Etymologies and
where to find them

by
David Mead

2010

Sulang Language Data and Working Papers:
Topics in Lexicography, no. 1

Sulawesi Language Alliance

LANGUAGES

Language of materials : English

ABSTRACT

This paper is an annotated bibliography of thirty-eight resources where etymologies of Indonesian words can be found, arranged according to source (e.g. Arabic loanwords, Chinese loanwords, Indic loanwords, words of Austronesian provenance, etc.). Compiled over several years and completed in 2005, this paper has largely, but not completely, been superseded by the 2007 publication of Russell Jones's *Loan-words in Indonesian and Malay*.

TABLE OF CONTENTS

General; Arabic loanwords; Chinese loanwords; Indic loanwords; European loanwords; Other loanwords; Words of Austronesian provenance; Other resources; References.

VERSION HISTORY

Version 1   [28 December 2010]   Completed 2005; amended for publication, 2010.

Etymologies and where to find them

by David Mead

[N.B.: This paper was largely completed prior to the publication
of Russell Jones's 2007 *Loan-words in Indonesian and Malay*.]

There's no doubt about it, tracking down etymologies is a scholarly endeavor, and we can
be grateful for the bookworm-scholars who have preceded us. For English, the Oxford
English Dictionary is the classic source, but Webster's Unabridged and many other
dictionaries include etymological information. Even the modest Random House
dictionary that our family uses at home includes etymologies.

In the Indonesian context, however, resources are considerably thinner. As yet there is no
single resource to which one can turn to find etymologies for Malay and Indonesian
words, let alone for the myriads of words in the various vernacular languages of
Indonesia. Consequently, for lack of interest, lack of resources, and lack of time, most of
us choose to ignore etymological concerns—except for perhaps the most bald-faced
loans—when compiling indigenous language dictionaries.

The hope, of course, is that some day there will appear a scholarly and comprehensive
etymological dictionary for Indonesian. To this end, the Indonesian Etymological Project
was founded in 1973. Whilst some notable works have already resulted (see below no.s 5,
9, 11, 14 and 18), a final compilation into a single reference has yet to be completed.[1]

For the interim, then, I lay out some of the most accessible resources, which can help you
work more efficiently. This article makes no claim to completeness, though I hope I have
hit on all the principal works. Most of the cited works have lists of references of their
own, to guide the scholar who wants to drink deeply from the well of knowledge. This
article is essentially an annotated bibliography. Full bibliographic citations appear at the
end.

General

1. Wilkinson's classic 1959 *A Malay-English Dictionary* (first published 1901-1903) is
unparalleled by any dictionary on the market today in the amount of etymological
information it contains. As an informal test, I randomly opened his dictionary and
counted off the first 100 entries I came to. Of these, a full twenty contained etymological
information. (We would be doing well if our own dictionaries were so richly supplied!)

In many cases Wilkinson indicates simply the source language, as in:

---

[1] Dr. Russell Jones, one of the few remaining active members of the committee, writes (2005:pers.comm.)
"I am working hard to complete *Loan-words in Indonesian*, carrying all the loan-words (from outside
Nusantara) which we could find, and reasonably complete. The last problem is presenting the Chinese
characters correctly. Hopefully it will be out next year." In fact this work was published in 2007.

bĕndera   Port.   Flag (of European design).

Sometimes, however, he gives a fuller amount of information, in which case he sets the etymological information off in square brackets:

bĕndi   [Hind. and Telegu bandi]   Gig, two-wheeled cart or carriage.

When a word is not general Malay, Wilkinson gives information about where the word is spoken. This regional information, which he encloses in parentheses, is not to be confused with his etymological information, though it may sometimes give clues (or even explicit information) as to source language, compare:

begok   (Min.)   Coarse; rough; dull.

bonteng   (Batav. from Sund.)   Cucumber.

Whilst the language and regional abbreviations which Wilkinson uses were for the most part readily interpretable, some (like Pk., N.I., Lask.) we unclear to me, but the 1959 edition didn't have any table in which to look up what these abbreviations stood for. A second drawback is that Wilkinson's dictionary contains etymological information only for borrowed words, and does not indicate which words might represent inherited Austronesian vocabulary.

2. Collin's 2001 *Mukadimah Ilmu Etimologi* is a scholarly introduction to issues involved in how to do etymological research, with special reference to the Malay context. This book would be a very useful resource for anyone who wants to teach in Indonesian about sound change, borrowing, phonological adaptation of loanwords, and other aspects about how the word stock of a language can change and develop. Along the way Collins gives the fascinating history of a number of Malay words (the *urat ~ otot* example given below, for example, comes from here). Speakers of Indonesian who 'just want to know about words' will find many gems within this book's pages. Unfortunately it does not have a word index, so it less useful in cases where you already have a particular word in mind, for which you are trying to chase down an etymology.

3. Ngajenan's 1987 *Kamus Etimologi Bahasa Indonesia* is rather small, containing only around 1900 entries of Indonesian words and their etymologies (definitions for the Indonesian words are not given). This book has a number of deficiencies, not the least of which is the lack of attention to detail. For example Sanskrit transliterations, which in this work never have macrons over vowels, are sometimes also misspelled or incorrectly cited, compare **duka** < Skt dukha (correct: Skt duhkha) and **bagi** < Skt bhaga (correct: Skt bhāgī). Another lack is that even when a word is correctly identified as Austronesian, the 'Austronesia Kuno' (Proto Austronesian) form which Ngajenan cites looks decidedly like 'Malay'. One also suspects that Ngajenan must be a Javanese speaker, because a number of words are incorrectly given as coming from or through Javanese or Old Javanese, for example **hujan** < Jw udan < AK hujan (correct: hujan < AK *quZaN). Because of the many errors, the best possible use of this book is as a first stop only, in order to guide you to one of the more scholarly works reviewed below.

4. Jones's 1984 article "Loan-words in contemporary Indonesian" is devoted to 'understanding the phenomenon of borrowing in the Indonesian context', and is probably the best article available if one simply wants an overview of the subject, including an historical perspective. Separate sections are devoted to Indic languages (Sanskrit, Middle Indo-Aryan, Hindi, Tamil), Arabic (along with Persian), Chinese, Japanese, and, finally, European languages (Portuguese, Dutch, English). In an article only twenty-eight pages long, Jones aims for breadth rather than depth, but he also manages to discuss (in brief) sound changes which occurred in the process of borrowing, loan translations, and the semantic and grammatical classes of words (sometimes even affixes) which were borrowed.

Arabic loanwords

5. Jones's 1978 *Arabic loan-words in Indonesian* was the first publication to result from the Indonesian Etymological Project. As the subtitle indicates, it is a checklist. Each entry gives the Indonesian word, a one or two word English gloss of the Indonesian word, and the Arabic or Persian source word. No gloss is given for the Arabic; rather, the reader is referred to the page number in a particular Arabic-English dictionary for more information. Jones's checklist is thorough, containing over 4,700 entries. Three appendices list Indonesian loan words according to the Arabic root; Indonesian words of Persian origin; and Indonesian words ultimately of Arabic origin, but which entered the language through Dutch or English.

6. In the present-day lexicon of Indonesian, most words of Arabic provenance have forms which suggest they came from Classical Arabic. The purpose of Van Dam's 2010 article "Arabic loanwords in Indonesian revisited" is to ask: can we discern in the form of some Arabic loanwords, perhaps representing an older stratum, the influence of a colloquial Arabic (particularly South Arabia or Hadramaut),[2] or—as some have suggested—that some loanwords of Arabic origin actually reached Indonesia via Persian? These same questions intrigued other scholars long before Van Dam's work, and his article is a useful entry point to the extended literature on this topic.

Chinese loanwords

7. Leo's 1975 *Chinese loanwords spoken by the inhabitants of the city of Jakarta* contains some interesting material such as Indonesian *becak* < Ch. *bé-ts'ia* 'horse cart'. The booklet is short—it contains only about 300 entries altogether—leaving one with the impression that Leo is just scratching the surface. Nonetheless, it is more complete than older alternatives,[3] and its 'checklist' format makes it usable for the casual lexicographer. Unlike Kong's (1993) longer list (reviewed below no. 8), Leo includes both tone marks as well as glosses for the Chinese. In my brief perusal, I noted one glaring error, Malay

---

[2] That is, today's Yemen.

[3] By comparison, Schliegel (1891) treated only ninety-seven Malay words of Chinese (or probable Chinese) origin, while Hamilton (1924) included 147 words of definite Chinese origin, twenty-two of probable Chinese origin, and eighteen of possible Chinese origin.

duit which was incorrectly ascribed to Chinese *lui*. As Schlegel had noted eighty years earlier (1891:392), Chinese *lui* as well as Malay *duit* both originate from Dutch *duit*.

8. In his 1987 article, "A study of Chinese loanwords (from South Fujian dialects) in the Malay and Indonesian languages", Kong concludes that nearly 90% of all Chinese loanwords originate from southern Fujian dialects (e.g. Hokkien), and goes into history to explain why this is so. He also covers domains to which the borrowings relate (e.g. names of vegetables, numerals and measures, etc.) along with some semantic shifts which occurred in the borrowing process (e.g. *tsabɔ* 'woman' > *cabo* 'courtesan, prostitute'). Whilst Kong worked with a database of 507 Chinese loanwords, only a smattering of them make it into his article by way of example.

9. Kong's 1993 article "Kata pinjaman bahasa Cina dalam bahasa Melayu" is the most complete checklist of Chinese borrowings in Indonesian published to date. After a brief, three-page introduction, the remainder of the article lists 1054 loan words from Chinese, ordered alphabetically according to the Malay. Other columns are devoted to (a) a romanization of the Chinese source word, <u>without</u> indication of tone; (b) the Chinese character (both in the local Chinese dialect and standard Chinese), (c) a brief gloss of the Malay term, and (d) indication of the source language (F for Fujian = Hokkien; G for Guangdong = Cantonese; K for Kejia = Hakka; C for Chaozhou = Teochew = Swatow). The meaning of the Chinese source word is not given.

10. Santa Maria's 1996 article "I prestiti lessicali cinesi in malese-indonesiano nel loro contesto storico e sociale" follows a familiar outline: after a review of the available literature, Santa Maria discusses the history of Chinese immigration (including Chinese 'dialects'), semantic domains of borrowing, and phonological adaptation of loan words. As the title indicates, however, the article is in Italian, and the lack of any checklist makes is less useful for the casual lexicographer (in the article itself only a sprinkling of Chinese loan words are brought up by way of example).

11. In 1999 Russell Jones submitted the script for his *Chinese Loan-words in Malay and Indonesian: a Background Study* for publication by the University of Malaysia. Unfortunately, this important work has yet to appear in print, and my own familiarity with the work comes only from a prepublication draft of the introduction (cited in the references as Jones 1996). Whilst the draft contains neither maps, figures, references or checklist, nonetheless this 'preview' shows the work to be of the highest caliber. Jones even goes into evidence from tombstones as to what Chinese dialects were spoken in pre-European Indonesia, yet he has made his writings accessible to people who, like me, have little knowledge of the Chinese language situation. I especially enjoyed Jones's description of Chinese dialects. Is Foochow the same thing as Fuzhou? Is Guangzhou the same thing as Guangfu? Other articles had left me wondering, but Jones lays everything out very clearly!

12. I myself have not had access to Mashudi and Yeong's 1989 article "Pengaruh bahasa Cina dalam perbendaharaan kata bahasa Malaysia", but we learn in Jones (1996) that these authors worked with a database of 341 Chinese loanwords. Jones further writes that Mashudi and Yeong "provide probably one of the best published linguistic analyses of

phonological changes in Chinese loan-words as they enter Malay… They also deal with semantic changes; finally their survey discusses the relatively few loan-words which are subjected to morphological change after they have been accepted by Indonesian" (Jones 1996:10).

Indic loanwords

13. The principle work on Indic loanwords is Gonda's 1952 *Sanskrit in Indonesia* (2nd edition, 1973). This is a scholarly work of the highest caliber, detailing the background of how Sanskrit (and other Indic languages) came to Indonesia, and devoting sometimes long discussions to particular words, the paths they traveled to reach Indonesia, and the phonological and semantic changes they underwent in process. Specific chapters discuss words in various semantic domains, such as "Death and the life hereafter", "Medicine" and "Botanical names". Sanskrit words are helpfully cited in both Devanagari script and in romanized transliteration. An appendix includes an index of source words, arranged according to whether they are Sanskrit words or Middle or new Indo-Aryan words,[4] and a second index of words in Indonesian languages (primarily Malay and Javanese) which are of Indic origin.

14. De Casparis's 1997 *Sanskrit loan-words in Indonesian* (along with some amendments later proposed by Mahdi 2000) is a much simpler guide to Indonesian words originating from Sanskrit. Alphabetized according to Indonesian, it is essentially a checklist: Indonesian word, gloss, Sanskrit source word, gloss, with comments here and there provided as endnotes. This list has also been reproduced electronically by the Center for Research in Computational Linguistics, and can be accessed online at http://crcl.th.net/indic/indo.htm (beware, however, the non-standard romanized Devanagari employed in the online version, e.g. grave accent above vowels instead of macron among others).

15. In their section on 'Indonesia' in *Sanskrit Words in Southeast Asian Languages*, Sedyawati and Somvir (2005) have compiled a list of over 1900 Sanskrit borrowings which are found in Old Javanese, Balinese, and/or Indonesian. Arrangement is alphabetical according to the Old Javanese. The following entries are typical:

> **Gopāla** (Oj) – Gembala Shepherd
> Gembāla (BI)
> Skt. गोपाल Gopāla – Cowherd
>
> **Maṇik** (Oj) – Permata Precious stone
> Mānik (BI)
> Manik (BL)
> Skt. मणिक Maṇika – Jewel, precious stone, gem

---

[4] Middle Indo-Aryan is sometimes refered to as Prakrit. The relationship of Prakrit to Sanskrit is similar to that of vulgar Latin to classical Latin, indeed the word Prakrit originates from Sanskrit *prākṛta* 'oridinary, usual, common'.

The retention of macron and of dot underneath in the spelling of Old Javanese, Balinese and Indonesian words apparently follows the authors' source texts (such as the Rāmāyaṇa, written in the Pallava script).

16. Tamil loanwords in Indonesian were exploded in a series of articles by Van Ronkel (1902, 1903a, 1903b). Unlike Sanskrit and Hindi, which are of Indo-Aryan (Indo-European) stock, Tamil is one of the major Dravidian languages of southern India,[5] and the Dravidian language which was in closest contact with the Malays, who had established themselves on the island of Ceylon by the sixth century AD. Tamil mediation of Sanskrit loanwords is also explored by Gonda (1952, 1973).

17. I have not had the opportunity to review Francisco (1964), which is primarily concerned with Indic loan words in Philippine languages.

European loanwords

18. Grijns, de Vries and Santa Maria's 1983 *European loan-words in Indonesian* is a scholarly compilation of approximately 7,500 Indonesian words which entered the language from European sources. It is a checklist in three column format, the first column giving the Indonesian, the second column a single-word English gloss (sometimes further explained in an endnote), and the third column giving the source form, in most cases a Portuguese, Dutch or English word. Source words are further divided out according to language and listed alphabetically in various appendices.

19. In his 1988 "Dutch loanwords in Indonesian", de Vries first of all tackles the history of Dutch in Indonesia, the reasons for borrowing, and kinds of borrowings (loan shifts, loan translations, and loan blends among others). He also discusses the early preference of Indonesian language planners for Dutch loans—which later gave way to preferences for English, Sanskrit and even Javanese—as they sought to expand the word stock of the national language. Also included is a section on how Dutch loanwords were both phonemicized to and affected the sound system of Indonesian. De Vries estimates that of some 5,400 borrowings from Dutch, nouns constitute approximately 85% of the total, verbs only 2%. Whilst the article is sprinkled with many examples which will engage and fascinate the reader, in the span of sixteen pages the author can, of course, give only the most general outline. The scholar who consults the references section will discover leads to yet other articles which have not been reviewed here.

20. In Van Minde's article "European loanwords in Ambonese Malay" the author has gathered "all Portuguese, Dutch and English loan words occurring in my fieldnotes" (Van Minde 2002:197), which amounts to around 120, 350 and 30 words respectively. Despite the small size and considerable overlap with Grijns, et al. (1983), nonetheless Van Minde hits upon some loan words not included in the larger work, for example *baileu* 'village ceremonial hall' (< Portuguese *baileu*), *hodeng* 'sturdy, burly, macho' (< Dutch *houding*) and *hibop* 'pull up!' (< English *heave up*).

---

[5] Other major Dravidian languages of India are Telugu, Kannada and Malayalam.

21. Robson's 2002 article "Dutch loan-translations in Indonesian" tackles the little-covered topic of loan translations: complex words or phrases which are translated literally from one language to another. Robson, for example, makes the case that Indonesian *terdiri dari* is based rather literally on Dutch *bestaand uit*, and so with thirty-eight other Indonesian words and phrases which he covers. Some of Robson's connections seem conjectural, but reading his article could open your mind to thinking about loan translations in the vernacular language you are researching.

Other loanwords

Apart from two paragraphs in Jones (1984:21), I do not know of any resource which specifically treat loanwords from Japanese. According to this same author, there are no known Indonesian words of modern Cambodian, Burmese or Vietnamese provenance, and almost none from Thai. However, words from an older strata of Austroasiatic (e.g. Old Mon, Old Khmer) languages have made their way into Malay and other languages of Indonesia.

22. As early as 1994, Waruno Mahdi discussed some words of probable Austroasiatic provenance, including *perak* 'silver', *emas* 'gold', *merak* 'peacock', *kerbau*, *karabau* 'water buffalo' and *kembar* 'twin' (Mahdi 1994a:168, 1994b:453). Mahdi's major work on the subject to date, however, was introduced in a 2009 conference paper titled "Some obscure Austroasiatic borrowings in Indonesian and Old Malay". Mahdi dates the earliest Austroasiatic loanwords to 2500 BC, and continuing for at least 4000 years.

23. Uri Tadmor has also produced a conference paper on Mom-Khmer loan words in Malay (Tadmor 2009), but I have not been able to obtain a copy of his paper for review.

Words of Austronesian provenance

When investigating what appear to be native 'Austronesian'[6] words in an indigenous language of Indonesia, one of the most important things is to be able to distinguish bona fide inherited forms (that is, words which can be traced directly back to Proto Austronesian), from forms which were borrowed at a later stage from some neighboring Austronesian language.[7] To do this, you must know what sound changes were 'regular' for the language under investigation.

For example, Indonesian has two terms, *urat* 'vein, tendon, sinew' and *otot* 'muscle', both of which can ultimately be traced back to a Proto Austronesian form which scholars reconstruct as *uRat. However, only the first of these, *urat*, reflects regular sound development in Malay (in the ancestor to all Malayic languages, PAN *R, perhaps a uvular trill, merged with *r). The form *otot*, on the other hand, which is actually from

---

[6] Whilst this section revolves around Austronesian, nonetheless many of the general principles apply to Papuan languages as well. Unfortunately, I am unqualified to speak specifically to the Papuan situation.

[7] In general, the older the borrowing, or the more closely related the neighboring language, the harder it is to distinguish borrowings from inherited vocabulary.

reduplicated *uRat-uRat, exhibits PAN *R > zero and subsequent merger of *ua as *o. Since both of these changes would be irregular for Malayic languages, but happen to be regular in Javanese, the conclusion is that *otot*—even though it 'looks' like it could have been inherited—must be considered a borrowing (from Javanese). By coincidence, Javanese also has both *urat* and *otot*, but here the situation is exactly reversed: in Javanese, *otot* is the inherited form, while *urat* is a borrowed word (from Malay, of course)!

One of the most exciting things to me, when ten years ago I began studying historical sound change in the Bungku-Tolaki languages of Sulawesi, was to see the patterns of regular sound change 'emerge', as it were, out of an initial confusion of often contradictory data. With growing confidence, I began to be able to recognize borrowed terms: Aha, this form exhibits all the regular sound changes, therefore it must be inherited; but that form is irregular in its development—it must be borrowed.

While you yourself might not be motivated by the excitement of uncovering such heady insights, it is possible that some other linguist has. If there is already a study reconstructing historical sound changes in the language area where you work, then this is where you will want to begin. You may find it even has a list of low level reconstructions—an invaluable source for etymologies. In point of fact, nine out of the ten language groups of Sulawesi have now been covered by one or more studies in historical reconstruction, and progress is being made in other regions of Indonesia as well.

If unfortunately there is no such study covering the language area where you work, and thus you have to search for Proto Austronesian forms on your own, I recommend the following sources for the beginner.[8]

24. Tryon (1995) is a very accessible overview of the classification of Austronesian languages as is presently known, and his discussions include a number of reconstructed forms. What is meant by 'Western Malayo-Polynesian'? Which language groups of Indonesia are included under 'Central-Eastern Malayo-Polynesian'? You'll find the answers here.

25. Blust (1999), the appendix (pp. 82-87) contains a very helpful list of 200 Proto Austronesian and Proto Malayo-Polynesian[9] basic vocabulary items.

26. Blust (2002), a compilation of reconstructed faunal terms with commentary.

---

[8] Lists of reconstructed etyma can also be found in Dyen and McFarland (1970), Zorc and Charles (1971), and Wurm and Wilson (1975), but these resources must be used with caution on account of the variant ways that Proto Austronesian phonemes have been symbolized.

[9] Proto Austronesian (PAN) is the supposed ancestor language of all Austronesian languages, from Madagascar to Hawaii, and is thought to have been spoken in Formosa (Taiwan) about 6000 years ago. Proto Malayo-Polynesian (PMP) is a daughter language that developed on Formosa out of Proto Austronesian about a thousand years later, and is the supposed ancestor of all Austronesian languages *outside of Formosa*. Of the two, Proto Malayo-Polynesian is thus the more immediate ancestor to the Austronesian languages of Indonesia.

27. Blust (2003), an article devoted to aspects of the prefix *ka-, but an appendix (pp. 471-475) contains a list of all reconstructed Proto Austronesian affixes and clitics.

28. Blust and Trussel (2010) *Austronesian Comparative Dictionary*. This is one of Blust's lifetime works, a compilation of all his reconstructions along with supporting evidence. If printed this electronic database would run to over twelve hundred pages. Unfortunately it does not include etyma which others have reconstructed.

29. Dahl (1981), a specific treatment of numerals. He discusses inherited terms and also some irregular replacements (such as Malay *delapan* from *dua-alap-an*).

30. Zorc (1995), a long list of reconstructions amounting to several pages, along with an English finder list. Zorc uses a slightly different notation system than Blust, and he also uses the throwback terms PHF (Proto Hesperonesian-Formosan, viz. reflexes are found in Formosa and Western Malayo-Polynesian languages), and PHN (Proto Hesperonesian, viz. reflexes are found only in Western Malayo-Polynesian languages).

31. John Wolff's 2010 work, *Proto-Austronesian Phonology with Glossary*, is published in two volumes comprising over a thousand pages, and it is breathtaking in its scope, including (from the publisher's description) "to examine exhaustively the forms that can be reconstructed for Proto-Austronesian and also for the earliest stages after the Austronesian languages began to spread southward from Taiwan." However, because of his unique interpretation of the Proto Austronesian sound system, Wolff employs a system for representing proto-phonemes which is different from that endorsed by most Austronesianists. Unless proper care is taken, the casual lexicographer who 'mixes and matches' protoforms from Wolff with those from other authors is likely to introduce inconsistencies into his or her own work.

32. At the time of its publication in 1938, the third volume of Dempwolff's *Vergleichende Lautlehre des austronesischen Wortschatzes* was a formidable compilation of Proto Austronesian reconstructions, many of which are still accepted today. However, because of subsequent changes in spelling conventions for Proto Austronesian etyma which have overtaken the field, there is a daunting learning curve to climb before his work becomes accessible. If you have only a casual interest in Proto Austronesian reconstructions, you should use others of the above-mentioned works rather than this one. I mention it here mainly by way of completeness.

Those who work in languages of the northern peninsula of Sulawesi should be aware of the following resources:

33. Zorc (1986) proposed ninety-eight lexical innovations in support of a Philippines group which includes the Sangir, Minahasan and Gorontalo-Mongondow languages. According to Blust (2005:56), external evidence has been found for twelve of these, which can no longer be considered Proto Philippine innovations.

34. Blust (1991) includes a list of ninety-four Proto Greater Central Philippines innovations. The Greater Central Philippines group includes the Gorontalo-Mongondow languages, but excludes the Sangir and Minahasan languages.

35. In an appendix to his "The Linguistic Macrohistory of the Philippines", Blust (2005) presents a provisional list of 241 additional Proto Philippine reconstructions. He notes (page 35) that the coverage is very uneven, and that a final list of 'Philippines only' innovations may well exceed one thousand in number.

Those who work in Austronesian languages of the Moluccas or further east may appreciate the following resources.

36. Blust (1978) discusses sound change in the Austronesian languages of southwestern Halmahera and western New Guinea, and their classification. An appendix lists fifty-six lexical and semantic innovations which help define Proto Eastern Malayo-Polynesian as a valid subgroup.

37. Blust (1993) discusses phonological, lexical and semantic innovations in the Central-Eastern Malayo-Polynesian and Central Malayo-Polynesian branches of the Austronesian family tree. An appendix traces the development 200 basic vocabulary items from PMP into PCEMP, PCMP, and finally Proto Oceanic.

38. Chuck Grimes (1990) is a 389-item wordlist arranged for comparative (not survey) purposes. Each entry gives a reconstructed etymon followed by representative reflexes, along with a suggested meaning or range of meanings. For example:

> ***ma-dingding**, **dingin**, **diding**, **ringring**, **riri**, **lingi**, **lilingi**, **eb-ridi-n**, **ringi**; 'be cold, cool, fresh, refreshing, chills (from fever)': _____
> _____

Etyma are not identified as to level (Proto Malayo-Polynesian, Proto Central-Eastern Malayo-Polynesian, etc.) so this information would have to be retrieved from elsewhere, e.g. in the above example one would normally cite PMP *ma-diŋdiŋ.

Other resources

Finally, don't forget to check if there is an indigenous language dictionary for your area which gives etymologies. I'm extremely fortunate, for example, to have at hand René van den Berg's (1996) *Muna-English Dictionary*, among other reasons that he has brief etymologies for a number of his entries. There will always be gaps in our work—for example, Blust (1997:196) rattled off some forty additional etymologies which Van den Berg had missed—but we shouldn't allow the thought of criticism to deter us from doing our best.

References

Adelaar, K. Alexander; and Robert Blust (eds.) 2002. *Between worlds: linguistic papers in memory of David John Prentice*. (Pacific Linguistics, 529.) Canberra: Australian National University.

Berg, René van den. 1996. *Muna-English dictionary*. Leiden: KITLV Press.

Blust, Robert A. 1978. Eastern Malayo-Polynesian: a subgrouping argument. *Second International Conference on Austronesian linguistics: proceedings*: fascicle 1, *Western Austronesian*. (Pacific Linguistics, C-61), edited by S. A. Wurm and Lois Carrington, 181–234. Canberra: Australian National University.

Blust, Robert. 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30:73–129.

Blust, Robert. 1993. Central and Central-Eastern Malayo-Polynesian. *Oceanic Linguistics* 32:241–293.

Blust, Robert. 1997. Book review: René van den Berg (1996, Muna-English dictionary). *Oceanic Linguistics* 36:194–198.

Blust, Robert. 1999. Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. *Selected papers from the Eighth International Conference on Austronesian Linguistics* (Symposium Series of the Institute of Linguistics, 1), edited by Elizabeth Zeitoun and Paul Jen-kuei Li, 31–94. Taipei: Academia Sinica.

Blust, Robert. 2002. The history of faunal terms in Austronesian languages. *Oceanic Linguistics* 41:89–139.

Blust, Robert. 2003. Three notes on early Austronesian morphology. *Oceanic Linguistics* 42:438–478.

Blust, Robert. 2005. The linguistic macrohistory of the Philippines: some speculations. *Current Issues in Philippine linguistics and anthropology: parangal kay Lawrence A. Reid*, edited by Hsiu-chuan Liao and Carl R. Galvez Rubino, 31–68. Manila: Linguistic Society of the Philippines and SIL Philippines.

Blust, Robert; and Stephen Trussel. 2010. *Austronesian comparative dictionary, web edition*. Online. URL: http://www.trussel2.com/acd/.

Casparis, J. G. de. 1997. *Sanskrit loan-words in Indonesian: an annotated check-list of words from Sanskrit in Indonesian and traditional Malay*. (NUSA Linguistic Studies of Indonesian and Other Languages in Indonesia, 41). Jakarta: Badan Penyelenggara Seri NUSA, Universitas Katolik Indonesia Atma Jaya.

Center for Research in Computational Linguistics. n.d. *Indic roots in Indonesian / Malaysian*, maintained by Doug Cooper. Bangkok: CRCL. Available at: http://crcl.th.net/indic/indo.htm (accessed May 20, 2004).

Dahl, Otto Christian. 1981. Austronesian numerals. *Historical linguistics in Indonesia,* part 1 (NUSA Linguistic Studies in Indonesian and Languages in Indonesia, 10), edited by Robert Blust, 46–58. Jakarta: Universitas Katolik Indonesia Atma Jaya.

Dam, Nikolaos van. 2010. Arabic loanwords in Indonesia revisited. *Bijdragen tot de Taal-, Land- en Volkenkunde* 166:218–243.

Dempwolff, Otto. 1938. *Vergleichende Lautlehre des austronesischen Wortschatzes*, volume 3: *Austronesisches Wörterverzeichnis*. (Zeitschrift für Eingeborenen Sprachen, Beihefte 19.) Berlin: Dietrich Reimer.

Dyen, Isidore; and Curtis McFarland. 1970. Proto-Austronesian etyma constituting an Austronesian cognate finder list. Unpublished typescript, 31 pp. Yale University.

Francisco, Juan R. 1964. *Indian influences in the Philippines, with special reference to language and literature*. Quezon City: University of the Philippines.Gonda, J. 1952. *Sanskrit in Indonesia*. Nagpur: International Academy of Indian Culture. [2nd edition 1973; reprinted 1998, New Dehli: International Academy of Indian Culture.]

Grijns, C. D.; J. W. de Vries, and L. Santa Maria (1983) *European loanwords in Indonesian*. Leiden: published for the Indonesian Etymological Project by the Koninklijk Instituut voor Taal-, Land, en Volkenkunde.

Grimes, Charles. 1990. Comparative wordlist for Austronesian languages of Maluku, Nusatenggara and Irian Jaya. Ambon: Universitas Pattimura and SIL.

Hamilton, A. W. 1924. Chinese loan-words in Malay. *Journal of the Malayan Branch of the Royal Asiatic Society*, 2:48–56.

Jones, Russell. 1978. *Arabic loan-words in Indonesian: a check-list of words of Arabic and Persian origin in Bahasa Indonesia and traditional Malay, in the reformed spelling.* London: School of Oriental and African Studies, University of London.

Jones, Russell. 1984. Loan-words in contemporary Indonesian. *Towards a description of contemporary Indonesian: preliminary studies II* (NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia, 19), edited by John W. M. Verhaar, 1–38. Jakarta: Badan Penyelenggara Seri Nusa, Universitas Katolik Indonesia Atma Jaya.

Jones, Russell. 1996. Chinese loan-words in Malay and Indonesian, a background study: introduction. Cornwall: prepublication draft, 40 pp. [available at SOAS Library, London, England.]

Jones, Russell. In press. *Chinese loan-words in Malay and Indonesian: a background study*. Kuala Lumpur: University of Malaysia. [ISBN: 983-100-074-9]

Jones, Russell (ed.) 2007. *Loan-words in Indonesian and Malay*. Leiden: KITLV Press.

Kong Yuan Zhi. 1987. A study of Chinese loanwords (from South Fujian dialects) in the Malay and Indonesian languages. *Bijdragen tot de Taal-, Land- en Volkenkunde* 143:452–467.

Kong Yuan Zhi. 1993. Kata pinjaman bahasa Cina dalam bahasa Melayu, 2 parts. *Jurnal Dewan Bahasa* 37:676–702, 772–793. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Leo, Philip. 1975. *Chinese loanwords spoken by the inhabitants of the city of Jakarta*. Jakarta: Lembaga Research Kebudayaan Nasional, LIPI.

Mahdi, Waruno. 1994a. Some Austronesian maverick protoforms with culture-historical implications—I. *Oceanic Linguistics* 33:167–229.

Mahdi, Waruno. 1994b. Some Austronesian maverick protoforms with culture-historical implications—II. *Oceanic Linguistics* 33:431–490.

Mahdi, Waruno, 2000, Review of J.G. de Casparis (1997, Sanskrit loan-words in Indonesian). *Bijdragen tot de Taal-, Land- en Volkenkunde* 156:844–852.

Mahdi, Waruno. 2009. Some obscure Austroasiatic borrowings in Indonesian and Old Malay. Paper presented at the 11th International Conference on Austronesian Linguistics, Aussois, France, June 22–26.

Mashudi, Kader; and Yeong Ah Lok. 1989. Pengaruh bahasa Cina dalam perbendaharaan kata bahasa Malaysia. *Dewan Bahasa* 33:236–267,

Minde, Don van. 2002. European loanwords in Ambonese Malay. In K. Alexander Adelaar and Robert Blust, eds., 2002, pp. 193–216.

Ngajenan, Mohamad. 1987. *Kamus etimologi bahasa Indonesia*. Semarang: Dahara Prize.

Robson, Stuart. 2002. Dutch loan translations in Indonesian. In K. Alexander Adelaar and Robert Blust, eds., 2002, pp. 139–148.

Ronkel, Ph. S. van. 1902. Het Tamil-element in het Maleisch. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 45:97–119.

Ronkel, Ph. S. van. 1903a. De oorsrpong van het Maleische woord *tjuma*. Over den oorsprong van het Maleische woord *binara*. Over de herkomst van het Maleische woord *djodo*. De oorsprong van het Maleishe woord *bagai*. Over den oorsprong van het Maleische woord *pawai*. De oorsprong van het Maleische woord *Kamandikai*. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 46:71–72, 92–94, 128–131, 241–242, 469–471, 476–477.

Ronkel, Ph. S. van. 1903b. Tamilwoorden in Maleisch gewaad. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 46:532–557.

Santa Maria, Luigi. 1996. I prestiti lessicali cinesi in malese-indonesiano nel loro contesto storico e sociale. *Studi in onore di Lionello Lanciotti* (Istituto Universitario Orientale, Dipartimento di Studi Asiatici, Series Minor 51), edited by Sandra Carletti,

Maurizia Sacchetti and Paolo Santangelo, 1171–1213. Napoli: Istituto Universitario Orientale.

Schlegel, Gustav. 1891. Chinese loanwords in the Malay language. *T'oung Pao* 1:391–405. Leiden: Brill.

Sedyawati, Edi; and Somvir. 2005. Sanskrit in Indonesia. *Sanskrit words in Southeast Asian languages*, edited by Satva Vrat Shastri, 1–220. Mumbai: Somaiya Publications.

Tadmor, Uri. 2009. Mon-Khmer loanwords in Malay-Indonesian: Linguistic and historical implications. Paper presented at the 11th International Conference on Austronesian Linguistics, Aussois, France, June 22–26.

Tryon, Darrell. 1995. Proto-Austronesian and the major Austronesian subgroups. *The Austronesians: historical and comparative perspectives*, edited by Peter Bellwood, James J. Fox and Darrell Tryon, 17–38. Canberra: Australian National University.

Vries, J. W. de. 1988. Dutch loanwords in Indonesian. *The sociolinguistics of Dutch* (International Journal of the Sociology of Language, 73), edited by Judith Stalpers and Florian Coulmas, 121–136. Berlin: Mouton de Gruyter.

Wolff, John U. 2010. *Proto-Austronesian phonology with glossary*. 2 vols. Ithaca, NY: Southeast Asia Program, Cornell University.

Wilkinson, R. J. 1959. *A Malay-English dictionary (Romanised)*. 2 parts. London: Macmillan.

Wurm, S. A. and B. Wilson. 1975. *English finderlist of reconstructions in Austronesian languages (post-Brandstetter).* (Pacific Linguistics, C-33.) Canberra: Australian National University.

Zorc, R. David. 1986. The genetic relationships of Philippine languages. *FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics* (Pacific Linguistics, C-94),, edited by Paul Geraghty, Lois Carrington, and S. A. Wurm, 147–173. Canberrra: Australian National University.

Zorc, R. David. 1995. A glossary of Austronesian reconstructions. *Comparative Austronesian dictionary: an introduction to Austronesian studies* (Trends in Linguistics, Documentation 10), part 1, edited by Darrel T. Tryon, 1105–1197. Berlin: Mouton de Gruyter.

Zorc, R. David and Matthew Charles. 1971. Proto-Philippine finder list. Unpublished typescript, 122 pp. Ithaca: Cornell University.