STUART CAMPBELL

The Distribution of *-at* and *-ah* Endings in Malay Loanwords from Arabic[1]

Introduction

Bahasa Malaysia (BM) and Bahasa Indonesia (BI) have many loanwords that derive ultimately from Arabic. An interesting subset of these loans comprises those words that have the *tā' marbūṭah* suffix in Arabic, which usually indicates natural or grammatical feminine gender in that language. In BM and BI, reflexes of these words may end in *-at* or *-ah*, e.g., *berkat* 'blessing' and *jenazah* 'corpse'. A rare variant of BM and BI *-ah* is *-a*, which occurs in words like *kahwa* 'coffee' and *menara* 'minaret'. Although *tā' marbūṭah* in Arabic has the *-at/-ah* variants, the distribution is different from that in BM/BI. In Arabic the variation is motivated by syntax, and in BM and BI by a complex of etymological processes. This article reviews previous research on the question, and attempts to put forward a principled explanation for the reasons for the BM/BI reflexes. The term 'Malay' is used to speak of 'BM' and 'BI' in general, while BM and BI will be used only to refer to points relating specifically to the modern forms of the language.

The explanation comprises three elements: a mixed historical source for Arabic loanwords is proposed, which includes a substantial Persianized Arabic element. This is overlaid by a standardizing tendency and an Arabizing tendency. The discussion opens with a brief account of *tā' marbūṭah* in Arabic, then moves to the methodological issues involved in establishing a checklist of *-at/-ah* words and some remarks on the con-

[1] The research reported here was carried out at the Language Acquisition Research Centre, University of Sydney, during Staff Development Leave granted to me by the University of Western Sydney Macarthur. The assistance of both universities is gratefully acknowledged. The advice of individuals who read and commented on the work is also gratefully acknowledged. They include Dr Geoffrey Hull, Dr Russell Jones and Dr Colin Yallop.

STUART CAMPBELL is Associate Professor and Acting Director of the Language Acquisition Research Centre at the University of Western Sydney Macarthur and took his PhD at the University of Sydney. Specializing in Arabic linguistics, he is the author of 'Towards a model of translation competence', *Meta* 36-2/3:329-43, and principal author of *Unlocking Australia's language potential. Profiles of 9 key languages in Australia: Arabic*, Melbourne, The National Languages and Literacy Institute of Australia. Professor Campbell may be contacted at the Faculty of Education, University of Western Sydney Macarthur, PO Box 555, Campbelltown, NSW 2560, Australia.

sistency and reliability of the checklist. Previous treatments of the *-at/-ah* question are then dealt with from synchronic and diachronic viewpoints. With the standardizing and Arabizing tendencies described in a preliminary fashion, the checklist is examined again to show the striking similarity between the distribution of *-at/-ah* in Malay and in Modern Persian. Finally the three elements are brought together in a summary of the explanation. It should be pointed out that this article does not intend to fix the etymology of large numbers of individual loanwords; instead, it seeks to put forward a framework that can be used for etymological work.

Tā' marbūṭah in Arabic

The Arabic term *tā' marbūṭah* means 'tied t', as does its BM translation *ta simpul*.[2] The Arabic letter that indicates the suffix has the shape of the letter *h*, with two dots written above (ۃ). These dots are otherwise found on the ordinary letter for *t*, one of the boat-shaped characters (ت) which, it is traditionally held, has its ends bent upwards and tied to form the tied *t*. Hence, the very form of the Arabic *tā' marbūṭah* contains the seeds of ambiguity – the *h* shape and the *t* dots. In fact, in Arabic both *-ah* and *-at* are possible pronunciations of *tā' marbūṭah*; as a noun suffix, it is pronounced *-ah* unless the word that bears it occurs as the non-final term in the *'iḍāfah* (possessed-possessor) construction, or unless the speaker is using an elevated form of speech, in which case inflections are added. In these cases the *-t* serves as the consonantal prop for the vocalic case inflections, as in the following examples.

(a) Neither first term of *'iḍāfah* nor case-marked:
'as-sayyārah huna
the-car (is) here
'The car is here.'

(b) As first term of *'iḍāfah*, not case-marked:
sayyārat 'al-mudīr huna
car the-manager here
'The manager's car is here.'

(c) In case-marked speech:
'as-sayyārat-u huna
the-car-nominative here
'The car is here.'

---

[2] Lewis (1958:47) uses the terms *ta bersimpul* and *ta merbuta*. For consistency *ta simpul* is used henceforth in the present article, and *ta panjang* is used for the regular open *t*.

(d) As first term of *'iḍāfah*, case-marked:
sayyārat-u l-mudīr huna
car-nominative the-manager here
'The manager's car is here.'

When *tā' marbūṭah* is used as a feminine adjective suffix, it occurs as *-ah* in non case-marked speech, and as *-at* in case-marked speech. There is also a rare construction known as the false *'iḍāfah*, in which the adjective can occur as the first element, and in this case, the *-at* ending appears just as it does in the nominal *'iḍāfah*. The various possibilities for feminine adjectives are as follows.

(a) With no case marking:
'al-sayyārah 'al-jadīd-ah
the-car the-new-feminine
'The new car.'

(b) In case-marked speech:
'al-sayyārat-u l-jadīd-at-u
the-car-nominative the-new-feminine-nominative
'The new car.'

(c) As first term of false *'iḍāfah*, case-marked:
jamīl-at-u l-wajh
beautiful-feminine-nominative the-face
'Having a beautiful face.'

Methodological issues

A major difficulty is that of establishing a checklist of these loanwords. The ideal source would be a large selection of texts which could be electronically searched, as is the practice in modern corpus linguistics. The first difficulty is that no parallel to, say, the Lancaster-Oslo-Bergen Corpus of British English exists for Malay. Even though such a corpus might eventually be compiled from texts in Roman script by electronic scanning, Jawi texts pose problems; while electronically produced Jawi could be made comprehensible to a corpus analysis program, older texts produced by hand or by poor printing techniques are probably not amenable to scanning. Even if these technical difficulties were overcome, Jawi texts would prove poor sources for this particular kind of study because it is seldom possible to know whether a *ta simpul* ending is pronounced as *-ah* or *-at*. The only exceptions would be provided by fully vowelled texts (which are rare) or by cases where the pronunciation could be inferred from a rhyme in poetry (which would yield too few examples for an empirical study). In short, while a text-based study would be ideal, the

basic tools do not exist.

In this case, the researcher's next recourse is the efforts of language experts to record usage in compendia and dictionaries. For the purpose of this research, accounts of loanwords by two native speaker authorities were used as sources, and two major dictionaries were used as confirmation.

It is extremely difficult to reconcile the various authorities on the question of the extent and form of Arabic loanwords. Beg mentions numerous attempts to list Arabic loanwords in Malay and tabulates eight such works that appeared between 1801 and 1975, in which estimates vary from 150 to 1,125 loanwords (Beg 1979:81). He estimates that there are perhaps 1,000 such loans in current usage (Beg 1979:78). A more recent list is that of Jones (1978), which lists 2750 Arabic and Persian loans[3], while the Indonesian etymological dictionary of Ngajenan (1987) lists a considerable number. For the purpose of this study a methodology was used that aimed at compiling a checklist from a range of sources of very different types: first, *-at/-ah* words were extracted from the lists contained in two recent studies of Arabic loanwords in BM, namely Beg (1979) and Kasimin (1987). The advantage of using the two studies lies in their difference in approach. Beg has an orientation that leans towards modern linguistic method, and is perhaps more conservative. Kasimin has a much more liberal view on Arabic loanwords, including in his list many apparently recent and unassimilated loans. As a result, the two lists vary considerably. The lists were then checked against BM and BI dictionaries, this time to ensure a wider Malay perspective beyond the Johore/Riau split. *Kamus lengkap* (Hairul and Khan 1982) and *Kamus besar Bahasa Indonesia* (Moeliono et al. 1989) were the yardsticks. It may be asked why Jones (1978) was not used as a source, given its comprehensiveness. In fact, its comprehensiveness, systematic lexicographic approach and range of sources argued against its use; the value of Beg and Kasimin lies in the differences between them, and what these reveal about the instability and uncertainty of Arabic loanwords in Malay.

Establishing a checklist of -ah *and* -at *words*

Combination of the lists of Kasimin and Beg yielded 408 items (including a handful of two-word combinations) – 343 from Kasimin and 251 from Beg; the 408 included some items that appeared in both lists, some given only by Kasimin, and some given only by Beg. Of the 408, *Kamus besar* confirmed 286, and *Kamus lengkap* confirmed 300. Bear in mind that the two dictionaries certainly contain *-at/-ah* words that do not appear in Kasimin's or Beg's list, and that all four sources may fail to list some

[3] After the publication of this list, Dr Jones had three young university lecturers in Indonesia examine the list independently; they were familiar with only about 10% of the total number of words (personal communication).

current items: Beg, for example, does not list *surat* 'letter' (although see below for a comment on the origin of this word). The checklist is intended to be a broadly representative database rather than an exhaustive one.

Of the 408 items, there was a core of 176 items that were attested by all four authorities. The origin of the items in the list is not to be taken as confirmed. There are numerous doubtful cases and even conflicts between authorities. For instance the *kalimah/kalimat* doublet raises the question of whether *kalimat* 'sentence' might be derived from the Arabic plural *kalimāt* 'words' and *kalimah* 'creed' from the Arabic singular *kalimah* 'word'. Similarly, *sejarah* 'history' and *syajrat* 'genealogy' appear in a single entry; Kasimin derives *sejarah* from *sīrah* 'biography, history', Beg derives *syajrat* from *šajarah* 'tree'. Jones, meanwhile, states (1984:17) that *šajarah* has contributed both *sejarah* and *syajarah* (the *-ah* variant of Beg's *syajrat*). As mentioned above, Beg does not include *surat* 'letter'; does this imply agreement with Jones (1984:16) that *menyurat* 'to write' is a Malay word, even though it is apparently a regular derivative of *surat*, to which Ngajenan (1987) and Kasimin both assign an Arabic origin? The items *alkisah* and *kisah* seem to be considered separate despite their common origin in *qiṣṣah* 'story', and are thus listed separately. The entangled *sunah/sunat/sunnah/sunnat* group is considered as a single item despite two meanings, 'circumcision, orthodox law', and four spellings. On the whole, decisions to list items uniquely or together probably cancel each other out and do not seriously affect the calculations made below.

Consistency of the sources for the checklist

Some simple calculations were made to establish the consistency of the four sources for the checklist. These revealed the following facts.

227 items (55.63% of the total 408 items) were attested as *-ah* words by at least one authority with no disagreements or alternatives. These had the following measure of agreement:

All four authorities agreed on 69 items (30.40%)
Three authorities agreed on 42 items (18.50%)
Two authorities agreed on 28 items (12.33%)
One authority only agreed on 88 items (38.77%)

103 items (25.25% of the total 408 items) were attested as *-at* words by at least one authority with no disagreements or alternatives, with this degree of agreement:

All four authorities agreed on 63 items (61.17%%)
Three authorities agreed on 28 items (27.18%)
Two authorities agreed on 10 items (9.71%)
One authority only agreed on 2 items (1.94%)

There was disagreement by at least one authority, or alternatives were attested by at least one authority, in the case of 78 items (19.12% of the total 408 items).

An immediate observation is that there is considerable disagreement among the authorities about the *-ah* words. This disagreement is largely accounted for by the 103 items that were attested by one authority only. In fact, most of these are apparently unassimilated Arabic loans from Kasimin's list – he has 80 items that are not attested by the other authorities.

An Arabizing tendency

This rather simple observation has deeper significance for the explanation of *-at* and *-ah*, namely that recent loans tend to be given the unmarked Arabic *-ah* rather than *-at*; this is an aspect of an Arabizing tendency, an important element in Malay loanword etymology.[4] This tendency may be stronger in BM than in BI. As Alisjahbana remarks, 'In Malaysia the tendency is to choose Arabic words because of the Islamic character of the Malay culture: iktisad for ekonomi (economy), intikad for keritik (critic) ...' (Alisjahbana 1976:120). Kasimin's 80 unique words are telling in this regard: all but two have the *-ah* ending.

Nevertheless, the Arabizing tendency is not absent in BI. Abas, in discussing the terminology work of the Commission on Indonesian established by the Japanese military administration in 1942, notes that 'In most cases, the decision regarding a new term depended upon the composition of the members present at a particular meeting ... devout Muslims tended to prefer words of Arabic origin' (Abas 1987:45).[5]

One implication of this aspect of the Arabizing tendency is that *-at* words are likely to be less recent loans, and another simple analysis will go some way to confirm this. A not unreasonable assumption is that older loans are more likely to have accretions of derived forms. The list was checked against the two dictionaries for such derivatives. Of the 408 items, 125 were recorded by at least one dictionary as having derived forms. Of these, 35 (28.00%) were *-ah* words (attested with no disagreement by at least one authority), and 62 (49.60%) were *-at* words (similarly attested). This difference is a clear pointer that *-at* words are likely to be older loans

[4] The presence of unassimilated Arabic loans is not only the case in Malay: Al-Harbi, in a study of the phonology of Arabic loanwords in Acehnese, purged his database of many 'unassimilated loans, in the sense that they are used exclusively in hikayat (written narrative stories in verse)' (Al-Harbi 1993:102).

[5] Between 1951 and 1959 the Ministry of Information of the Republic of Indonesia published a magazine called *Al-Shu'oun Al-Indonesiyah* in Arabic every two months. It contained social and political articles as well as letters from readers in Indonesia and beyond, including Hadramaut. Whether this publication had any role in facilitating Arabic loans is unknown.

than *-ah* words. At this point, the discussion will move to some wider issues before the checklist is brought back into this account.

Previous accounts of the distribution of -ah *and* -at *words*

The *-at/-ah* suffix is mentioned peripherally by a number of writers on Malay linguistics, but there seem to be no systematic treatments either from a synchronic or from a diachronic point of view. There are a number of studies that do not mention *-at/-ah* at all, but contain insights that will help in the present attempt to explain their distribution. What is intriguing is that the topic seems to have evaded a number of seminal works or works by leading authors in the field. For example, Van Ronkel 1899, usually cited as a significant work on the influence of Arabic on Malay, does not mention *-at/-ah*. While the article is mainly concerned with the influence of Arabic syntax on Malay, it is a little surprising that the intrusion of the Arabic *'iḍāfah* construction is not mentioned, since it still has echoes in BM and BI and is at least a matter of phrase level syntax. Similarly, Abdullah Hassan (1974) does not discuss *-at/-ah* in his work on the morphology of Malay, nor does Teeuw (1962) in his treatment of word classes. The extensive bibliography of Soviet works on Indonesian linguistics in Demidyuk and Makarenko (1980) does not contain a title dealing explicitly with the subject, although a thorough search of this literature still needs to be made. Stokhof (1975) discusses phonological doublets of the *f-p* variety without mentioning *-at/-ah*. Neither are they mentioned by Mahdi (1981) in his very extensive treatment of doublets. Three important works on Jawi spelling – Shellabear 1901, Zainal-Abidin bin Ahmad 1928 and Winstedt 1941 – ignore *ta simpul* altogether, concentrating largely on the problem of reconciling a spelling system developed to cope with a long/short vowel distinction with a language that has no such distinction. The last sentence of Winstedt's article warns that 'The spelling of Arabic loanwords requires special study' (Winstedt 1941:233). The guide to Indonesian spelling by Badudu (1984) does not tackle the issue of *-at/-ah* doublets. Drewes (1971) is also silent on the matter. Indeed, one is inclined to agree with Teeuw, who remarks that on the matter of Arabic influence '... not much more than a start has been made' (Teeuw 1959:154).

So much for the works that do not mention the topic. Turning to those that do mention it or that can throw some indirect light on it, it is useful to divide these studies into synchronic and diachronic treatments.

From a synchronic point of view there are useful studies dealing with affixation, orthography and standardization. Verhaar takes the view that, because of the connection between Islam and Arabic, Indonesians are more conscious of the origins of Arabic than of Sanskrit loans, and that this consciousness extends to loan affixes (Verhaar 1984:21). He claims that there may be a higher awareness of Arabic affixes in Muslims, so that, while

a non-Muslim may say *kaum muslim* 'the Muslim community', a Muslim might add the Arabic plural affix, thus: *kaum muslimin* (Verhaar 1984:22). A similar difference is claimed for the suffix *-at*, which '... seems to be more limited to Muslim speakers of Indonesian' in, for example, the feminine plural *muslimat* 'female Muslim' (Verhaar 1984:22). Alisjahbana (1963:54) also discusses the *muslimin* and *muslimat* types. The reader with little knowledge of Arabic should, however, beware: the *-at* referred to is a reflex of the Arabic feminine plural *-āt*, and not *tā' marbūṭah*. Where reflexes of *tā' marbūṭah* are discussed, Verhaar only mentions *-ah*, which 'occurs as a feminine ending in a few words' and is unproductive. Examples are *almarhumah* 'female deceased' and *syarifah* 'female descendant of the Prophet' (Verhaar 1984:22). The suffix *-iah* is also mentioned as a productive adjectivizer, as in *ilmiah* 'scientific' (from *ilmu*) and *harafiah* 'literal' (from *huruf*) (Verhaar 1984:22). This latter example is indeed a reflex of *tā' marbūṭah*; in Arabic it is added to the adjectivizing suffix *-īyy* to make feminine adjectives with *-īyyah*. In summary, what Verhaar has done is to isolate two reflexes of *tā' marbūṭah* which can be considered productive in Indonesian, and to implicitly assume all other manifestations of *-at/-ah* to be unproductive. In terms of the explanation for the distribution of *-at/-ah* which this article seeks to put forward, Verhaar's productive suffixes are examples of the Arabizing tendency, in that they build Arabic morphology onto existing Arabic loans or borrow Arabic morphology direct.

The confounding factor of orthography

The matter of orthography adds an extra tangle to the question of *-at/-ah*. Although Jawi script is virtually extinct in Indonesia, it is still current for some purposes in Malaysia. Instruction in the Kitab Jawi, for example, still forms an important part of the education of many Malays, who will read these instructional books on Islamic theology in Malay written in Jawi script (see Nor bin Ngah 1983 for a discussion of Kitab Jawi in general, and Matheson and Hooker 1988 for an account of Patani Kitab Jawi). In the transition from Jawi script in both Indonesia and Malaysia, the representation of Arabic phonology and morphology has become less and less of a concern. In the Van Ophuijsen orthography introduced into Indonesia at the turn of the century, there were attempts to identify Arabic phonemes, so that the glottal stop received the symbol ['] (Abas 1987:105). Similarly, a diacritic ["] was used to show that the letter bearing it was the initial letter of the second part of an Arabic compound word, e.g., *aläswad* (Abas 1987:85). Indeed, one of the recommendations of the 1968 Seminar on Indonesian in Jakarta was 'To suggest [to the committee to formulate the way to write Arabic words using the Latin alphabet in Indonesia] to pay special attention to the need of diacritics for differentiating the pronunciation of words which have two different meanings' (Abas 1987:75).

That this is a forlorn hope is now obvious; in the revised new orthography common to Indonesia and Malaysia dating from 1972 there is virtually no provision for retrieving the phonology of Arabic, let alone its morphology. Jawi script is, in fact, apparently rather careless of reflecting *tā' marbūṭah*, and this may have contributed to the present vagueness with which it is sometimes dealt with in Latin script. Lewis gives *ta simpul* as a 'variation' of the regular open *t* (ت) in word-final position (Lewis 1958:56), and also as a variant of the regular *h* symbol (ه), even in indigenous Malay words, which '... is sometimes written with two dots over it at the end of a word. When so written it should be pronounced as a t, e.g. itu ايتة, serta سرتة.' (Lewis 1958:47.) This matter will be raised again in the discussion of the diachronic aspect of *-at/-ah*.

Meanwhile, Kasimin claims that 'Perkataan Arab yang terdiri daripada nama yang berakhir huruf ta simpul (ة), biasanya digantikan dengan huruf ta' (ت), ini berlaku pada pertuturan dan tulisan, misalnya: sifat, keramat, alamat ...' (Kasimin 1987:18).[6] However, the issue is somewhat clouded by the fact that, while *keramat* is given as an example, another reflex, *karamah*, is given on page 76.

De Hollander, in his discussion of Jawi letter shapes, remarks on the fact that *ta simpul* is sometimes pronounced *-at* and sometimes *h* (De Hollander 1984:9). He goes on to give an account of the difference between *ta simpul* and *ta panjang* ('long *t*') that is surely very ambitious. He claims that 'Huruf (ة) yang terdapat dalam akhiran jenis wanita at pada kata-kata Arab, diucapkan sebagai t bila diikuti oleh atribut' (De Hollander 1984:18).[7]

Thus, for De Hollander, (كعبة) is pronounced *kabah* when it is not followed by an attribute, but *kabat* when it is, thus *kabatoe'llah* (De Hollander 1984:18). In fact, De Hollander is following an aspect of the distribution of *-at/-ah* in Arabic, and there is some truth in what he claims – Kasimin has four pairs of loans where there is an *-ah* ending when the word is in isolation and an *-at* one when it is followed by a possessor, as in:

> jannah/jannatulnaim 'garden/garden of happiness, paradise', lailah/lailatulqadar 'night/night of the 26th-27th of the month of fasting', silah/silatur rahim 'bond/bond of friendship', walimah/walimatulurus 'wedding party/wedding party of the groom'.

---

[6] 'In Arabic words consisting of nouns that end in a *ta simpul* (ة), this is usually replaced with a *ta'* (ت). This is the case in speech and writing, for example: *sifat, keramat, alamat* ...'

[7] 'The letter (ة) which is found in the feminine ending *at* in Arabic words is pronounced *t* when it is followed by an attribute.'

However, it would be ambitious indeed to claim that this is a productive pattern that extends beyond a handful of barely assimilated loan phrases. To postulate such a rule for BM or BI is undoubtedly another example of the Arabizing tendency mentioned previously.

To sum up these comments on orthography, it is reasonable to claim that synchronically, the use of *ta simpul* in contemporary Jawi script is somewhat unsystematic despite some attempts to impose on it Arabic morphological rules. *Daftar ejaan*, which must be considered the final arbiter on contemporary Jawi spelling in BM, gives, for example, the dotted *ta simpul* for *bidaah* and *adat*, and the undotted Jawi *h* for *akidah*. In view of the earlier comments on the productivity of *-ah*, it can be said that any synchronic processes at work will contribute only part of the explanation for the distribution of *-at/-ah*.

Standardization

The final aspect of synchronic interest is standardization. Mahdi discusses phonological doublets of the *pikir/fikir* type where an assimilated and unassimilated loan exist side by side. He suggests that there is a tendency for speakers to settle for one or the other: 'For /f-p/ doublets, [Metropolitan Indonesian] exhibits increasing standardisation whereby either the /f/- or the /p/- mode is ultimately retained' (Mahdi 1981:402).

While Mahdi does not mention *-at/-ah* doublets, their existence does beg the question of preferred forms. In fact, Sarumpaet gives some guidance on this question. This book lists numerous words about which there are questions of usage, and gives hints to the student on how to overcome these problems. Scattered among these are some *-at/-ah* words. For example, in the case of *derajat/ darajat/ derjat/ darjah/ derjah/ daranja*, Sarumpaet says 'All these are variant spellings for degree, stage, level, but use *derajat'*. He adds that 'In Malaysia *darjah* means class ...' (Sarumpaet 1980:56). What makes it clear that this is a process of standardization (and not, for example, Arabization) is that either the *-at* or *-ah* ending may be the preferred form, thus in the case of *ziarah/ ziarat/ jiarah*, *'Ziarah* is now used for pilgrimage to the grave of a holy man or hero' (Sarumpaet 1980:223). Whether the ordering of *-at/-ah* alternatives in dictionaries indicates preferred forms is difficult to say; on the whole, *-ah* forms seem to be listed before *-at* forms, but this may simply be a matter of alphabetical ordering for lexicographic consistency.

This standardizing tendency can now be added to the Arabizing tendency as constituting the two main dynamic influences on the distribution of *-at/-ah* and two of the principles in the explanation attempted here: in synchronic terms, the distribution of *-at/-ah* is in the main not motivated by productive rules. Where doublets exist, the Arabizing tendency may attempt to impose Arabic rules (i.e., *-ah* in

isolation, *-at* in phrases). At the same time the standardizing tendency will fix on one or the other as a preferred form without regard to the syntactic environment. The discussion now turns to the diachronic aspect of the subject.

History of the influence of Arabic on Malay

This article is not the place to deal at any length with the history of the influence of the Arabic language on Malay. In fact, the topic has barely been dealt with at all. The cognate area of the historical impact of Islam on the Malay world is better, if imperfectly, understood. Drewes (1968) presents a critical treatment of theories on the origins of Indonesian Islam, in particular disposing of the earlier idea that the Arab world, rather than India, was its source – apparently the view behind the claim by Shellabear that 'It is generally agreed that the Arabs gained ascendancy over the Malays during the 13th century, and that it was from them that the Malays received their present written character' (Shellabear 1901:75).

Ricklefs presents a thorough account of the primary sources for the introduction of Islam – inscriptions and travellers' accounts – and a critical discussion of the various theories, noting that '... the records of Islamisation which survive are so few, and often so uninformative' (Ricklefs 1981:3).

The prehistorian Bellwood summarizes historical and linguistic evidence to paint a picture of early Arab and Persian penetration by traders followed by Islamization from India. Arabic and Persian loanwords are said to have a mainly Indian source (Bellwood 1985:140-1).

Using historical theories about the coming of Islam, the linguist is faced with a plethora of possibilities for the route by which the Arabic language accompanied the Muslim faith. While the weight of argument favours India, other scholars have suggested in addition a northern source: Fatimi, in his discussion of a connection between Champa and Malaysian Islam, traces a line of Islamic influence '... all through the eastern coast facing the China Sea: Phan-rang, Patani, Trengganu, Pahang and Leran' (Fatimi 1963:67). A very recent report of an Arabic gravestone in Brunei suggests that the stone was made in Quangshou around 1301 AD and then transported to Brunei (Chen Da-Sheng 1992).

The idea that Islam came to the region by a number of routes is paralleled by the idea that Arabic was also carried by a number of routes – perhaps secular and religious.[8] In his treatment of Persian words in Malay,

[8] Dr Geoffrey Hull has pointed out to me that there is a parallel with the influence of Portuguese on the Tetum language spoken in Timor. While Christian speakers in Dili use Portuguese loans to refer to Catholic culture, rural animist speakers use cultural loans from outside the domain of religion. The idea of parallel religious and secular sources for the introduction of Arabic into Malay is an attractive one, which may help to explain the persistence of Persianized spellings of Arabic loans, when the

Bausani reflects this notion, saying: '... en ce qui concerne les termes culturels arabes, il n'est pas toujours facile de dire s'ils sont entrés directement de l'arabe ... ou par la persano-indien' (Bausani 1974:349).[9]

The terms that have come direct from Arabic may themselves have come via numerous routes, for example, 'par d'anciens marchands ou religieux, ou plus récemment, par l'influence de l'Egypte ou de Hedjaz et de l'Hadramuth' (Bausani 1974:349).[10] The advice offered by Jones is similarly that '... further investigation will be needed into the finer points, such as the variety of Arabic (or Persian) from which [loanwords] may have come, whether from the Kuran, or classical Arabic or a dialect of Arabic' (Jones 1984:12).

Bausani goes on to remind the reader of the fact that the Malayo-Indonesian contingent of pilgrims to Mecca is among the biggest each year, and that Malay books in Jawi script are still printed in Mecca (Bausani 1974:349). This point is echoed by Kasimin, who stresses the importance of the direct influence of the Arab world on BM during this century. He writes:

> 'Bilangan orang Melayu yang pergi ke Mekah ketika harga getah melambung tinggi (1903-1910) begitu besar, melebihi bilangan bangsa-bangsa lain. Banyak di antara mereka yang belajar di Mekah dan kembali ke tanah air mengembangkan akidah Islam setelah beberapa tahun disana. Usaha mempelajari ilmu keagamaan itu bertambah pesat dengan pemergian setengah-setengah penduduk setempat ke Mesir dan ke negeri-negeri lain di Timur Tengah.' (Kasimin 1987:13.)[11]

Jones, like Bausani, notes the difficulty of distinguishing direct Arabic loans from those that have come through Persian. In addition, 'A further complication rests in the fact that many loanwords from Arabic may have come indirectly via one of the "Muslim" languages of India' (Jones 1984:12).

If Arabic came to Malay via many routes, then it is likely that a potentially unstable linguistic element such as *tā' marbūṭah* will be somewhat travel-worn; its modern reflexes may carry traces of the various journeys they have made. For example, there is a striking similarity

---

presumably ubiquitous text of the Holy Qur'an would have always been a source of 'correction'; the simultaneous existence of Persianized and Arabic spellings surely points to a separation between the 'religious' and 'secular' spheres of life, in the face of Islam's insistence on there being no such separation.

[9] '... with regard to Arabic cultural terms, it is not always easy to say whether they have come direct from Arabic ... or via Perso-Indian.'

[10] 'Via the merchants or preachers of ancient times, or, more recently, by way of the influence of Egypt, Hejaz or Hadramaut.'

[11] 'The number of Malays who went to Mecca when the price of latex soared (1903-1910) was very high, exceeding the number from other countries. Many of those who studied in Mecca returned to the homeland to develop the Islamic faith after having spent some years there. The effort to study theology accelerated as some local inhabitants went to Egypt and to other countries in the Middle East.'

between the distribution of *-at/-ah* endings in Malay and that of their equivalents in Modern Persian. This third element in the explanation of the distribution of *-at/-ah* is what has been referred to above as the Persianized Arabic element.

To complete this discussion of the diachronic dimension, the question of Jawi script is briefly revisited. Historically, Jawi script writes *-ah* endings with the *h* symbol (ه), and sometimes with *ta simpul*, and appears to have allowed variation in the writing of *-at* endings. A check of the text samples in Lewis (1958), which date from 1602 (Text 36) to the middle of this century, reveals that in a single text there are some *-at* words spelt with *ta simpul* (ۃ) and some with *ta panjang* (ت). In a few cases, the same word appears with both variants, and it is possible to detect trends from one text to another: one might use mostly *ta simpul* and occasionally *ta panjang*, while another might show the reverse trend. There are also examples where a letter form is used that is midway between the two. This appears to be the case in 19th- and 20th-century texts, and a similar picture appears in the 17th-century texts in Gallop (1991). This raises a number of issues. One is that, given that *ta simpul* appears to be a variant of both *h* (ه) and *ta panjang*, there may have been *-at/-ah* pronunciation variants of some words for a very long time; the ultimate choice between the two may have been a result of the standardizing tendency in earlier times. A second issue is that – although the data are too scanty for a reliable statistical analysis, since each text contains only a handful of these words – it may be that individual scribes have been more or less inclined to spell Arabic loanwords with their original Arabic spelling, that is, giving *-at* words *ta simpul* rather than *ta panjang*. Where *ta panjang* is used, one might suggest that scribes are treating these words as assimilated loans. A modern example (Lewis's Text 31) is the spelling (حيكا يت) for *hikayat* 'story', where the short Arabic vowel *i* has been lengthened and the *-at* written as *ta panjang*.

A third possible explanation is that some scribes used Persianized spellings of *-at* words, that is, with *ta panjang*, the regular method of spelling *-at* in Persian, which has not adopted the Arabic *tā' marbūṭah* symbol. This would certainly conflict with the view of Shellabear, who suggests '... that for many years, perhaps for centuries, the art of writing may have been almost entirely confined to those Arabs who had learned the Malay language' (Shellabear 1901:77).

In the same article Van de(r) Wall[12] is cited as saying that the original Jawi spelling is known as 'Achinese spelling' (Shellabear 1901:89) – the same 'spelling of the MSS. of the 17th century' that displays 'remarkable uniformity' (Shellabear 1901:77). In fact, the treatment of *-at* words

[12] The name is spelt differently in two different places. Dr Russell Jones suggests to me that Shellabear is probably referring to A.F. von de Wall.

appears not to be uniform in all 17th-century texts. This is not to deny the fact that places such as Aceh and Patani have had, and still have, considerable direct intercourse with the Arabs of the Arabian Peninsula, especially Hadramaut. Fatimi mentions the Arabs of Hadramaut as a source of influence of the Shafi'i school of law (Fatimi 1963:34); Van der Meulen and Von Wissman (1932) describe thriving Malay colonies in the Hadramaut in the early years of this century; and Matheson and Hooker (1988:13-4) mention the large numbers of Patani Malays who are resident in Saudi Arabia. The Jawi spelling of *-at* words does, however, raise the strong possibility of a heavy Persianized Arabic influence at some point in the coming of Arabic to the Malay world. This rather speculative view is a convenient point for the return of the discussion to the checklist and to the development of a more solid argument for a Persianized Arabic source.

Persian as a vector of Arabic

Arabic script was adopted by Persian earlier than it was by Malay. According to Windfuhr, 'Book Pahlavi ... was still used in the early centuries of Islam, until the 9th century and locally even later. Then the Arabic script began to win over and spread ...' (Windfuhr 1979:149.)

While one of 'the earliest [Persian] manuscripts in Arabic script is Movaffaq's pharmacological compendium dating from 1055 A.D.', and 'the number of texts from the 10th century onward is quite extensive' (Windfuhr 1979:171), the Trengganu Stone dating from the 14th century is usually claimed as the earliest Malay document in Arabic script (Jones 1984:12).

Given this fact, Persian must be considered a possible vector for Arabic loanwords in Malay, and this can be shown very easily by comparing Malay reflexes of *-at/-ah* words with their reflexes in Modern Persian.[13] Before this comparison is made, though, a little discussion of Persian and the Persian data sources is necessary.

Modern Persian reflects Arabic *tā' marbūṭah* in two ways. In some words, the ending is written with the *h* symbol (ه) and pronounced *-e*. In others it is written with the regular *t* symbol (ت) and pronounced *-at*. Lambton explains that the tendency is for *tā' marbūṭah* to be written with (ت) in certain Arabic-derived forms, and with (ه) in other Arabic-derived forms and when it occurs as a feminine suffix (clearly natural gender is

[13] This article omits to explore the semantics of Arabic loans, which is a potentially rich source of study in itself. While Beg (1979), Jones (1984), Tham Seong Chee (1990) and others attempt semantic classifications of these loans, a Persianized Arabic theory will require attention to the semantic substrate. Telegdi (1974) argues that the Arab invasion of Persia led to a state of diglossia, with Arabic as the high variety; Arabic loans into Persian were largely 'mots savants' (learned words) and few verbs. Any semantic classification of Persianized Arabic loans in Malay must be made in the light of the semantics of this putative substrate.

meant here, since Persian does not have grammatical gender). A small number of *-at/-e* doublets exists in Persian, e.g., *erade* 'will', *eradat* 'respect for (someone)'[14], and *amniye* 'gendarmerie', *amniyat* 'security' (Lambton 1960:13-4). Older dictionaries do not all reflect the *-at/-e* distinction clearly, perhaps because of the prominence given the Arabic element. In fact, Steingass's *Comprehensive Persian-English dictionary* seems to treat Arabic words as unassimilated loans, spelling almost all *tā' marbūṭah* words exactly as they are spelt in Arabic. In the Romanized transcriptions most of the cognates of Malay *-at* have *-at*, while among the cognates of Malay *-ah* many have both *-at* and *-a*, some have *-at*, and a few have *-a*. Steingass' lineage goes back well into the nineteenth century, when, one suspects, Persianists may have been somewhat mesmerized by the Arabic influence. Meanwhile, Lambton's 1961 *Persian vocabulary*, a companion to her 1960 *Persian grammar*, clearly distinguishes between *-at* and *-e* forms in spelling (ۃ versus ہ) and transcription (*-at* versus *-e*). It was for this reason that her Vocabulary was chosen for matching Persian and Malay cognates. In case there might be the slightest suspicion that the data source was chosen to fit the hypothesis, an independent check was made against Hindustani cognates in Craven's 1911 *New royal dictionary*, which gives a very similar distribution of *-at/-ah=-e* forms to Lambton's Vocabulary. This dictionary includes only a Romanized transcription[15], perhaps a guarantee that the compilers have not been swayed by the Arabic spellings. Indeed, if the vectors of Persianized Arabic words included Persianized Indian languages, then nineteenth-century Hindustani ought to be considered a useful yardstick.[16] A final word

---

[14] Lambton's Vocabulary lists only *erade*, although the Grammar gives both *erade* and *eradat*.

[15] Modern Turkish represents another useful data source, since it has not been written in Arabic script since 1928. It shares the Modern Persian *-at/-ah* distribution and spells the endings as *-at/-et* and *-a/-e*, with the *a/e* alternation being explained largely by vowel harmony. Old grammars of Turkish show that the Persian spellings (ۃ) and (ہ) were used. See, for example, Barker 1854.

[16] This is not to suggest that Hindustani was necessarily a vector of Arabic words. Tracing the provenance of such words in Indian languages is probably a hopelessly tangled task. For one, the fluidity of boundaries between Indian languages in the fourteenth century was undoubtedly no less than it is today, when '... from the Panjab to Bangladesh, there is not one sharply defined linguistic boundary ...' (Zograph 1982:22). The Persianization of Hindustani took place during various periods and at various places. Brass uses the term Urdu for the precursor of Hindustani, saying: 'As Muslim rule spread over north India, especially after the eleventh century, Urdu went with it as the court language and interacted with local dialects, introducing Persian words into them and adopting indigenous vocabulary from them, but gradually developing a more or less standard spoken, urban form distinct from the local dialects' (Brass 1974:128). According to Zograph, the use of Hindustani as a literary language dates only from the sixteenth or seventeenth century, in the form of the Dakhnī dialect, used in the Muslim courts of South India. The Persianization of the Rēkhtā form (later called Urdū) in North India began from the eighteenth century

should be said in defence of Lambton's Vocabulary, which only contains about 5,000 Persian items and is after all a companion to her Grammar, with some supplementation, and not a comprehensive dictionary: for the present study, its significance lies in the proportion of correspondences between Malay and Persian, not the absolute number; what is sought is a pattern rather than individual etymologies. The Vocabulary contained sufficient cognates to provide this.

The checklist was compared with Lambton's *Persian vocabulary*. Of the 408 items in the checklist, 95 (23.28%) have Modern Persian cognates with *-at*, and 86 (21.08%) have cognates with *-e*. There are 227 items (55.64%) with no Modern Persian cognates. Comparison of the distributions, however, reveals an interesting picture.

Of the 86 Persian *-e* words, 63 (73.26%) had an uncontested *-ah* reflex in Malay, distributed as follows:

all four authorities agreed on 37 items
three authorities agreed on 10 items
two authorities agreed on 4 items
one authority only agreed on 12 items.

Of these *-e* words, 9 (10.47%) had an uncontested *-at* reflex in Malay, with this distribution:

all four authorities agreed on 8 items
three authorities agreed on 1 item.

Finally, 14 (16.28%) of the Persian *-e* words had Malay reflexes that either showed disagreement among the authorities, or had alternatives cited by at least one authority.

The picture for the 95 Persian *-at* words was as follows.

49 (51.58%) had an uncontested *-at* reflex in Malay, distributed as follows:

all four authorities agreed on 33 items
three authorities agreed on 9 items
two authorities agreed on 7 items.

(Zograph 1982:32). These dates are rather late for borrowing by Malay, of course. Bengali is another candidate; Chatterji says that 'Arabic words have come into Bengali, after these were naturalised in that language and had conformed to its phonetics' (Chatterji 1970:557). Perhaps Chatterji should have the last word on this tangle: 'A large percentage of Persian words have been borrowed from Hindōstāni ... it is now impossible to determine what words were directly borrowed from Persian and what words were borrowed from Hindōstāni...' (Chatterji 1970:560-1).

19 (20.00%) had an uncontested *-ah* reflex in Malay, with this distribution:
all four authorities agreed on 4 items
three authorities agreed on 5 items
two authorities agreed on 3 items
one authority only agreed on 7 items.

27 (28.42%) of the Persian *-at* words had Malay reflexes that either showed disagreement among the authorities, or had alternatives cited by at least one authority.

If the weakly attested items (i.e., those cited by only one or two authorities) are eliminated, so that a very conservative analysis results, the complementary distribution is still very obvious.

	Malay *-at*	Malay *-ah*	Malay contradictory
Persian *-e*	9	46	13
Persian *-at*	41	8	26

These items are listed in Appendices 1 and 2.

Each cell in this matrix raises questions, but the reason for the striking distribution of the two leftmost columns especially bears brief discussion. There are only two possible reasons for the complementary distribution. One is that both languages independently borrowed these words from the same source – Arabic – and that identical internal processes took place that resulted in this distribution. This is an implausible hypothesis, which could only be justified if a similar process were in train in Arabic itself at the time of the borrowing of these words; if that were the case, the results of the process would have been obvious in modern forms of Arabic – there would today be ossified *-ah* and *-at* forms, which there are not. This leaves the inescapable conclusion that there was borrowing of Arabic words via Persian or a Persianized language. The matrix should not be taken to mean that all the Malay reflexes have come from Persian Arabic sources; Modern Persian may have lost some older Arabic loans that are still current in Malay, while some Arabic loans in Persian may have arrived later than the period of borrowing by Malay. In the light of these general observations, three specific points can now be raised.

The first point is that the above matrix only provides conclusive evidence of a Persianized source for the *-at* words in Malay. As was explained earlier, *-at* is the marked form in Arabic, and direct loans from Arabic arrive in BM and BI with the *-ah* ending. Where *-ah* appears in Malay, it may be due either to a Persianized loan or to a direct loan.

The second point concerns those Malay words that have an uncon-

tested reflex contrary to the Persian word, that is, Persian *-e* for Malay *-at*, and Persian *-at* for Malay *-ah*. No obvious explanation applies in the former case, but one possibility is that words have been borrowed direct from Arabic as the first part of an *'iḍāfah*, carrying the required Arabic *-at* ending. In the latter case, a number of explanations are possible, including an Arabic rather than Persianized origin or a remodelling of a Persianized borrowing on Arabic lines. In either case, the ambiguity of *ta simpul* in Jawi script could be responsible.

The third point concerns the contradictory cases in Malay, that is, where both endings are attested for a single word. Any Malay *-at/-ah* doublets for Persian *-e* are difficult to explain, but doublets for Persian *-at* are more easily explicable: they may be due to the same word being borrowed from both Persianized and Arabic sources, or to Arabization of Persianized forms. In some cases, a choice of *-ah* or *-at* may have been made from earlier alternatives as a result of standardization. Again, the ambiguity of *ta simpul* in Jawi script is a possible explanation. The small number of *-at/-e* doublets in Persian may occasionally provide a solution.

Summary

With this statistical analysis complete, the explanation for the distribution of *-at/-ah* can now be summarized. The distribution of *-ah* and *-at* endings in BM and BI can be explained within a general framework that comprises a diachronic and a synchronic element. Diachronically, *-at* words are likely to be older loanwords than *-ah* words. At the same time, it can be shown statistically that a considerable core of *-at* words must have been borrowed not direct from Arabic but from a Persianized source. Historical evidence of a Perso-Indian rather than Arab origin for early Islam in the Malay world helps to independently confirm this. It is logical to assume that many *-ah* words must have been borrowed from the same source, but this cannot be conclusively shown through the comparative technique. From a synchronic standpoint, there exists in modern Malay (and probably more strongly in BM) an Arabizing tendency, such that items are borrowed virtually unmodified direct from Arabic with the unmarked *-ah* ending; Arabic morphosyntactic rules are sometimes also borrowed to yield, for example, *'iḍāfah*-like phrases where the *-at* ending occurs in the first word. Arabic-like adjective and gender endings are also borrowed. At the same time, there is a standardizing tendency in both BM and BI which causes one of an *-at/-ah* doublet to be favoured without regard to the history of the word or to Arabic morphosyntax. Some of these *-at/-ah* doublets may result from ambiguities of pronunciation of *ta simpul* in Jawi script.

In conclusion, the findings offered here surely concur with Jones' view that 'The study of etymology ... provides its own academic justification; the study of Indonesian etymology can also be justified for the light it throws on the cultural history of the region' (Jones 1984:33). The

interested reader is now invited to consider the historical implications of the fact that a Persianized Arabic origin for many Arabic loanwords can be shown statistically, and to test the general explanation offered here for the distribution of *-ah* and *-at* endings in BM and BI.

REFERENCES

Abas, Husen, 1987, *Indonesian as a unifying language of wider communication; A historical and sociolinguistic perspective*, Canberra: Australian National University, Research School of Pacific Studies. [Materials in Languages of Indonesia 37. W.A.L. Stokhof, Series Editor. Pacific Linguistics Series D, 73.]

Abdullah Hassan, 1974, *The morphology of Malay*, Kuala Lumpur: Dewan Bahasa dan Pustaka.

Al-Harbi, Awwad Ahmad Al-Ahmadi, 1993, 'Arabic loanwords in Acehnese', in: Bernard Comrie and Mushira Eid (eds), *Perspectives on Arabic linguistics III; Papers from the third annual symposium on Arabic linguistics*, pp. 93-117, Amsterdam: Benjamins.

Alisjahbana, Sutan Takdir, 1963, *Tatabahasa baru Melayu/Indonesia* II, Kuala Lumpur: Bi-Karya.

–, 1976, *Language planning for modernisation; The case of Indonesian and Malaysian*, The Hague: Mouton.

Badudu, Yus, 1984, *Ejaan Bahasa Indonesia*, Bandung: Penerbit CV Pustaka Prima.

Barker, W. Burckhardt, 1854, *Practical grammar of the Turkish language with dialogues and vocabulary*, London: Bernard Quaritch.

Bausani, Alessandro, 1974, 'Notes sur les mots persans en malayo-indonésien', *Acta Iranica*, Première série, 2:347-79.

Beg, Muhammad Abdul Jabbar, 1979, *Arabic loan-words in Malay*, Kuala Lumpur: The University of Malaya Press.

Bellwood, Peter, 1985, *Prehistory of the Indo-Malaysian Archipelago*, Sydney: Academic Press.

Brass, Paul R., 1974, *Language, religion and politics in North India*, Cambridge: Cambridge University Press.

Chatterji, Suniti Kumar, 1970, *The origin and development of the Bengali language; Part I, Introduction, phonology. Part II, Morphology, Bengali index*, London: Allen and Unwin. [First published in 1926 by the Calcutta University Press.]

Chen Da-Sheng, 1992, 'A Brunei Sultan in the early 14th century; Study of an Arabic gravestone', *Journal of Southeast Asian Studies* 23-1:1-13.

Craven, Thomas, 1911, *The new royal dictionary; English into Hindustani and Hindustani into English*, Lucknow: Methodist Publishing House.

Daftar ejaan Rumi-Jawi, 1994, Kuala Lumpur: Dewan Bahasa dan Pustaka. [Fourth Edition.]

Demidyuk, L.N., and V.A. Makarenko, 1980, 'Indonesian linguistics in the Soviet Union in the 60's and 70's', *Bijdragen tot de Taal-, Land- en Volkenkunde* 136:440-62.

Drewes, G.W.J., 1968, 'New light on the coming of Islam to Indonesia?', *Bijdragen tot de Taal-, Land- en Volkenkunde* 124:433-59.

–, 1971, 'The study of Arabic grammar in Indonesia', in: P.W. Pestman (ed.), *Acta Orientalia Neerlandica; Proceedings of the Congress of the Dutch Oriental Society*, pp. 61-70, Leiden: Brill.

Fatimi, S.Q., 1963, *Islam comes to Malaysia*, Singapore: Malaysian Sociological Research Institute. [Edited by Shirle Gordon.]

Gallop, Annabel Teh, 1991, *Golden letters; Writing traditions of Indonesia / Surat emas; Budaya tulis di Indonesia*, London: The British Library.

Hairul, Awang Sudjai, and Yusoff Khan, 1982, *Kamus lengkap*, Selangor: Pustaka Zaman.

Hollander, J.J. de, 1984, *Pedoman bahasa dan sastra Melayu*, Jakarta: Balai Pustaka.

Jones, R., 1978, *Arabic loan-words in Indonesian*, London: Indonesian Etymological Project, Paris: *Archipel*. [Indonesian Etymological Project III.]

–, 1984, 'Loan words in contemporary Indonesian', in: J.W.M. Verhaar (ed.), *Towards a description of contemporary Indonesian; Preliminary studies, Part 2*, pp. 1-38, Jakarta: Universitas Katolik Indonesia Atma Jaya. [NUSA Linguistic Studies in Indonesian and Languages in Indonesia 19.]

Kasimin, Amran, 1987, *Perbendaharan kata Arab dalam Bahasa Melayu*, Bangi: Penerbit Universiti Kebangsaan Malaysia.

Lambton, Ann K.S., 1960, *Persian grammar*, Cambridge: Cambridge University Press.

–, 1961, *Persian vocabulary*, Cambridge: Cambridge University Press.

Lewis, M.B., 1958, *A handbook of Malay script*, London: Macmillan.

Mahdi, W., 1981, 'Some problems of the phonology of Metropolitan Indonesian', *Bijdragen tot de Taal-, Land- en Volkenkunde* 137:399-418.

Matheson, Virginia, and M.B. Hooker, 1988, 'Jawi literature in Patani; The maintenance of an Islamic tradition', *Journal of the Malaysian Branch of the Royal Asiatic Society* 61-1:1-86.

Meulen, D. van der, and H. Von Wissman, 1932, *Hadramaut; Some of its mysteries revealed*, Leiden: Brill.

Moeliono, Anton M., et al., 1989, *Kamus besar Bahasa Indonesia*, Jakarta: Balai Pustaka. [Second Edition.]

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

Nor bin Ngah, Mohd., 1983, *Kitab Jawi; Islamic thought of the Malay Muslim scholars*, Singapore: Institute of Southeast Asian Studies. [Research Notes and Discussion Paper 33.]

Ricklefs, M.C., 1981, *A history of modern Indonesia, c.1300 to the present*, London: Macmillan.

Ronkel, Ph.S. van, 1899, 'Over invloed der Arabische syntaxis op de Maleische', *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 41:498-528.

Sarumpaet, Jan Pieter, 1980, *Modern usage in Bahasa Indonesia*, Carlton: Pitman Australia.

Shellabear, Rev. W.G., 1901, 'The evolution of Malay spelling', *Journal of the Royal Asiatic Society, Straits Branch* 36, July:75-135.

Steingass, F., 1957, *A comprehensive Persian-English dictionary*, London: Routledge and Kegan Paul. [Fourth Impression.]

Stokhof, W.A.L., 1975, 'On the phonology of Bahasa Indonesia', *Bijdragen tot de Taal-, Land- en Volkenkunde* 131:254-69.

Teeuw, A., 1959, 'The history of the Malay language; A preliminary survey', *Bijdragen tot de Taal-, Land- en Volkenkunde* 115:138-61.

–, 1962, 'Some problems in the study of word classes in Bahasa Indonesia', *Lingua* 11:409-21.

Telegdi, Zsigmund, 1974, 'Remarques sur les emprunts arabes en persan', *Acta Iranica*, Première série, 2:337-45.

Tham Seong Chee, 1990, *A study of the evolution of the Malay language; Social change and cognitive development*, Singapore: Singapore University Press.

Verhaar, J.W.M., 1984, 'Affixation in contemporary Indonesian', in: Bambang Kaswanti Purwo (ed.), *Towards a description of contemporary Indonesian; Preliminary studies*, Part 1, pp. 1-26, Jakarta: Universitas Atma Jaya. [NUSA Linguistic Studies in Indonesian and Languages in Indonesia 18.]

Windfuhr, Gernot L., 1979, *Persian grammar; History and state of its study*, The Hague: Mouton.

Winstedt, R.O., 1941, 'Jawi spelling', *Journal of the Malayan Branch of the Royal Asiatic Society* 19:227-33.

Zainal-Abidin bin Ahmad, 1928, 'Jawi spelling', *Journal of the Malayan Branch of the Royal Asiatic Society* 6-2:81-104.

Zograph, G.A., 1982, *Languages of South Asia*, London: Routledge and Kegan Paul. [Translated by G.L. Campbell. Languages of Asia and Africa 3.]

APPENDIX 1: WORDS WITH PERSIAN REFLEX IN *-E*, ATTESTED BY AT LEAST THREE AUTHORITIES

Glosses are intended for identification only; they do not purport to indicate exact or preferred meaning. Spellings primarily follow *Kamus besar Bahasa Indonesia* (Moeliono et al. 1989), and the other authorities where *Kamus besar* lacks a particular form.

akidah	faith	*kuliah*	school
alkisah	story	*kuria*	village
daerah	region	*madrasah*	school
darjah/derajat	grade	*mahkamah*	court
faedah	profit	*majalah*	magazine
falsafah	philosophy	*makalah*	working paper
fitnah	slander	*martabat*	rank
hadiah	prize	*masalah*	problem
halkah	bracelet	*mukadimah/t*	introduction
handasah	surveying	*mukjizat*	miracle
helah/t	deception	*mutalaah*	study
hijrah/t	Hijrah	*muzakarah*	discussion
idah	waiting-period before remarriage	*naskhah*	manuscript
		natijah	core issue
		nekara	drum
ijarah	price	*noktah*	dot
ijazah	permission	*rabitah/t*	league
iradah/t	will	*rasuah*	bribery
istikharah/t	pray for guidance	*risalah/t*	essay
jamiah/t	assembly	*sufrah*	table cloth
jazirah/t	peninsula	*sahifah/t*	document
jumaat	Friday	*salsilah*	genealogy
jumlah	total	*sejarah/syajrat*	history
kemah	tent	*surah*	verse of the Holy Kuran
kaidah	rule		
kafilah	caravan	*syubahat*	dubious
kahwa	coffee	*taifah*	group
kalimah	creed	*tajribah*	attempt
kanisah	temple	*tarekat*	path
kasidah	song	*tobat*	repent
khalifah	caliph	*terjemah*	translation
khazanah	property	*tuhfah/t*	gift
khurafat	superstition	*wasitah*	procuress
kiblat	direction of Meccah	Zulhijah	twelfth month of Muslim calendar
kisah	story	Zulkaedah	eleventh month of Muslim calendar
kuat	strong		

APPENDIX 2: WORDS WITH PERSIAN REFLEX IN *-AT*, ATTESTED BY AT LEAST THREE AUTHORITIES

Glosses are intended for identification only; they do not purport to indicate exact or preferred meaning. Spellings primarily follow *Kamus besar Bahasa Indonesia* (Moeliono et al. 1989), and the other authorities where *Kamus besar* lacks a particular form.

adalah/t	justice	*maruah*	self-respect
adat	custom	*mesyuarah/t*	meeting
akibat	result	*muafakat/pakat*	agreement
alamat	sign	*mudarat*	loss
amanah/amanat	security	*muhibah*	goodwill
azmat	awe-inspiring	*munasabah/t*	reasonable
azza	glorious	*musibah/t*	calamity
balaghah	fluent	*muslihat*	trick
barakah/berkat	blessing	*najasah/t*	filth
dakwah	propaganda	*nasihat*	advice
darurat	compulsion	*nawbah/nobat*	drum
daulat	good fortune	*nikmat*	comfort
fadilat	virtue	*nisbah*	family name
fahrasat	index	*rahmat*	mercy
hadrat	presence	*rakyat*	subjects
hajat	wish	*riwayat*	legend
hakikat	truth	*saadah/t*	majesty
harakah/t	vowel points	*saat*	moment
hemah/t	thrifty	*sakarat*	death agony
hikayat	story	*selamat*	safety
hikmah/t	knowledge	*serbat*	sherbet
ibadah/t	devotion	*siasah/t*	politics
ibarat	example	*sifat*	attribute
ifah	abstinence	*sihat*	health
ikamah/t	last call to prayer	*sobat*	friend
istirahat	rest	*surat*	letter
izzah/t	honour	*syafaat*	intercession
jemaah/t	group	*syafakah/t*	compassion
kaifiat	character	*syahadah/t*	testimony
kamat	call to begin prayer	*syariah/t*	Islamic law
kanaah/t	contentment	*tabiat*	character
khianat	treachery	*tarbiah*	education
khidmah/t	service	*umat*	people
kiamat	day of resurrection	*wahdah/t*	unity
kimah/t	value	*wakalah/t*	agency
kodrat	power	*wasiat*	exhortation
laknat	curse	*wilayah/t*	district
loghat	word	*ziarah/t*	visit to shrine
makrifat	wisdom		