Chapter 27

Loanwords in Indonesian*

Uri Tadmor

1. The language and its speakers

1.1. Name and classification

Indonesian (or *Bahasa Indonesia* 'the Indonesian language') is a form of Malay that serves as the national language of Indonesia. It is a member of the Malayic subgroup of Western Malayo-Polynesian, a branch of the Austronesian language family. Other Malayic languages include Minangkabau (spoken on Sumatra), Iban (northern Borneo), and Banjar (southern Borneo). No satisfactory internal classification of Malayic languages has been proposed so far, in part because no linguistic criteria have been established to distinguish between Malay dialects and Malayic languages.

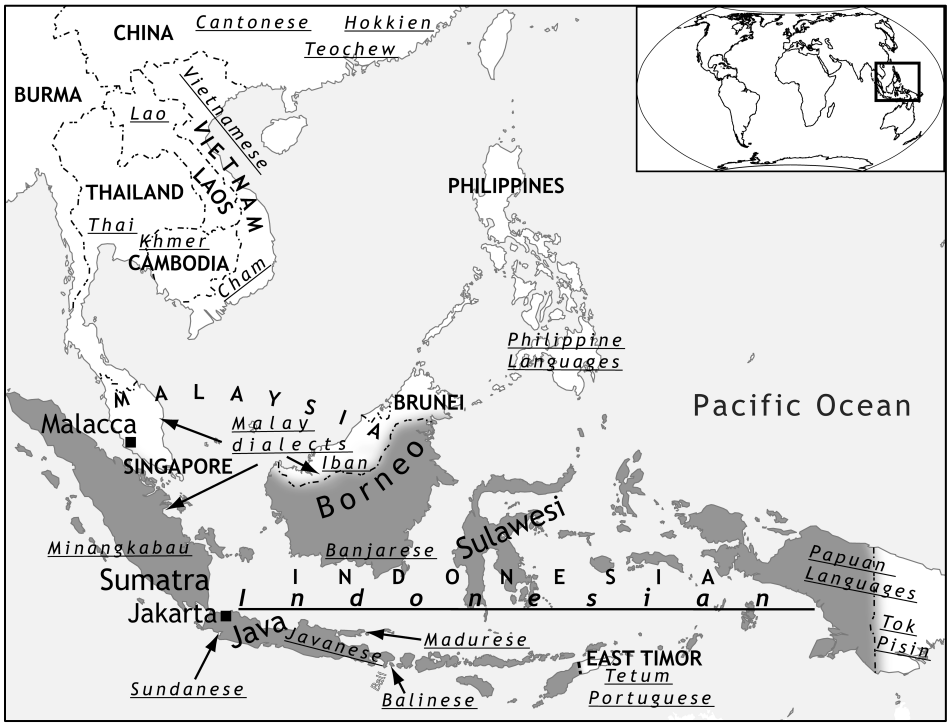
It would not be possible to discuss all varieties of Indonesian here, not only because of space limitations, but also because many of them are poorly documented. The discussion therefore focuses on standard Indonesian, the most widely used variety. Whenever the term "Malay-Indonesian" is used, it will refer to the language as a whole (especially in historical perspective, when it is not possible to make a distinction between Malay and Indonesian). "Indonesian" will refer specifically to contemporary standard Indonesian, while "Malaysian" will refer to standard Malay as used in Malaysia. In citing Malay-Indonesian words the standard orthography is used, with one exception: the mid front vowel /e/ is written *é*, to distinguish it from the mid central vowel /ə/, written *e*. In the standard orthography, both are written *e*.

1.2. Sociolinguistic position

Once used only by ethnic Malays, the Malay language assumed a role as a regional lingua franca at an early date. Currently ethnic Malays constitute only a small part of the total number of speakers of Malay-Indonesian, although they may still con-

* The subdatabase of the World Loanword Database that accompanies this chapter is available online at http://wold.livingsources.org. It is a separate electronic publication that should be cited as: Tadmor, Uri. 2009. Indonesian vocabulary. In Haspelmath, Martin & Tadmor, Uri (eds.) *World Loanword Database*. Munich: Max Planck Digital Library, 1947 entries. <http://wold.livingsources.org/vocabulary/27>

stitute a majority among native speakers. There are two different standardized varieties, one used in Malaysia (with very similar varieties also used in Brunei and Singapore) and the other in Indonesia.



Map 1: Geographic situation of Indonesian

The total number of speakers of Malay-Indonesian is estimated at about 250 million, making it by far the most widely spoken language in Southeast Asia, as well as the most widely spoken Austronesian language. Most Indonesians know at least some Indonesian and use it on a regular basis. However, the standard language is not acquired as a first language. Where Indonesian is used as a home language, it is in the form of a local colloquial variety. In this sense, all speakers of standard Indonesian are at least bidialectal[1]. Children acquire the standard language early on from its use on television and in school. In recent years the Jakarta dialect has been making inroads into areas that have previously been the sole domain of standard Indonesian, such as advertisements and television interviews. It is also widely used in youth magazines, on the Internet, and in text messaging. However, in more formal situations the use of standard Indonesian is still the overwhelming choice.

[1] In a diglossic situation where speakers use Standard Indonesian in more formal situations and a colloquial variety of Malay-Indonesian as a home language, the two can be said to form the two ends of a continuum.

As the sole official language, Indonesian is used in all government communication, both oral and written. Practically anything published in Indonesia (books, newspapers, magazines) is in Indonesian, as are almost all product markings and public signs. Spoken Indonesian is also used as a lingua franca among people who belong to different ethnolinguistic groups.

2. Sources of data

In addition to the author's personal knowledge, a large number of dictionaries were consulted. The major dictionaries of Indonesian used were *Kamus Besar Bahasa Indonesia* (2002), Echols & Shadily (1975), Echols & Shadily (1998), and Stevens & Schmidgall-Tellings (2004). Among dictionaries of Malaysian, the following deserve special mention: Wilkinson (1959), *Kamus Dewan* (1991), and *Kamus Inggris-Melayu Dewan* (1992). In addition, many dictionaries of the various donor languages were also consulted. Of the few methodological studies of loanwords in Malay-Indonesian, Jones (1984) provides a good overview, and Gonda (1952) remains the classic work on words of Sanskrit origin in the languages of Indonesia. Particularly useful were the three publications of the Indonesian Etymological Project, which consist of lists of Indonesian loanwords that ultimately originate from Arabic and Persian (Jones 1978), European languages (Grjins et al. 1983), and Sanskrit (de Casparis 1997).[2]

3. Contact situations

3.1. Languages of India

The earliest foreigners known to have had significant influence on the Malay-Indonesian archipelago were Indians. However, practically all evidence of this early contact – dating back to the first millennium CE – is secondary, in the form of Indian cultural, religious, and linguistic traits. Therefore, we do not know exactly who introduced Indian civilization to the area, although it is assumed that it was mostly visiting Indian traders, scholars, and missionaries, rather than through large-scale immigration or political domination. The oldest Malay inscriptions (7th century CE) contain parts in Sanskrit, and even the Malay sections of these inscriptions include many Sanskrit loanwords. Sanskrit continued to be used in the Malay-speaking world for centuries as a liturgical language (for both Hinduism and Buddhism) as well as a literary language. This has made Sanskrit a major donor

[2] The results of the Indonesian Etymological Project, which investigated lexical borrowing into Malay-Indonesian from languages outside the Malay-Indonesian Archipelago, were summarized in Jones (2007). This book, consisting of a glossary of words borrowed into Malay-Indonesian from ten foreign languages, was published after the data collection for this chapter had been practically completed.

language for Malay-Indonesian and the source of many common words such as *kepala* 'head' (< *kapāla* 'cup, skull'), *cahaya* 'light' (< *chāya* 'reflection, light'), *nama* 'name' (< *nāma(n)* 'name'), *kerja* 'work' (< *kārya* 'duty, work'), *semua* 'all' (< *samūha* 'multitude'), and *karena* 'because' (< *kāraṇa* 'cause').

Later Indic languages such as Hindi-Urdu also contributed some loanwords to Indonesian, but it is often difficult to point to the precise source words, which may have been early or dialectal forms. Some of these loanwords are *roti* 'bread' (cf. Hindi-Urdu *roṭī* 'bread'), *celana* 'trousers' (cf. Hindi-Urdu *charnā* 'half-trousers'), *topi* 'hat' (cf. Hindi-Urdu *ṭopī* 'hat'), and *kunci* 'key' (cf. Hindi-Urdu *kuṁjī* 'key').

Dravidian languages of southern India – mostly Tamil – were also in contact with Malay-Indonesian, and have left traces in the form of numerous loanwords. Tamil loanwords in Indonesian include *kapal* 'ship' (< *kappal* 'ship'), *teman* 'friend' (< *tamaṉ* 'male relative or friend'), *nelayan* 'fisherman' (< *nulaiyan* 'seashore dweller, fisherman'), and *kedai* 'shop' (< *kaṭai* 'shop, market'; this loanword is more commonly used in Malaysia). Some of the source words, such as *nulaiyan*, are obsolete in modern Tamil (E. Annamalai, p.c.), testifying to the antiquity of these loanwords in Malay-Indonesian. Many words of Indic (Indo-European) origin also show evidence of having been borrowed via speakers of Dravidian languages (see §4.1.1).

3.2. Chinese languages

Chinese pilgrims and traders have been visiting Indonesia for well over a thousand years, in the past often on their way to India. Chinese communities have existed throughout the archipelago for many centuries. Various Chinese languages were (and still are) spoken in Indonesia, and have influenced colloquial varieties of Indonesian, although in the standard language their influence has been limited and purely lexical. Most Chinese immigrants to the Malay-Indonesian archipelago were speakers of Southern Min varieties, and by far the most important Chinese donor language was Hokkien (also called Amoy). There are also some loanwords from Teochew (another Southern Min language) and a handful from Mandarin and Cantonese. Examples of Chinese loanwords in Indonesian include *cat* 'paint' (< Amoy 漆 *chhat* 'paint'), *toko* 'store' (< Amoy 土庫 *thó' khò'* 'store, warehouse'), *giwang* 'earring' (< Foochow 耳環 *ngi*^{7} *hwang*^{5} 'earring'), and *téh* 'tea' (< Amoy 茶 *tê* 'tea').

3.3. Near East languages

Travelers from the Near East first arrived in Indonesia during the second half of the first millennium CE. Eventually the Arabic and Persian languages were to have a strong impact on Malay-Indonesian. However, this did not take place until centuries later, when local inhabitants began converting to Islam. The lexical influence of Arabic has been especially strong. Many words of Arabic origin did not enter

Malay-Indonesian from spoken Arabic, but rather through Arabic literature or through Persian literature (where Arabic loanwords abound); see §4.1.1. Loanwords of Arabic origin, which are very numerous, include *dunia* 'world' (< *dunyā* 'world'), *badan* 'body' (< *badan* 'body'), *kuat* 'strong' (< *qūwat-* 'strength, power'), *kursi* 'chair' (< *kursī* 'chair'), *waktu* 'time' (< *waqt* 'time, period'), *pikir* 'think' (< *fikr* 'thinking, cognition'), and *jawab* 'answer' (< *jawāb* 'answer'). Words of ultimate Persian (non-Arabic) origin are far fewer, and include *kawin* 'marry' (< *kāwīn* 'dowry'), *domba* 'sheep' (< *duṃba* 'a kind of sheep with a thick tail'), *anggur* 'grape, wine' (< *angūr* 'grape, raisin'), and *gandum* 'wheat' (< *gandum* 'wheat').

3.4. Portuguese

The earliest Europeans with a significant presence in Indonesia were the Portuguese, who first arrived in the early 16th century. There are numerous loanwords of Portuguese origin in Indonesian, but most were not borrowed directly but rather via the Portuguese-based creole that once served as a lingua franca in Batavia (capital of the Dutch East Indies, now Jakarta). This creole was spread to Batavia from Malacca (in the Malay Peninsula) by slaves captured from the Portuguese after Malacca fell to the Dutch in 1641. Batavian Portuguese Creole died out in the 20th century, but Malaccan Portuguese Creole (known as Kristang) is still used by small groups in Malacca itself and in Singapore. The Indonesian words for many everyday objects are of ultimately Portuguese origin, such as *garpu* 'fork' (< *garfo*), *keméja* 'shirt' (< *camisa*), *sepatu* 'shoes' (< *sapato*), *méja* 'table' (< *mesa*), *roda* 'wheel' (< *roda*), *bola* 'ball' (< *bola*), and *jendéla* 'window' (< *janela*). A hitherto overlooked Portuguese loanword is *kaléng* 'tin, can', recorded as *calaim* (with various other spellings) in Asian Portuguese sources by Yule and Burnell (1903:145–6). The ultimate source of this loanword is Turkish *kalay* 'tin'.

Some Indonesian words previously considered as Dutch loanwords are analyzed in the present study as having been borrowed via Portuguese Creole. These include *lampu* 'lamp' (< Portuguese Creole *lampu* < Dutch *lamp;* cf. Kristang *lampu*) and *buku* 'book' (< Portuguese Creole *buku* < Dutch *boek*; cf. Kristang *buku*). In these words, indirect borrowing explains the presence of the unexpected final *-u*, a common phonological strategy in Portuguese Creole[3] but not in Malay-Indonesian. Other words also betray indirect borrowing by their unusual phonological or semantic correspondence pattern. Indonesian *pompa* 'pump' is ultimately from Portuguese *bomba* 'pump' (cf. Kristang *bomba*); however, its phonology has been influenced by Dutch *pomp* 'pump'. (The Malaysian counterpart of this word did not undergo the change and has remained *bomba*.) Indonesian *pipa* '(water) pipe' is ultimately from Portuguese *pipa* 'barrel', influenced by the semantics of Dutch *pijp*

[3] Still used until relatively recently – cf. *waistu* 'waist' in Singapore Kristang, from English *waist*.

'pipe'[4]. The unusual syncopation exhibited by Indonesian *taflak* 'tablecloth' (ultimately from Dutch *tafellaken*) is also explained by borrowing via Portuguese Creole (cf. Kristang *taflak*). An interesting doublet is represented by the two Indonesian words for 'bullet', *peluru* and *pélor* 'bullet'. *Peluru* was borrowed from Kristang *piloru* while *pélor* would have been borrowed from Batavian Portuguese Creole **pilor*[5].

3.5. Dutch

After the Portuguese, the next Europeans to send expeditions to Indonesia were the Dutch, who first came towards the end of the 16th century. Eventually the Dutch came to control all of present-day Indonesia until the mid-20th century. The use of Dutch in Indonesia was limited, however, and only a small fraction of the indigenous population ever gained fluency in the language. Nevertheless, since the ruling class spoke Dutch (and the few Indonesians who spoke Dutch belonged to the influential elite), Dutch had a strong impact on the Indonesian lexicon, and some impact on its grammar as well. Interestingly, many of the Dutch source words were themselves loanwords, mostly from French. Dutch loanwords in Indonesian are very numerous, and include *kamar* 'room' (< *kamer* 'room'), *kopi* 'coffee' (< *koffie* 'coffee'), *duit* 'money' (< *duit*, the name of an old Dutch coin), *mobil* 'car' (< *automobiel* 'car'), *setir* 'driving wheel, drive' (< *stuur* 'driving wheel), *lat / telat* 'late' (< *laat* 'late'), and *koran* 'newspaper' (< *courant*, the older form of *krant* 'newspaper').

3.6. English

Following full independence in late 1949, English quickly became the most widely taught foreign language in Indonesia. Members of the educated elite generally have a good knowledge of English and frequently code-switch between English and Indonesian. English is also heard daily on television and in movie theaters, so most Indonesians have had at least some exposure to it. English loanwords in Indonesian include *flu* 'cold, flu', *koin* 'coin', *gaun* 'dress' (< *gown*), *bolpoin* 'pen' (< *ballpoint*), *mall* [mol] 'shopping center' (< *mall*), and *bil* 'check, bill'.

3.7. Other languages of the Malay-Indonesian archipelago

In addition to coming in contact with languages from outside the region, Malay-Indonesian has been in contact with many local languages, principally via its role as

[4] For an example of Malay *pipa* used with the sense 'barrel' (in a Malay letter of 1797) see Mu'jizah 2009: 25.

[5] In Batavian (but not Malaccan) Portuguese Creole, final *-u* is deleted after *r*.

a lingua franca throughout the archipelago. The most influential of these local languages overall has been Javanese, in contact with Malay-Indonesian for well over a millennium. Today, native speakers of Javanese form the largest group among speakers of Indonesian. Javanese loanwords entered Indonesian through at least two distinct contact situations. There is evidence for borrowing from Old Javanese into a very early form of Malay. Such loanwords are typically characterized by their presence in Classical Malay manuscripts of the 16th and early 17th centuries (see §4.1.1). The second contact situation was between modern Javanese and Indonesian (or its precursor) since the mid-17th century. Javanese has had a strong impact on Java Malay and on Betawi (the variety used in Jakarta), and through them on the standard language as well.

The two other local languages which have strongly influenced the lexicon of modern Indonesian are Balinese and Sundanese. Balinese people (mostly slaves) once constituted the largest ethnic group in Batavia, and it is there (rather than in Bali) that most lexical transfer from Balinese into Indonesian took place. Sundanese was the language of Batavia's rural hinterland (and indeed is still used in the areas surrounding Jakarta). After natives of Java were permitted to settle in the city, there was a massive inflow of migrants from the Sundanese-speaking hinterland, who brought their language with them.

Indonesian words for which possible source words exist in all three languages (Javanese, Balinese, and Sundanese) include *bébék* 'duck', *mulus* 'smooth', *sabuk* 'belt', *pusar* 'navel', *kocok* 'shake, mix', *keponakan* 'niece/nephew', *tuding* 'to accuse', *ajak* 'to invite', *sepi* 'quiet', and *mirip* 'similar'.

Finally, another local language that has had some lexical influence on Indonesian is Minangkabau, a Malayic language of western Sumatra. Many Indonesian authors and educators, especially those active in the early formative years of modern standard Indonesian, were native speakers of Minangkabau. Their writings contain numerous Minangkabau words, some of which have become part of the general vocabulary of Indonesian. Because Minangkabau is a Malayic language very closely related to Malay-Indonesian, it is difficult to distinguish between shared retentions and loanwords, let alone to determine the direction of borrowing. The following words appear to have been borrowed from Minangkabau into Malay-Indonesian: *kalian* 'you (pl.)', *gadis* 'girl', *dangkal* 'shallow', *datar* 'flat', *bersua* 'meet', *bertikai* 'fight', and *pidato* 'speech'. Some of these words display final consonants whose realization has changed in modern Minangkabau, changes not reflected in the Arabic-based writing system formerly used for writing Minangkabau.

3.8. Other neighboring languages

In addition to numerous and mostly recent examples of borrowing from other languages of the Malay-Indonesian archipelago, Malay-Indonesian also appears to share numerous vocabulary items with other languages of the region, especially Austroasiatic and Tai languages. It is often difficult to pinpoint the immediate

donor language and the direction of borrowing, because the same etymon may be represented in several languages in each family. Some of the loanwords in question are very old and have undergone subsequent sound changes in the recipient languages, further obscuring their origin.

Words of definite Tai origin are mostly used in Malaysian rather than in Indonesian, e.g. *bomoh* 'shaman' (< Old Thai *bɔ:*[1] 'father (used as an epithet)' + *hmɔ:* 'shaman'), *natang* 'kind of window' (< Thai *nâ:ta:ng* 'window'), and *wau* 'kite' (< Thai *wâ:w* 'kite')[6]. Words of Austroasiatic origin are more numerous and include words used in Indonesian (as well as Malaysian), such as *ketam* 'kind of crab' (cf. Proto Mon-Khmer **kt‿aam*) and *sekam* 'husk' (cf. Proto Mon-Khmer **skaam?* 'husk')[7]. Later loanwords came in from the languages of the major Austroasiatic civilizations of Southeast Asia, Mon and Khmer.

A large group of etyma occurs in Mon-Khmer and Tai as well as in Malay-Indonesian. Most are due to borrowing from a common source (principally Sanskrit/Pali), while some others are clearly the result of borrowing from Malay-Indonesian rather than into Malay-Indonesian. But there remains a large number of words which are not of Indic origin and where the direction of borrowing seems to be into Malay-Indonesian; some examples are provided in Table 1. Without getting into the possibility of ancient genealogical connections between Austronesian, Austroasiatic, and Tai-Kadai, I believe that further research would show that most items in Table 1 are the result of borrowing from Khmer into Malay-Indonesian, either directly or via Tai. Indeed, quite a few of the etyma have reflexes in Old Khmer and/or have reconstructed proto-forms in Proto Mon-Khmer.

4. Numbers and kinds of loanwords

4.1. Loanwords by donor language

4.1.1. Challenges in identifying immediate donor languages

Identifying the immediate donor language of many loanwords in Indonesian proved difficult and even impossible. Some particularly challenging groups of loanwords are discussed below.

Javanese, Balinese, and Sundanese are closely related to each other and share a significant part of their lexicon. This is due not only to shared retentions but also to the fact that Balinese and Sundanese have both borrowed from Javanese. It is therefore often difficult to tell whether a particular word was borrowed into Indonesian from (modern) Javanese, Balinese, or Sundanese. This is one of the reasons why these three languages are grouped together as "Languages of the Java Area" in Tables 3, 4, and 5. Loanwords from Old Javanese are easier to identify because they

[6] By a strange coincidence, the Dutch word for *kite* (the bird, not the toy) is *wouw*, which led Wilkinson (1959:1282) to wrongly cite it as the source for Malay *wau*.

[7] Proto Mon-Khmer forms cited in this chapter are from Shorto 2006.

are attested in Classical Malay, which was used before Balinese and Sundanese had any significant influence on Malay. Some are also used in Malay dialects of Borneo and the Malay Peninsula, where the direct influence of Balinese and Sundanese has been minimal.

Table 1: Some shared etyma in Khmer, Thai, and Malay-Indonesian

Khmer	**Thai**	**Malay-Indonesian**
bɔŋci: 'a register'	*banchɔ:i* 'account, list, register'	*banci* 'census'
cam 'remember'	*cam* 'remember, recall'	*cam* 'recognize, be able to recall'
tiən 'candle'	*thian* 'candle'	*dian* 'candle, oil lamp'[8]
tɔən 'manage to, be in time for'	*than* 'manage to, be in time for'	*dan* 'manage to, be in time for'
pù:ək 'group'	*phûak* 'group'	*puak* 'group'
krɛ: 'bed'	*khrê:* 'litter, light bed/seat'	*gerai* 'platform, stall'
Proto Mon-Khmer **dga:m* 'molar tooth'	*kra:m* 'molar tooth'	*geraham* 'molar tooth'
khtɤ̀:y 'hermaphrodite'	*kàthə:y* 'hermaphrodite'	*kedi* 'hermaphrodite'
Proto Mon-Khmer **[t]ruŋ* 'cage', **kruŋ* 'to confine'	*krong* 'cage'	*kurung* 'cage, to confine'
krəbɤ̀y 'water buffalo'	*krà bɯ:* 'water buffalo'	*kerbau* 'water buffalo'
lŋɔ̀: 'sesame'	*ngaa* 'sesame'	*lenga* 'sesame'
sbay 'muslin'	*sàbay* 'shawl'	*sebai* 'shawl'
srəmaoc 'ant'	*mót* 'ant'	*semut* 'ant'
srənok 'pleasant'	*sànùk* 'enjoyable'	*seronok* 'pleasant, enjoyable'
tra: 'seal, stamp'	*tra: 'seal, stamp'*	*tera* 'seal, stamp'
thùən 'endure'	*thon* 'endure'	*tahan* 'endure'

Loanwords of Sanskrit origin also presented some problems. Although most such words show a close enough phonetic and semantic resemblance to Sanskrit and seem to have originated from Sanskrit literature, it is also possible that some have come into Malay not directly from Sanskrit, but via a Prakrit or a later vernacular. An interesting feature of some loanwords of Sanskrit origin is intervocalic voicing, which can be seen in words such as *kuda* 'horse' (cf. Sanskrit *ghoṭa*), *gergaji* 'saw' (cf. Sanskrit *krakaca;* the initial consonant may have undergone voicing by assimilation), *ajar* 'teach/learn' (cf. Sanskrit *ācārya*), *bijaksana* 'wise' (cf. Sanskrit *vicakṣaṇa*), *curiga* 'suspect' (cf. Sanskrit *churikā*), and *segala* 'all' (cf. Sanskrit *sakala*). Intervocalic voicing was never a feature of Malay-Indonesian phonology, but it is a hallmark of Tamil and other Dravidian languages. This indicates that these words (and probably others which happen not to have intervocalic voiceless stops) may have been borrowed via a Dravidian language or were learned from native speakers of a Dravidian language. A loanword where this development is clear is Indonesian *tiga* 'three', ultimately from Sanskrit *trika* 'triple'; indeed, Telugu exhibits an identical form, *tiga* 'three'[9]. The fact that Malay-Indonesian borrowed a lower numeral from

[8] This etymon appears to have entered Old Khmer from Chinese, where it meant 'oil lamp'.
[9] This was pointed out to me by Waruno Mahdi.

or via Dravidian speakers is strong testimony to the latter group's strong influence in ancient Indonesia. This analysis matches historical evidence that early Indian influence on Indonesia emanated from southern (Dravidian-speaking) India. To overcome the problem of identifying the immediate donor language, all words ultimately originating in languages of India are grouped together for statistical purposes.

Regarding words of Arabic origin, it is seems probably that many came into Malay-Indonesian via intermediate languages. Campbell (1996) investigated loanwords of ultimate Arabic origin in Malay which in Classical Arabic contained the feminine suffix *-at-* (followed by a case ending, e.g. nominative indefinite *-un*). This suffix is represented in words of Arabic origin in Malay-Indonesian by *-ah* or *-at*, e.g. *fitnah* 'slander' and *adat* 'custom' (both words have the same ending in Arabic)[10]. Campbell compared well-attested Arabic etyma that occur both in Persian and in Malay whose source words in Classical Arabic have the feminine suffix *-at-*. His findings indicate that more than half the words that end in *-at* in Persian also end in *-at* in Indonesian, while only a fifth end in *-ah*. Even more striking, nearly three quarters of the words ending in *-e* in Persian end in *-ah* in Malay, while only about 10% end in *-at*. Campbell's results are summarized in table 2.

Table 2: Correspondence rate between Persian *-e* and *-at* and Malay *-ah* and *-at* (after Campbell 1996: 38–39)

	Malay *-at*	**Malay *-ah***	**Problematic**[11]
Persian *-at*	51.58%	20.00%	28.42%
Persian *-e*	10.47%	73.26%	16.28%

These rates of correspondence in table 2 cannot be the product of chance, and seem to indicate that the words in question were borrowed into Malay via Persian. Since there is no reason to assume that only words with these endings were borrowed into Malay, it may be inferred that most Arabic loanwords in Malay-Indonesian in general were borrowed principally via Persian. The possibility remains, of course, that the direct donor language was a Persianized language of India (such as Urdu). It also remains to be explained why Persian *-é* would be borrowed as Malay *-ah*; perhaps this reflects a spelling pronunciation or an earlier Persian pronunciation. Moreover, a recent study by van Dam (2009) has shown that if older and more extensive lexicographical sources are used, different statistical results are obtained,

[10] Superficially, it would appear as though words in *-ah* reflect the Classical Arabic pausal form, while words with *-at* reflect the construct state form. However, there is a good explanation why these marked forms would be the ones borrowed rather than the absolute forms. Moreover, as will be explained below, it is highly probable that most of these words did not enter Malay-Indonesian directly from Arabic.

[11] "Problematic" words are those that according to Campbell "either showed disagreement among the authorities, or had alternatives cited by at least one authority" (Campbell 1996: 39).

possibly indicating a much stronger place for direct borrowing from Arabic. Whichever the case might be, it is clear that words were borrowed with *-at* only at earlier stages of Malay-Indonesian; modern borrowings exhibit exclusively *-ah*, e.g. *kuliah* 'to attend university', *nasabah* 'bank customer', and *majalah* 'magazine'.

It is possible that this study might help settle an old debate among historians of Islam in Southeast Asia. By the last quarter of the 19th century, European historians had realized that Islam has spread to Southeast Asia principally via India, rather than directly from Arabia. At first Gujarat in northwestern India was thought to have been the more specific locus whence Islam was introduced to the Malay-Indonesian archipelago, although later scholars also theorized that southern India was a more probable locus. The debate raged for a century (for a summary see Meuleman 2005: 24–25).

This study did not find any Indic loanwords in Malay-Indonesian that can definitely be shown to have originated from Gujarati. Moreover, none of the ultimately Arabic loanwords show signs of having been borrowed via Gujarati. On the other hand, there are numerous loanwords of obvious Dravidian origin in Malay-Indonesian (including ultimately Indic loanwords borrowed via Dravidian or Dravidian speakers). Even more importantly, certain Arabic loanwords in Malay-Indonesian show signs of having been borrowed via a Dravidian language (or from native speakers of a Dravidian language). Specifically, these are words that in standard Arabic end in a cluster. Some cluster-final words exhibit a final *-u* which eliminates the unphonotactic (in Malay-Indonesian) cluster, e.g. *Sabtu* 'Saturday' (< Arabic *sabt*), *waktu* 'time' (< Arabic *waqt*), *salju* 'snow' (< Arabic *ṯalj*), *perlu* 'need' (< Arabic *farḍ*). The final *-u* appears to reflect the Classical Arabic nominative ending. However, such forms occur in Arabic itself only in the construct state (roughly, when the noun is the head of a genitive construction) or following a definite article, and it would be difficult to explain why such highly marked forms would be borrowed (without the genitive noun or definite article) rather than the simple unmarked forms. Moreover, as noted in §5.2 below, Arabic words with final clusters are normally integrated into Malay-Indonesian with an echo vowel inserted between the final two consonants: *subuh* 'dawn' < Arabic *ṣubḥ*, *jisim* 'body' < Arabic *jism*, *rajam* 'stoning' < Arabic *rajm*. So the origin of final *-u* in these words seems to originate neither in Arabic itself nor in the Malay-Indonesian integration pattern. An alternative explanation would be that these forms were learned from native speakers of a Dravidian language such as Tamil or Telugu, where *-u* is regularly appended to consonant-final loanwords. The linguistic evidence therefore supports the theory of the introduction of Islam to the Malay-Indonesian archipelago from southern India.

Finally, a word of caution is also in order regarding Dutch and English loanwords. It is often impossible to tell simply by looking at a word whether it was borrowed directly from Dutch or from English. This is especially true for words ultimately derived from Greek and Latin, many of which would have the same shape in Indonesian regardless of whether they were borrowed from Dutch or from English. For example, Dutch *connectie* and English *connection* would both be bor-

rowed as Indonesian *konéksi*, and Dutch *kwantiteit* as well as English *quantity* would both be borrowed as Indonesian *kuantitas* (see §5 and §6.2). This has the potential of making newer English loanwords appear to be older Dutch ones. However, after carefully examining all potential cases in the database, I concluded that the number of such words in the Indonesian subdatabase is very small.

4.1.2. Grouping donor languages

The number of donor languages (including dialects and language groups) that have contributed loanwords to Indonesian is very large. For ease of presentation and discussion, they have been grouped into eight groups of donor languages (plus "Miscellaneous" and "Unidentified source"), as in Tables 3, 4, and 5. The grouping was based on socio-historical as well as practical grounds. Languages participating in broadly the same contact situation were generally grouped together (e.g. languages of India, Arabic/Persian). In grouping together the languages of the Java area, a practical consideration was used: as already mentioned, it was often impossible to tell if a particular word was from Javanese, Sundanese, or Balinese. Fortunately, this practical approach did not conflict with sociolinguistic and historical considerations. The percentages of words in the database originating from the various donor languages and donor language groups are presented in Table 3.

In previous studies of loanwords in Indonesian, loanwords from local languages have been excluded from the discussion[12]. As the present study shows, this approach is wholly unwarranted, as such loanwords form an important part of the lexicon. Words from languages of the greater Java area (Java, Bali, and Madura) constitute the most numerous group of loanwords, with 8.9% of the total number of words in the database. Loanwords from these languages are characterized by their common, everyday nature, e.g. *samping* 'side, next to' (< Balinese), *keriput* 'wrinkled' (< Sundanese), and *ketombé* 'dandruff' (< Javanese). These words first entered Indonesia via second-language speakers who were transferring words from their first languages into Indonesian. This process (technically known as "imposition") is rather different from the transfer of words into one's first language from another language ("adoption") which accounts for most Indonesian loanwords of other sources. Different types of borrowing processes are discussed in §8.

Although Sanskrit is the earliest recorded donor language for Malay-Indonesian and has not been used by speakers of Malay-Indonesian for many centuries, loanwords from languages of India (consisting mostly of words of Sanskrit origin) still constitute the second largest category, with 8.4% of all words in the subdatabase. This is especially remarkable considering that many Sanskrit loanwords that appear in early writings (and doubtlessly many unrecorded ones) have long become obsolete and are therefore not included in the count. These numerous and enduring Sanskrit loanwords testify to the tremendous impact that Indian cultures, religions, and languages have had on the Malay-Indonesian speaking world.

[12] Including the recent *Loan-Words in Indonesian and Malay* (Jones 2007).

Table 3:
Loanwords in Indonesian by donor languages/language groups (percentages)

Donor language	Proportion of all words in Indonesian database
Languages of the Java area	8.9
Languages of India	8.4
Dutch	6.4
Arabic/Persian	5.7
Portuguese (including Creole)	1.4
English	1.2
Chinese languages	0.7
Languages of Sumatra	0.4
Unidentified source	0.5
Miscellaneous languages	0.4
Total loanwords	**34.0**

Words of Indian origin are particularly well represented in the domain of religion, and constitute about a quarter of all words in this category (as represented in the database). Only a small fraction of Indonesia's population still adheres to religions that originated from India. However, many religious terms of Sanskrit origin persist, and are now applied to Muslim and Christian concepts, e.g. *agama* 'religion', *surga* 'heaven', *neraka* 'hell', *pahala* '(religious) merit', *puasa* 'fasting', and *pendéta* 'priest, Protestant minister'. It is also remarkable that over 11% of the Indonesian function words in the database are of Indian origin. It should be noted, however, that many of these were not function words in the donor language (see §4.2). Expectedly, the field least impacted by Sanskrit was the Modern World. Although words of Sanskrit origin are still used for coining neologisms, these were not counted as loanwords, as they were created in Indonesian.

Dutch words were borrowed into Indonesian throughout the colonial period, which started in 1619 with the occupation and destruction of Jayakarta (modern Jakarta). The era of Dutch colonialism ended with the transfer of sovereignty in 1949. For all intents and purposes, the borrowing of Dutch words into Indonesian also ended then, because Dutch never became an auxiliary language in Indonesia, as other European languages have become in many former colonies. Since then, numerous Dutch loanwords have become obsolete, but many others are still well established in the lexicon. With 6.5% of the total number of words in the database, the number of Dutch loanwords even exceeds the number of loanwords of Arabic and Persian origin combined. The difference between Dutch and English, the major conduit of Western linguistic influence in Indonesia since independence, is also striking. English loanwords constitute only 1.2% of the total number of words in the database. While the proportion is certain to rise on current sociolinguistic trends, it will be a long time before English loanwords can eclipse the large number of Dutch loanwords adopted throughout centuries of colonialism.

Finally, Portuguese (including Batavian Portuguese Creole) and Chinese languages (mostly Hokkien) also played a significant role as donor languages, although the number of loanwords they contributed to standard Indonesian is relatively small (respectively 1.5% and 0.9% of the total number of words in the Indonesian database).

4.2. Semantic word classes

The breakdown of loanwords according to semantic word classes is summarized in Table 4. The figures for Indonesian conform to the general trend of borrowing proportionally more nouns than verbs (see Chapter III). Indeed, fully 43.7% of all nouns in the Indonesian database are loanwords. Verbs were borrowed far less frequently, although loanwords still constitute a significant proportion (17.2%) of the verbs in the database. If Indonesian borrowed more verbs than other languages, this was probably because this is easily done Indonesian, which is agglutinative and has almost no inflectional morphology. Compare this to Semitic languages like Arabic or Hebrew, for example, which have complex inflectional verbal morphology and therefore present substantial challenges to the borrowing of verbs (see discussion in Chapter III). The only morphological condition on the borrowing of verbs into Indonesian is that the citation form must contain the prefix *meng-* (for transitive verbs and a subcategory of stative verbs) or *ber-* (for intransitive verbs)[13].

Table 4:
Loanwords in Indonesian by donor language group and semantic word class

	Languages of the Java area	Languages of India	Dutch	Arabic/Persian	Portuguese (inc. Creole)	English	Chinese languages	Languages of Sumatra	Unidentified source	Miscellaneous languages	Total loanwords	Non-loanwords
Nouns	10.0	10.1	9.6	7.4	2.1	2.0	1.0	0.3	0.8	0.6	43.7	56.3
Verbs	6.8	3.8	1.9	3.2	0.4	0.2	0.4	0.2	0.2	–	17.2	82.8
Adjectives	10.9	6.8	0.6	5.3	–	–	–	1.2	–	–	24.8	75.2
Adverbs	–	15.4	–	–	–	–	–	–	–	–	15.4	84.6
Function words	4.0	11.3	2.4	1.2	–	–	–	0.8	–	–	19.8	80.2
All words	**8.9**	**8.4**	**6.4**	**5.7**	**1.4**	**1.2**	**0.7**	**0.4**	**0.5**	**0.4**	**34.0**	**66.0**

Two findings that stand out require some explanation. The first is that languages of India contributed a high proportion (15.4%) of adverbs, while no other language contributed any adverbs at all. This is easily explained by the fact that the LWT meaning list only contains 7 meanings classified as 'adverbs'. While just one of the

[13] With inherited bases, these prefixes behave much less regularly.

words corresponding to an adverbial meaning is of probable Sanskrit origin (*laju* 'fast' from Sanskrit *laghu* 'light, swift, quick'), it is enough to appear as significant. The second striking finding is the high proportion of borrowed function words. Nearly one of five function words in the Indonesian database is a loanword, a high proportion for this category into which it is considered difficult to borrow. However, a detailed examination of the data reveals that most of the source words did not constitute function words in the donor languages. They were probably borrowed as content words and underwent grammaticalization later. Thus *saya* 'I' derives from the Sanskrit noun *sahaya* 'companion'. It was first borrowed into Malay in the sense of 'royal companion', and was also used for self-reference by some court officials when addressing a monarch. Later, use of the word expanded to general polite self-reference, and today *saya* functions as a first person pronoun, its original meaning as a noun having been lost. In cases where there is no evidence that the word was ever used as a content word in Indonesian, it is also possible that grammaticalization was part of the borrowing process. Examples include *bahwa* 'complementizer' (< Sanskrit *bhāva* 'being, state'), *bila* 'when, if' (< Sanskrit *velā* 'time'), and *karena* 'because' (< Sanskrit *kāraṇa* 'cause'), which are only attested in Malay-Indonesian as function words, never with the nominal meanings of the source words.

4.3. Semantic fields

Indonesian exhibits considerable variation in the rates of borrowing into different semantic fields, as can be seen in Table 5.

Four semantic fields consist mostly of loanwords: *Religion and belief* (70.0% borrowed vocabulary), the *Modern world* (66.4%), *Clothing and grooming* (55.6%), and *Law* (51.4%). It is fairly obvious why modern world terms would be borrowed, especially in a civilization prone to cultural borrowing. However, why the sphere of religion and belief should be even more susceptible to borrowing requires further explanation. Well over 90% of Indonesia's population adhere to an introduced monotheistic faith (Islam or Christianity), and many of the remaining population follow other non-indigenous faiths (Buddhism or Hinduism). Only a tiny fraction of the population still adheres to an indigenous religion, officially at least. With the disappearance of indigenous religions, much of the vocabulary associated with them also disappeared, or was replaced by the vocabulary of the introduced religions. A similar process affected the indigenous, pre-contact clothing of Indonesia, which is now mostly reserved for ceremonial purposes, and in few areas at that. In their daily lives, most Indonesians use clothes patterned after those of India and the West, which were borrowed along with their names: *celana* 'trousers' and *topi* 'hat' from Hindi; *keméja* '(button-down) shirt' and *sepatu* 'shoes' from Portuguese; *rok* 'skirt' and *kaos* 'socks' from Dutch; *daster* 'house dress' and *bot* 'boots' from English. Finally, the Indonesian legal system is also based on systems introduced from abroad. In ancient times the sources were India and the Middle East, while the modern

Indonesian legal system is based on the Dutch one. Hence the large number of borrowed legal terms.

Table 5:
Loanwords in Indonesian by donor languages and semantic fields (percentages)

		Languages of the Java area	Languages of India	Dutch	Arabic/Persian	Portuguese (inc. Creole)	English	Chinese languages	Languages of Sumatra	Unidentified source	Miscellaneous languages	Total loanwords	Total non-loanwords
1	The physical world	10.2	16.5	1.2	3.9	–	0.4	–	–	–	0.8	33.1	66.9
2	Kinship	15.3	12.9	8.2	3.5	–	–	–	2.4	–	–	42.4	57.6
3	Animals	15.5	8.2	2.3	5.5	–	4.1	0.9	–	1.8	0.9	39.3	60.7
4	The body	10.0	2.1	2.6	7.9	–	1.1	–	–	–	0.5	24.2	75.8
5	Food and drink	14.1	5.0	6.0	5.0	2.5	3.0	2.0	–	–	–	37.7	62.3
6	Clothing and grooming	3.7	8.1	20.0	8.8	6.3	5.0	2.5	–	1.3	–	55.6	44.4
7	The house	12.3	3.5	12.3	5.3	7.0	–	–	–	1.8	–	42.1	57.9
8	Agriculture and vegetation	17.2	6.1	6.1	1.8	1.8	1.2	–	–	–	2.5	36.8	63.2
9	Basic actions and technology	6.2	5.6	4.6	2.1	1.0	0.5	3.1	–	1.0	1.0	25.1	74.9
10	Motion	9.1	3.2	3.2	1.2	2.4	–	–	–	–	–	19.0	81.0
11	Possession	3.9	10.2	7.8	4.7	–	4.7	3.1	–	–	–	34.4	65.6
12	Spatial relations	7.3	3.7	2.1	–	1.0	–	–	1.0	–	–	15.2	84.8
13	Quantity	2.3	11.5	4.6	–	–	–	–	–	–	–	18.4	81.6
14	Time	1.2	11.2	3.7	18.6	3.7	–	–	–	–	–	38.5	61.5
15	Sense perception	8.0	4.7	1.3	–	–	–	–	1.3	1.3	–	16.7	83.3
16	Emotions and values	5.6	14.7	1.1	7.9	–	–	2.3	–	–	–	31.6	68.4
17	Cognition	14.7	12.9	1.2	11.0	1.2	–	–	–	–	–	41.1	58.9
18	Speech and language	6.7	9.2	3.3	5.0	3.3	1.7	–	1.7	1.7	–	32.5	67.5
19	Social and political relations	7.6	10.7	1.5	6.1	–	1.5	1.5	1.5	–	–	30.5	69.5
20	Warfare and hunting	9.2	10.8	4.6	3.1	1.5	–	–	1.5	3.1	–	33.8	66.2
21	Law	2.9	14.3	17.1	17.1	–	–	–	–	–	–	51.4	48.6
22	Religion and belief	2.5	26.3	7.5	31.3	2.5	–	–	–	–	–	70.0	30.0
23	Modern world	2.8	2.8	49.1	1.4	1.4	4.8	1.4	–	1.4	1.4	66.4	33.6
24	Miscellaneous function words	4.7	7.0	4.7	–	–	–	–	–	–	–	16.3	83.7
	All words	**8.9**	**8.4**	**6.4**	**5.7**	**1.4**	**1.2**	**0.7**	**0.4**	**0.5**	**0.4**	**34.0**	**66.0**

Equally interesting are the semantic fields least affected by borrowing. In the case of Indonesian, these are *Spatial relations* (15.2% borrowed), *Sense perception* (16.7%),

Quantity (18.4%), and *Motion* (19%)[14]. Apparently Indonesian speakers and their predecessors felt no objective necessity to borrow extensively into these fields, because they already contained most of the words necessary to convey the required concepts. Nor did they have motivation for copious cultural borrowing here, because many of the concepts in these fields are relatively culture-free.

5. Integration of loanwords

5.1. Morphological integration

Most words borrowed into Malay-Indonesian did not undergo any morphological integration, because none was necessary. The language has little inflection, and most roots may occur as words without any modification. As already mentioned (§4.2), loan verbs constitute a notable exception, in that their citation forms must contain a verbal prefix. For example, the English noun *access* was borrowed into Indonesian as *akses* without any morphological modification. However, all citation forms of verbs deriving from the English verb *(to) access* contain a prefix, e.g. *mengakses* 'to access', *mengakseskan* 'to access on someone else's behalf', *berakses* 'to have access', and *terakses* 'to be accessed / accessible'.

Certain types of complex words had a fixed pattern of integration into Indonesian. In contemporary Indonesian this process principally involves two types of borrowed English nouns. The first consists of abstract nouns whose source forms in English end in *-ation or -ization*. The integration pattern of these words is based on an earlier pattern of borrowing similar Dutch words ending in *-atie* [asi] and *-isatie* [isasi]. For example Dutch *proclamatie* 'proclamation' was borrowed as Indonesian *proklamasi*, and Dutch *modernisatie* 'modernization' was borrowed as Indonesian *modernisasi*. The same integration pattern is now applied to English loanwords which end in *-(iz)ation*. Thus English *stagflation* was borrowed as *stagflasi* and English *globalization* was borrowed as *globalisasi*. The second type is also based on an earlier pattern of borrowing abstract nouns from Dutch, in this case those ending in *-iteit*, but with an added twist. Such Dutch words were initially integrated into Indonesian with the ending *-itet* or *-iteit*, e.g. Indonesian *kualitet* or *kualiteit* 'quality' from Dutch *kwaliteit*. After independence, the ending *-ite(i)t* was viewed as too Dutch-sounding by Indonesia's language planners, who replaced it with the Latin ending from which it ultimately derived, *-itas*. Thus in contemporary Indonesian, the word for 'quality' is *kualitas*. This pattern is now used to integrate English words ending with *-ity*, e.g. *integritas* from *integrity*. The elements *-(is)asi* and *-itas* are not phonological adaptations of English *-(iz)ation* and *-ity*, but result from the mechanical application of an established integration pat-

[14] Function words are not included in this discussion because the semantic field *Miscellaneous function words* only contains items not already included in one of the other semantic fields. A more precise count is based on semantic word class, see §4.2.

tern. The elements *-isasi* and *-itas* have been so well integrated into Indonesian that they are used productively to derive new words; see §6.2 below.

5.2. Phonological integration

At its earlier stages, Malay-Indonesian had a relatively small inventory of phonemes as well as a restrictive syllable structure, so borrowing words from other languages often necessitated considerable phonological integration (see Tadmor 2007: 304–308). Initially, loanwords from any language were assimilated to the existing phonological structure. For example, Sanskrit *śīghra* [çīg^{h}ra] 'quick' was borrowed as Malay *segera* [səgəra] 'immediate'. Sanskrit *ś* was represented by the closest Malay phoneme, *s*; *gh* was likewise replaced by the closest Malay phoneme, *g*; vowel length was disregarded, since it is not distinctive in Malay; since Malay did not allow clusters, a schwa was inserted between the *g* and *r*; this resulted in a trisyllabic word, so the initial vowel *i* was reduced to schwa as required by early Malay phonology (antepenult reduction).

Other processes which affected loanwords (some of which are still productive) include the following:

- Devoicing of final voiced consonants (e.g. *masjid* (/masɟit/) 'mosque' < Arabic *masjid*, *sebab* (/səbap/) 'reason' < Arabic *sabab*);
- Reduction of complex vocalic nuclei of closed syllables by monophthongization (especially in polysyllabic words), e.g. *héran* 'surprised' < Arabic *ḥairān*, *tobat* 'repentance' < Arabic *taubat*, or by turning the diphthong into two syllables with hiatus (especially in monosyllables), e.g. *kaos* 'socks' < Dutch *kous* /kaws/, *wain* 'wine' < Dutch *wijn* or English *wine* /wajn/;
- Assimilation of the place of articulation of nasals to that of following stops (e.g. *mungkin* 'maybe' < Arabic *mumkin*, *amplop* 'envelope' < Dutch *enveloppe*);
- Reduction of some final clusters by schwa epenthesis (in loanwords from European languages, when both consonants are sonorants; e.g. *filem* 'film' < Dutch *film*, *modéren* 'modern' < Dutch *modern*), or by deleting the second consonant (in loanwords from European languages, when the two consonants are not both sonorants, e.g. *arsiték* 'architect' < Dutch *architect*, *ban* 'tire' < Dutch *band*), or echo-vowel epenthesis (in Arabic loanwords, e.g. *subuh* 'dawn' < Arabic *ṣubḥ*, *jisim* 'body' < Arabic *jism*, *rajam* 'stoning' < Arabic *rajm*).

Due to extensive borrowing from languages with different phonemic inventories and phonotactics, the phonology of Malay-Indonesian underwent changes (see §6.1) which now allow borrowed morphemes to be adopted into Indonesian with far fewer modifications than before.

6. Structural borrowing

This section contains a brief overview of structural borrowing in Indonesian. For a more detailed discussion, see Tadmor (2007).

6.1. Phonological borrowing

When borrowing from a particular language was extensive (as was the case with words from Sanskrit, Arabic, and Dutch), this eventually led to changes in phonotactics, and even to the introduction of new phonemes. For example, in older stages of Malay, the semivowels *w* and *y* were allophones of the vowels *u* and *i*. Under the influence of loanwords, they have fully phonemicized. In addition, Indonesian has several loan phonemes which were borrowed outright, such as *f* and *ç* (spelled *sy*).

Even more than the inventory of phonemes, the syllable structure of Indonesian has been profoundly affected by borrowing. In Proto Malayic, the syllable shape was (C)V(C). Due to massive lexical borrowing, Indonesian now allows up to three consonants in the onset and coda, so that the syllable shape is (C)(C)(C)V(C)(C)(C).

6.2. Morphosyntactic borrowing

6.2.1. Borrowed bound morphemes

It is difficult to establish clear-cut criteria in Indonesian for distinguishing between affixes and clitics, and it is equally difficult to distinguish between inflection and derivation. Clitics as well as affixes are treated together here under the cover term *bound morphemes*. Nevertheless, this does not mean that all are equally borrowable, and some generalizations can be made regarding the borrowability of different types of bound morphemes.

Generally speaking, bound morphemes which are more affix-like and whose function is more grammatical are less borrowable, while bound morphemes which are more clitic-like and whose function is more semantic are more borrowable. Thus *meng-* (which forms active verbs), *-i* (which derives transitive verbs), and *-an* (which derives nouns) are definitely not borrowed. At the other end of the spectrum are "pseudo-affixes", bound morphemes whose function is semantic and which do not affect the base's syntactic properties, and are as borrowable in Indonesian as free content morphemes (words). These are usually (though not always) written as separate words, unlike the true affixes mentioned above, which are never written as separate words, even by uneducated speakers. This may reflect a perception of the former as part of the same word and of the latter as separate words (even though they do not occur in isolation). Examples of borrowed pseudo-affixes in Indonesian are *non-* (< Dutch / English, e.g. *non pemerintah* 'nongovernmental'), *ekstra-* (< Dutch / English, e.g. *ekstra ketat* 'extremely strict') *super-* (< Dutch /

English, e.g. *super murah* 'super-cheap'), *maha-* (< Sanskrit, e.g. *maha penting* 'extremely important'), *pra-* (< Sanskrit, e.g. *prabayar* 'prepaid'), *pasca-* (< Sanskrit, e.g. *pasca perang* 'post-war'). In all of these examples, the "pseudo-affix" is borrowed but the base is not.

A few class-changing morphemes were borrowed into Indonesian. Two of them have already been discussed (§5.1): the Latinate affixes *-(is)asi* and *-itas* which derive abstract nouns, and became productive in Indonesian after extensive lexical borrowing from Dutch. These affixes are used to derive new words in Indonesian, although still mostly from loanword bases (interestingly including loanwords of non-European origin). The suffix *-(is)asi* was used, for example, to derive the word *swastanisasi* 'privatization' from the Sanskrit-derived base *swasta* 'private' (the *-n-* is epenthetic); the suffix *-itas* was used to derive *koneksitas* 'connectivity' from the base *koneksi* 'connection' (the word was coined in Indonesian and does not occur in Dutch); the suffix *-isme* was used to derive the word *koncoisme* 'cronyism' (the base *konco* is from Javanese).

In addition to these Dutch-derived, ultimately Latin affixes, Indonesian also borrowed a few affixes from other languages. The suffix *-awi* , of Arabic origin, derives adjectives from nouns, such as *manusiawi* 'humane' (from the base *manusia* 'human being', of Sanskrit origin) and *geréjawi* 'ecclesiastical' (from the base *geréja* 'church', of Portuguese origin). The suffix *-wan*, of Sanskrit origin, derives agent nouns, e.g. *ilmuwan* 'scientist' (from the base *ilmu* 'knowledge', of Arabic origin) and *jomblowan* 'bachelor' (from the base *jomblo* 'unmarried', of Austronesian origin).

6.2.2. Syntactic borrowing

The syntax of Malay-Indonesian has undergone many changes during its intermittently documented history. However, it is difficult to point to borrowing as the direct cause of specific changes, and it is more probable that a mixture of internal and external factors have been at work. Among features whose origin may be due at least in part to borrowing are: the broadening of the function of pronominal enclitics from purely genitive to accusative (as well as genitive); the development of copulative-like constructions; and the emergence of locative relative clauses. For more details, see Tadmor (2007).

7. Lexical adoption vs. imposition and borrowing through writing vs. borrowing through speech

Some degree of bilingualism is a precondition for linguistic interference, which may result in contact-induced language change, including lexical borrowing. However, in the case of Malay-Indonesian, the principal agents of change were not members of bilingual communities whose languages underwent linguistic interference. Sanskrit has profoundly influenced the lexicon of Malay-Indonesian, yet there was never a bilingual community that used Sanskrit and Malay. We have no reason to

assume that more than a small minority of ancient speakers of Malay ever had any knowledge of Sanskrit, and it would not be correct to say that they "spoke" it, because it was used as a literary and liturgical language. However, this minority of scholars and clergy was very influential and constituted the cultural and religious (if not political) elite. Once this elite adopted Sanskrit elements, the rest of society imitated their prestigious speech, thus incorporating Sanskrit words into their language without actually knowing – much less speaking – any Sanskrit. Similarly, the great majority of Arabic loanwords in Indonesian, whether borrowed directly or via Persian or an Indian language, were adopted from writing, and were initially used by a small elite minority before spreading to the language of the general population.

Dutch words were borrowed principally from speech and not from writing. As already mentioned, only a small minority of Indonesians were ever fluent in Dutch, although many Dutch residents of Indonesia did become fluent in some variety of Malay-Indonesian. Thus the conduits of lexical borrowing from Dutch to Indonesian were members of the Dutch-educated indigenous elite and Dutch speakers of Malay-Indonesian. Both these groups were very small numerically but very influential sociolinguistically.

The cases of Sanskrit and Arabic on the one hand, and of Dutch on the other hand, contrast with a third scenario: indigenous Indonesians borrowing lexical items from various local languages into Malay-Indonesian. The agents of this lexical borrowing were imposing words from their indigenous vocabulary on their second language rather than adopting words from another language into their native language. This view is espoused by van Coetsem (1988) as well as by Thomason & Kaufman (1988), although they used different terminologies to describe it. In this study all transfer of words from one language to another, regardless of the sociolinguistic circumstances, is subsumed under the cover term "borrowing".

No significant differences were found in Indonesian between the results of lexical imposition (shift-induced lexical change) and lexical adoption (borrowing in a maintenance situation). Both types of lexical transfer have been common in the language's history, and both have affected various semantic categories and fields in similar ways. However, two important differences were observed between borrowing through speech and borrowing through writing. On the whole, sound correspondence between source words and loanwords tended to be more regular when the borrowing was through writing. Moreover, loanwords borrowed through speech tended to be of a more colloquial and everyday nature, while loanwords borrowed through writing tended to be more literary or formal. Of course, if enough time has elapsed since the borrowing such distinctions can get somewhat blurred; colloquial words may work their way up into more literary genres, while literary words may percolate down to more colloquial styles. However, as the findings in Table 6 show, even in the case of the oldest documented borrowing situation (Indianization) the overall distinction is maintained.

Each word in the Indonesian database was tagged as "colloquial", "formal", or "general". Table 6 compares the numbers of loanwords of ultimate Sanskrit to

Arabic origin (which were borrowed almost exclusively through writing[15]) and loanwords from languages of Java and its environs (specifically Javanese, Balinese, Sundanese, and Madurese) and Dutch, which were borrowed through speech.

Table 6: Effects of borrowing through writing and through speech (percentages)

Type of borrowing	Donor language	Number of loans by style		
		Colloquial	General	Formal
Through writing	Arabic	-	79	21
	Sanskrit	-	64	36
Through speech	Java area	20	71	8
	Dutch	12	83	4

The difference in the stylistic distribution of loanwords from written and spoken sources is quite striking. The database contains no examples at all of exclusively colloquial loanwords ultimately borrowed from Arabic or Sanskrit, despite their large number and long presence in the language. On the other hand, loanwords originating in local languages of the greater Java area and in Dutch include many purely colloquial words and relatively few formal ones.

8. Speakers' attitudes

Are speakers aware of the origin of the words that they use? Generally speaking, a child acquiring the vocabulary of his first language regards all words are equally "native". However, as children grow older and acquire more formal education, they may become aware of the fact that some words in their language were borrowed from other languages, and this knowledge may affect their linguistic behavior.

Because speakers are more conscious of their vocabulary than they are of their grammar, they have a higher degree of control over their vocabulary than they do over their grammar use (see Tadmor 1995: 37ff, Tadmor 2000). This can have various effects on the use of loanwords. In the case of Indonesian, speakers who wish to emphasize their pious Islamic background may increase their use of Arabic of loanwords, while speakers who wish to show their sophisticated worldliness may exaggerate the use of English loanwords. The opposite can also be true: purists or nationalists may consciously avoid using loanwords when speaking Indonesian.

Many loanwords have nothing about their shape to indicate their foreign origin. Others, however, have loan phonemes or other features which mark them as having a foreign origin. Thus, educated speakers of Indonesian know that words

[15] A handful of Arabic loanwords, representing a tiny proportion of the total, appear to have been borrowed through speech rather than through writing. An example is the word *raib* 'disappear' (< Arabic *ɣaib*), where the Arabic *ɣ* is represented by Malay-Indonesian *r*, whereas normally it is represented by *g*. Another example is the word *menara* 'tower'; if borrowed through writing, it would end with *-h* (or *-t*).

containing the consonants /f/ or /z/ are loanwords. They also know that nouns ending in *-ah* / *-at* are of Arabic origin, and that adjectives ending in *-if* are of European (Dutch or English) origin.

As mentioned above, Indonesian speakers productively use several borrowed derivational affixes, such as the abstract noun forming *-asi*, *-isasi*, and *-itas*, and the adjective forming *-awi* (see §6.2.1). Interestingly, these affixes are used to create new words almost exclusively from loanword bases (regardless of their origin). This is an indication that loanwords constitute a distinct category in the speakers' minds.

A well-known morphophonological process of Indonesian is the sandhi rule that assimilates the active prefix *meng-* (and the agentive prefix *peng-*) to the initial consonant of the base. Interestingly, despite protestations from *Pusat Bahasa* (Indonesia's language planning body), speakers often disregard this rule when they are aware that the base is a loanword. This is another indication that loanwords enjoy a special (or different) status in the speakers' minds. Table 7 presents a few examples of forms produced by speakers, compared to normative forms[16].

Table 7: Standard and nonstandard forms of words with the prefix *meng-* and a borrowed base

Nonstandard form	**Standardized form**	**Gloss**	**Base**	**Origin**
mensubsidi	*menyubsidi*	to subsidize	*subsidi*	Dutch
mentaati	*menaati*	to obey	*taat*	Arabic
memprotés	*memrotés*	to protest	*protés*	Dutch / English
mengcopy / *mengkopi*	*mengopi*	to copy	*copy/kopi*	English

Orthography is another area where knowledge of a word's origin can come into play. In Indonesian, voiced consonants do not occur in final position. Loanwords which end in a voiced consonant are invariably assimilated into Indonesian by undergoing final devoicing. However, their spelling often reflects the final voiced consonant of the source word, for example in *tertib* [tərtip] 'orderly' (< Javanese), *masjid* [masɟit] 'mosque' (< Arabic), *uleg* [uləg] 'beat in a mortar' (< Javanese), and *iméj* [imec] 'image' (< English).

9. Conclusion

About one third of the words in the Indonesian database were identified as probable or certain loanwords. This relatively high proportion is explained by the long history of contacts between speakers of Malay-Indonesian and other cultures, as well as by the long-standing role of Malay-Indonesian as a regional lingua franca. Specific

[16] It should be noted that older, well-assimilated loanwords are no longer perceived as borrowed and therefore do undergo the expected sandhi rule. Moreover, words from local languages in which similar rules operate also undergo them after being borrowed into Indonesian.

historical processes, such as Indianization, Islamization, colonization, and globalization, have all resulted in considerable lexical borrowing into Malay-Indonesian.

Previous studies have pointed out the importance of distinguishing between imposition and adoption. The present study did not find significant differences between the impacts of these two different processes on the lexicon of Indonesian. However, some important differences between the results of borrowing through speech and borrowing through writing were observed. An area that has not been satisfactorily investigated yet, and about which some ideas are presented here, is speakers' attitudes and role in contact-induced lexical change. Whether and how speakers perceive loanwords can significantly influence borrowing processes and their linguistic outcomes.

A specific contribution of this study was providing the first systematic discussion of words borrowed into Indonesian from other languages of Southeast Asia, including languages indigenous to the Malay-Indonesian archipelago as well as neighboring languages. As a preliminary effort, it probably contains numerous errors (of omission as well as of commission), and leaves much room for future research. With regard to languages of Europe, the Near East, South Asia, and China, the present study has built upon extensive previous work, and its efforts have been focused on identifying previously unrecognized loanwords, correcting errors in the identification of donor languages and source words, and linking between historical events and their linguistic outcomes.

Acknowledgment

The author wishes to express his thanks to Waruno Mahdi, who has read an earlier draft of this chapter and has provided many helpful comments.

References

Alwi, Hasan (chief ed.). 2002. *Kamus Besar Bahasa Indonesia (KBBI) [Unabridged Indonesian Dictionary]*. 3rd edn. Jakarta: Pusat Bahasa.

Campbell, Stuart. 1996. The distribution of -at and -ah endings in Malay loanwords from Arabic. *Bijdragen tot de Taal-, Land- en Volkenkunde* 152:23–44.

de Casparis, J. G. 1997. *Sanskrit Loan-words in Indonesian*. Nusa 41 (monograph). Jakarta: Atma Jaya University.

Echols, John M. & Shadily, Hassan. 1975. *Kamus Inggris-Indonesia: An English-Indonesian Dictionary*. Ithaca: Cornell University Press / Jakarta: Gramedia.

Echols, John M. & Shadily, Hassan. 1998. *Kamus Indonesia-Inggris: An Indonesian-English Dictionary*. 3rd edn. Ithaca: Cornell University Press / Jakarta: Gramedia.

Gonda, Jan. 1952. *Sanskrit in Indonesia*. Den Haag: Oriental Bookshop.

Grijns, D. J. & de Vries, J. W. & Santa Maria, L. 1983. *European Loan-words in Indonesian*. Indonesian Etymological Project 5. Leiden: KITLV.

Johns, A. H. & Prentice, D. J. (eds.-in-chief). 1992. *Kamus Inggris-Melayu Dewan: An English-Malay Dictionary*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Jones, Russell. 1978. *Arabic Loan-words in Indonesian*. Indonesian Etymological Dictionary 3. London: School of Oriental and African Studies.

Jones, Russell. 1984. Loan-Words in Contemporary Indonesian. *Nusa* 19:1–38.

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

Kamus Besar Bahasa Indonesia. 2002. see Alwi (2002).

Kamus Dewan. 1991: see Othman bin Sheikh Salim (1991).

Kamus Inggris-Melayu Dewan. 1992: see Johns & Prentice (1992).

Meuleman, Johan H. 2005. History of Islam in Southeast Asia: Some questions and debates. In Nathan, K. S. & Kamali, Muhammad Hashiun (eds.), *Islam in Southeast Asia: Political, Social and Strategic Challenges for the 21st Century*, 2nd edn. 22–44. Singapore: Institute of Southeast Asian Studies.

Mu'jizah. 2009. *Iluminasi dalam Surat-Surat Melayu Abad ke-18 dan ke-19 [Illuminations in Malay letters from the 18th and 19th centuries]*. Jakarta: Kepustakaan Populer Gramedia, École française d'Extrême-Orient, Pusat Bahasa – Departemen Pendidikan Nasional / KITLV-Jakarta.

Othman bin Sheikh Salim, Sheik (chief ed.). 1991. *Kamus Dewan: Edisi Baru [Dewan dictionary: New edition]*.

Stevens, Alan M. & Schmidgall-Tellings, A. Ed. 2004. *Kamus Lengkap Indonesia-Inggris: A Comprehensive Indonesian-English Dictionary*. Athens: Ohio University Press / Bandung: Mizan.

Tadmor, Uri. 1995. Language Contact and Systemic Restructuring: The Malay Dialect of Nonthaburi, Central Thailand. Ph.D. dissertation. University of Hawaii.

Tadmor, Uri. 2000. *Can speakers control contact-induced language change?* Paper presented at Sociolinguistics Symposium 2000, University of the West of England, Bristol, 27th–29th April, 2000.

Tadmor, Uri. 2007. Grammatical borrowing in Indonesian. In Matras, Yaron & Sakel, Jeanette (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*, 301–328. Berlin: Mouton.

van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

van Dam, Nikolaos. 2009. Arabic loan-words in Indonesian revisited. Unpublished manuscript. Jakarta.

Wilkinson, R. J. 1959. *A Malay-English Dictionary (Romanised)*. 2 vols. London: MacMillan & Co.

Yule, Henry & Burnell, A. C. 1903. *Hobson-Jobson: A Glossary of Colloquial Anglo-Indian Words and Phrases, and of Kindred Terms, Etymological, Historical, Geographical and Discursive.* Crooke, William (ed.). London: Routledge & Kegan Paul.

Loanword Appendix

Sanskrit

jagat	world (lit.)
buana	world (lit.)
bumi	land
gua	cave
samudra	sea, ocean (lit.)
gempa	earthquake
angkasa	sky (lit.)
surya	sun (lit.)
candra	moon (lit.)
cahaya	light
udara	air, weather
bayu	wind (lit.)
méga	cloud (lit.)
cuaca	weather
manusia	human being
pria	man, (human) male (formal)
putra	boy, son (formal)
teruna	young man (lit.)
putri	girl, daughter (formal)
suami	husband
istri	wife
janda	widow
saudara	relatives
saya	I
gembala	herdsman
kuda	horse
angsa	goose
rajawali	eagle
(burung) merpati	dove
serigala	wolf, jackal
singa	lion
gajah	elephant
kepala	head
muka	front, face
bahu	shoulder
selesma, selésma	the cold (lit.)
kendi	jug, pitcher
cabé	chili pepper
madu	honey
gula	sugar
busana	clothing, clothes (lit.)
kapas	cotton
sutera	silk
manik-manik	beads
lépa	mortar (substance)
menenggala [meng-tenggala]	to plough/ plow (lit.)
biji	seed, grain
belia	youth (lit.)
cemara	conifer, esp. the casuarina
labu	gourd, pumpkin, squash
kerja	work
mencuci [meng-cuci]	to wash
kencana	gold (lit.)
kaca	glass (material)
mengendarai [meng-kendara-i]	to drive (lit.)
marga (1)	road (lit.)
benda	thing
memelihara [meng-pelihara]	to preserve
papa (1)	poor (lit.)
harga	price
membagi [meng-bagi]	to share, to divide
sisa	remains
utara	north
daksina	south (lit.)
semua	all
segala	all (lit.)
pertama	first
masa	time, period
usia	age (formal)
dini	early (lit.)
segera	immediately, soon
laju	fast
mula	beginning
sedia	ready (lit.)
sentiasa, senantiasa	always (lit.)
berasa [ber-rasa]	to taste (intr.)
suara	sound, noise, voice
sunyi	quiet
jiwa	soul, spirit
celaka	bad luck
bahagia	happy, content
gembira	glad, happy
cinta	to love
murka	(monarch's) anger (lit.)
bahaya	danger
setia	faithful
berdusta [ber-dusta]	to lie, tell a lie (lit.)
puji	praise
loba	greedy (lit.)
pandai	clever
percaya	to believe
menerka [meng-terka]	to guess (lit.)
bijaksana	wise
guru	teacher
rahasia	secret
menyangka [meng-sangka]	to suspect
cara	manner
karena	because
atau	or
bila	when? (lit.)
berbicara [ber-bicara]	to speak, talk

bahasa	language
kata	word
nama	name
membaca [meng-baca]	to read
negara	country
kota	town
raja	king
setru	enemy (lit.)
berkelahi [ber-kelahi]	to fight
tentara	army, soldier
senjata	weapons
jala	fishnet
saksi	witness
pidana	punishment (legal)
denda	fine
penjara	prison
agama	religion
déwa	god
pendéta	priest
suci	holy
berpuasa [ber-puasa]	to fast
surga	heaven
neraka	hell
berhala	idol
bidadari	fairy
menteri	minister
angka	number, digit
sama	same

Hindi-Urdu

unta, onta	camel
roti	bread
celana	trousers
topi	hat, cap
kunci	lock
tembakau	tobacco
tembaga	copper
jam	hour, clock
mencuri [meng-curi]	to steal

Tamil

badai	storm
keledai	donkey
mengacu [meng-acu]	to cast (metal) (lit.)
kapal	ship
pasar	market
kedai	shop, store
cuma	only
teman	friend
handai	friend (lit.)
taulan	friend (lit.)
bedil	gun
perisai	shield (lit.)
nelayan	fisherman
kuil	temple

Telugu (?)

tiga	three

Dutch

és	ice
papa (2)*, papi*	father
mama, mami	mother
opa	grandfather
oma	grandmother
om, oom	uncle
tante	aunt
famili	relatives
kelinci	rabbit, hare
kangguru	kangaroo
léver	(human) liver (col.)
maag	stomach (col.)
pénis	penis (lit.)
vagina	vagina (lit.)
dokter	physician
open	oven
gelas	cup, glass
tang	tongs
sosis	sausage
sop, sup	soup
bir	beer
wol	wool
linen	linen
katun	cotton
mantel	coat
blus	blouse
kaos	T-shirt
kerah	collar
rok	skirt
kaos (kaki)	sock, stocking
péci	hat, cap
(kain) lap	handkerchief, rag
handuk	towel
salep	ointment
kamar	room
selot	lock, latch, door-bolt
kompor	stove
rak	shelf
balok	beam
semén	cement, mortar
kamp	camp
got	ditch
sekop	shovel
havermut	oats
palem	palm tree
mengelap [meng-lap]	to wipe
mengebor [meng-bor]	to bore
lém	glue
karpét	carpet, rug
bumerang	boomerang
menyupir [meng-supir]	to drive
menyetir [meng-setir]	to drive
as	axle
kano	canoe
duit	money
rékening	bill, invoice
bon	bill, check
gaji	wages
puing	remains
huk	(street) corner
nol	zero
massa	crowd
lat	late, to be late
telat	late, to be late
arloji	clock
jelék	bad, ugly
idé	idea
pulpén	pen
trompét	horn, trumpet
komplot	plot
ketapél	sling
senapan, senapang	gun
hélem	helmet
memvonis [meng-vonis]	to condemn, sentence
setrap	punishment (in school)

penalti	penalty (in soccer)
vonis	penalty, (legal) punishment
bui	prison
altar	altar (in a church)
pastor, pastur	priest (Roman Catholic)
radio	radio
télepon, télpon, telpon, télefon, telfon	telephone
mobil	car
bis	bus
listrik	electricity
baterai	battery
mengerém [meng-rém]	to brake
mesin	motor, machine
motor	motor
suster	nurse, nun
pil, pél	pill, tablet
inyéksi	injection
présidén	president
polisi	police
plat (mobil)	license plate
nomor	number
pos	post, mail
prangko, perangko	postage stamp
bank	bank (financial institution)
kran, keran	tap, faucet
wastafel	sink
toilét, toalét	toilet
klosét	toilet
WC	toilet
sekrup, sekerup	screw
permén	candy, sweets
plastik	plastic
bom	bomb
béngkél	workshop
rokok	cigarette
koran	newspaper
kaléndér	calendar
filem	film, movie
musik	music
kopi	coffee
nihil	nothing (lit.)
silét	razor

English

kormoran	cormorant
(burung) tukan	toucan
oposum	opossum
jaguar	jaguar
tapir	tapir
flu	cold
picer	jug, pitcher
olive	olive
wine	wine
gaun	(woman's) dress (long)
tunik	(woman's) dress (short)
daster	(woman's) dress, robe
(sepatu) bot	boot
tato	tattoo
laso	lasso
koin	coin
membarter [meng-barter]	to barter
bolpoin	ballpoint pen
konspirasi	plot
télevisi	television
TV	television
injéksi	injection

English or Dutch

laguna	lagoon
botol	bottle

Portuguese (incl. Creole)

garpu	fork, pitchfork
kéju	cheese
mentéga	butter
keméja	shirt
sepatu	shoe
saku	pocket
peniti	pin
tuala	towel (lit.)
ténda	tent
jendéla	window
méja	table
terigu	wheat
berdansa [ber-dansa]	to dance
keréta	cart, wagon
roda	wheel
bola	ball
témpo	time
minggu	week
(hari) Minggu	Sunday
sekolah	school
péna	pen
serdadu	soldier (lit.)
geréja	church
lampu	lamp
menyéka [meng-séka]	to wipe
buku	book
kaléng	tin, can
pipa	pipe

Spanish

sabana	savanna

French

sepéda	bicycle

Latin

pinus	pine

Arabic

dunia	world
alam	world
médan	plain, field (lit.)
salju	snow
hawa	weather, air (lit.)
bagal	mule
(burung) nuri	parrot
arnab	rabbit (lit.)
badan	body
wajah	face
rahim	womb
bernafas [ber-nafas]	to breathe
lahir	to be born
hamil	pregnant
mayat	corpse
jasad	corpse (lit.)
jenazah	corpse
kubur, kuburan	grave
makam	grave
kuat	strong
séhat	healthy

beristirahat [ber-istirahat]	to rest
(buah) zaitun	olive
jubah	cloak
jilbab	headband, headdress
sabun	soap
pondok	hut
kémah	tent
kursi	chair
alat	tool
sejedah, sajadah	(prayer) rug
raib	to disappear (mysteriously)
menyelamatkan [meng-selamat-kan]	to rescue, save
miskin	poor
waktu	time, when (conj.)
umur	age
awal	early, beginning
akhir	end (temporal)
tamat	end (temporal, lit.)
fajar	dawn (lit.)
subuh	dawn
(hari) Ahad	Sunday (lit.)
Senin	Monday
Selasa	Tuesday
Rabu	Wednesday
Kamis	Thursday
Jum'at	Friday
Sabtu	Saturday
musim	season
arwah	soul, spirit
roh	soul, spirit
héran	surprised, astonished
amarah	anger (lit.)
serakah	greedy
tamak	greedy
akal	mind
berpikir [ber-pikir]	to think
murid	pupil
yakin	certain, convinced
maksud	intention, meaning
niat	intention
sebab	cause, because
mengkhianati [meng-khianat-i]	to betray
menjawab [meng-jawab]	to answer
kertas	paper
kalam	pen (lit.)
rakyat	people
jiran	neighbor (lit.)
adat	custom
menara	tower
hukum	law
mahkamah	court
hakim	judge
mendakwa [meng-dakwa]	to accuse (in court)
menghukum [meng-hukum]	to punish
zina, zinah	adultery
masjid, mesjid	mosque
korban	sacrifice, victim
beribadah [ber-ibadah]	to worship
berdoa [ber-doa]	to pray (for something)
sholat	to pray (formulaic, Islamic)
imam	clergyman (chiefly Islamic)
kudus	holy (lit.)
berkhotbah [ber-khotbah]	to preach
sétan	demon
sihir	magic
alamat	address

Persian

domba	sheep
rubah	fox
piring	dish, plate
pinggan	dish
anggur	grape, wine
seluar	trousers
cadar	veil
serban, sorban	headband, headdress
destar	man's headdress
bandar	ditch, port
gandum	wheat
saudagar	merchant (lit.)
laskar	soldier (lit.)
peri	fairy

Old Javanese

bapa, bapak	father
ibu	mother
meréka	they
jawawut	millet
mengantar [meng-antar]	to bring, take
merusak [meng-rusak]	to damage
warna	color
mengajar [meng-ajar]	to teach
pasti	certain
ratu	queen
Javanese	
kali	river, stream, canal
rembulan	moon (lit.)
menyulut [meng-sulut]	to light
bapak mertua	father-in-law
ibu mertua	mother-in-law
wanita	woman
bocah	child (young human)
leluhur	ancestors
caplak	flea
jénggot	beard
bréwok	beard
ketombé	dandruff
kerongkongan	throat
kéték	armpit
payudara	breast (formal)
usus	intestines, guts
cérét	kettle
blangkon	man's headdress
jagung	maize, corn

singkong	cassava, manioc
jamur	mushroom
paron	anvil
timbal	lead
terpelését [ter-pelését]	to slide, slip
merintih [meng-rintih]	to groan
mengerti	to understand
gagasan	idea
konyol	stupid
enggak, nggak, gak, ga	no (col.)
kapan	when?
warga(negara)	citizen
lonté	prostitute

Sundanese

situ	man-made lake
monyét	monkey
kelapa	coconut

Javanese or Sundanese

pesisir	shore
bayi	baby
mertua	parents-in-law
sapi	cattle
banténg	ox
(ayam) jago	cock, rooster
kalong	bat
(burung) bétét	parrot
tawon	bee, wasp
rayap	termites
kéong	snail
kodok	frog
kadal	lizard
jambut	pubic hair
buntut	tail
kontol	penis
pilek	cold
capék, capai	tired
paceklik	famine
panci	pot, pan
séndok	spoon
sarapan	breakfast
adonan	dough
soto	meat soup
kentang	potato
gubuk	hut
gapura	gate (lit.)
gréndél	latch, door-bolt
obor	torch
pacul	spade
membabat [meng-babat]	to mow
panén	harvest
jeruk	citrus fruit
menggebrak [meng-gebrak]	to pound with fist
memencét [meng-pencét]	to press
menétés [meng-tétés]	to drip
menyemprot [meng-semprot]	to splash, spray
mengguyur [meng-guyur]	to splash, spray, drench
berjogét [ber-jogét]	to dance
menggéndong [meng-géndong]	to carry
pojok	corner
separo	half
kecut	unpleasantly sour
énténg	light (of weight)
kagét	surprised, astonished
merangkul [meng-rangkul]	to embrace
pintar	clever
goblok	stupid
bégo	stupid
gampang	easy
gara-gara	because
ngobrol [ng-obrol]	to speak, talk (col.)
mencegah [meng-cegah]	to prevent
prajurit	soldier
bénténg	fortress
romo	(Catholic) priest
kembang gula	candy

Balinese

bianglala	rainbow (lit.)
lindung	freshwater eel
menénténg [meng-ténténg]	to carry in hand
samping	side

Balinese/Javanese/Sundanese

rawa	swamp
paman	uncle
bibi	aunt
keponakan	sibling's child
duda	widower
céléng	boar
bébék	duck
bunglon	chameleon
iga	rib (of animals)
jempol	thumb
pusar	navel
sabuk	belt
témbok	wall
arit	sickle
mengecor [meng-cor]	to cast
kembang	flower
mengocok [meng-kocok]	to shake
mengungsi [meng-ungsi]	to flee
pajak	tax
garis	line
mirip	similar
soré	afternoon
wangi	fragrant
sepi	quiet
empuk	soft
ngomong [ng-omong]	to speak, talk (col.)
mengoméli [meng-oméli]	to scold
désa	village
taméng	shield
menuding [meng-tuding]	to accuse
tanpa	without

Madurese (modern or earlier)

celurit, clurit	sickle

Minangkabau (modern or earlier)

gadis	girl, young woman
dangkal	shallow
pidato	speech
kalian	you (plural)
kumuh	dirty

Toba Batak

marga	clan

Hokkien Chinese

kecoa	cockroach
téko	kettle, teapot
giwang	earring
képang	plait, braid
cat	paint
uang	money
toko	shop, store
hoki	good luck
sué	bad luck
cabo	prostitute
téh	tea

Mandarin Chinese

cawan	cup (lit.)

Khmer (modern or earlier)

mas, emas	gold
semut	ant
kuk	yoke
dian	(oil) lamp

Unknown origin

binatang	animal
pusut	awl
remaja	youth
permadani	rug (lit.)
nokén	net bag
mengendus [meng-endus]	to sniff
pujangga	poet (lit.)
pesakitan [pe-sakit-an]	captive, prisoner
obéng	screwdriver