of changes which distinguish MLG sharply from other Southeast-Barito languages, and which can be explained through contact with Bantu languages (more particularly Comoran). Sometime after the migration of the ancestors of the Malagasy, their language underwent the influence of this “substratum,” which involved, among other things, the fricativization of a number of occlusives and semivowels, the acquisition of a murmured vowel after final consonants, and the change of ⁺d, ⁺t, and ⁺r, into *tr* (a postalveolar affricate with trilled release) (Dahl 1954). Since none of the affected sounds is present in the case of *sombily*, this Bantu “substratum” did not have (and could not have had) any effect on the shape of the word, and it is consequently not possible to determine whether it was borrowed before or after the migration of the ancestors of the Malagasy.

With regard to the dating of borrowings, *setra* is more enlightening. The form *setra* means ‘opposition, resistance; facing the wind or the flow; tempting fate, disparaging the law; fighting the impossible’, and it is ultimately borrowed from Arabic (cf. *silr* ‘sorcery, witchcraft’). It underwent the influence of the Bantu substratum; it could have entered the language before the migration, because we find correspondences in other Barito languages (Maanyan *pañeher* ‘magician’; Ngaju Dayak *seher* ‘magic spell’). The Barito languages must have borrowed it from ML (cf. ML *ilmu seher* ‘black magic’, where *ilmu* means ‘science, knowledge’), and ML in turn borrowed it from Arabic. Postulating *setra* as a loanword ultimately borrowed from Arabic and predating the Bantu “substratum” has implications for the dating of the Malagasy migration. These will be discussed in § 3.1.

1.6 MLG *landaizana* ‘anvil’ (Richardson 1885; Betsileo MLG *landeza*, Dalmond 1842). Dempwolff reconstructed PMP **lanDasan* ‘anvil’ on the basis of JV *landəsan*, Toba Batak, ML *landasan*, and MLG *landaizana*. He considered the *a* in the JV reflex as irregular. But his interpretation of the data was wrong. JV *landəsan* means ‘support, (hard) surface; chopping

block', and it is derived regularly from *landas* 'firm ground, surface'. It is only rarely used to denote 'anvil', which must be a calque-translation from ML: the usual JV word for 'anvil' is *paron*. Toba Batak *landasan* means 'flat surface' and is derived from *landas* 'flat, even'; it means 'anvil' in the phrase *landasan ni panompa*, lit. 'block with a flat surface used by a blacksmith'. Malay *landas* means 'solid block on which heavy work can be done', and *landasan* (also *palandas*) 'an anvil; a chopping block'. So, the meaning 'anvil' is only a specific use of a much wider meaning of derivations of *landas*/*landas*. Correspondences of ML *landasan* which have this meaning are probably either ML borrowings or (as in the case of JV) inherited forms with a calque-translation from ML. MLG *landaizana* is the only correspondence with the specific meaning of 'anvil'.⁸ Moreover, in spite of its polymorphemic structure, the latter is not part of a nominal paradigm, as are the other correspondences. The consonant cluster *-nd-* is irregular, as PMP *-n(d, D)- as a rule yields MLG *-ndr-*. In summary, there are semantic, structural, and phonological reasons to consider *landaizana* as a borrowing, probably from ML.

However, the *-aiz-* sequence is rather curious, and cannot be explained through borrowing from ML. Dempwolff attributed it to dialectal variation in Malagasy (Dempwolff 1937:98), but there is little support for this assumption. On the other hand, it is possible that *landaizana* was borrowed from a vernacular form of ML, where lexeme-final ⁺s became *h*, and where a preceding ⁺a was diphthongized to *ay* or heightened to *e*. The change from ⁺as to *-ayh* or *-eh* is seen in a large number of Peninsular and Sumatran Malay dialects, or Minangkabau (which is very closely related to ML; cf. Minangkabau *landehan* 'chopping block; anvil'). In MLG, the ⁺*-h-* of a borrowing would be lost, and the palatal glide on the morpheme boundary between ⁺*landay(h)* or ⁺*lande(h)* and ⁺*-an* would yield *z* (Dahl 1951: 56). Another tentative explanation is that the word was borrowed at a time when the MLG changes from PMP *-əs, *-as to *-y* (yielding *-az-* and *-ez-* in combination with *-ana*) were still in process. The effect of such processes may have worked out differently in a polymorphemic loanword as *landaizana*. But the explanation remains uncertain. Nonetheless, the phonological reflexes of this word and its striking agreement in form and meaning with ML *landasan* (or its Minangkabau correspondence) leave little doubt about its borrowed origin.

1.7 MLG *harafesina* 'rust'. In connection with *landaizana*, it is relevant to point out another term related to the use of iron: *harafesina* 'rust'. It must be borrowed from ML, cf. ML *karat bəsi* 'rust of iron'. The *-f-* and *-na* are unexpected: the cluster *-t* and *b-* at the morpheme boundary between ML

karat and *bəsi* may first have been perceived as a ⁺p, which later on developed into f. The final -na is sometimes attached to a stem which historically ended in *r, *q, *R, or Ø (see § 2.1.7).

1.8 MLG *soratra* ‘writing; any drawing, lines, marks, figures, or painting’. The evaluation of this lexical item involves some reflection on a more general problem, namely, whether PAN *surat ‘to write; writing’ (first reconstructed for PMP by Dempwolff) is a valid reconstruction, and whether the early Austronesians were therefore acquainted with writing. On the one hand, reflexes of *surat are found in all Austronesian-speaking areas except Oceania (Blust 1977:28, 33). But on the other hand, all Austronesian peoples with writing systems of their own derived these systems from the Pallava script, a South Indian variety of the Brahmi syllabary. Apart from these, Latin and Arabic writing are used, the only exception being the Rongorongo characters of Rapanui (Easter Island). The latter was probably also known to other Eastern Polynesian peoples at some earlier stage, but is still a very local development. Moreover, as stated above, the languages of Oceania have no reflex of *surat.

There are other very strong arguments against a PAN term for ‘writing’. Other high-order protolanguages such as Proto-Indo-European have no etymon for ‘writing’ either, although they are ancestral to languages with long-established writing traditions. Nor is there, for instance, a Proto-Germanic etymon for this notion. Terms for ‘writing’ in such closely related languages as English and Dutch are not related, and as far as cognates of these terms go, they often refer to scratching, carving, painting, or drawing. Greek *grafein* ‘write’ is a cognate of ‘carve’, English ‘write’ is a cognate of Dutch *rijten* and German *reissen* ‘rip, tear’; Dutch *schrijven* and German *schreiben* ‘write’ are borrowed from Latin *scribere* ‘write’; and Russian *pisat'* ‘write’ is a cognate of Latin *pingere* ‘paint; draw, sketch; embroider’ (cf. French *peindre*, English *paint*). The same irregular picture is shown by Semitic and Caucasian languages, and probably by many other language groups. Comparison of terms for ‘reading’ in Austronesian languages or other linguistic families yields analogous results.

Projecting a Proto-Austronesian writing system which was later lost in some places as a consequence of cultural impoverishment or lack of need is pointless, as there are no other indications for it than the hypothetical etymon *surat. The reason why most Austronesianists⁹ have not rejected it is probably that its alleged reflexes are so widespread: they are found in Madagascar, East and West Indonesia, the Philippines, Taiwan, and on mainland Southeast Asia. Most Austronesianists hitherto assumed that MLG (after the departure from Southeast Asia) and the languages of Taiwan have not been subject to influence from any other Austronesian

language. As far as Malagasy is concerned, the present article is an attempt to provide lexical evidence for such influence in support of the extralinguistic evidence already at hand (Dahl 1951: 365–372; Ottino 1974). As far as Taiwan is concerned, it was presumably more accessible for early mariners from insular Southeast Asia than Madagascar, as it is much closer and can be reached by island-hopping. Moreover, there are some indications of contacts between Malays and Taiwanese.¹⁰ Finally, some Austronesianists (e.g. Wolff, see § 2.2.3) dispute the validity of *r as a PAN/PMP phoneme. Their main argument is that many of the etyma in which *r occurs appear to be based on borrowing, usually from ML or JV into another Indonesian or Philippine language. The dismissal of PAN *surat, with its large number of reflexes, would of course strengthen this argument. In light of the above, I discard *surat from the PAN and PMP vocabularies. This word must have spread through borrowing, apparently from ML or JV. Wilkinson (1959) gives for ML *surat* only ‘thing written, letter, epistle’, but Klinkert (1916) also gives ‘row, layer of stones in masonry’. Gericke and Roorda (1901) give for JV *surat* ‘thin line of light or of colors’, and for *sərat* ‘1. line, vein, thread (in meat); 2. (high register term for) write; letter’. Pigeaud (1938) gives *surat* ‘1. (with a) streak (of color); 2. a letter, a few lines (in administration)’, and *sərat* (high register) ‘book, letter; haircomb’, *dipunsərat* ‘written; batiked’. Zoetmulder (1982) gives Old JV *surat* ‘(drawn) line, writing, drawing, letter’. So, ‘writing’ is clearly a specific application of the originally much more general meaning of *surat*. Consequently, the languages that have a correspondence of *surat* with the specific meaning ‘writing’ are likely to have borrowed it from ML or JV.

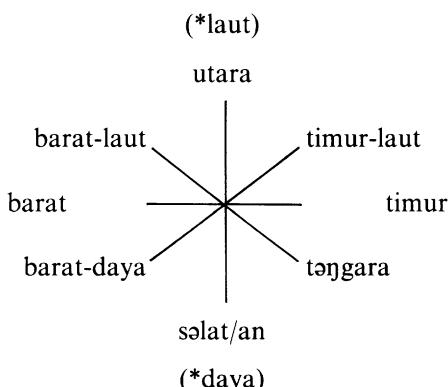
To return to MLG *soratra* in light of the above, this must consequently also be considered a loanword from ML or from JV.¹¹ Although *soratra* also has a wider meaning than only ‘writing’, this may be seen as evidence for an independent semantic development, and there is no evidence for a native Malagasy script. The only forms of writing on the island are Arabic script (traditionally called Sorabe) and Latin script. Furthermore, *soratra* reflects both s and r in correspondence to ML/JV s and r respectively, which suggests borrowing (cf. § 2.1.4 and § 2.2.3).

1.9 MLG NAMES OF WINDS AND WIND DIRECTIONS. MLG *avaratra* ‘north’ and *atsimo* ‘south’ do not have the same meaning as their correspondences in other languages, cf. ML *barat* ‘west’ and *timur* ‘east’. This is due to the direction of the monsoon winds in the different areas where these languages are spoken. PAN *qabaRat ‘northwest monsoon’ and *qaCimuR ‘southeast monsoon’ were names of winds, and only later did their names become associated with the directions from which they blew, which may differ from one Austronesian area to another. For MLG

the directions differ ninety degrees clockwise from those for ML: MLG *avaratra* and *atsimo* became associated with north and south, whereas ML *barat* and *timur* became associated with west and east. (The initial *a*- in the Malagasy forms is a locative marker.) In light of these associations, it is remarkable that in the Betsimisaraka dialect of MLG, *tsimilotru* means ‘north wind’ and *anghi varatrazza* ‘south wind’ (Sakalava MLG *varatrazza* ‘east wind’; cf. Dalmond 1842).¹² Formally these terms correspond to ML *timur-laut* ‘northeast’ and *barat-daya* ‘southwest’, but semantically they did not follow the ninety degrees clockwise change of *avaratra* and *atsimo*. One would expect a meaning more associated with ‘north’ for Betsimisaraka MLG *anghi varatrazza* and Sakalava MLG *varatrazza*, and one more associated with ‘south’ for Betsimisaraka MLG *tsimilotru*, but this is not so. In MLG dialects the meanings of the terms *varatrazza* and *tsimilotru* are no longer internally consistent with *avaratra* and *atsimo*. Moreover, the reflexes of *avaratra* and *varatrazza* are irregular. Dahl (1977a: 95) has already pointed out that the *r* in *avaratra* must be due to borrowing from ML or Madurese.¹³

In *varatrazza*, apart from the first *r*, the final syllable is also irregular and should have become ⁺-y (cf. PAN **buqaya* ‘crocodile’ > MLG *vuay*). If *avaratra* is borrowed from ML, there is all the more reason to suppose that *varatrazza* and *tsimilotru* also have the same origin.

There is also evidence that all three are borrowed from a kind of ML that was spoken outside Kalimantan. The ML terms for wind directions betray an older terminology which provides a clue to the place where these terms were established, and hence, where the early Malagasy borrowed them. Compare the following system:



H. Kern (1917: 120) took the fact that *sølat/an* means ‘south’ as well as ‘straits’ as evidence that the Malay Peninsula is the original homeland of the Malays. From the point of view of the Malay Peninsula, the Straits of

Malacca are in the south. But judging from the internal structure of the above system, the fact that *utara* is borrowed from Sanskrit, and the fact that *səlat/an* has a basic meaning ‘straits’ (cf. *səlat* ‘narrow space between two things’), we should reconstruct an earlier system which shows a quite different picture. Both *barat-laut* and *timur-laut* show that *utara* replaced an earlier **laut* (which still occurs as *laut* ‘open sea’ < PMP **lahud* ‘towards the sea’), and likewise, *barat-daya* and *təŋgara* (apparently a loan, but its origin is unclear) show that *selat/an* replaced an earlier **daya* (cf. PMP **Daya* ‘towards the interior’, and cf. the general term *Dayak* ‘inhabitant of the interior [of Kalimantan]’). In other words, in the system underlying the present ML system of cardinal points, the term for ‘north’ is associated with the open sea, and the term for ‘south’ is associated with the interior. The underlying system does not apply specifically to the Malay Peninsula, nor, for instance, to South Kalimantan, where the city of Banjarmasin (close to the Maanyan area) is located. But it works well for the Jambi and Palembang area in South Sumatra, where Srivijaya may have been located. This fact supports the hypothesis that the early Malagasy underwent ML influence through contacts with Sumatran Malays, and not only through contacts with Banjarese Malays (see § 3.1). Correspondences of the directional terms *barat*, *barat-daya*, and *timur-laut*, do not occur in the lexical sources for Banjarese ML,¹⁴ and we are led to the conclusion that MLG *avaratra*, *varatrazza* and *tsimilotru* were borrowed from a South Sumatran kind of ML.

1.10 MLG NAMES FOR PARTS OF THE BODY. There is a striking agreement between MLG and ML terms for external body parts. Dempwolff took such terms for inherited vocabulary and used them for PMP reconstructions. In some cases, as with MLG *tánana* ~ ML *tajan* ‘hand’, MLG *hoho* ~ ML *kuku* ‘nail’, MLG *fify* ~ ML *pipi* ‘cheek’, MLG *voto* ~ ML *butuh* ‘penis’, this is warranted on account of their correspondences in many other Austronesian languages. But in other cases MLG must have borrowed from ML. Compare the following correspondences:

- (a) MLG *voavitsy* ‘calf of leg’, ML *buah bətis* ‘id.’, (Maanyan *wisis*),
PMP **bitis*
- (b) MLG *molotra* ‘lip’, ML *mulut* ‘mouth’ (Maanyan *wiwi* ‘lip’,
aruβaβa? ‘mouth’), PMP **mulut* ‘mouth’
- (c) MLG *sofina* ‘outer ear’ (occurs along with *tadiny* ‘inner ear’), ML
cupig ‘ear-lobe’ (Maanyan *silu?* ‘ear’), PMP **cupin*
‘ear-lobe’, **taliŋa* ‘ear’
- (d) MLG *valahana* ‘loins’, ML *bəlakaj* ‘(lower part of the) back’,
(Maanyan *pugug* ‘back’),¹⁵ PMP **balakaj* ‘back’,
**likuj* ‘id.’

- (e) MLG *tratra* ‘chest’, ML *dada* (Maanyan *dadaP*), PMP *DaDa ‘chest’
- (f) MLG *lamosina* (Sakalava MLG *lambosy*) ‘back’, Minangkabau ML *lambusi*⁹ ‘shoulder of a cow’, Sundanese *lam(b)usir* ‘shoulder, rib of a slaughtering animal’, Old Javanese *lamujsir* ‘back, piece (of meat) from the back’

In all these cases, MLG must have borrowed from ML.

- (a–b) According to Dahl (1951:304) Tidung (a Murutic language spoken in Northeast Kalimantan) also has a compound with *buah* (lit. ‘fruit’) to denote ‘calf of a leg’, but elsewhere no such compound is found. He also says that the only other languages with a correspondence of *mulut*, *molotra* are Sibuyau, Lemanak, Milikin (with *mulut* ‘lip’) and Salakau (*mulut* ‘mouth’). These are all Malayic-Dayak languages, and are closely related to ML (Hudson 1970).¹⁶
- (c) ML *cupiq* originally means ‘torn piece of something’ and it is also used for nostril (in the compound *cupiq hiduj*). Correspondences of *cupiq* are found in Karo Batak (Neumann 1951) and in Lampung (both having *cupiq* ‘ear’), and Toba Batak (having *supiq* ‘ear-lobe’). It is clear that the Batak, Lampung, and MLG correspondences are borrowed and reflect the secondary meaning of ML *cupiq*. I have found correspondences only in Sumatran languages. The initial consonant in *sofina* is unexpected, as MLG usually has *ts* corresponding to ML *c*.
- (d) Dempwolff reconstructed **balakanj* on the basis of Tagalog *balakáj* (‘hip’), ML *bəlakanj*, and MLG *valahana*. Its reconstructed meaning ‘back’ is actually ill chosen: ‘(the space) behind (a person)’ would be more likely to have generated the semantic diversity of the correspondences. Compare also Antaisaka MLG *valaha* ‘haut des cuisses’ (Deschamps 1936), Sakalava MLG *am/balak* ‘south’ (Dalmond 1842), and from de Houtman (1603) (Southeast MLG) *valahan* ‘the buttocks’, *valahang* ‘anus’. But also from other points of view this reconstruction is very weak. Tagalog *balakáj* has no correspondences in other Philippine languages, and it has stress on the last syllable, which is unusual for Tagalog inherited vocabulary in words not containing a consonant-cluster. Other correspondences of *valahana*, *bəlakanj*, etc. are very few, and they are suspect of borrowing. I found Madurese *balakanj* ‘back’, in which *b* is not the inherited correspondence of ML *b* (Nothofer 1975: 136). I reject **balakanj*, and I consider the correspondences of *bəlakanj* as ML loans.

- (e) Correspondences to *tratra/dada* are more widespread, and as far as Acehnese and Chamic correspondences go, they must be inherited (cf. Acehnese *dada*, Cham *taṭa* [Moussay 1971], Jarai *tɔ̄da* ‘chest’ [Lafont 1968]). But they do not occur outside the Sumatran and Southeast Asian mainland area, except for JV *dāda*, which must be borrowed on account of its retroflex *d*’s, and MLG *tratra*, Maanyan *dada?*. Dahl (1951:302) considers *tratra* and Maanyan *dada?* as borrowings predating the Malagasy migration. He probably means that they were borrowed from ML, which is the most likely solution. (Apart from *tr* for expected *r*, the *-a* in *tratra* is also irregular [*-a > -y].) A Fijian cognate provided by Dempwolff, *vaka/rara/kombi* ‘cross arms over one’s chest’, is too far-fetched.
- (f) A possible connection between *lamosina* and its correspondences has already been pointed out.¹⁷ These correspondences are restricted to ML, JV, and Sundanese, and they may derive from original compounds constituted of (ML, JV) *lambug* ‘side, flank’ and (a syncopated form of) JV *sisir* ‘side, edge’ (ML *məñisir* ‘walk along something’). None of the other Barito languages seems to have a corresponding form (cf. Hudson 1967). The form *lamosina* must be borrowed, and the absence of an ending in *-tra* points to Minangkabau ML as a source language (see § 2.1.7).

Most of the above sets have correspondences outside MLG and ML, and some of these correspondences are even inherited. But it is clear that *sofina* and *tratra* are ML borrowings, and given the circumstance that the other terms belong to small sets of correspondences that are often suspect of borrowing and that are not attested for Maanyan and most other East Barito languages (Dahl 1951: 300–304; Hudson 1967: 77ff.), they too must be borrowed from ML. How then does it happen that MLG has such a high score of borrowed terms for parts of the body? A possible explanation is that these terms were subject to the working of taboo. A recent taboo phenomenon in West Indonesia is the development of a formal and polite vocabulary in addition to the everyday language unmarked for register. Sundanese, Balinese, Madurese, and particularly JV, have sophisticated systems of socially stratified speech styles which also affect the morphology. The Malay of Palembang in South Sumatra has a highly marked court vocabulary which is strongly influenced by JV (de Clercq 1876: 526–527; P. D. Dunggio et al. 1983). A limited court vocabulary was also found in the literary Malay of the sultanate of Riau and Johore (Prentice pers. comm.). It is likely that the high-register speech styles of JV, Sundanese, Balinese, and Madurese, developed from such a court vocabulary. If the early Malagasy had contacts with the Indianized Malays in Sumatra (as

will be argued in § 3.1), they may have begun to develop a “polite language” which particularly affected terms for parts of the body. Another possibility is that they used ML vocabulary (which was apparently more prestigious) to relexify these terms, which in many cultures are often subject to social devaluation.

1.11 DOUBLETS. In some cases MLG has doublets for a PMP lexeme. In the following examples, the first correspondence of each pair exhibits the expected form of a reflex of the protolexeme, whereas the other is closer in form and/or meaning to ML, and is explained as being borrowed from the latter:

- (1) *tany* ‘earth, ground’ < *tanəq ‘id.’
tanána ‘city, village’; cf. ML *tanah* ‘1. earth, ground; 2. country’
- (2) *voto* ‘penis’ < *butuq ‘id.’
boto ‘young lad’; cf. ML *butuh* ‘penis’, Banjarese ML *butuh* ‘penis’,
butuh, utuh, tuh ‘(term of address for young boys)’ (W. Kern n.d.);
cf. also Maanyan *butoh*
- (3) *tsidika* ‘put one’s head through a door or a window; pay a short visit’ < *tilik ‘look closely at something, observe’
tilik/ambo ‘(look-out) tower’ (*ambony* ‘high’); cf. ML *tilik* ‘looking hard and fixedly at anything’
- (4) MLG (eastern dialect) *taiky* ‘sea’ (Dahl 1951:317) < *tasik ‘id.’
I/tasy ‘(name of a large lake)’; cf. ML *tasik* ‘lake; sea’; cf. also Maanyan *tahik* ‘sea’ (Dahl 1951:317)

MLG *boto* must be borrowed from Banjarese ML on account of its phonological shape and its derived meaning. (The same holds for Maanyan *butoh*, which did not undergo the regular Maanyan sound changes; cf. PMP *b > Maanyan w; PMP *q > Maanyan Ø). Austronesian peoples have a custom of giving small children unpleasant or obscene names in order to keep bad spirits away.¹⁸

In MLG *I/tasy*, the final *k was lost. The original name of Lake Itasy was *I/tasih/anaka* (‘small sea’); due to back-formation it was reanalyzed as *I/tasi/hanaka* (‘extensive sea’) and thence shortened to its present form (Dahl 1951:316). I consider *taiky* as the regular reflex of *tasik, and *I/tasy* as a borrowing.

2. LINGUISTIC INFERENCES. The fact that MLG borrowed from ML has some important consequences for the phonological history of MLG: some of the alleged MLG reflexes of PMP protophonemes turn out to occur only in loanwords (§ 2.1). But it also has consequences for PMP phonology. A number of PMP protophonemes have a rather weak status, and some Austronesianists prefer to eliminate them from the PMP pho-

neme inventory. Now that it is clear that MLG underwent considerable lexical influence from ML, the MLG evidence for some of these doubtful protophonemes can in large part be eliminated as loanwords or false reflexes (§ 2.1).

2.1 MLG SOUND-CORRESPONDENCES REFLECTING BORROWING. MLG often has several reflexes for one PMP protophoneme. The occurrence of these different reflexes is not conditioned, that is, there is no rule predicting which of several alternative reflexes occurs in which position. For instance, in MLG, *g becomes *h* or *g*; (nonfinal) *d and *D become *r*, *tr*, or *d*; *R becomes *r*, *z* or \emptyset ; *s becomes \emptyset or *s*; *ə becomes a front vowel or *a*; PMP *-a(R, s, \emptyset) becomes -*y* or -*a*; nonfinal stops as a rule became fricativized, but sometimes remain stops.¹⁹ It is appropriate to relate the occurrence of several unconditioned reflexes for one protophoneme to the fact that MLG borrowed from ML. It may turn out that one reflex (or even several reflexes) can only be found in loanwords. Having more than one unconditioned reflex for a protophoneme is in itself no reason to conclude that only one reflex is inherited, and that the other reflexes are due to borrowing: sound correspondences by no means have to occur on a one-to-one basis. But in MLG, it is very likely that borrowing accounts for at least part of the cases in which there is more than one reflex.

The following paragraphs contain examples from Dempwolff (1938), Dahl (1951), and Blust (1970, 1980). All examples quoted are Dempwolff's, unless indicated otherwise. Page numbers are given if the source has no systematic arrangement of the examples (as in Dahl 1951). In the few cases where I present an etymon from Blust, the MLG reflex is provided by me.

2.1.1 *g > *h*, (in a few cases) *g*. In MLG, PMP *g usually becomes *h*. In the sample, there are also seven instances exemplifying the change PMP *g > MLG *g*. From Abinal and Malzac's dictionary (A/M) it appears that *g* is not a very frequent phoneme in MLG, and that only eight out of 876 pages are occupied by entries with initial *g*. A large proportion of these entries are of a sound-symbolic nature or are French or English loans. The seven instances exemplifying PMP *g > MLG *g* are:

- (a) *əŋgan̥ ‘hornbill’ > *engana* ‘wings of birds’ (Dempwolff; cf. ML *əŋgan̥*, Toba Batak *ongan̥* ‘hornbill’)
- (b) *gagak ‘bird’s voice’ > *gaga* ‘crow’ (cf. ML *gagak*)
- (c) *garunggan̥ ‘name of a tree’ > *harongana* ‘Haronga madagascar’
(although not given by Dempwolff, ML has *gəroŋŋay*, *gəroŋgoŋ* ‘a tree, *Cratoxylon formosum*’ [Wilkinson])

- (d) *girik ‘bore, pierce’ > *girigirika* ‘point’ (Dahl 1951:47–48; cf. ML *girik* ‘bored, pierced’; cf. also MLG *hirika* ‘pierced, bored’, Dempwolff)
- (e) *guam ‘cutaneous eruption’ > *goana* ‘blind in one eye, having a black eye’ (Dahl 1951:48; cf. ML *guam* ‘thrush in children’)
- (f) *pingan ‘saucer’ > *finga* ‘plate’ (cf. ML *piŋgan* ‘plate’, which is borrowed from Persian [cf. Persian *piŋga:n* ‘cup’])
- (g) *pingan ‘loins’ > *fingana* ‘trip a person’ (cf. ML *pingan* ‘loin’)

These instances do not reflect PMP *g.

- (a) From a semantic point of view, *engana* is hard to reconcile with *əŋgan̩. Furthermore, ML dialects have a constraint on initial (h)ə- sequences followed by a stop, and they insert a nasal in this position. This constraint, and the consequent insertion of a nasal are reflected in ML *əŋgay*; the constraint does not apply to MLG and Toba Batak lexemes. For instance, the PMP etyma *ə(N)tut ‘fart’ and *ə(N)pat ‘four’ become ML *k/əntut* and *əmpat* (cf. MLG (*k*)*etotra* ‘fart’, MLG *efatra*, Toba Batak *opat* ‘four’; Adelaar 1985:122 and n. 1).
- (b) *gaga* is of an onomatopoetic character and does not reflect the final *k of *gagak.
- (c) *harongana* is suspect of borrowing on account of its four syllables (one syllable more than the canonic number in MLG) and its being a botanical term.
- (d) *girigirika* differs considerably in meaning from *girik, and is probably not related to it. The doublet *hirika* must be borrowed (see § 2.2.3.5).
- (e) Semantically, *goana* does not match well with *guam.
- (f) *finga* is borrowed from Persian.²⁰
- (g) There is no obvious semantic connection between *fingana* and *pingan̩.²¹

The above correspondences cannot constitute an argument for the development PMP *g > MLG g. Rather, they support Wolff’s assertion that *g is a false protophoneme in the first place (Wolff 1974). There are two other interesting MLG forms which need mention, namely, *tongolany* ‘prop, support (e.g. under a stone or a fishing net, to keep it up)’ (cf. A/M) and *godàna* ‘large, vast’. The form *tongolany* must be from ML *tungul-an; the form as such is not found in ML dictionaries, but ML has *tungul* ‘tree stump’ and *pənuggul* ‘hobble’. (Cf. also Iskandar’s [1970] gloss *tungul* ‘1. tree stump; 2. rice-stalks left standing after harvest, stubble; 3. stake, pole [for mooring boats, etc.]’) *godàna* corresponds with Toba Batak *godàŋ* ‘big; much’ (cf. Minangkabau *gaday*, other Sumatran ML dialects *gəday* ‘big’).

According to Dempwolff, *godàŋ* derives from PMP *gudaŋ, but this is not the case. If *godàna* is related to *godàŋ*, it can only be through borrowing, because Toba Batak *o* derives from PMP *ə (and PMP *ə as a rule yields MLG *e* or *i*). The stress on the second syllable is in striking agreement with *godàŋ*, which follows the regular stress pattern for predicatively used and underived adjectives in Toba Batak.

2.1.2 Initial and intervocalic *d, *D > r, tr, d. A similar but more complicated picture is obtained from the MLG lexemes with nonfinal *tr* and *d* from PMP *d and *D. As a rule, *d and *D yield MLG *r* in initial and intervocalic position, *dr* in postnasal position, and *-tra* in final position. However, consider the following list of correspondence sets, where initial and intervocalic *d or *D become MLG *tr*:

*d, *D > *tr*:

- (a) *bidaj²² 'be large' > *vitrana* 'sew together two pieces of textile, join two beams' (ML *bidaj* 'large', also [Klinkert 1916] 'numeral classifier for flat, thin objects [paper, cloth, mats], or flat surfaces'; JV *bedaj* 'large rim')
- (b) *Dəku 'bent' > (Betsileo MLG) *treko* 'obstinate' (Dahl 1951:46; JV *dəku* 'stoop, lean down')
- (c) *kuDug 'boom, thunder' > *kotroka* 'thunder' (JV *gludug*, *krudug* 'boom, thunder')
- (d) *dibbib 'suckle' > *tritry* 'id.' (Dempwolff), (Bara MLG *tritriky* 'id.', Dahl 1951:46)
- (e) *qadaŋ²² 'prepare an attack' > *atrana* 'load a gun, arm oneself' (ML *hadaj* 'lie in ambush')
- (f) *qadəp²² 'front(side)' > *atrika* 'front' (Sakalava MLG *ap-atrefa* 'in front'; ML (*h*)*adap* 'id.')
- (g) *DaDa 'chest' > *tratra* (§ 1.10)
- (h) Sanskrit *doṣa* 'sin' > *trosa* 'debt' (Dahl 1951:343; ML *dosa*, JV *dosa* 'sin')
- (i) *(dD) *anaw* 'shed, house' > *trano* 'house' (ML *dagaw* 'field shed')
- (j) *paDaŋ²² 'prairie' > *fandra* 'plain', *fatrana* 'clearing, glade' (ML *padaj*, JV *padaj* 'prairie')
- (k) *suDu²² 'spoon' > *sotro* 'id.' (ML *sudu* 'id.')
- (l) *Duyuŋ 'sea cow' > *trozona* 'whale' (ML *duyuŋ*, JV *duyuŋ* 'sea cow')

I reject the sound change PMP *d/D > MLG *tr*. Most of the above 12 sets are based on loanwords, and the remainder are weakly attested correspondences which in large part can be explained through etymology or borrowing.

- (a–b) The correspondences **bidaŋ* > *vitrana* and **Dəku* > *treko* have a rather arbitrary semantic connection.
- (c) **kuDug* and *kotroka* are of a sound-symbolic character.
- (d) **dibbib* is clearly an erroneous reconstruction: Dempwolff's evidence for it is MLG *trity* and Tagalog *dibbib*, for which he gives the gloss '*Brust*'. German *Brust* means both 'breast' and 'chest', but the Tagalog dictionaries leave no doubt as to its meaning 'chest'. It is very unlikely that the notion of 'sucking from a breast' would be linked with that of 'chest'.
- (e–f) *atrana* and *atrika* are probably borrowed. But the semantic connection between *atrana* and its correspondences is not so obvious, and it may therefore not be related to **qadaŋ*. The form *atrika* has a penultimate *i*, which points to a schwa in the protolanguage or in the lending language. ML has (*h*)*adap* 'front; in front of' and cannot be the lending language, although it is striking that the original schwa of (*h*)*adap* was maintained in the lexicalized derivation *dəp/an* 'ahead, in front of' (< Proto-Malayic ⁺*hadəp-an*, Adelaar 1985). The source language may have been JV (cf. JV *adəp* 'in front of, facing').
- (g–h) *tratra* is borrowed from ML, while *trosa* is borrowed from Sanskrit.
- (i) Correspondences of ML *dayaw* are very much limited to Kalmantan languages and to MLG. These languages must have borrowed their correspondences from ML at a period prior to the Malagasy migration.
- (j–l) *fatrana*, *sotro*, and *trozona*, must also be borrowed from ML. There are admittedly no decisive arguments for their being borrowed, but their number is too small to consider them as evidence for PMP **d/*D* > MLG *tr*.

In a small number of cases, PMP **d* and **D* become MLG *d*:

- **buDu* 'be afraid' > *bodo* 'infantile' (cf. ML *bodo(h)* 'stupid')
- **lanDəsan* 'anvil' > *landaizana* (§ 1.6)
- **li(N)Dun* 'protection' > *tan/dindona* (Dempwolff), (dialectally) *lindona* 'shadow' (Dahl 1951:347; cf. ML *lindug* 'protection; protected')
- **disdis* 'cut, slice' > *didy* 'a cut' (Blust 1980; cf. ML *didis* 'cut into thin slices')

One of these four examples is based on a mistaken etymology. The form *didy* developed regularly from **lily* (cf. Sakalava MLG *lily* 'cut, command'), where the occlusivization of **l* before **i* has not taken place.²³

Two other examples must be borrowed from ML. MLG *bodo* also has an irregular *b*, and the distribution of its correspondences is limited to ML

and JV, and languages that have been intensively influenced by them.²⁴ For the ML origin of *landaizana*, see § 1.7.

The *tandindona/lindona* example is the only one that cannot be dismissed offhand. But it could easily be borrowed, and it would not make sense to posit a sound correspondence on the basis of this one example. I reject the correspondence PMP *d, *D > MLG *d*.

But when does MLG reflect *tr*, and when *d*, for *d and *D? The conditioning factor must be time. This is also suggested by Dahl. He explains the phonetic realization of *tr* through the Bantu “substratum” (Dahl 1954; cf. § 1.5). He considers lexemes with *d* corresponding to ML *d* as loanwords having entered MLG after the period of Bantu influence (Dahl pers. comm.).

2.1.3 *R > r, z, Ø. For PMP *R there are three reflexes (*r*, *z* and Ø), but Dahl (1977a:86–96) demonstrates that *z* and Ø are regular reflexes, and that lexemes with *r* < *R are due to borrowing.

2.1.4 (nonfinal) *s > Ø, s. In other cases it is much harder, if not impossible, to show that one of the MLG reflexes is a loan-phoneme, but still there is often one reflex that usually occurs in borrowings, and one that seems more authentic. Nonfinal *s corresponds to MLG Ø in lexemes that are usually above suspicion of borrowing, whereas this *s corresponds to MLG *s* in lexemes that are usually not. Nevertheless, as already pointed out by Dahl (1951:42), there are a few lexemes with *s* < *s which cannot be eliminated as being loanwords: *sivy* ‘nine’ < *siwa (Dahl 1951:42), *sesika* ‘stuffed’ < *səksək ‘stuff(v)’, and *solo* ‘substitute, deputy’ < *sulur ‘sprout, shoot’. It is probable that the change *s > Ø affected a large part, but not all, of the early MLG (or Southeast Barito) vocabulary before it ceased to be effective; later on many lexemes with *s* were introduced through borrowing.

2.1.5 *ə > e/i, a. The MLG correspondences *e*, *i*, and *a*, for PMP *ə continue to form another problem. Generally speaking, penultimate *ə is rendered as *e*, and final-syllable *ə is rendered as *i* in inherited vocabulary, for example, *ənəm ‘six’ > *enina*; *pusəj ‘navel’ > *foitra*; *təlu ‘three’ > *telo*; *bəlum ‘alive’ > *velona*. There are also loanwords with *e* for *ə: *harafesina* ‘rust’ (< ML *karat bəsi*, § 1.7); *vesatra* ‘heavy’ (< ML *bəsar* ‘big’, cf. § 2.2.3.5).²⁵

But PMP *ə can also correspond to *a*, and where this is the case, borrowing from ML seems to be more likely than heredity, for example,

**pə(nñ)u* ‘turtle’ and *fano* ‘id.’ (cf. ML *pəñu*); **kəmbar* ‘twins’ and *kambana* ‘id.’ (cf. ML *kəmbar*); **qampalas* ‘tree, the leaves of which are used for polishing’ and *ampaly* ‘shrub, the leaves of which are used for polishing’ (cf. ML *əmpəlas*); **səmbaq* ‘gesture of homage’ > *sambasamba* ‘expression of gratitude to God, benediction’ (cf. ML *səmbah* ‘gesture of worship or homage’); **cəcak* ‘k.o. lizard’ > *tsatsaka* ‘id.’ (cf. ML *cəcak* ‘id.’); **kənciŋ* ‘stench of urine’ > *hantsina* ‘stench’ (cf. ML *kəñciŋ* ‘urine’). Moreover, in most cases where PMP **ə* corresponds to MLG *a*, there are also other phonological indications of borrowing. The ending *-na* of *kambana* points to borrowing (see § 2.1.7), as does the fact that *sambasamba* has maintained +*s-* (see § 2.1.4), and that *tsatsaka* and *hantsina* both have *ts* (see § 2.2.1).

But why should these loanwords from ML have a double correspondence (*a* and *e*) for an alleged shwa in the lending language? The answer is that the double correspondence is due partly to the position of the sound in the loanword in question, and partly to borrowing from different ML source dialects.

An original +*ə* in antepenultimate syllables of ML loanwords seems to yield *a* in MLG (cf. *taranaka* ‘posteriority’ [<ML *tər(a)nak* ‘cattle; breed of any locality; native’, § 1.1], or *valahana* ‘loins’ [<ML *bəlakəŋ* ‘back’, § 1.10]).

But MLG lexemes with penultimate *a* corresponding to ML *ə* must be borrowed from a ML dialect where PMP penultimate **ə* became *a*. In most Malayic languages and dialects, **a* and **ə* merged in the last syllable; in a few of them (e.g. Banjarese ML, Minangkabau ML, and Brunei ML) this merger also took place in the penultimate syllable (Adelaar 1985). It is possible that Banjarese ML was the lending language for the MLG forms with *a*. Provided that South Kalimantan was already inhabited by ML speakers, and that their language had already undergone a merger of penultimate **a* and **ə* to *a* before the early Malagasy left their original homeland, one may conclude that pre-Malagasy first borrowed from Banjarese ML, and subsequently (after the early Malagasy left Kalimantan?) from ML dialects which still maintained **ə* in nonfinal syllables. In the latter case, ML *ə* is rendered as *e*. But the problem with this assumption is that not all forms in question have a Banjarese ML correspondence, and I am not able to give a satisfactory explanation for all MLG cases with *a* reflecting PMP **ə*. For instance, there exist Banjarese ML forms *kambar*, *hampalas*, *sambah*, *cacak*, with the same meanings as in Standard ML; but there are no Banjarese ML forms corresponding to *fano* or *hantsina*.²⁶ The missing Banjarese ML correspondences may have been replaced by other lexical items after borrowing into MLG had taken place. Their nonoccurrence in the Banjarese lexical sources (Abdul Hapip 1976, W. Kern n.d.) may also be due to incompleteness of the latter. Finally, we cannot exclude the possibility that some of the forms with *a* were borrowed from other

ML dialects, or that at a certain stage ML *a* was heard as *a*. Many uncertainties remain.

2.1.6 *a(R, s, Ø) > -y, -a. I have not been able to find much regularity in the reflexes of *a before word-final *R, *s, or zero. MLG usually shows -y for *-a in vocabulary that is unlikely to be borrowed, such as *pija ‘how much’ > *firy*, *lima ‘five’ > *dify*, *siwa ‘nine’ > *sivy*. But there is also *tunda ‘drag’ > *tundra* ‘load, freight’, which shows an inherited cluster *-ndr-* for *-nd-, but has a final -a. Compare also *tsiary*, *tsiadry* ‘no, not’: this could be borrowed from ML (cf. ML *tiada* ‘id.’, see § 3.4), which would mean that -y can also originate from a borrowed +a.

MLG shows both -y and -a for *-as, as in *qampəlas > *ampaly* (§ 2.1.5), *hery* ‘power, strength’ (according to Dempwolff from *keras, but probably borrowed, see § 2.2.3.5), *alas ‘forest’ > *ala* ‘id.’, *panas ‘hot’ > *fana* ‘id.’, *ləpas ‘free, loose’ > *lefa* ‘escape, shun, avoid’ (*lefasana* ‘deliverance’).

MLG has -y for *-aR in *i(N)paR ‘on the opposite side’ > *tan/dr/ify* ‘go to the opposite side’, *damaR ‘resin’ > *ramy* ‘id.’; but it has -a- in *DataR ‘flat, even’ > *ratana* ‘level country, a plain’ (cf. § 2.1.7).

2.1.7 MLG -na for expected Ø or +tra. In some cases MLG has -na corresponding to PMP *-q, *-r, *-R *-s, or *-Ø. Examples follow:

- *DataR ‘flat, even’ cf. *ratana* ‘level country, a plain’ (cf. ML *rata* ‘level, flat’, which is originally from JV)
- *kə(N)bar ‘twins’ cf. *kambana* (Sakalava MLG *hamba*) ‘id.’ (cf. ML *kəmbar*, Minangkabau *kamba*, Banjarese ML *kambar*)
- *lurus ‘slide down’ cf. Merina MLG *rorona* ‘go down the social scale’ (Dempwolff), Betsileo MLG *roroya* ‘go down’ (Dahl 1951: 73) (cf. ML *luruh* ‘dropping, being shed’; see Adelaar 1985: 105–106 for *-s > -h)
- *pañcar ‘jet of water’ > *fantsana* ‘rinse one’s mouth’ (cf. ML *pañcar* ‘gushing out; emanation’)
- *pañcur ‘jet of water’ > *fantsona* ‘waste-pipe’ (cf. ML *pañcur* ‘flowing along a conduit or pipe’)
- *sisiq ‘edge, side’ > *sisina* ‘edge, border’ (cf. ML *sisi* ‘side, edge’, but also ML *sisir/məñisir* ‘walk along something’)

Compare also:

- harafesina* ‘rust’ and ML *karat basi* ‘id.’ (§ 1.7)
- fasana* ‘grave’ and Maanyan *pasar-an* ‘id.’, *na-pasar* ‘bury’
- lamosina* ‘back’ and Minangkabau ML *lambusi*^a etc. (§ 1.10)

sakana ‘anything put across, any obstacle or impediment’ and ML *saŋkar* ‘diagonal, line across or athwart’

A tentative explanation of the *-na* in these examples is that it was acquired in some loanwords in the process of adaptation to MLG phonology. Nine of the ten instances presented here have correspondences in ML.²⁷

2.2 FALSE EVIDENCE FROM MLG FOR SOME DOUBTFUL PAN/PMP PHONEMES? Wolff (1974, 1982) articulated the doubts of some Austronesianists about the status of Dempwolff's *r, *c, *T, *g, *d, and *z. While I sympathize with the general idea of Wolff's argument, I cannot approve of his somewhat sweeping dismissal of Dempwolff's evidence for these protophonemes. A combination of detailed research into each of the languages apparently reflecting these protophonemes and the use of culture-historical and ecological data is required in order to decide on the legitimacy of *r, *c, *g, *d, and *z; *T has been rejected by others as well (Dahl 1977a:55–69).

Although it may be too early to take Wolff's views as proven, I would like to show that the MLG material lends much support to them. It is clear that MLG g cannot testify for *g, nor are there separate MLG reflexes for *d and *D (both have merged with *j to yield *r*, *-tra*, §2.1.2). In the case of *c, *z, and *r, there remain a few examples that cannot as yet be dismissed as loanwords or secondary developments, but their number is not very significant. It is not realistic to make the rejection of a protophoneme dependent upon the elimination of all possible evidence for this protophoneme, and it seems very likely that *c, *z, and *r are not reflected at all in MLG. Nevertheless, such a conclusion can best be made in conjunction with the investigation of reflexes of these protophonemes in other languages.

2.2.1 *c. Dempwolff reconstructed 23 etyma containing *c which have a reflex in MLG.²⁸ A large number of these etyma are given in Wolff (1982:3–6). But Wolff does not include etyma with *c as a component of a consonant cluster because he considers this to be an allophone of *s. Apart from Dempwolff, three instances of *c with a MLG reflex are found in Blust (1970) and one instance in Blust (1980).

Out of 27 instances of PMP *c corresponding to MLG *ts* (or *s*), the following ten must be borrowed from ML:

- *cœak ‘kind of lizard’ cf. *tsatsaka* ‘id.’ (§ 2.1.5)
- *cincinj ‘pull/tuck up’ cf. *tsintsina* ‘lift up’ (cf. ML *cigciŋ* ‘tuck or roll up [sleeves, curtains]’; Dempwolff gave ML *siŋsiŋ* as a correspondence)
- *cirit ‘discharge in a squirming manner’ cf. *siritra* ‘stench’ (A/M: ‘urine’) (cf. ML *cirit*, *ceret* ‘have diarrhoea’)
- *cuŋkil ‘poke’ cf. *tsoky* ‘pointed’ (cf. ML *cukkil* ‘prising up with a pointed instrument; digging up with a pointed stake’)
- *cupiŋ ‘ear-lobe’ cf. *sofina* ‘outer ear’ (§ 1.10)
- *guñcaŋ ‘shake’ cf. *hontsana* ‘rinsing (of bottles)’ (cf. ML *goñcaŋ* ‘shake’)
- *kəñciŋ ‘stench of urine’ cf. *hantsina* ‘stench’ (§ 2.1.5)
- *licak ‘soaked, drenched’ cf. *ditsaka* ‘id.’ (ML *lečak* ‘id.’)
- *pañcar ‘jet of water’ cf. *fantsana* ‘rinse one’s mouth’ (§ 2.1.7)
- *pañcur ‘jet of water’ cf. *fantsona* ‘waste-pipe’ (§ 2.1.7)

The following four instances are based on lexemes of a sound-symbolic nature which are not likely to be connected through inheritance or borrowing; these lexemes must be considered independent developments.

- *cit ‘cheep! chirrup!’ cf. *tsetra* ‘a click (of the tongue)’
- *cicit ‘cheep, chirrup (the voice of a bird)’ cf. *tsetsetra* ‘click one’s tongue (in approval)’
- *ciak, *ciap both with the meaning ‘cheep, chirp’ (Blust 1970) cf. *tsiaka* ‘the noise of something being torn; the chirping of little birds’

In seven cases, either the semantic connections between the correspondences are not convincing or the sound correspondences do not match, or both:

- *ca(N)gap ‘skim off’ > *tsaka* ‘scoop, fill with water’; the loss of *-p is irregular; the only other reflex given by Dempwolff is Tagalog *sagap* ‘skim off; a skimmer’
- *camuk ‘eat with caution’ cf. *tsamo/tsamo* ‘make noise while eating’
- *cinDuk ‘ladle; ladle out (v.)’ cf. *tsindroka* ‘food taken with the beak, lumps, fragments picked from the ground’
- *ca(N)kup ‘snap/bite at’ cf. *tsako* ‘chew (v.)’ (cf. ML *cakup* ‘catching in the open mouth (as a dog catches a biscuit thrown at it)’)
- *cium ‘kiss by sniffing; sniff’ cf. *tsio/tsiona* ‘name of a fragrant plant’ (cf. ML *cium* ‘kiss by sniffing; sniff’)²⁹
- *pœcaq ‘in pieces’ cf. *fetsaka* ‘collapsed, thrown down’ (cf. ML *pœkah* ‘in pieces’)
- *pœcut ‘whip’ > *fitsoka* (Dempwolff) ‘the sound of whipping’, (Wolff 1982) ‘horsewhip’ (cf. ML *pœcut* ‘whip’)

The six remaining cases do not have clear origins. The examples given may not be connected, or they may only be connected through borrowing. As a whole, they do not provide a strong argument that PMP *c was inherited in MLG, but some of them cannot be explained away (as with PMP *cəñcən > *tsentsina*, which has inherited reflexes for *ə). Compare:

- *cacaq ‘in pieces’ > *mi-tsatsà* ‘divided in equal parts’ (cf. JV, ML [but from JV] *cacah* ‘chopping up, cutting up into small pieces’)
- *cəlak ‘shine, beam’ > *tselaka* ‘flare up, sparkle’ (cf. ML, JV *cəlak* ‘eye cosmetic’, but also Minangkabau ML *bər-cəlak* ‘shining, glistening’ (Wilkinson 1959))
- *cəñcən ‘plugged, stopped up’ > *tsentsina* ‘id.’
- *cu(l)cul ‘insert’ (Blust 1980) > *solo/sòso* ‘slip through/inside with impudence’ (cf. Jakartanese ML *cocol* ‘dip (food into sauces etc.)’ (Blust 1980))
- *lac(aə)k ‘destroy’ (Blust 1970) > *latsaka* ‘fallen, poured; placed, put’ (the reflexes of *lac(aə)k given by Blust are from Philippine languages only)
- *lucak ‘mud deposit’ cf. *lotsaka* ‘submerged and about to dissolve’ (cf. ML *locak* ‘be disturbed [of water], slap about (of liquid in a moving vessel)’)

2.2.2 *z. All instances in the sample in which PMP *z is reflected by MLG z, j, are from Dempwolff. They are eleven in number (excluding those instances where *z was reinterpreted as *Z in Dyen (1951) and in later publications). I list them together with the ML correspondences wherever these are available:

- (a) *añzaŋ ‘long’ > MLG (Dempwolff) *anjana* ‘impressive, beautiful (stature)’, (Dahl 1951, Wolff 1982) *andrindrana* ‘erectness, uprightness of men’ (cf. ML *pañjaŋ*, *añjaŋ* ‘long’)
- (b) *qizaw ‘unripe’ > MLG *ma/itso* ‘green’ (cf. ML *hijaw* ‘green’)
- (c) *tañzuŋ ‘cape’ > MLG *tanjona* ‘id.’ (cf. ML *tañjuŋ* ‘id.’)
- (d) *tiñzak ‘tread on’ > MLG *tsinjaka* ‘dance with stamping of feet’ (cf. ML *tiñjak* ‘tread on’)
- (e) *zabi ‘kind of ficus tree’ > MLG *zavi* ‘id.’ (cf. ML *jawi/jawi*, *jə/jawi* ‘id.’, which is not given in Dempwolff or Wolff [1982])
- (f) *zahit ‘sew, put together’ > MLG *zaitra* ‘sewn’ (cf. ML *jait* ‘sew’)
- (g) *zəlag ‘brightness’ > MLG *jelaka* ‘glaringly bright’
- (h) *zəŋkal ‘span’ > MLG *zehy* ‘id.’ (cf. ML *jəŋkal* ‘id.’)
- (i) *zu(N)kuk ‘with bent head’ > MLG *joko* ‘approach quietly, stooped down’ (cf. ML *joŋkok* ‘squat, crouch’;

Dempwolff connected *zu(N)kuk with *jokoka* ‘stoop’, but this form is not found in A/M; Dahl (1951) and Wolff (1982) give *joko*

- (j) *zuluk ‘bore, pierce’ > MLG *joloka* ‘enter, force into’ (cf. ML *juluk* ‘bore, pierce’)
- (k) *zuŋzuŋ ‘support on head’ > MLG *jonjona* ‘proud’ (cf. ML *juŋjuŋ* ‘support on head’)

Wolff rejects the evidence for *z. With regard to c, d, e, f, j, i, I agree with him; *tanjona*, *tsinjaka*, *zavy*, *zaitra*, and *joloka* must be borrowed from ML, and *joko* is not connected with *zu(N)kuk. For b and k Wolff considers *ma/itso* and *jonjona* as unconnected to the supposed correspondences in other languages. Although he may be right, I would not exclude borrowing in these cases. With respect to g, I agree with Wolff that *zəlag is a false etymon. Its alleged MLG reflex *jelaka* ‘glaringly bright’ may be borrowed from ML (cf. ML *jəlah* ‘clear, unobstructed, of a view; distinct, of colouring’), although I do not have an explanation for its irregular ending. However, (a) *anjana* and (h) *zehy* remain a problem. The form *anjana* may be borrowed from ML, but this is somewhat contradicted by the fact that in ML *pañjañ* is (at least nowadays) far more usual. As far as *zehy* is concerned, it is hard to see how it can be borrowed, as Wolff suggests, since its *h* does not correspond with the medial cluster in the reflexes of *zəŋkal in other languages (cf. Tagalog *daykal*, Toba Batak *joŋkal*, ML *jəŋkal* ‘id.’; PMP *k > MLG *h*; PMP *ŋk > MLG *k*).

But, *anjana* and *zehy* notwithstanding, the MLG evidence is too scanty to serve as an argument for the reconstruction of PMP *z.

2.2.3 *r. After close inspection of the MLG lexical evidence for Dempwolff’s *r in our sample, it appears that only four instances (4.5%) out of 89 cannot be disproved, namely, *ga[r]aŋ, *rapət, *surak, and *ukur (see 2.2.3.6).³⁰ The MLG lexemes with *r* or *-tra* supposedly reflecting *r can be categorized into six groups.

2.2.3.1 In some cases *r/-tra* actually reflects *d, *D, or *j.

- *ela[r] ‘feathers’ > *elatra* ‘wing’ (cf. *əlaj, Dahl 1951:71–72; *qola(dDj), Charles 1973)
- *pagə[r] ‘fence, enclosure’ > *fahitra* ‘enclosure for cattle’
- *putə[r] ‘turn around’ > *fotitra* ‘reverse (v); incest’
- *waka[r] ‘root’ > *vahatra* ‘nascent (biological) colony’ (A/M: ‘kind of shrub’) (cf. *(w)akaD, Dahl 1951:71–72; *wakat, *əakaD, Zorc 1971)

*ruwaŋ ‘opening, hole’ > *roana* ‘abyss’ (cf. *Du(qw)aŋ, Wolff 1974: 83–84; note that *roana* could also be borrowed from ML and ultimately derived from *Ruqaŋ ‘intervening space’)

In these cases Dempwolff wrote final *r between square brackets because the evidence did not unequivocally point to *r. As a matter of fact, other comparative linguists have found additional evidence for altering the final consonant of *əla[r] and *waka[r]. Compare also Maanyan *elat* ‘wing’, *paket* ‘closed with a fence and prohibited to outsiders (of a village)’, and *wakat* ‘root’. For *putə[r] there is no evidence for another final consonant instead of *-r, but here too this consonant is ambivalent for *-r and *D.³¹

The following protolexemes probably also did not contain *r. They are based either only on MLG and Ngaju Dayak, or only on MLG and Tagalog. If inherited at all from PMP, they must reflect *d, *D, or *j. (The protophonemes *j and medial *d or *D become *r* in MLG and Ngaju Dayak, and medial *d or *D become *l* in Tagalog.)

- *ba[r]i ‘k.o. rice-dish’ > Ngaju Dayak *bari* ‘id.’, MLG *vary* ‘rice’ (cf. other Barito languages: Katingan *bari* ‘cooked rice’, Durdje Durasid et al. 1986:100; Bakumpai *bari* ‘id.’, Djantera Kawi 1985:125 sub *untal*)
- *ta[r]uj ‘information, announcement’ > Ngaju Dayak, *taronj* ‘news’, MLG *tarona* ‘interview, conversation’
- *qi[r]a ‘sing’ > Tagalog *hilaP*, MLG *h/ira*; this is a very badly attested etymon: MLG has an unexplained *h*-, and there seem to be no other correspondences for this set

2.2.3.2 In some cases Dempwolff reconstructed *r in etyma where it is now reinterpreted as *R (by Wolff 1974).

- *butir ‘grain’ > *votsy*, *botsy* ‘wart’ (cf. *but(ə)liR ‘wart’, Wolff 1974:99)
- *qiri ‘desire, envy’ > *iry* ‘desire’ (cf. *qiRi, Wolff 1974:86)
- *parapara ‘(under-)frame’ > *farafara* ‘bed-frame’ (cf. *paRa, Wolff 1974:80)
- *qampa[r] ‘spread, stretch’ > *ampatra* ‘stretched (of limbs)’ (cf. *qampaR, Dahl 1977a:48)
- *ribu ‘a thousand’ > *a/rivo* ‘id.’ (cf. *Ribu, Wolff 1974:80)
- *sulur ‘shoot, sprig’ > *solo* ‘substitute’ (cf. *suluR, Wolff 1974:99)
- *zuru ‘end point’ which, combined with *duRu ‘corner’, yields *ZuRu (Wolff 1974:92) > *zoro*, *andzoro* ‘corner’

PMP *R became MLG Ø or z (Dahl 1977a:87). MLG *iry*, *farafara*, *ampatra*, *rivo*, and (*and*)*zoro* are very similar in form and meaning to ML

iri ‘spite, malice’, *parapara* ‘shelf or framework built over the fireplace as a receptacle for cooking-pots’, *hampar* ‘spread’, *ribu* ‘thousand’, and *pəññ/ juru* ‘corner’, respectively, and they must be borrowed from ML.

The protoform *i(m)pa[r] ‘on the other (opposite) side’ is reinterpreted as *i(N)paR, with a variant form *d-i(N)paR (Dahl 1977a:107). The MLG reflex of the latter is *tan/drify* ‘vis-à-vis, on the opposite side’, which Dempwolff erroneously considered a reflex of *qaDep.

2.2.3.3 Dempwolff’s *r is not reflected by MLG *r/-tra*.

- *kə(N)bar ‘twins’, cf. *kambana* ‘id.’ (§ 2.1.7)
- *kun(D)ur ‘k.o. gourd’ cf. *hondro* ‘calabash’, *a/kondro* ‘banana’
- *lamba[r] ‘a textile’ cf. *lamba* ‘id.’
- *pañca[r] ‘jet of water’ cf. *fantsana* ‘rinse one’s mouth’ (§ 2.1.7)
- *pañcu[r] ‘jet of water’ cf. *fantsona* ‘waste-pipe, tube’ (§ 2.1.7)
- *qañir ‘fat, greasy’ cf. *l/ani* ‘smell of fish’, *m/ani* ‘stench’ (besides their irregular phonological correspondence, the meanings of the MLG items seem not to be connected)
- *ta(N)bi[r] ‘rim, flat dish’ cf. *tavy* ‘wash-basin’
- *Ta(N)kur ‘scrape with the hand’ cf. *tango* ‘cut (e.g. leaves) with the fingers or a knife, pull out (e.g. legs of insects) with the fingers’; *tango* is probably not related, as its medial cluster is irregular (usually *ŋk > k, and *k > h)

2.2.3.4 The MLG forms are not inherited from the alleged protoforms.

- *arak ‘intoxicating drink, arrack’ is a false protoform: ML *arak* and MLG *araka* are borrowed from Arabic, which has *Paraq* ‘sweat; juice; arrack’
- *qampir ‘be close’ cf. *ampirina* ‘arrangement of each thing in its right place; arranged next to each other’; Dahl (pers. comm.) points out that there is no connection: Dempwolff erroneously interpreted the ending of *ampirina* as an object-focus marker, whereas *ampirina* is underived (its object-focus form being *ampirim-ina*)
- *raba ‘touch, fumble’ cf. *rapa/rapa* ‘groping in the dark’: the sound correspondences are irregular, as *b > v, and *-a > -y (cf. also Wolff 1974:89, who finds the meanings unlikely to be connected); the Ngaju Dayak reflex *hawe* ‘have a presentiment’ points to a protoform *Raba
- *surud ‘recede’ cf. *tsororoka* ‘pour down from an elevated point, come down perpendicularly’: the meanings are difficult to reconcile, nor do the sound correspondences match, as *s > Ø or s, and *-d > -tra (the diagnostic reflex of *surud would have one syllable less: +(s)o(ro)tra)

- *surup ‘penetrate’ cf. *tsorofoka* ‘enter directly, force oneself into a place’: here too the sound correspondences and the number of syllables do not match (the diagnostic shape of a reflex would be ⁺(s)o(ro)ka); Dahl rejects this correspondence and considers *tsorofoka* as a derivation of MLG *tsofoka* ‘push/drive/knock in’ (Dahl 1951:68)
- *[t]ari ‘dance (v.)’ cf. *tary* ‘small drum on which a boy is placed at his circumcision’ (Dempwolff’s gloss for *tary*, ‘small dancing drum’ is wrong)
- *urub ‘burn’ cf. *oro* (cf. ML *m/urup* ‘bright red’, JV *urub* ‘be in flames’); this correspondence is irregular, as *-b would normally yield ⁺-ka; according to A/M *oro* is related to *doro* ‘set fire to’; a connection with *m/urup* and *urub* seems very unlikely

Note, too, *rumbia ‘sago palm’ cf. *rofia* ‘id.’ In MLG, *-mb- became -mb- (and *-a became -y), and the diagnostic reflex of *rumbia would have been ⁺(r)omby. Wolff pointed out that the sago palm is quite well known in some areas, but that it is only sporadically found in many other areas of Indonesia and the Philippines, and that it is denoted by a large number of doublets and analogical reformations (Wolff 1974:83). He rightly concludes that *rumbia cannot be evidence for *r. Dempwolff reconstructed *paraqu ‘kind of boat’, but its supposed reflexes (ML *pərahu* ‘id.’) are ultimately borrowed from Tamil (Wolff 1976:362; cf. also Burrow and Emeneau 1961:lemma 3175). In Flacourt (1658) Tanosy MLG *paraho* ‘id.’ is found. If this form really occurs, it must be a post-“substratum” term which entered the Tanosy dialect via ML.

2.2.3.5 There are lexemes that are, or could be, borrowed from ML. More than half of the lexemes allegedly reflecting *r belong to this category.

- *bəsar ‘big’ cf. *vesatra* ‘heavy’ (cf. ML *bəsar* ‘big’ (§ 2.1.5))
- *bu(N)kar ‘dig, uproot’ cf. *vokatra* ‘harvest, crop’, Sakalava MLG *vokatse* ‘brought to the surface, dug up (e.g. from the grave)’ (Dahl 1951:336; ML *boŋkar* ‘heaving up, turning up the soil’)
- *burit ‘rear’ cf. *voritra* ‘perineum of cattle’ (cf. ML, JV *burit* ‘hind-part; stern’)
- *buruŋ ‘bird’ cf. *vorona* (cf. ML *buruŋ*); the spread of correspondences is very much restricted to Kalimantan languages, ML, and MLG³²
- *burut ‘hydrocele, hernia’ cf. *vorot/ina* ‘hydrocele, swelling of the testicles’ (ML *burut* ‘hernia, hydrocele’)
- *cirit ‘discharge in a squirting manner’ cf. *siritra* ‘stench’ (§ 2.2.1)
- *garunγaŋ ‘k.o. tree’ cf. *harongana* ‘Haronga madagascar’ (see § 2.1.1)
- *girik ‘pierce, bore’ cf. *hirika* ‘pierced, bored’ (§ 2.1.1)

- *guris ‘scratch’ cf. Betsileo MLG *hory* ‘gnaw’ (Dahl 1951:55; ML *gores*, *guris* ‘scratch’, *koris* ‘scratch with a point’)
- *guruq ‘boom, thunder’ cf. *koro* ‘confused’ (Dempwolff), *horo/horo* ‘trepidation, fear’ (Dahl 1951:348; ML *guruuh* ‘thunder; thundering noise’)
- *i[r]in ‘accompany’ cf. MLG (provincial) *irina* ‘fringe, rim (of a textile)’ (Dahl 1951:337; cf. ML *irij* ‘accompany’)
- *karaj ‘coral-reef, coral-rock’ cf. *harana* (cf. ML *karaŋ*)
- *karaq ‘shell’ cf. *hara* ‘mother-of-pearl’ (cf. ML *karah* ‘scale, metal sediment’)
- *karun ‘bag’ cf. *harona* ‘basket’ (cf. ML *karug* ‘bag’)
- *keras ‘hard’ cf. *hery* ‘power, strength’ (cf. ML *kəras* ‘hard, tough’)
- *kikir ‘file, stinginess’ cf. *hihitra* ‘stinginess’ (ML *kikir* ‘file; stinginess’); ML *kikir* is the only correspondence that combines the meanings ‘file’ and ‘stinginess’: except for MLG, the other languages used by Dempwolff have ‘file’; the meaning ‘stinginess’ is a secondary semantic development in ML; this is another indication that *hihitra* is a ML borrowing
- *kurun ‘cage’ cf. *horona* ‘closed in’ (cf. ML *kurug* ‘encaged’)
- *lampar ‘spread, stretched’ cf. *mi/lampatra* ‘spread legs or arms’, *va-lampatra* ‘with spread legs’ (Dahl 1951:329; cf. ML *lampar*, Maanyan *ba/lampar* ‘be spread’); Dahl reconstructed *lampar as a parallel form of *qampar; the alleged reflexes are obviously borrowed from ML, where *lampar* originated through regular regressive dissimilation of *r to l and subsequent back-formation (thus, ⁺bər + amp̥ar > ⁺bəlamp̥ar > bə-lampar); MLG *va/lampatra*, Maanyan *ba/lampar* are borrowed versions of the ML derived form *bəlampar*³³
- *laraŋ ‘prohibit’ cf. *rarana*, *rara* ‘prohibited’ (cf. ML *laraŋ*)
- *laru ‘means to purify fermenting liquids’ cf. *laro* ‘a tree, the juice of which is used as a fish-poison’ (cf. ML *laru* ‘tree-extract used to purify palm-wine’)
- *luar ‘outside’ cf. *loatra* ‘extract (v.)’ (Dahl 1951:54; ML *luar* ‘outside’)
- *lurus ‘slide down’ cf. Merina MLG *rorona* ‘go down the social scale’, Betsileo MLG *roroŋa* ‘go down’ (§ 2.1.7)
- *pa[r]ij ‘fortification’ cf. *faritra* ‘mark, sign to delimit something; frame, delimitation, outline’ (Dahl 1951:330; cf. ML *parit* ‘groove, cutting moat’)
- *pirak ‘silver’ cf. *firaka* ‘tin’ (cf. ML *perak* ‘silver’)
- *radrad ‘expand’ cf. *raratra* ‘expansion’ (cf. ML *larat* ‘become larger’)
- *ra(N)gas ‘defoliate’ cf. *raka/raka* ‘undo (one’s hair)’ (cf. ML *ragas*, *ranggas*, ‘defoliated’); besides possible borrowing from ML, the irregular final ending (*-as usually becomes -y) and the semantic difference between *raka/raka* and *ra(N)gas are also strong arguments for rejecting a relationship through common inheritance
- *rambu ‘tassel, fringe’ cf. *rambo* ‘id.’ (cf. ML *rambu* ‘id.’)

- *ra(N)bus 'tear, snatch away' > *rambos/ana* (with a variant form *rambon/ana*) 'what is quickly torn or snatched away' (cf. ML *bər-ambus* 'leave hurriedly, move away fast' (cf. Iskandar 1970; not found in Dempwolff); the origin of *rambos/ana* was probably a back-formation through reinterpretation of *bər-ambus* as ⁺*bə*-*rambus*
- *ra(N)pak 'crumble' cf. *rapaka* 'put fragments into the mouth, seize fragments with the mouth' (a somewhat arbitrary semantic connection) (cf. ML *rapak* 'fragile, brittle, crumbling')
- *raprap 'take off, strip' cf. *raraka* 'id.' (cf. ML *larap* [with regular regressive dissimilation] 'in great demand, selling well; easily disposed of', also *larap-an* 'rest (for the chisel-holder)', *mə-larap-kan* 'open out (the strands of a rope)')
- *raqup 'collect, pick up by hand' cf. *raoka* 'collected by hand', *raofina* 'things to be picked' (cf. ML *raup* 'both hands full', *mə-raup* 'scoop with both hands')
- *rasras 'break/tear open' cf. *raratra* 'loosen, let go, break away' (cf. ML *laras* [with regressive dissimilation] 'fall, as leaves from a tree, also of a remark that a man has "let slip" without thinking'); the ending (-*tra*) of *raratra* does not reflect PMP *-s, and the form may therefore not be related
- *raut 'tear off, split, cleave' cf. *raotra* 'pull out grass' (cf. ML *raut* 'smoothing, cutting off the asperities of anything' (Dempwolff: 'splitting knife'))
- *repuk 'crumble (v.)' cf. *refo* 'brittleness' (Dempwolff), *ka/repoka* 'cracking noise' (Dahl 1951:67; cf. ML [Klinkert 1916] *rəpoh*, *rapuh* 'broken up easily, brittle, fragile', also *rapeh*, *rəpas*, *rəpuy* 'brittle, fragile', and *rapok* 'flop, pop; to go pop')
- *retak 'burst (v.)' cf. *retaka* 'collapse' (cf. ML *rətak* 'crack [n.]')
- *riak 'stirred water' cf. *riaka* 'pouring rain, torrent' (cf. ML *riak* 'ripples on water')
- *riam 'a rapid' cf. *riana* 'id.' (cf. ML *riam* 'id.')
- *ribut 'storm' cf. *riovtra* 'storm, wind' (cf. ML *ribut* 'id.')
- *rumbak 'destroyed, broken', cf. Sakalava MLG *robaka* 'destroyed' (Dahl 1951:338; cf. ML *rombak* 'destroyed, broken')
- *rumput 'grass', cf. *rompotra* 'manioc leaves' (Richardson 1885), 'pull out [weeds] by hand' [A/M]; cf. ML *rumput* 'grass, weeds'
- *rusak 'destroy' cf. *rotsaka* 'fall downwards (etc.)' (§ 1.3)
- *sandar, 'reclining, resting on a support' cf. *sandratra* 'elevated' (Dahl 1951:71; cf. ML *sandar* 'reclining, resting on a support')
- *sarak 'separated' cf. *saraka* 'separated, abandoned' (cf. ML *sarak* 'separated')
- *sarun 'sheathe' cf. *sarona* (cf. ML *saruŋ*)
- *surat 'write, writing' cf. *soratra* (§ 1.8)
- *surun 'push/shove forward' cf. *sorona* 'shove into', *ant/sorona* 'push forward' (Dempwolff), *sorona* 'stir up a fire (= push forward

- the embers), sacrifice, offering to the ancestors (= push forward food to them)’ (Dahl 1951: 330; cf. ML *soroy* ‘pushing forward along the ground, pushing out’)
- *tarik ‘pull’ cf. *tarika* ‘id.’ (cf. ML *tarik* ‘id.’)
 - *[t]aruk ‘shoot, sprig’ cf. *taroka* ‘sprout, shoot’ (cf. ML *taruk* ‘shoot, sprig’)
 - *urut ‘rub’ cf. MLG (provincial) *orostra* ‘pull out (leaves, grass) with the hand’ (cf. ML *urut* ‘massage’)
 - *zambuara[q] ‘whale’ cf. *lambuara* ‘a species of fish’ (cf. ML *ləmbuara* ‘a giant fish [possibly a whale]’; cf. also Old Javanese *ləmbwara*, *ləmbora* ‘a very large fish [whale? porpoise?]’)

There are other correspondence sets (from Dahl 1951) which, according to Dempwolff’s sound correspondences, attest PMP *r:

- MLG *sara* ‘pay, rental, wage’ cf. ML *sara* ‘livelihood, means of support, provisions’ and *sərah* ‘making over to another, surrendering’, and Ngaju Dayak *sarah* ‘refund, restitution’, Maanyan *nasarah* ‘refunded, restituted’ (Dahl 1951: 335)³⁴
- MLG *toraka* ‘throw, launch’ cf. ML *torak* ‘spool, rolling-pin’ (Dahl 1951: 334)

2.2.3.6 There are MLG reflexes of *r that cannot be disproved. In the following instances there is neither sufficient semantic nor sufficient phonological agreement between the MLG and ML correspondences for a borrowing hypothesis.

- *ga[r]aŋ ‘hot’ > *karana* ‘dried out’ (Dahl 1951: 47; ML *garaj* ‘fierce, turbulent (of colour), loud or glaring’); the meanings of *karana* and *garaj* seem hard to reconcile (*karana* agrees more with JV *garaj* ‘dried, roasted’)
- *rapət ‘join together’ > *rafitra* ‘trade, craft’ (cf. ML *rapat* ‘tight-fitting’); the different meanings of *rafitra* and *rapat*, and the *i* in the former, seem to exclude borrowing
- *surak ‘cheer, shout’ > *oraka* ‘cry, shout (n.)’ (Dempwolff), *horaka* ‘a shout (of astonishment, admiration, repulsion’ (Dahl 1951: 340; according to Dahl, *horaka* lost *s- and subsequently pre-fixed *ka- (*surak > *urak > *ka- + *urak(a) > *horaka*) (cf. ML *sorak* ‘cheering, applause’)
- *ukur ‘measure (v.)’ > *ohotra* ‘preparation’ (Dempwolff’s gloss: ‘(Ermessen =) Vorbereitung’); Blust (1970) reconstructed *(q)ukur ‘luck, fortune’; a gloss ‘calculation, estimation’ would cover the scale of meanings reconstructed by Dempwolff and Blust; in view of the different meanings of *ohotra* and *ukur* ‘linear measurement’, borrowing from ML is not very obvious

3.0 CULTURE-HISTORICAL INFERENCES

3.1 DATE OF THE MIGRATION. It is rather striking that out of the 30 MLG lexemes identified as Sanskrit loans by Dahl, there is only one that is not also found in ML or JV (Dahl 1951:97).³⁵ These lexemes often underwent the same semantic and phonological alterations as they did in ML or JV. Dahl (1951:367) concludes from these facts that the ancestors of the Malagasy people must have left Kalimantan at a time when Indian influence was just beginning to make itself felt on the island, which is roughly 400 A.D.

But the evidence should be interpreted in another way. Since almost all Sanskrit loanwords in MLG also occur in ML or JV, and since ML and JV had a strong influence on MLG, the conclusion presents itself that the Sanskrit loanwords in MLG were directly borrowed from ML and JV. This fits in with the picture of contacts of the early Malagasy people with South Sumatra, where an Indianized Malay kingdom was established (see below). It also fits in with a similar pattern for borrowing found in Philippine languages, where Sanskrit loanwords were apparently also introduced through ML along with many other loanwords (Wolff 1976:347). Malay must have been an important language with much influence on other Southeast Asian vernaculars at a very early date. A comparison of the Malay loanwords in Philippine languages, Old Javanese and MLG, will certainly contribute to a better insight into the nature of the cultural influence of the Malays on other Southeast Asian ethnic groups in pre-colonial times.³⁶ Finally, Bernard-Thierry (1959:340) notices that the few Sanskrit loanwords in MLG do not belong to any particular semantic subset, and that this is in contradistinction to Indian influence on Southeast Asian languages. In the latter, Sanskrit loanwords often pertain to distinct semantic fields such as religion, art, philosophy, or science. On the assumption that Sanskrit loanwords in MLG entered the language through ML (and JV), the cause of the random spread of these “secondhand” loanwords becomes evident.

As a consequence of borrowing from Sanskrit via ML and JV, Dahl's criterion for dating the migration of the early Malagasy does not hold. It is not the first Indian influence in South Kalimantan which is significant for this dating, but the period of Malay influence on early MLG. The migration must have taken place at a period when the Malays had already undergone Indian influence, and when they had already established contacts with other insular Southeast Asian peoples, including the Dayaks of the Barito area. It is as yet impossible to locate this period with great exactitude. From the Old Malay inscriptions of South Sumatra, we know that Malay had already undergone Sanskrit influence by the end of the seventh century A.D. But this linguistic influence may have started earlier, as

the first traces of Indian influence in Indonesia are from the fifth century A.D.

Other indications for a migration date are the South Sumatran inscriptions, and some Arabic and JV loanwords in MLG. But the problem with both these inscriptions and loanwords is that they point to Malay–Malagasy contacts which could have taken place either before or after the migration.

In the South Sumatran inscriptions, there are, along with the Old Malay text, a few lines in an unknown language which shows some striking agreements with the Southeast Barito languages and MLG. Moreover, the Old Malay lines of the inscriptions also have some lexical agreements with MLG. But since these agreements may also point to postmigratory contacts, or even to contacts between South Sumatran Malays and Southeast Barito speakers other than the early Malagasy, the South Sumatran inscriptions cannot be considered as decisive evidence for dating the migrations (see § 3.4 for a further account of the inscriptions).

The forms *sombily* and *setra* are Arabic loanwords which were borrowed into MLG through ML (§ 1.5), and *rotsaka* was ultimately borrowed from JV (§ 1.3). If these loanwords were already part of the lexicon of the early Malagasy migrants, this would imply that their migrations took place after the rise of an influential Javanese culture (which was probably in the ninth century A.D.; Pigeaud 1967: 1) and even after the introduction of Islam in Indonesia (probably at the end of the 13th century A.D.; Drewes 1968: 457). But irrespective of the relative arbitrariness of these dates, these loanwords could as well have entered MLG after the migration. There is no way of measuring the relative antiquity of *sombily* and *rotsaka*, because none of the diagnostic sound changes due to the Bantu “substratum” apply to them. The form *setra* does exhibit the effects of the Bantu substratum, as is shown in the last consonant (*tr*), and therefore it must be a relatively old loanword. In fact it could already have been borrowed before the early Malagasy left Kalimantan, as corresponding forms are found in Maanyan and Ngaju Dayak. This would be evidence for Arabic influence in the language of the early Malagasy migrants. But it is rather speculative to establish a migration date on the basis of one etymology only, and besides, there is no way of determining with any precision when the Bantu influence was operative. If it was active at, or right after, the arrival of the early Malagasy in East Africa, *setra* may have been part of their lexicon. But if the Bantu influence is a relatively recent phenomenon, *setra* can only be proven to have entered the language some time previous to this influence (which borrowing may have occurred before or after the migration), and to be relatively older than words that did not undergo it.

In conclusion, then, there is no clear indication as to when the early Malagasy left Kalimantan. Although the fifth century A.D. as a migration

date is not excluded, the earliest traces of Indian influence in Indonesia cannot be taken as critical. Loanwords from Sumatran Malay, Arabic, or JV, and agreements between the languages of the South Sumatran inscriptions and Southeast Barito languages or MLG are not critical either. All that can be said with any certainty is that the migration must have taken place after Indianized Malays had begun a mercantile and/or political expansion (in the process of which they came into contact with the Barito area) and that it is in principle not excluded that it took place at a time when Javanese culture and Islam had begun to play an important role in insular Southeast Asia. This covers a time span of eight centuries (from the fifth till the 13th century A.D.).³⁷

There is, however, an argument in favor of positing the migration in the seventh century A.D. at the earliest, although the evidence is admittedly not crucial. The argument is that the early Malagasy must have been in contact with Srivijaya, and that they may have learnt about the existence of Madagascar via these contacts. The ancient kingdom of Srivijaya, which was situated in South Sumatra (from about the seventh to the 13th century A.D.), is an important argument in favor of a migration date in the seventh century A.D. at the earliest. Srivijaya was an important trade center between India and China. It is one of the oldest recorded Indianized kingdoms in insular Southeast Asia, and it is likely that the early Malagasy were in contact with it. Sumatra was one of the geographical obstacles on their way to East Africa, and they had to circumnavigate it through either the Sunda Straits or the Straits of Malacca in order to reach the Indian Ocean. As it is extremely unlikely that they headed for East Africa without knowledge of their final destination, they must have obtained information from people who knew the Indian Ocean. Srivijaya was a cosmopolitan and maritime power with contacts and settlements along the Indian Ocean, and the Srivijayans must have known of the places and regions lying at the southwestern extreme of the Indian Ocean trade route. The presence in MLG of ML loanwords which are typical for the cultural and geographical circumstances in Sumatra (namely, *ranto* and names of winds and wind-directions, § 1) support the theory that the early Malagasy had contacts with South Sumatra.³⁸

A migration between the seventh and the 13th century A.D. is also in line with anthropological insights regarding this matter. Paul Ottino estimates the migration date between the eighth/ninth century and the 12th/13th century A.D. He believes that the occurrence of Indianized influence in Malagasy culture (as found, for instance, in the organization of power and concepts of the State) along with authentic Indonesian aspects (death rituals, sexual symbolism, eschatological beliefs) are due to two different forms of culture which the early Malagasy brought with them, an aristocratic and a popular one. Rather than attributing these forms to different migration waves (as Ferrand [1932] and Deschamps [1936] did), Ottino

assumes that there was a socially stratified group of immigrants to Madagascar (Ottino 1974: 14–19).

With the few data at our disposal, a migration in or after the seventh century A.D. seems more probable than one before that date. This assumption, however, remains hypothetical.

3.2 POSTMIGRATORY CONTACTS. However, what the data (the inscriptions and loanwords) do demonstrate is that Malagasy remained in touch with Southeast Asia for a considerable time after their migration. This is shown by the fact that there are ML loanwords which did not undergo the Bantu “substratum” (such as *toggolany* [§ 2.1.1], *bodo*, *landaizana*, and *lindona* [§ 2.1.2]). The continued (or renewed?) contacts covered a timespan long enough for a set of sound changes to have affected the Malagasy lexicon (the Bantu “substratum”). Since it is not known when the Bantu “substratum” was operative, there is no way of dating loanwords on the basis of whether or not they have undergone the Bantu influence. On the other hand, in view of the Arabic loanwords which entered MLG via ML, the postmigratory contacts must have lasted till after the introduction of Islam to Indonesia. There is good reason to pay more attention to non-linguistic evidence for continued contacts between Madagascar and Southeast Asia until the arrival of the Portuguese.³⁹ This is an important point, because many Austronesianists take it for granted that MLG is a kind of fossil language which had never undergone any significant borrowing prior to its migration, and which afterwards developed in complete isolation from other Austronesian languages.

The data also demonstrate that there must have been contacts with South Sumatra (see previous paragraph).

3.3 THE NATURE OF ML LOANWORDS. As a preliminary observation, it appears that ML loanwords in MLG often pertain to a maritime environment (which includes names of winds and directional terms) and terms for body-parts.

The loanwords treated here were selected on the basis of the unequivocal evidence of borrowing which they show, or on their alleged reflexes of PMP/PAN protophonemes. A study of the semantic fields covered by these borrowings has yet to be made. Such a study would include dialect research in Madagascar and language research in South Kalimantan.

3.4 EVIDENCE FROM THE SOUTH SUMATRAN INSCRIPTIONS. The Old Malay inscriptions represent a language which is the oldest form of ML. It has not undergone any Arabic influence (unlike all later forms

of ML), and it is recorded from a number of inscriptions found in South Sumatra, on the island of Bangka, and (some later ones) in Java. The inscriptions of Kota Kapur and Karang Brahi, two almost identical inscriptions from 686 A.D. (one on Bangka, the other in the province of Jambi, South Sumatra), contain a few lines in an unidentified language. Aichele pointed out that this language shows some striking agreements with Maanyan (Aichele 1936; Dahl 1951: 158; cf. also Damais 1968). As a matter of fact, if we allow for the scanty material and if (as is implied by Coedès [1931] and other scholars) *paihumpaan* means 'imprecation' and is related to PMP *sumpaq 'id.', *vinunu* is an object-oriented verb form related to PMP *bunuq 'kill', and *ulun* means 'human being', then this language shows *h* for PMP *s, Ø for PMP *q, and a replacement of an original PMP *tau(-mataq) 'human being' by *ulun*. It has these developments in common with the Southeast Barito languages and with MLG. Compare MLG *ompa* 'imprecation'; Maanyan *wuno*, MLG *vono* 'killed'; Maanyan *olon*, MLG *olona* 'human being'. Coedès (1931: 50) has already remarked on the absence of *s* and the abundance of *h* in these lines. In connection with *paihumpaan*, Dahl (1951: 159) notices that according to the regular sound changes shown in Maanyan, a Maanyan form corresponding to ML *pərsumpahan* would appear as ⁺*pajhumpaan* (the ⁺j here designates a palatal semivowel). The lines also contain two *-ndr-* clusters (in *tandrun* and *kandra*): these clusters also occur in MLG, where they are the regular reflex of PMP *-N(dDj)- (Dahl 1951: 52).

Before Aichele, Ferrand (1932) collected linguistic data from the Old Malay inscriptions in order to show a direct genetic line from Old Malay to MLG. He treated the material in a rather random way, and he failed to make a distinction between the language of the first lines of the Kota Kapur and Karang Brahi inscriptions, and the Old Malay of the other lines. But in spite of his own basically wrong theory, he collected a number of Old Malay words, some of which have correspondences only in MLG. The correspondence sets in question are *sa:mvaw* 'boat', MLG *sambo* 'k.o. boat'; (-*ta:pik*) *mana:pik* 'to organize a punitive expedition', MLG *tafika* 'army, expedition'; and *ti:da* 'no, not', MLG *tsy* 'no, not', *tsi/ary* 'no, nothing'. The origin of *sa:mvaw/sambo* has been traced back to Khmer (Porée-Maspero 1986: 76). I disagree with de Casparis (1956: 352) on a relation between *mana:pik* and Minangkabau *manape* 'to defend oneself against, to resist' (rather than between *mana:pik* and *tafika*) because of their phonological and semantic incompatibility (cf. also ML *tapik* 'to strive' [Howison 1801]). As far as *ti:da* is concerned, it is also used in Standard Malay, witness *tida/k* 'no, not' (*tiada* '[there is] not'), but the Malay vernaculars almost never have *ti/-* as the first sequence of a negator.⁴⁰ They have forms like *nda*, *ənda*, *nda?*, *ənda?*, *da?*, *ŋga?*, *ka/da*, *ta?*, *t(a)ra*. On the other hand, *tsy* (< early MLG ⁺*ti*) is the usual MLG

negator. The forms *ti/-* and *tsy* (or *tsi/-*) have no correspondences in other languages, and they must originally be the same lexeme. The forms *sambo*, *tafika*, and *tsy* (*tsi/-*) may be explained as Old Malay words which only survived as loanwords in MLG—although in view of the lack of correspondences in modern forms of ML, borrowing in the opposite direction should not a priori be ruled out.⁴¹

NOTES

1. This applies to any other language known in 1951. In a later publication (Dahl 1977b), Dahl concludes that Malagasy is most closely related to the Southeast Barito language group (to which Maanyan belongs), but that it cannot be decided which Southeast Barito language in particular is closest to Malagasy.
2. Dahl refers to three stones with Sanskrit inscriptions from c. 400 A.D. which were found in Muara Kaman, near Kutei (Dahl 1951:368, referring to H. Kern 1882 and Vogel 1918).
3. Dahl (pers. comm.) now believes in postmigratory borrowing. J. U. Wolff (1974:80; 1982:4 and n. 7) assumes that Malagasy borrowed from Malay, but he does so as a result of his ideas on PAN phonology, and he never studied the matter as such. Furthermore, he does not attempt to give a chronology of the borrowing in question.
4. I am very grateful to Prof. Dahl for his useful comments and criticism on earlier drafts of this publication. I would also like to thank Jim Collins, Jack Prentice, and Hein Steinhauer for their careful reading and comments.

The following symbols are used in this article:

- * (at the upper left corner of a word) = belonging to a protolanguage
- + (at the upper left corner of a word) = belonging to any stage between the attested language and the protolanguage
- > = “developed into”
- < = “developed from”
- / (in a quoted lexeme) indicates diachronic morpheme boundaries

Phonemes, morphemes and words quoted from languages under discussion are italicized, except when they are reconstructed.

The terms “correspond,” “correspondence,” and “corresponding” refer to a common origin, either through inheritance or through borrowing; if two or more lexemes are known to have a common origin through inheritance, they are referred to as cognates.

PAN and PMP reconstructions are represented in Dyen’s orthography as modified by Blust.

Except when indicated otherwise, the following language sources are used in this article: Acehnese (Hoesein 1934), Arabic (Wehr 1976), Banjarese ML (Abdul 1976), Fijian (Dempwolff 1934–1938), Iban (Richards 1981), (modern) Javanese (Jansz 1913), Lampung (Junaiyah 1985), Maanyan (Hudson 1967; Dahl 1951), Malagasy (Abinal and Malzac 1963, Dempwolff 1934–1938, Dahl

1951), Malay (Wilkinson 1959), Minangkabau (M. Thaib 1935), Ngaju Dayak (Hardeland 1859), Old Javanese (Zoetmulder 1982), Persian (Steingass 1930), Sanskrit (Gonda 1973), Sundanese (Eringa 1984), Tagalog (Panganiban 1966), Toba Batak (Warneck 1977).

5. Wolff (1982:4, n. 7) has already pointed out that MLG *ranto* must be borrowed from ML on account of its extended meaning.
6. Although no such compound is attested for Old Javanese (cf. Zoetmulder 1982).
7. Cf. Antilla (1972:68); cf. also Adelaar (1985:90, 191) for this phenomenon in Malayic dialects and languages.
8. Blust (1977:31) adds to Dempwolff's evidence Kadazan *handasan* (Antonissen 1956) and Cebuano *landasan* (Wolff 1972), but these correspondences are subject to the same semantic and phonotactic suspicions as *landaizana*. Both of them only mean 'anvil', and neither of them is derived, whereas the great majority of lexemes is disyllabic in the languages in question. Correspondences of ML *landasan* seem to prevail in languages that underwent ML influence and that are spoken by peoples without indigenous iron work. Conversely, in cultures where iron work is an independent development, terms pertaining to iron work are usually very different from the ML ones (cf. Maloh [West Kalimantan] *lalajan* 'anvil' [my own field material]).
9. Dempwolff's PMP *surat was never questioned by Dyen, and Blust (1977:33) took this etymon as evidence for the existence of writing independent of Indian influence in West Indonesia. Dahl (1977a) offered evidence for PAN *surat from Taiwanese languages (see note 10), but Dahl (pers. comm.) specifies that this etymon must have had a meaning (such as drawing or scribbling) which was later on associated with writing in the Austronesian daughter languages. Wolff (1974:86) rejects *surat on account of its sound correspondences.
10. From the recently published diaries of two merchants who visited the village of Soulang in 1623, it seems that the Siraya-speakers there attributed the occurrence of Malay-like words in their language to foreign sailors who established themselves in their area (Blussé and Roessingh 1984). These sailors were apparently from Johore, on the Malay peninsula. Consider the following quotation from one of the journals (Blussé and Roessingh 1984:74–75, text B):

They use a language that is foreign in itself, mixed with Malay and other foreign or outlandish words, as *Babij*, *tacot*, *Boesock*, *maccan*, *Ican* etc., also *Maccselo* and *Mapilo*, which are used in the island of Talao, situated about Sangij.⁽²²⁾ What sort of dialect this is, is unknown to me, but I understood that it was derived from the *Joristen*,⁽²³⁾ sailing from these islands and coming ashore here by accident, who stayed here and created progeny. But whether this is true or not, I refer to veracity.

(In ML, the above "foreign and outlandish words" occur as *babi* 'pig', *takut* 'afraid', *busuk* 'rotten', *makan* 'eat', and *ikan* 'fish'. The words Talao and Sangij refer to the Talaud and Sangir islands, north of Sulawesi. Note 22 says that *Maccselo* is possibly a corruption of Sinkan/Siraya *mackdiou* 'to make friends', and that *Mapilo* may be Sinkan/Siraya *mapil* 'to return'. Note 23 tentatively explains *Joristen* as Malay sailors from Johore.)

Although one should be critical regarding the details they contain, these diaries are an indication that there were contacts between Taiwanese and Malays. As the Austronesian-speaking peoples of Taiwan originally were not acquainted with writing, they may well have first heard of it from Malays.

Another possibility is that correspondences of *surat* in Taiwan were introduced by Dutch Calvinist missionaries. These introduced writing and put much emphasis on catechization, and furthermore, some had preached in the Moluccas before they came to Taiwan.

Dahl (1977a:62) gives correspondences of **surat* from Paiwan, Rukai, Saaroa, and Siraya. A look at the linguistic map of Taiwan (Tsuchida 1976:xxx) shows that these languages are all spoken in the Southern part of the island, and Tsuchida believes that the Saaroa, Rukai (and Kankanabu) correspondences are borrowed (from Siraya, judging from the *s* they exhibit) (Tsuchida 1976:180). Siraya must in turn have borrowed it from a non-Formosan Austronesian language. Blust (1977:33) gives Kavalan *s-əm-uläl*, but PAN *r seems not to be attested in this language (Li 1982: passim) and the final consonant does not correspond (PAN *t, *T, *C > Kavalan *t*, Li 1982:488). Therefore, *s-əm-uläl*, if related at all, cannot be related through common inheritance.

11. Blust (1977:33) takes MLG *soratra* as evidence for the presence of reflexes of **surat* at the time of the MLG migration. If by "reflexes" he means inherited lexical items testifying to the existence of ancestral forms in the protolanguage, I do not agree. Nor do I agree with Blust's statement that "reflexes of **surat* exclusively and invariably refer to writing" (Blust 1977:33): the ML, JV, and MLG correspondences have a much wider meaning.
12. Flacourt (1658) gives Tanosy MLG *simoulots* 'northeast', which, in Ferrand's comments in his 1905 re-edition of Flacourt, becomes *tsimilotra* 'north wind'.
13. Dahl (1977a:95) points out that besides ML, Madurese may also have had influence on the languages of South Kalimantan, as South Kalimantan was easily accessible for such a maritime nation as the Madurese. There are, however, no important historical or linguistic arguments for Madurese influence. Moreover, Madurese *bharat* would probably have yielded *a/baratra instead of *a/varatra*, since Sanskrit loanwords in MLG reflect Sanskrit *b* as *v*, and Sanskrit *bh* as *b* (cf. Dahl 1951:104).
14. Banjarese (= South Kalimantanese) ML *barat* means 'heavy', and neither Abdul (1976), Ras (1968), W. Kern (n.d.), nor Durdje Durasid and Djantera Kawi (in the final part of their 1978 outline of Hulu Banjarese) has correspondences for the directional terms *barat*, *barat-daya*, or *timur-laut*.
15. Hudson (1970:70); Dahl (1951:53) has *pongon*.
16. Hudson (1970) does not mention Lemanak, but from Ray (1913:7) it is clear that this is a variety of Iban, which is a member of Hudson's Malayic-Dayak subgroup.
17. Richardson relates *lamosina* to ML *lamusir* 'the flesh of an animal's back', and Aichele (n.d.:9) relates it to Sundanese *lambusir* and Old Javanese *lamujsir*.
18. Compare also JV *kontol* 'scrotum' and *tol/e* 'my boy, my lad' (< **kontol-e*); ML *kulup* 'prepuce; familiar name often given to an eldest son of a good family';

- Proto-South-Sulawesi **laso* 'penis', which also became a term of address for young boys in the daughter languages (Mills 1975:752).
19. **b* had already become ⁺*w*, while **d*, **D*, and **j* had already become ⁺*r* in Proto-Southeast Barito (cf. Dahl 1977b). In this article, I will not give a systematic treatment of all double MLG reflexes of single PMP protophonemes (such as the double MLG reflexes of original stops, or the different endings *-ka* and *-tra* reflecting PMP *-p or *-b).
20. *finga* is a Persian loan, but may have been borrowed via ML. If it was, it would be relevant for the dating of ML influence on MLG.
21. However, *fingana* and *piggaj* may be related if one considers the derivations MLG *mpamiiggana* 'one who thwarts or hampers' and Minangkabau *pamingaj* 'beam attached to the middle of standing poles or stakes to hold them together'. If they are related, it is probably by borrowing.
22. Dempwolff actually reconstructed *bi(N)Daŋ, *pa(N)Daŋ, *qa(N)daŋ, *qa(N)Dəp, *su(N)Du, because he supposed that intervocalic *tr* reflected PMP *-nD- clusters. As it appears that the MLG correspondences are not inherited, there is no point in reconstructing *N.
23. Dahl (pers. comm.); Dahl specifies that when the discussion of the king's councillors had come to an end and agreement was obtained, the king *cut* the discussion by *commanding* what had been agreed upon.
24. ML *bodo(h)* and its correspondences may ultimately derive from Macassarese, where *bodo* 'short, small' occurs in idiomatic phrases such as *bodo-pammuntulu?* '(short of knowledge =) be ignorant about something', *bodo-pattjini?* '(short of sight =) not pay attention to what happens around oneself', *bodo-pallaŋgere?* '(short of hearing =) not realize what is being said' (Cense 1979).
25. Dahl rejects Dempwolff's correspondence PMP **basar* 'big' > MLG *vesatra* 'heavy'. According to him, **bəsar* yielded Maanyan *wahai* 'numerous', MLG *be* 'big, numerous', while **bəRat* yielded Maanyan *weat* 'weight, heaviness' and MLG *vesatra* (cf. also Sakalava MLG *vezatse*, Dahl 1951:349–350). I disagree with Dahl, because the *s* in *vesatra* and the *b* in *be* are irregular. In some MLG dialects, *vesatra* (or a corresponding form) is used with reference to pregnancy; in this context, the semantic change 'big' to 'heavy' is evident. Dempwolff's **bəsar* must be rejected on account of the fact that its reflexes are found only in ML and in a very few languages influenced by ML. MLG *vesatra* must be a ML loan.
26. To complicate the matter, ML does have *hanciq* 'smell vilely, e.g., as the smell of a urinal'.
27. Six of them (or seven, if *sisina* is also counted) have correspondences (in ML or Maanyan) with final *r*. At a certain point, final **d*, **D*, and **j* merged with final **t* to produce MLG *-tra*, and it is conceivable that ⁺*-r* was replaced by *-na* in loanwords that entered MLG after this merger. But this explanation obviously does not apply to cases with *-na* corresponding to forms lacking final *r* in other (lending) languages.
28. In one case, Dempwolff reconstructed a PMP etymon with **c*, but he did not give a MLG reflex, namely, *[r]ancaŋ 'stuck/inserted into something'

(Dempwolff: '*Hineinstecken*') (cf. MLG *rantsana* 'a branch or something of the like'). The form *rantsana* must be borrowed from ML (cf. *rañcay* 'pointed upwards; vertical stake or rod').

29. Even if one were to accept a semantic connection between *cium* and *tsio/tsiona*, this would not lead to a reconstruction **cium*. ML *cium* is borrowed from Northern Indian languages (cf. Bengali *cum(a)*, Hindi *chumma* 'a kiss') (in neither case is the final vowel pronounced; Adelaar [to appear]). Correspondences in other Austronesian languages are borrowed from ML.
30. Dyen (1953) and Blust (1980) have no etyma containing *r with a reflex in MLG. Blust (1970) has 48 etyma with *r. For two of these 48 etyma, I found MLG correspondences which in Dempwolff's interpretation of the PMP sound system would be considered as inherited reflexes: *rikə(Ct) 'sticky' > *rihitra* 'thick liquid', and *(q)ukur 'luck, fortune' > *ohotra* 'preparation'. The form *rihitra* may be borrowed from Banjarese ML (cf. *rikit* 'glutinous'), and *(q)ukur is actually Dempwolff's *ukur (cf. § 2.2.3.6). Blust's *səŋkar 'cross seat in a boat' may seem to be related to *sakana* 'anything put across, any obstacle or impediment' (but cf. § 2.1.7).
31. Mills (1975:808) has some doubt about the final consonant of *putə[r] on account of its South Sulawesi reflexes.
32. Rejang, on Sumatra, has *burug* as a general term for 'bird', but this is a borrowing, since r is a loan-phoneme in Rejang (Aichele, in Jaspan 1984). Kambera, a language from Sumba (Onvlee 1984), has two apparently related forms meaning 'bird', but *mahawurugu* is derived from *hawurugu* 'fly' and *wuru* 'rustle; fly', and *mburugu* is only used for heraldic birds on coins. Most Austronesian languages reflect PAN *manuk for 'bird'.
33. Borrowing from ML may also explain the *va-* sequence in *va/lampatra* and the *ba-* sequence in Maanyan *ba/lampar*. Regressive dissimilation from *r to l is typical for ML and JV. It is not for MLG and Maanyan, where the reverse phenomenon is regularly observed, that is, regressive assimilation from *r to l (cf. *Zalan 'road' > *ralan > MLG *lälana*, Maanyan *lalan* 'id.'), and regressive assimilation from *l to r (cf. *ludaq 'spittle (n.)' > *lura > MLG *rora*, Maanyan *rura* 'id.').
34. Both ML *sərah* and ML *sara* agree with MLG *sara*. However, although ML *sara* is closer in meaning (and form) to MLG *sara*, it does not match with Ngaju Dayak *sarah* and Maanyan Dayak *nasarah* (which, again, must be borrowed from ML in view of its final h).
35. The only one is *sakarivo* 'ginger' < Sanskrit *çṛṅgavera* 'id.'. Two more lexemes can be added to Dahl's number of 30 Sanskrit loanwords: *belaka* 'open, fresh (of a face)' and *sahaza* 'suitable, fitting, sufficient'. The form *belaka* must derive from ML *bəlaka* 'quite, altogether, without exception' which, according to Gonda (1973), derives from Sanskrit *ba:laka* 'boy or youth' (also 'minor' and 'fool, simpleton'; see Gonda for the semantic developments in ML *bəlaka*). The form *sahaza* 'suitable, fitting, sufficient' must derive from ML *sahaja* '1. intentionally 2. on one's own 3. only, simply' which, according to Gonda, derives from Sanskrit *sahaja-* 'innate, by nature'. Possibly some other loanwords can be added, but these do not affect my main argument, as they all seem to occur in ML or JV as well. Cf. Bernard-Thierry 1959; a possible exception is MLG *faria* 'k.o. wet rice-field': it has no correspondences in ML or JV, and

Bernard-Thierry derives it from Sanskrit *va:ri* ‘water’, *va:riayana* ‘lake, pond’. This derivation, however, seems doubtful.

36. However, ML influence on Philippine languages includes many lexemes which are ultimately borrowed from Arabic, and it must therefore have lasted longer than ML influence on Old Javanese or MLG.
37. Vérin, Kottack, and Gorlin (1969:61–63) on the basis of glottochronology locate the migration around the beginning of the Christian era. But glottochronological evidence, if evidence at all, is no more than a vague way of measuring time depth. Where possible, it should be tested against qualitative evidence, and it should always yield to it.
38. It is striking that Lampung, another South Sumatran language, has (*h*)*ulun* for ‘human being’; the only other languages that have correspondences of MLG *olona* ‘human being’ are Kalimantan languages. Note that *godàna* and *landaizana*, which may be loanwords from Toba Batak (2.1.1) and from Minangkabau (1.6) respectively, do not constitute evidence for contacts between early Malagasy emigrants and Sumatra, because they were not affected by the Bantu substratum, and must have entered MLG after the (first?) migration.
39. See Dahl (1951:357–366) for a survey of these contacts.
40. Possible exceptions are Deli ML (East Sumatra) *tei* ‘no(t)’ (Roolvink 1953: 28, *passim*) and Urak Lawoi’ (South Thailand’s West Coast) *tet* ‘no(t)’ (Hogan 1972:226; Proto-Malayic *-n > Urak Lawoi’ -t).
41. It is possible that *tidak* and *tiada* obtained their first syllable through borrowing from an early form of MLG, and that pre-MLG ⁺*ti* plus early Malay ⁺*da* would have yielded the compounds *ti/dak*, *ti/ada* (Old Malay *ti:da*). At any rate, the origin of the *ti-* sequence in Standard Malay *ti/dak*, *ti/ada*, remains a problem in Malayic historical linguistics (Adelaar 1985:161, 168).

REFERENCES

LIST OF ABBREVIATIONS

- OL Oceanic Linguistics*, University of Hawaii, Honolulu
PL Pacific Linguistics, The Australian National University, Research School of Pacific Studies, Department of Linguistics, Canberra.
 (Series A: Papers; Series B: Monographs; Series C: Books; Series D: Special Publications.)
Nusa Nusa linguistic studies of Indonesian and other languages in Indonesia, Badan Penyelenggara Seri Nusa, Jakarta
 PPPB Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan, Jakarta

Abdul Djebar Hapip. 1976. *Kamus Banjar—Indonesia*. Jakarta: PPPB.
 Abinal and [V.] Malzac. 1963. *Dictionnaire malgache-français*. Paris: Éditions maritimes et d’Outre-mer.

- Adelaar, K. A. 1981. Reconstruction of Proto-Batak phonology. *Nusa* 10:1–20.
- . 1985. *Proto-Malayic: The reconstruction of its phonology and part of its morphology and lexicon*. Ph.D. dissertation, University of Leiden. Alblasser-dam: Kanters.
- . 1988. More on Proto-Malayic. In *Rekonstruksi dan cabang-cabang Bahasa Melayu Induk*. (Mohd. Thani Ahmad and Zaini Mohamed Zain, eds. Siri monograf sejarah Bahasa Melayu. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Aichele, W. 1936. Ein neu erschlossene frühindonesische Literatursprache in ihrem Einfluss auf das Altjavanische. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 90: back matter, pp. 18–19.
- . n.d. Einführung in das Madagassische. Ms.
- Anttila, Raimo. 1972. *An introduction to historical and comparative linguistics*. New York: Macmillan.
- Antonissen, A. 1956. *Kadazan–English and English–Kadazan dictionary*. Canberra: Government Printing Office.
- Barthélemy, A. 1935. *Dictionnaire arabe–français (dialectes de Syrie: Alep, Damas, Liban, Jérusalem)*. Paris: Librairie orientaliste Paul Geuthner.
- Bernard-Thierry, Solange. 1959. À propos des emprunts sanskrits en Malgache. *Journal Asiatique* 247(3): 311–348.
- Blussé, Leonard, and Marius P. H. Roessingh. 1984. A visit to the past: Soulang, a Formosan village anno 1623. *Archipel* 27:63–80.
- Blust, Robert A. 1970. Proto-Austronesian addenda, *OL* 9(2): 104–162.
- . 1977. Austronesian culture-history: Some linguistic inferences and their relations to the archeological record. *Nusa* 4:25–37.
- . 1980. Austronesian etymologies. *OL* 19 (1–2).
- . 1982. An overlooked feature of Malay historical phonology. *Bulletin of the School of Oriental and African Studies* 45(2): 284–299.
- Burrow, T., and M. B. Emeneau. 1961. *A Dravidian etymological dictionary*. Oxford: Clarendon Press.
- Casparis, J. G. de. 1956. *Selected inscriptions from the 7th to the 9th century A.D. Prasasti Indonesia II*. Bandung: Masa Baru.
- Cense, A. A., in collaboration with Abdoerrahim. 1979. *Makassaars–Nederlands woordenboek*. The Hague: Nijhoff.
- Charles, Matthew. 1973. [Unpublished PAN/PMP etyma listed in Wurm and Wilson 1975.]
- Çoedès, G. 1930. Les inscriptions Malaises de Śrīwijaya. *Bulletin de l'Ecole Française d'Extrême Orient* 30: 29–80. Paris.
- Dahl, Otto Christian. 1951. *Malgache et Maanjan. Une comparaison linguistique*. Avhandlinger utgitt av Instituttet 3. Oslo: Egede Instituttet.
- . 1954. Le substrat bantou en Malgache. *Norsk Tidsskrift for Språkvidenskap* 17: 325–362.
- . 1977a. *Proto-Austronesian*. 2d rev. ed. Scandinavian Institute of Asian Studies Monograph Series, no. 15. Lund: Studentlitteratur, Curzon Press.
- . 1977b. La subdivision de la famille Barito et la place du malgache. *Acta Orientalia* 38: 77–134. Copenhagen.
- Dalmond. 1842. *Vocabulaire et grammaire pour les langues Malgaches Sakalave et Betsimisara*. Saint Denis, Ile de Bourbon.
- Damais, Louis-Charles. 1968. Etudes soumatraine III. La langue B des inscriptions de Sri Wijaya, *Bulletin de l'Ecole Française d'Extrême-Orient* 56: 523–566.
- Dempwolff, Otto. 1934–1938. *Vergleichende Lautlehre des Austronesischen Wortschatzes*, I [1934], *Induktiver Aufbau einer Indonesischen Ursprache*, II [1937],

- Deduktive Anwendung des Urindonesischen auf Austronesische Einzelsprachen*, III [1938]. *Austronesisches Wörterverzeichnis. Zeitschrift für Eingeborenen sprachen*, supplements 15, 17, 19. Berlin: Dietrich Reimer.
- de Clercq, F. S. A. 1876. Eenige bijzonderheden over het Maleisch van Palembang. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 23: 517–554.
- Deschamps, Hubert. 1936. *Le dialecte antaisaka*. Tananarive: Pitot de la Beaujardière.
- Djajadiningsrat: see Hoesin Djajadiningsrat.
- Djantera Kawi. 1985. *Kamus Bakumpai Indonesia*. Jakarta: PPPB.
- Drewes, G. W. J. 1968. New light on the coming of Islam in Indonesia. *Bijdragen tot de Taal-, Land- en Volkenkunde* 124: 433–459.
- Dunggio: see P. D. Dunggio.
- Durdje Durasid, Aries Djinal, and Djantera Kawi. 1986. *Morfosintaksis bahasa Katingan*. Jakarta: PPPB.
- Durdje Durasid and Djantera Kawi. 1978. *Bahasa Banjar Hulu*. Jakarta: PPPB.
- Dyen, I. 1951. Proto-Malayo-Polynesian *Z. *Language* 27: 534–540.
- . 1953. *The Proto-Malayo-Polynesian laryngeals*. Baltimore: The Linguistic Society of America.
- Eringa, F. S. 1984. *Soendaas–Nederlands woordenboek*. Dordrecht: Foris.
- Ferrand, G. 1932. Quatre textes épigraphiques malayo-sanskrits de Sumatra et de Bangka. *Journal Asiatique* 221: 271–326.
- Flacourt, E. de. 1658. *Dictionnaire de la langue de Madagascar*. (Revised by G. Ferrand and republished in 1905. Paris: Publications de l'Ecole des Lettres d'Alger. *Bulletin de correspondance africaine* 33.)
- Gericke, J. F. C., and T. Roorda. 1901. *Javaansch–Nederlandsch handwoordenboek*. (Revised and expanded by A. C. Vreede.) Amsterdam: Johannes Muller.
- Gonda, J. 1973. *Sanskrit in Indonesia*. New Delhi: International Academy of Indian Culture.
- Hapip: see Abdul Jebar Hapip.
- Hardeland, A. 1859. *Dajacksch–Deutsches Wörterbuch*. Amsterdam: Frederik Muller.
- Hazeu, G. A. J. 1907. *Gajösch–Nederlandsch woordenboek*. Batavia: Landsdrukkerij.
- Hoesin Djajadiningsrat. 1934. *Atjéhsch–Nederlandsch woordenboek*, 2 vols. Batavia: Landsdrukkerij.
- Hogan, David. 1972. Men of the sea: Coastal tribes of South Thailand's West Coast. *Journal of the Siam Society* 60(1): 205–235.
- Houtman, Frederick de. 1603. *Spraek ende woord-boeck, inde Maleysche ende Madagaskarsche talen*. Amsterdam: Cloppenburgh.
- Howison, James. 1801. *A dictionary of the Malay tongue*. London: Arabic and Persian Press.
- Hudson, Alfred B. 1967. *The Barito isolects of Borneo*. Southeast Asia Program, Dept. of Asian Studies, Data Paper no. 68. Ithaca: Cornell University Press.
- . 1970. A note on Selako: Malayan Dayak and Land Dayak languages in Western Borneo. *Sarawak Museum Journal* 18(36–37): 301–318.
- Iskandar, Teuku. 1970. *Kamus Umum*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Jansz, P. 1913. *Practisch Javaansch–Nederlandsch woorden boek*. Semarang, Surabaya, and The Hague: Van Dorp.
- Jaspan, M. A. 1984. *Materials for a Rejang–Indonesian–English dictionary*. (With a fragmentary sketch of the Rejang language by W. Aichele, and a preface and additional annotations by P. Voorhoeve.) PL D59.
- Junaiyah H. M. et al. 1985. *Kamus Lampung—Indonesia*. Jakarta: PPPB.
- Kern, H. 1882. Over de opschriften uit Koetei in verband met de geschiedenis van het schrift in den Indischen Archipel. *Verslagen en Mededeelingen der Koninklijke*

- Akademie van Wetenschappen, Afdeeling Letterkunde*, 2d series, 11: 182–203. Amsterdam. (Reprinted in *Verspreide Geschriften* 7: 55–76, The Hague, 1917.)
- . 1889. Taalkundige gegevens ter bepaling van het stamland der Maleisch-Polynesische volkeren. *Verslagen en Mededeelingen der Koninklijke Akademie van Wetenschappen, Afdeeling Letterkunde*, 3d series, 6: 270–287. Amsterdam. (Reprinted in *Verspreide Geschriften* 6: 105–120, The Hague, 1917.)
- Kern, W. n.d. Banjarese-Dutch wordlist. Ms., Leiden University Library, cod. or. 8332.
- Klinkert, H. C. 1916. *Nieuw Maleisch–Nederlandsch woordenboek*, 3d rev. and exp. ed. Leiden: Brill.
- Lafont, Pierre Bernard. 1968. *Lexique Français–Jarai–Vietnamien*. Paris: Publications de l'Ecole Française d'Extrême Orient, Vol.
- Li, Paul Jen-Kuei. 1982. Kavalan phonology: Synchronic and diachronic. In: *Gava': Studies in Austronesian languages and cultures*, edited by Reiner Carle, Martina Heinschke, Peter W. Pink, Christel Rost, and Karen Stoldtlander. Berlin: Dietrich Reimer.
- M. Thaib gl. St. Pamoentjak. 1935. *Kamoës bahasa Minangkabau–bahasa Melajoe-Riau*. Batavia: Balai Poestaka.
- Mills, Roger F. 1975. *Proto-South-Sulawesi and Proto-Austronesian phonology*. Ph.D. dissertation, University of Michigan, Ann Arbor. (University Microfilms International, 1978.)
- Moussay, Gérard. 1971. *Dictionnaire cam–vietnamien–français*. Phanrang: Centre Culturel cam.
- Neumann, J. H. 1951. *Karo-Bataks–Nederlands woordenboek*. Jakarta: Lembaga Kebudayaan Indonésia.
- Nothofer, Bernd. 1975. *The reconstruction of Proto-Malayo-Javanic*. The Hague: Nijhoff.
- Onvlee, L., in collaboration with Oe. H. Kapita and P. J. Luijendijk. 1984. *Kamberra (Oost-Soembaas)–Nederlands woordenboek*. Dordrecht: Foris.
- Ottino, Paul. 1974. *Madagascar, les Comores et le Sud-Ouest de l'Océan Indien*. Tananarive: Université de Madagascar, Publications du Centre d'Anthropologie Culturelle et Sociale.
- P. D. Dunggio, Suwarni N., Asnah S., Nur Indones. 1983. *Struktur bahasa Melayu Palembang*. Jakarta: PPPB.
- Pamoentjak: see M. Thaib
- Panganiban, Jose Villa. 1966. *Talahuluganang Philipino–Ingles*. Maynila: Kawanihan ng Palimbagan.
- Pigeaud, Th. G. Th. 1938. *Javaans–Nederlands handwoordenboek*. Groningen: Wolters-Noordhoff.
- . 1967. *Literature of Java, vol. I, Synopsis of Javanese literature*. Codices Manuscripti IX. Leiden: The University Library.
- Porée-Maspero, Eveline. 1986. Jonques et po, sampou et sampān. *Archipel* 32: 65–85.
- Ras, J. J. 1968. *Hikajat Bandjar*. The Hague: Nijhoff.
- Ray, Sidney H. 1913. The languages of Borneo. *Sarawak Museum Journal* 1(4). Kuching.
- Richards, A. 1981. *An Iban–English dictionary*. Oxford: Clarendon Press.
- Richardson, J. 1885. *A new Malagasy–English dictionary*. Tananarivo.
- Roovlink, R. 1953. Dialek Melaju di Deli. *Bahasa dan Budaja* 3: 3–40.
- Steingass, F. 1930. *Persian–English dictionary*. London: Kegan Paul, Trench, Trübner.
- Tamsin Medan. 1980. Dialek-dialek Minangkabau di daerah Minangkabau/ Sumatera Barat (Suatu pemerian dialektologis). Ms. Jakarta: PPPB.

- Tendeloo, H. J. E. 1901. *Maleische grammatica*. Leiden.
- Thaib: see M. Thaib.
- Tsuchida, Shigeru. 1976. *Reconstruction of Proto-Tsouic phonology*. Study of Languages and Cultures of Asia and Africa Monograph Series, no. 5. Tokyo.
- Van der Tuuk, H. N. 1865. Outlines of a grammar of the Malagasy language. *The Journal of the Royal Asiatic Society of Great Britain and Ireland* 1 (n.s.): 419–446.
- Vérin, Pierre, Conrad P. Kottack, and Peter Gorlin. 1969. The glottochronology of Malagasy speech communities. *OL* 8(1): 26–83.
- Vogel, J. Ph. 1918. The yupu inscriptions of King Mulawarman, from Koetei (East Borneo). *Bijdragen tot de Taal-, Land- en Volkenkunde* 74: 167–232.
- Warneck, Joh. 1977. *Toba-Batak–Deutsches Wörterbuch* (neu herausgegeben von R. Roolvink). The Hague: Nijhoff.
- Wehr, Hans. 1976. *A dictionary of modern written Arabic*, edited by J. Milton Cowan. Ithaca: Spoken Language Services.
- Wilkinson, R. J. 1959. *A Malay–English dictionary*. London: Macmillan.
- Wolff, J. U. 1972. *A dictionary of Cebuano Visayan*, 2 vols. Ithaca: Cornell University Southeast Asia Program and the Linguistic Society of the Philippines.
- . 1974. Proto-Austronesian *r and *d. *OL* 13: 77–121.
- . 1976. Malay borrowings in Tagalog. In: *Southeast-Asian history and historiography: Essays presented to D. G. H. Hall*, edited by C. D. Cowan and O. W. Wolters, pp. 345–367. Ithaca: Cornell University Press.
- . 1982. Proto-Austronesian *c, *z, *g, and *T. *PL* C75: 1–30.
- Wurm, S. A., and B. Wilson. 1975. *English finderlist of reconstructions in Austronesian languages (post-Brandstetter)*. *PL* C33.
- Zoetmulder, P. J., in collaboration with S. O. Robson. 1982. *Old Javanese–English dictionary*. The Hague: Nijhoff.
- Zorc, R. David, and Matthew Charles. 1971. *Proto-Philippine finderlist*. Ms. Ithaca: Cornell.
- . 1978. Proto-Philippine word accent: Innovation or Proto-Hesperonesian retention? *PL* C61: 67–119.