J. Sneddon

Diglossia in Indonesian

In: *Bijdragen tot de Taal-, Land- en Volkenkunde* 159 (2003), no: 4, Leiden, 519-549

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

J.N. SNEDDON

Diglossia in Indonesian

Introduction

Indonesian, the national language of Indonesia, so closely fits the concept of diglossia as originally defined by Charles Ferguson that it is surprising there has to date been so little study of the language in this context.¹ This is all the more surprising given that it is the sole national and official language of the fourth most populous nation in the world. This paper takes Ferguson's original *Diglossia* paper of 1959 as its starting point and compares the ways in which Indonesian conforms to his model and the areas in which it appears to differ.²

Ferguson (1959:336) defined diglossia as:

a relatively stable language situation in which [...] there is a very divergent, highly codified (often grammatically more complex) superposed variety [...] which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation.

He called the superposed variety the high variety, abbreviated to H; the variety used in everyday situations he called the low variety, abbreviated

¹ Research on which this paper is based was funded by an Australian Research Council Grant. The study has been greatly facilitated by the generous cooperation of Soenjono Dardjowidjojo, Anton Moeliono, Uri Tadmor, Bambang Kaswanti Purwo and Yassir Tjung. An earlier abbreviated version of the paper was read at the Sixth International Symposium on Malay/Indonesian Linguistics, Bintan, in August 2002. I wish to thank anonymous reviewers whose comments resulted in improvements to the paper.

² In his bibliographic review of the literature on diglossia Hudson (1992) cites just three papers dealing with Indonesian: Tanner (1967), Cartier (1980) and Errington (1986). None of these is centrally concerned with diglossia in the language and none refers to the continuum between H and L. Diglossia is also briefly mentioned by Wallace (1979:71, 85), Grijns (1991b:81) and Moeliono (1994a:377-8).

J.N. SNEDDON is Associate Professor at the School of Languages and Linguistics, at Griffith University and holds a PhD from the Australian National University. Specializing in the grammatical description and the sociolinguistic situation of the Indonesian language, he is the author of *Indonesian; A comprehensive grammar*, London, Routledge, 1996, and *The Indonesian language; Its history and role in modern society*, Sydney, UNSW Press, 2003. Professor Sneddon may be reached at the School of Languages and Linguistics, Nathan Campus, Griffith University, Brisbane, Queensland 4111, Australia.

to L. Following Ferguson, in this paper colloquial or informal varieties of Indonesian are called L and the superposed variety is called H, although the terms 'formal Indonesian', 'informal Indonesian' and 'colloquial Indonesian' are also used. For Indonesian H is recognized as the standard form of the language; there is no standard L variety, although colloquial Jakartan Indonesian is acquiring such a status.

As will be shown, Indonesian conforms to a remarkable degree to the concept as described by Ferguson in his examination of four 'defining' languages. It does, however, differ in one important way from diglossia as originally described by Ferguson: rather than two distinct forms of Indonesian with a clear boundary there is a continuum between the two extremes. As the social situation becomes more formal L features are gradually replaced by H features, though not at a consistent rate. Some characteristics of the L variety are replaced in semi-formal speech, while others persist even in quite formal situations. Moreover, there is considerable variation in usage between individuals of even quite homogeneous background. Although some scholars regard strict separation of the two varieties as an essential characteristic of diglossia, it is argued below that Indonesian, despite the continuum, should be regarded as existing in a diglossic state. In a later paper, Ferguson (1991) recognized a continuum between L and H in his four defining languages and this may be a general characteristic of diglossic languages.

The occurrence of many regional languages used alongside the national language adds a further dimension to diglossia in Indonesia. This paper deals only with variation within the national language, in line with Ferguson's original conception of diglossia as a single-language phenomenon. Regional languages are mentioned briefly when their role vis-à-vis the national language needs to be considered. The historical background of Indonesian, which emerged from varieties of Malay, is not discussed here.³

Since independence H Indonesian has become standardized to a significant degree, due to its use in education and in the mass media. Among L varieties there is significantly more regional variation than in the H variety, although colloquial Jakartan Indonesian, in the subvariety used by the educated classes, is increasingly influencing other informal varieties and developing towards standardization. This variety is sometimes confused with Jakarta Malay, which evolved from the lingua franca used originally between the various ethnic groups who came together after 1619 when the Dutch founded Batavia. Although Jakartan Indonesian has been greatly influenced by Jakarta Malay, it is nevertheless distinct from it. Jakarta Malay, also called Betawi

³ The early history of the language is discussed by Prentice (1978), Steinhauer (1980), Adelaar and Prentice (1996), Collins (1998), and Sneddon (2003). More detailed accounts of its development during the Dutch period are Hoffman (1979) and Maier (1993).

Malay, is the vernacular of the *anak Betawi* 'children of Batavia', inhabitants of 'the old kampungs' of Jakarta, whose families have usually lived in Jakarta for many generations, while the more general colloquial Jakartan Indonesian is used among the Indonesians who have flocked to the city since independence. Grijns (1981, 1983, 1991a) has most insightfully articulated the differences between Jakarta Malay and modern Jakartan Indonesian. Grijns shows that the *anak Betawi*, who constitute an ever-dwindling percentage of the overall Jakartan population, shift to informal Jakartan Indonesian in the presence of strangers; it is an 'in-group' code and rarely heard by people from outside the *anak Betawi* communities. Oetomo (1990: 69) points out that young families in Jakarta who speak Betawi Malay shift to Jakartan Indonesian 'when they rise socially or have aspirations in that direction'. This shows that the difference is clearly discernible and that colloquial Jakartan Indonesian is associated with a higher social status. It is the everyday language of the great majority of Jakartans, including the elite and the educated middle-class.

Diglossic characteristics of Indonesian

The characteristics of diglossia listed by Ferguson, with the exception of literary heritage and stability, are considered here in relation to Indonesian, although not necessarily in the order in which he presented them. Evidence for the existence of a continuum between the extremes of H and L is then given. A consideration of literary heritage lies beyond the scope of this paper, while stability is briefly considered in the conclusion.

Separate functions

One of the most important features of diglossia is the specialization of function for H and L. In one set of situations only H is appropriate and in another only L [...]. (Ferguson 1959:328.)

In Indonesia the H variety is the language of government, administration and the law and of formal situations, such as speeches and lectures. It is the language of literature, including poetry and most novels, and of most of the mass media. It is the medium of education at all levels. L is the language of the home and of casual conversation. Ferguson (1959:329) gives a list of possible contexts in which H and L are normally used. Indonesian conforms almost entirely with his listing, H being the language of sermons, speeches in parliament, university lectures, news broadcasts, newspaper editorials and news stories, and poetry. L is the language for instructions to servants, conversation with family, friends and colleagues, and captions on political cartoons.

In only a few minor ways does the situation in Indonesia differ from

Ferguson's listing. Ferguson's list has personal letters as normally in H while radio 'soap operas' and folk literature are normally in L. In Indonesian personal letters between young people are certain to contain many features of L. On the other hand, letters to parents may be in H; this is likely to be the case for instance among higher-class Javanese. Radio and television soap operas, which were once regularly in H, are now far more likely to be in L, reflecting natural conversation. Nevertheless, in many television serials and in most novels and stage dramas today characters still tend to use formal language, which is accepted in that context but would not occur in real life.⁴ One noticeable change in recent years has been the appearance of magazines and pulp novels for urban youths which tend to use L. Folk literature typically occurs in the language of the home, which for most Indonesians is a regional language. When such literature is translated into written Indonesian it is into the high variety, the variety associated with education, literature and writing in general. As L Indonesian becomes the language of the home for increasing numbers of people oral folk literature is carried over into it. Regional vernacular Malays have always had their folk literature and in earlier times there was also written folk literature in Classical Malay, almost certainly 'translated' from vernacular stories.

The inappropriateness of the use of L elements in strictly formal situations can be seen in the use of italics whenever an interviewee is quoted in the print media using L. With this device the publication dissociates itself from the particular word or phrase. Sometimes a publication will go further in dissociating itself from the use of informal language, as in the 'translation' of a stretch of colloquial speech into its formal equivalent, even though its meaning is perfectly clear to the reader, as in the following:

'Kami sudah pasrah. *Nggak* tahu mau apa lagi, [...] sekarang memang semua ditanggung Sumber Waras. Tapi, *abis gitu gimana* (setelah itu bagaimana)?' ujar Aminah [...]. (*Femina*, 18-24 June 1998, p. 24.)

'We are resigned to our fate. We don't know what else we can do [...] now everything is the responsibility of Sumber Waras [the name of a hospital]. But after that, who knows?' said Aminah [...]

In the passage the mother of an injured youth says in L Indonesian *abis gitu gimana*? 'After that who knows (what will happen)?', which is italicized and repeated in parentheses in formal style. The L word *nggak* 'not' is also italicized.⁵

⁴ Use of the distinctly literary or formal pronoun *engkau* 'you' is one characteristic of speech in television and stage dramas which are supposed to be about everyday life (George Quinn, personal communication).

⁵ Words from regional and foreign languages are likewise italicized in publications when they are felt not to be assimilated into Indonesian.

It is equally true that formal Indonesian is never associated with informal activities. Ferguson (1959:329) states: 'A member of the speech community who uses H in a purely conversational situation or in an informal activity like shopping is [...] an object of ridicule'. And (Ferguson 1959:337): 'no segment of the speech community in diglossia regularly uses H as a medium of ordinary conversation, and any attempt to do so is felt to be either pedantic and artificial [...] or else in some sense disloyal to the community [...]'. This observation is certainly true in the Indonesian situation. Nevertheless, because of the continuum between L and H extremes, H elements can occur in conversations, as shown below.

Acquisition

L is invariably learned by children in what may be regarded as the 'normal' way of learning one's mother tongue. H may be heard by children from time to time, but the actual learning of H is chiefly accomplished by the means of formal education. (Ferguson 1959:331.)

Most Indonesian children have little or no contact with formal Indonesian until they begin their education. The language of the home is L (although it may alternatively or in addition be a regional language). Formal Indonesian is the medium of instruction in all schools, with the exception of the use of a regional language in the first two or three years of education in some regions where there is a large predominant language, such as Javanese or Sundanese. Formal Indonesian is also a subject of study in the education system and is expected to be mastered by educated people. Proficiency in it is thus a mark of a person's level of education; someone having little or no education is very unlikely to have any proficiency in it.

Prestige

In all the defining languages the speakers regard H as superior to L in a number of respects (Ferguson 1959:329).

The fact that in Indonesian L is never associated with education results in negative attitudes towards it. Although teacher and students will speak to each other outside the classroom in L it is not heard within the classroom. As mentioned, able speakers of H are probably all well-educated and proficiency in it is a necessity for obtaining employment at a professional level. Its association with education, affluence and status contributes significantly to its prestige.

In addition, there are regular calls in the mass media for the use of *bahasa yang baik dan benar* 'good and correct language'. These calls come from people associated with language planning and education but also from others like

journalists and politicians (who are themselves rarely models of good usage). There is also a flourishing language advice industry, catering to people who recognize the advantages of improving their ability in H, with radio and television programs, newspaper columns and a constant stream of booklets on good usage. These deal solely with usage in formal situations; as informal situations are not dealt with aspects of pronunciation, vocabulary and grammar specifically associated with L are never mentioned as appropriate. The implicit message is that L is never *baik dan benar* 'good and correct'.

Because L is not associated with education it is usually regarded by native-speaking teachers as inappropriate for teaching to foreign learners of the language and mention of it rarely occurs in published Indonesian language courses.⁶ As just one example of the attitude of Indonesians who teach the language, Sarumpaet (1980:vi) writes: 'It is better for a foreign student to err on the correct or even formal side than to employ careless, incorrect or substandard usage'. Here 'correct' is equated with 'formal'; informal is careless, incorrect and substandard. This paper is not concerned with teaching of the language but the attitude to what is appropriate to teach to foreigners does highlight the lack of prestige of L.⁷

Ferguson (1959:329-30) notes that the feeling that H is superior is sometimes so strong that H alone is regarded as real and L is reported 'not to exist'. He refers to educated speakers of Arabic and Haitian Creole as frequently denying that they ever use L, whereas in fact they use it constantly in all ordinary conversation, a denial which he regards as 'almost a self-deception'.

The fact that formal Indonesian is taught and studied while informal varieties are not has resulted in a quite common assumption that 'Indonesian' refers solely to the formal language, and that consequently the language has no native speakers. In the words of one Indonesian linguist: 'The national language is considered a formal language, to be used in schools or at official functions where a range of ethnic groups is likely to be represented' (Lauder 1998:73). An Indonesian sociologist has written:

Bahasa Indonesia is a product of language planning, engineering, and Development programs *par excellence*.⁸ It does not evolve from communal activities in the ordinary lives of its speakers. It has not been a mother tongue to anyone. Speakers

⁶ Johns (1996) is probably the only published teaching text which specifically deals with L, contrasting features of it with H. Wolff, Oetomo and Fietkiewicz (1992) also includes some materials on informal language. One teaching text, entitled *Colloquial Indonesian* (Atmosumarto 1995), is misnamed, being a course devoted entirely to the formal variety of the language.

⁷ Problems associated with the teaching of informal Indonesian are discussed by Sneddon (2001).

⁸ The term 'Bahasa Indonesia' is sometimes used in English-language writings but is avoided here except when quoting others. In Indonesian the word *bahasa* 'language' is a common noun and is not spelled with a capital B. The term simply means 'the language of Indonesia'.

of Bahasa Indonesia learn it from authorized institutions and professionals as a language that their mothers do not speak. (Heryanto 1995:5.)

As mentioned, an ever-increasing number of people living in cities are now monolingual in Indonesian. The statements by Lauder and Heryanto clearly indicate that to them only H can be called Indonesian but the question of what then is the home language of such people is not addressed.

This failure to recognize anything but the most formal variety as Indonesian has frequently led to its being stigmatized as a 'soulless' and alienating language. One writer refers to Indonesian as 'a language of an alienated and bureaucratic intelligentsia' which seems to the Javanese, 'by comparison with their own rich language, peculiarly turgid, humourless, awkward, mechanical and bereft of emotion or sensuality' (Peacock 1973: 79). And another writes: 'Modern Indonesian has something curiously impersonal and neuter about it, which sets up psychological distances between its speakers' (Anderson 1966:105-6).

The fallacy common to such claims is that only very formal language is actually Indonesian. However, Indonesians use H (if at all) and L in appropriate settings. For most of their daily activities they speak informal variants of the language (if they are not speaking a regional language); Indonesian is the language of the market place as well as of the lecture hall. Anyone who has spent time with Indonesians can hardly be anything but puzzled at claims that their language sets up psychological barriers between its speakers or that it is turgid, humourless and bereft of emotion. These are perhaps characteristics of some varieties of formal language, such as the language of law (or rather of social situations in which such registers occur), but this is also the case in English and other languages. It is certainly not characteristic of the everyday language of Indonesians.

Linked to the view of H Indonesian as an austere and forbidding language is a feeling of alienation from it among many people. Ferguson does not refer to this phenomenon, although Valdman (1988:71) recognizes it in Haiti, among Creole speakers who do not control French. In Indonesia the belief of many less well-educated people that the formal language is something which will forever be beyond their ability leads to a significant degree of resentment and rejection. As a consequence there is among a sizeable proportion of the community little emotional attachment to it or motivation to improve competence. One Indonesian linguistics scholar, Harimurti Kridalaksana, has placed the blame for the situation on the attitude in the education system that standard formal language is the only variety deserving positive evaluation. He has stated that there should be a more positive view of non-standard forms of the language, calling for the claim that formal language is applicable in all situations to be done away with. Instead it

should be acknowledged that it has limited functions, being complementary to colloquial speech. Unless this occurs, the common attitude among school students that it is something impractical which has to be learnt by heart just to get a certificate, will not change.⁹

Description

In all the defining languages there is a strong tradition of grammatical study of the H form of the language. There are grammars, dictionaries, treatises on pronunciation, style, and so on. [...] By contrast, descriptive and normative studies of the L form are either non-existent or relatively recent and slight in quantity. (Ferguson 1959:331-2.)

Formal Indonesian has been the subject of considerable study and there have been numerous published descriptions of aspects of the grammar. The Language Centre (Pusat Bahasa), the national authority for propagation and codification of the national language (formerly Pusat Pembinaan dan Pengembangan Bahasa, National Centre for Language Cultivation and Development), produces dictionaries and grammars, sponsors research into the formal language, advises schools and conducts campaigns for the use of good language in the press, in television and radio programs, and so on.¹⁰

On the other hand, no study of L has been undertaken by Indonesians and the Language Centre has never shown interest in the cultivation of this variety (although it does sponsor research on regional languages, including regional varieties of Malay). The official grammar of Indonesian, *Tata bahasa baku bahasa Indonesia* 'Standard grammar of Indonesian' (Alwi et al. 1998) (hereafter *TBB*), includes no description of any aspects of informal language.

There has also been little research on L forms of the language by non-Indonesian linguists, although in recent years there has emerged an interest in this variety, with a number of published studies.¹¹ Ferguson (1959:332) comments that studies of the L form are often carried out first or chiefly by scholars outside the speech community and are written in other languages. In the case of Indonesian it can probably be said that to date they have been carried out only by scholars outside the speech community.¹²

⁹ Cited by Steinhauer (1980:368).

¹⁰ Work of the Pusat Bahasa and earlier language-planning authorities is discussed by Anwar (1980), Moeliono (1994b) and Dardjowidjojo (1998).

¹¹ Publications include studies of aspects of informal speech by Wouk (1998, 1999), Englebreton (2003), Ewing (forthcoming), and a study of Riau Indonesian by Gil (1994).

¹² Kaswanti Purwo (1997:61) has called for a commencement to study of informal Indonesian by Indonesian linguists.

Standardization

There is an established norm [in H] for pronunciation, grammar, and vocabulary which allows variation only within certain limits (Ferguson 1959:332).

Formal Indonesian is a relatively recent phenomenon and there is considerably more variation within it than in the standard forms of Western languages, with authorities sometimes giving conflicting advice on what forms are acceptable, in syntax, morphology and lexicon. Nevertheless, while differences may be conspicuous, in fact overall a remarkable degree of standardization has been achieved.

Compared with H there is a great deal of regional variation among L forms of the language. They have never been part of the language planning and standardization process nor of education. Ferguson (1959:327) writes: 'there is in no case a generally accepted orthography for L'. This again is true of L Indonesian, which is never taught and until recently rarely written; the fact that L varieties have been largely unwritten contributes to lack of standardization.

Despite this, there are signs of a standardized L variety emerging. Ferguson (1959:332) observes that:

In the case of relatively small speech communities with a single important centre of communication a kind of standard L may arise which speakers of other dialects imitate and which tends to spread like any standard variety except that it remains limited to the functions for which L is appropriate.

While the Indonesian speech community is not small and is spread through many islands over a vast area it does have one important centre and Ferguson's observation applies. Jakarta is not only the major urban population centre in Indonesia but colloquial Jakartan Indonesian is having an increasingly great influence on varieties throughout the country. Colloquial Indonesian varieties in other centres are more localized and none has anything other than a marginal influence outside its local area. Further, a homogeneous form of L Indonesian appears to have developed among educated speakers in Jakarta, particularly among those born in the city (Mahdi 1981). It is understood throughout Indonesia due to use in films and television serials and variety shows and through visits to Jakarta by people from other places. Oetomo (1990:71) gives some of the reasons for the influence of Jakartan Indonesian:

The role of Jakarta as the capital city where the most powerful, the most wealthy and the most attractive people are thought to live has been important in popularising the language. In addition, the fact that Jakarta is the centre for the mass media (television, film, publishing) has contributed greatly to popularising the lifestyle and values of prestigious social groups – values which are conveyed in Bahasa

Indonesia, often the Jakarta dialect. Witness how young people from the regions who stay only briefly in Jakarta nevertheless immediately strive to adapt their Bahasa Indonesia – changes which they carry with them when they return home.

Thus, although this is an L variety in a diglossic situation, it has gained considerable prestige. Anwar (1980:154) comments:

The use of Indonesian by educated people living in big towns such as Jakarta, Bandung, Medan, etc. has brought about the establishment of modern colloquial prestige Indonesian. The 'standard' variety of this colloquial Indonesian is undoubtedly connected with the speech prevalent in the capital.

A standardized L orthography may also be developing as colloquial Jakartan Indonesian becomes increasingly commonly used in writing, such as in magazines for urban youths.

Grammar

One of the most striking differences between H and L in the defining languages is in the grammatical structure [...]. It is certainly safe to say that in diglossia there are always extensive differences between the grammatical structures of H and L. (Ferguson 1959:333.)

The instances Ferguson gives are bound to his defining languages, such as the occurrence of cases, or a greater number of them, in H, and do not apply to Indonesian. Nevertheless, in general terms his comments are true of Indonesian. He states (Ferguson 1959:334) that in at least three of his four defining languages the grammatical structure of the L variety is simpler than that of its corresponding H, and there are areas in which H Indonesian displays greater complexity than L.

One difference is the greater level of semantic differentiation in H and the far more highly context-bound nature of the informal variety. Three examples are given here.

First, in H there is an inclusive-exclusive distinction in first person plural pronouns: *kita* 'we inclusive' and *kami* 'we exclusive'. In L this distinction does not occur, with *kita* being the general word for 'we', whether inclusive or exclusive. Although this has been noted by others for informal language it is usually not commented on in grammars. For instance, the official Indonesian grammar, *TBB*, contains no suggestion that the distinction is not always strictly maintained.¹³

¹³ *Kami* occurred a few times in recordings of natural conversations made for this study. However, when it did occur it, like *kita*, was used for both exclusive and inclusive first person plural (discussed in Sneddon 2002:11-2).

Second, transitive verbs in H have more complex morphology, occurring with applicative suffixes *-kan* or *-i* or no suffix. Some bases occur with all three possibilities while others occur with only one or two of the possibilities (no suffix, *-kan*, *-i*). The functions of the two suffixes are distinct; there are a few exceptions but generally if they can both occur with the same verb base they signal different semantic relations of the object to the verb.¹⁴ Thus, from the base *tulis* are derived transitive verbs *menulis* 'write (+ patient)', *menuliskan* 'write for (+ beneficiary)', *menulisi* 'write on (+ location)'. In L *-kan* and *-i* do not occur, their functions being merged in the suffix *-in*.

The following examples are taken from recordings made in Jakarta (further discussed below).

- (1) *Dan materi yang diajarin kan itu-itu aja.*
 and matter which taught you-see that-that just
 And the material taught, you see, is always the same old thing. (*diajarin* 'is taught' corresponds to H *diajarkan*.)
- (2) *Kita baik sama orang, gitu aja. Terus kita juga nggak bohongin orang.*
 we good to people thus just further we also not lie-to people
 We are kind to people, you know. What's more, we don't lie to people. (*bohongin* 'lie to, deceive' corresponds to H *membohongi*.)

This difference is used as one of the examples in the discussion of the continuum between the H and L extremes and is further discussed below.

Third, H has a considerable stock of prepositions, allowing a clear distinction of semantic relationships between verbs and oblique nouns in isolated sentences. Corresponding to some of them in L is a single preposition, *sama*. Expressing a variety of relationships *sama* exhibits semantic under-differentiation compared with the semantic precision of prepositions in H. *Sama* can also function as a coordinator.

Sama links an intransitive verb to a following complement. H has a number of prepositions, sometimes in free variation, although which ones can occur depends on the particular verb. The word-by-word translations gloss L constructions, with *sama* being labelled 'prep' (for 'preposition').

- (3) H: *Saya marah dengan/pada dia.*
 L: *Gua marah sama dia.*
 I angry prep him
 I'm angry at him.

¹⁴ The functions of *-kan* and *-i* are described in Sneddon (1996:69-98).

- (4) H: *Saya sudah kenal dengan/akan anak itu.*
 L: *Gua udah kenal sama itu anak.*
 I already know prep that child
 I know that kid.

Sama corresponds to *kepada* 'to' following a transitive verb to mark recipient:

- (5) H: *Saya menceritakan itu kepada nenek saya.*
 L: *Gua ceritain itu sama nenek gua.*
 I relate that prep mother my
 I told that to my grandmother.

Sama corresponds to *H dari* 'from':

- (6) H: *Si Astrid suka meminjam uang dari ayahnya.*
 L: *Si Astrid suka minjem duit sama bokapnya.¹⁵*
 Astrid like borrow money prep father-her
 Astrid is always borrowing money from her father.

Sama corresponds to *H dengan* 'with', *bersama dengan* 'together with', and *dan* 'and' where they indicate accompaniment or reciprocal relationship.

- (7) H: *Saya dan ayah saya tidak dekat.*
 L: *Gua sama bokap gua nggak dekat.*
 I prep father my not close
 My father and I aren't close.
- (8) H: *Kamu berjalan dengan saya, bukan?*
 L: *Elu jalan sama gua, kan?*
 You walk prep me tag
 You'll walk with me, won't you?

Sama corresponds to *H oleh* 'by' to indicate the agent following a passive verb:

- (9) H: *Saya tidak diterima oleh orangtuanya dan dia juga ditolak oleh keluarga saya.*
 L: *Gua enggak diterima sama orangtuanya dan dia pun, dia juga ditolak sama keluarga gua.*
 I not accepted prep parents-his and he also he
 also rejected prep family my
 I was not accepted by his parents and he too was rejected by my family.

¹⁵ The word *bokap* 'father', which is used frequently by younger speakers in the recordings, along with *nyokap* 'mother', is from 'Prokem', the argot common among Jakarta's youth (see Chambert-Loir 1984).

Since *sama* can identify both agent and recipient there occur many instances where, out of context, constructions with *sama* are ambiguous as to these two functions. In (10) *sama* marks the agent, although in other contexts it could mark recipient:

- (10) L: *Nah, itu dibilang sama temen gua, apa, 'Elu goblok banget'.*
 well that said prep friend my what you stupid very
 Well, my friend said, 'You're really stupid'.

Without the context being known *sama* in (10) could mean 'to' rather than 'by', *dibilang sama temen gua* being interpretable as 'told/said by my friend' or 'told/said to my friend'. The sentence could thus mean: 'Well, my friend was told, "You're really stupid"'. In H the distinction is maintained by the use of either *kepada* 'to' or *oleh* 'by'.

With some verbs ambiguity between the meanings 'by' and 'with' is possible:

- (11) L: *anak kecil umur dua atau tiga tahun, eh itu dibuang*
 child little age two or three year, eh that thrown-out
sama orangtuanya ...
 prep parent-its ...
 ...a young child of two or three years, eh ... is thrown out by its parents...

In (11) *sama* marks the agent, made clear by the context in which the sentence occurs. However, in another context *dibuang sama orangtuanya* could mean 'thrown out together with its parents'. Again, the distinction is explicit in H, either *oleh* 'by' or (*bersama*) *dengan* '(together) with' being used.

The above examples show that context is important for inferring the precise function of *sama* in informal Indonesian utterances. On the other hand, the prepositions *oleh* 'by', *kepada* 'to', *dari* 'from' and *dengan* 'with' in H make the speaker's intention clear even out of context.

Lexicon

A striking feature of diglossia is the existence of many paired items, one H one L, referring to fairly common concepts frequently used in both H and L, where the range of meanings of the two items is roughly the same, and the use of one or the other immediately stamps the utterance or written sequence as H or L (Ferguson 1959:334).

Many examples of such paired items occur in Indonesian, illustrated by the following short list, given in the manner Ferguson provided for his defining languages:

H		L (Jakarta)
<i>tidak</i>	no, not	<i>nggak</i>
<i>besar</i>	big	<i>gede</i>
<i>sekali</i>	very	<i>banget</i>
<i>saja</i>	just	<i>aja</i>
<i>bagaimana</i>	how	<i>gimana</i>
<i>seperti</i>	like	<i>kayak</i>
<i>ingin</i>	want	<i>pengen</i>
<i>membuat</i>	make	<i>bikin</i>
<i>mengatakan</i>	say	<i>bilang</i>
<i>memberi</i>	give	<i>kasi</i>

The continuum in Indonesian

In the original conception of diglossia the distinction between H and L is so clearcut that the community assigns different names to them, such as Katharevousa and Demotiki Greek, Schwyzertütsch and Standard German in Switzerland.

The H and L varieties in Indonesian are not so sharply differentiated that different names are given to them, as occurs in Ferguson's four defining languages. Nevertheless, Indonesians certainly recognize the differences between the H and L; one Indonesian linguist has remarked that they differ 'to such an extent that an attempt to account for both by a single unified approach would be extremely complex, if not impossible' (Halim 1981:7). In Indonesian the H code is referred to by such terms as *bahasa resmi* 'official language', *bahasa formal* 'formal language', *bahasa baku* 'standard language', and *bahasa halus* 'refined language', while the L variety is variously referred to as *bahasa sehari-hari* 'daily language', *bahasa non-baku* 'non-standard language', *bahasa percakapan* 'conversational language', *percakapan santai* 'relaxed conversation', and so on. These are impressionistic labels only, used unsystematically as people attempt to distinguish the types when the need arises; there are no official names to differentiate them.

One reason for this failure to distinguish H and L Indonesian by different names is that there is no clear dividing line between them. While the H and L varieties are associated with the most formal and informal social situations there are intermediate forms, associated with semi-formal situations. Speakers may make their colloquial speech more formal by incorporating some features of formal language. Likewise formal language is not homogeneous in all situations in which it is used.¹⁶

¹⁶ In regions where a post-creole variety of Malay is the L variety, such as Manado Malay

Ferguson (1959:328) implied that the difference in function between L and H is quite strict, with a clear-cut boundary between them, stating, 'In one set of situations only H is appropriate and in another only L, with the two sets overlapping only very slightly'. This separation of functions is taken by some scholars as an essential characteristic of diglossia. Wardhaugh (1998:88) writes: 'A key defining characteristic of diglossia is that the two varieties are kept quite apart in their functions'.

However, Ferguson himself does not view this separation as essential. In a later paper he wrote about the four defining languages:

In every one of the four cases the researcher can document a continuum of forms between the H and L varieties, and some linguists have denied the validity of my identification of the diglossia situation on this account. I recognized the existence of intermediate forms and mentioned them briefly in the article, but I felt then and still feel that in the diglossia case the analyst finds two poles in terms of which the intermediate varieties can be described; there is no third pole. (Ferguson 1991: 226.)

Kaye (2001:118, 127) refers to a continuum between Modern Standard Arabic and colloquial varieties but says very little on the nature of this continuum other than that there is 'continuous mixing' between the two. Of the situation in Arabic Eid (1982:55) writes: 'In some situations, which Ferguson labels 'semi-formal', both varieties are appropriate; and under these circumstances, speakers mix the two varieties using elements from both in a highly variable way'. Along with Ferguson, neither Kaye nor Eid recognizes strict functional separation of the two codes as a defining characteristic of diglossia. In his consideration of the nature of diglossia Hudson does not mention a continuum. Rather, he states, 'both the original and the extended versions of diglossia are characterized by pervasive and rigid functional compartmentalization of the diglossic speech varieties, at least to the extent that the elevated variety is never used by any member of the community for the purposes of within-group informal conversation' (Hudson 1991:13). He goes on to state of this rigid compartmentalization: 'This, perhaps, is less a defining characteristic of diglossia than it is a necessary prerequisite for long-term maintenance of significantly discrepant codes within a single speech community' (Hudson 1991:13-4). Thus Hudson also does not recognize distinct separation of functions as a defining characteristic, only that it is a prerequisite for long-term

in North Sulawesi and Ambon Malay in South Moluccas, there is a much more rigid compartmentalization of functions between L and H, although it is debatable whether such varieties of Malay can properly be assigned to the same language as formal Indonesian. Grimes (1991:145-6) regards Ambon Malay as mutually unintelligible with Standard Indonesian and a separate language from it.

maintenance of diglossia. Consequently the lack of a strict boundary between H and L in Indonesian does not appear to disqualify it from being regarded as existing in a diglossic state. The question of long-term maintenance is considered below.

The discussion above of H and L registers in Indonesian refers to extreme or 'pure' varieties. At the 'low' extreme only features of the L variety will occur. As the social situation shifts from very informal elements of the high variant will begin to appear and will occur with increasing frequency as the situation becomes more formal, until finally only H features will occur, as in impersonal situations like news broadcasts, newspaper editorials, legal documents and formal speeches. There is thus a continuum between most informal and most formal registers, in both grammar and lexicon.¹⁷

To demonstrate that in Indonesian a continuum exists between the L and H poles a limited number of variables are looked at in the speech of educated Jakartans in a number of different social situations. Statistical information is given to quantify the variation, at least as far as these particular variables are concerned.¹⁸ The variables (corresponding to Ferguson's paired items) each comprise variants typically associated with L and variants typically associated with H, in Jakartan Indonesian. The official Indonesian grammar, *TBB*, published by the Language Centre, is only concerned with formal Indonesian and recognizes none of the informal variants discussed here. The official dictionary, *Kamus besar bahasa Indonesia* 'Comprehensive Indonesian dictionary' (Tim Penyusun Kamus 1991) (hereafter *KB*), lists some of the informal variants as 'conversational' but does not list them all.

While there is a continuum, a series of discrete stages between extremes cannot be identified. Averaged over groups of speakers the shift from informal to formal variants occurs at significantly different rates from one variable to another as the social situation becomes more formal. Moreover, some individuals are much more inclined than others to use H variants in very similar social situations. Thus, not only does the L variant give way to the H variant at different stages of formality for different variables but for every variable the boundary is at different levels of formality for different individuals.

The variables are examined in three social situations involving unplanned spoken language: conversations between friends, interviews and meetings.

¹⁷ While 'pure' H exists, 'pure' L may be only an ideal, at least in the speech of educated people. However, it is likely that in the speech of people who have little or no facility in formal Indonesian H features are very infrequent. In a recording of 5,000 words of conversation among less well-educated speakers none of the H variants discussed here appears, except for one instance of *sudah* 'already' (as against 59 tokens of L *udah*).

¹⁸ This and Sneddon (2002) are the only quantitative studies of the continuum of which I am aware. Ewing (forthcoming) also discusses aspects of colloquial and formal Indonesian.

The conversations in general contain the most informal language as they occur in relaxed social situations. The interviews, where one person is asked questions and encouraged to talk about himself or herself, tend to exhibit more characteristics of formal language, being situations in which a certain amount of self-conscious linguistic behaviour might be expected. The meetings, between groups of academics, are the most formal situations. Nevertheless, the language in the meetings, in which there were no observers apart from the recorder, is less formal than would occur in public meetings and the context can perhaps best be described as semi-formal. None of the situations is formal; in strictly formal situations only the H variants occur.

The study is restricted to the speech of educated people, mostly university students or graduates, living in Jakarta, who were either born there or moved there at an early age. This restriction helps confine the study to examination of a controllable amount of variation, as the group can be regarded as homogeneous to a considerable extent in their linguistic habits. It also confines the study to the speech of the group whose colloquial language is becoming the model for informal usage throughout Indonesia.

All people recorded speak formal Indonesian with considerable proficiency, high proficiency in most cases. In informal situations they speak colloquial Jakartan Indonesian. They are not speakers of Jakarta (Betawi) Malay. The data are taken from texts totalling approximately 143,000 words, part of a larger corpus recorded in Jakarta in 2000-2001.¹⁹ (The L structures in examples (1) to (11) above are taken from these recordings.) Numbers in square brackets below identify individual texts.

Six linguistic variables were chosen for the study. One is a morphological variable in which there is a single L variant corresponding to two H variants, two are phonological variables and three are lexical variables. The number of times an item occurs is not in itself important; the larger the corpus the more times a particular form is likely to occur. What is of interest is the number of times an item occurs relative to the frequency of the other item within the same variable; for this reason percentages of occurrences are given.

In considering the H-L continuum the speech of a single individual, while interesting, can give a distorted view of the habits of the community in general. For instance, interviewees [016], [017] and [019] consistently used a greater percentage of H variants than other interviewees in apparently very similar circumstances. Such individual differences tend to cancel each other out over large numbers of texts of different speakers, giving a clearer view of the general tendencies within the community. (In the conversations

¹⁹ All recordings used in this study were made in Jakarta by Yulianti Kartika Sari and Tessa Yuditha, who also transcribed the conversations. I express my thanks for their meticulous work. Details of the texts are given in the Appendix.

and meetings individual variation is already to a certain extent lost in the averages.) Seven or eight texts of conversations and interviews were chosen, largely at random (and not for reasons connected with the variables being considered), although only two recordings of meetings were available at the time of this study.

The ages of participants in the conversations and interviews range from early 20s to late 40s and in the meetings from late 20s to early 60s. Details are given in the Appendix. While older speakers may be inclined to use formal variants more than younger people in a particular social situation, this is not true of the variables chosen for this study. Variation between age groups is probably not as great as between different individuals in the same age group. Conversation [036], a later recording of a conversation between four people aged from early 40s to early 50s, was added to conversations to see if it affected the results. The only other conversation among older people previously available was [002] (between 5 people aged between 39 and 49), a much shorter text than the other conversations (among participants in their 20s). In [002] the percentage of H variants is consistently higher than in any of the other conversations. However, in [036] this is not the case; percentages for use of the L variant are close to the average for all other conversations and are actually higher in four of the five cases (there being no instances of the *gede-besar* variable in [036]).

For each variable considered the number of tokens of each variant in each text is shown. The total number of tokens of each variant for each category (conversations, interviews, meetings) is then given, together with percentages.

Words with and without initial /s/

A limited number of words optionally commence with /s/, the variant without /s/ being favoured in L. These include *saja* ~ *aja* 'just', as in *satu aja/saja* 'just one', and *sudah* ~ *udah*, a temporal marker, glossed 'already' here. The data show that *aja* and *udah* are far more frequent than *saja* and *sudah*, which are the only variants recognized by KB and TBB. Conversations and interviews show the same tendency: overwhelming use of *aja* and very limited, sometimes zero, use of *saja*. Only in [016], an interview with a 25 year old female teacher, does *saja* occur an appreciable number of times. Somewhat surprising is the frequent use of *aja* in the meetings. Of all the texts, only in [022] is *saja* the more frequent variant, and even there *aja* is far from infrequent. Likewise, *udah* occurs far more frequently than *sudah* in conversations and interviews. In conversations *sudah* occurs relatively frequently only in [002]. In meetings, however, *sudah* occurs more frequently than *udah*.

Table 1a. Frequency of forms with and without initial /s/

	text	<i>aja</i>	<i>saja</i>	<i>udah</i>	<i>sudah</i>
conversations	002	20	3	17	9
	003	57	1	115	2
	004	90	1	70	4
	006	92	0	109	5
	007	67	0	141	3
	010	74	1	111	1
	011	50	0	134	2
	036	55	0	40	5
interviews	013	34	1	98	1
	015	59	2	78	2
	016	28	18	45	20
	017	23	3	28	10
	019	58	2	143	46
	021	28	0	110	2
	023	47	0	67	2
meetings	022	18	32	16	54
	031	27	5	17	20

Percentages of variants averaged over the texts are as follows:

Occurrences of *aja* ~ *saja*:

Table 1b. Totals and percentages for forms with and without initial /s/ in *aja* ~ *saja*

Conversations	Interviews	Meetings
Total occurrences: 511	Total occurrences: 303	Total occurrences: 82
<i>aja</i> 505; <i>saja</i> 6	<i>aja</i> 277; <i>saja</i> 26	<i>aja</i> 45; <i>saja</i> 37
Percentages:	Percentages:	Percentages:
<i>aja</i> 98.8	<i>aja</i> 91.4	<i>aja</i> 54.9
<i>saja</i> 1.2	<i>saja</i> 8.6	<i>saja</i> 45.1

Occurrences of *udah* ~ *sudah*:

Table 1c. Totals and percentages for forms with and without initial /s/ in *udah* ~ *sudah*

Conversations	Interviews	Meetings
Total occurrences: 768	Total occurrences: 652	Total occurrences: 107
<i>udah</i> 737; <i>sudah</i> 31	<i>udah</i> 569; <i>sudah</i> 83	<i>udah</i> 33; <i>sudah</i> 74
Percentages:	Percentages:	Percentages:
<i>udah</i> 96.0	<i>udah</i> 87.3	<i>udah</i> 30.8
<i>sudah</i> 4.0	<i>sudah</i> 12.7	<i>sudah</i> 69.2

Gede and besar 'big'

The common word for 'big' in H is *besar*. The word *gede* does not occur in H but is the more common word in the conversation of most people. The frequency of *besar* is greater in the meetings:

Table 2a. words for 'big'

	text	<i>gede</i>	<i>besar</i>
conversations	002	5	3
	003	8	3
	006	9	1
	007	10	3
	010	14	0
	011	11	0
	036	0	0
interviews	013	18	0
	015	2	1
	016	1	8
	017	3	11
	019	9	14
	021	7	0
	023	7	0
meetings	022	0	4
	031	2	19

Table 2b. Totals and percentages for variants meaning 'big'

Conversations		Interviews		Meetings	
Total occurrences: 67		Total occurrences: 81		Total occurrences: 25	
<i>gede</i> 57; <i>besar</i> 10		<i>gede</i> 47; <i>besar</i> 34		<i>gede</i> 2; <i>besar</i> 23	
Percentages:		Percentages:		Percentages:	
<i>gede</i>	85.1	<i>gede</i>	58.0	<i>gede</i>	8.0
<i>besar</i>	14.9	<i>besar</i>	42.0	<i>besar</i>	92.0

Kayak and seperti 'be like'

These variants mean 'be like', as in L *kayak gua* ~ H *seperti saya* 'like me, resembling me'. The percentages for L *kayak* were lower than for other L variants, with H *seperti* outnumbering it in one conversation and four interviews, as well as in the meetings.

Table 3a. words for 'like'

	text	kayak	seperti
conversations	002	6	9
	003	22	8
	004	40	6
	006	72	22
	007	69	3
	010	65	0
	011	55	1
	036	6	0
interviews	013	15	5
	015	22	32
	016	24	25
	017	14	40
	019	11	16
	021	21	0
	023	66	9
meetings	022	16	33
	031	9	27

Table 3b. Totals and percentages for words meaning 'like'

Conversations	Interviews	Meetings
Total occurrences: 384	Total occurrences: 300	Total occurrences: 85
<i>kayak</i> 335; <i>seperti</i> 49	<i>kayak</i> 173; <i>seperti</i> 127	<i>kayak</i> 25; <i>seperti</i> 60
Percentages:	Percentages:	Percentages:
<i>kayak</i> 87.2	<i>kayak</i> 57.7	<i>kayak</i> 29.4
<i>seperti</i> 12.8	<i>seperti</i> 42.3	<i>seperti</i> 70.6

Negatives

The negative with non-nominal predicates in H is *tidak*, as in *tidak baik* 'not good', *tidak pergi* 'not going'. It also occurs in the meaning 'no' in answer to a question. *Tidak* usually has extremely low frequency of occurrence in conversations. In interviews it is only common in [016], [017] and [019].

Far more frequent in L is *nggak*. This is often pronounced and sometimes spelled *enggak*, although the difference is not indicated here. Infrequently a reduced variant *gak* also occurs, again not distinguished here from *nggak*.²⁰

²⁰ While there is an official orthography for Standard Indonesian (formal language) there are no established conventions for informal varieties. Other spellings for this word include *ngga*'

In only one text, the meeting [022], does *tidak* occur more frequently than *nggak*, and even there *nggak* is far from rare. Despite its very low frequency in informal speech *tidak* is the only negative mentioned as occurring with non-nominals in *TBB*, which does not recognize *nggak*. It is given in *KB* (spelled *enggak*) as colloquial.

Frequency of the forms is shown in the following table:

Table 4a. forms meaning ‘not; no’

	text	<i>nggak</i>	<i>tidak</i>
conversations	002	39	3
	003	206	8
	004	199	3
	006	287	5
	007	265	5
	010	275	3
	011	245	6
	036	90	1
interviews	013	173	4
	015	171	5
	016	127	46
	017	106	71
	019	217	43
	021	203	2
	023	307	2
meetings	022	69	85
	031	91	60

Table 4b. Totals and percentages for negatives

Conversations		Interviews		Meetings	
Total occurrences: 1640		Total occurrences: 1477		Total occurrences: 305	
<i>nggak</i> 1606; <i>tidak</i> 34		<i>nggak</i> 1304; <i>tidak</i> 173		<i>nggak</i> 160; <i>tidak</i> 145	
Percentages:		Percentages:		Percentages:	
<i>nggak</i>	97.9	<i>nggak</i>	88.3	<i>nggak</i>	52.5
<i>tidak</i>	2.1	<i>tidak</i>	11.7	<i>tidak</i>	47.5

and *ngga*. *KB* includes the word, spelling it *enggak*. Two other *L* words, *kagak* and *ndak*, occur in the recordings, although they are infrequent and occur only in the speech of a few people. Their inclusion would marginally raise the percentage of *L* variants in conversations and interviews.

Verbal suffixes: -in vs -kan and -i

The transitive verbal suffixes are discussed above as an example of greater complexity in H.

There is a statistical bias towards L *-in* in the figures for two reasons. First, the suffix *-i* cannot occur after a base ending in /i/; verbs whose base ends in /i/ and which theoretically take suffix *-i*, such as *memberi* (*dia buku*) 'give (him a book)', are not counted as having a suffix. This constraint does not apply to *-in*, which can occur after base-final /i/, as in *buktiin* 'prove' (*bukti* + *-in*). Secondly, *-in* sometimes occurs where in H the verb takes no suffix:

- (12) *Cowok itu ngertiin kita.*
 man that understand us
 That man understands us. (*mengerti* 'understand' occurs in H.)

However, this bias is small and does not affect percentages significantly.

Suffixes *-kan* and *-i* on verbs with prefix *ter-* are not counted, nor are verbs with affixation *ber-...-kan*, because *-in* apparently never replaces the suffixes on these verbs. The suffix *-in* is frequent in all texts except the meeting [022], where four people use it just once each. The two H suffixes combined occur the same number of times as *-in* in one conversation and outnumber it in three interviews, as well as in both meetings.

Frequency of the suffixes is shown in the following table:

Table 5a. verbal suffixes

	text	<i>-in</i>	<i>-kan</i>	<i>-i</i>
conversations	002	15	11	4
	003	81	11	7
	004	97	34	7
	006	104	50	36
	007	59	13	3
	010	85	3	5
	011	91	16	2
	036	26	7	1
interviews	013	85	27	9
	015	71	23	2
	016	67	89	19
	017	31	147	38
	019	73	125	16
	021	51	7	1
	023	88	44	11
meetings	022	4	151	59
	031	39	102	31

Table 5b. Totals and percentages for verbal suffixes

Conversations	Interviews	Meetings
Total occurrences: 768	Total occurrences: 1015	Total occurrences: 386
-in 558;	-in 466;	-in 43;
-kan 145; -i 65	-kan 462; -i 87	-kan 253; -i 90
Percentages:	Percentages:	Percentages:
-in 72.7	-in 45.9	-in 11.1
-kan 18.9	-kan 45.5	-kan 65.5
-i 8.5	-i 8.6	-i 23.3

Averages for informal variants

For every variable the percentage of L variants is on average smaller in interviews than in conversations and smaller in meetings than in interviews. However, the rate of decrease varies; in some cases the decrease between conversations and interviews is large and in other cases small, although for every variable the drop between interviews and meetings is large. The percentage of informal variants in each of the three social situations is as follows:

Table 6. percentages for informal variants

	conversations	interviews	meetings
<i>aja</i>	98.8	91.4	54.9
<i>udah</i>	96.0	87.3	28.2
<i>gede</i>	85.1	58.0	8.0
<i>kayak</i>	87.2	57.7	29.4
<i>nggak</i>	97.9	88.3	52.5
<i>-in</i>	72.7	45.9	11.1

This is shown graphically in Graph 1.

Conclusion

It is shown above that all the characteristics of diglossia listed by Ferguson are applicable to Indonesian, except those relating to diachronic context (literary heritage and stability), which have not been considered.

The only difference is that in Indonesian there is a continuum between the H and L extremes, at least in the speech of educated Jakartans. While there is a continuum a series of discrete stages cannot be identified; rather, as the social situation becomes more formal H elements appear more and more fre-

Graph 1. Percentages for informal variants

quently, though at an unpredictable rate, there being considerable variation in the speech of different individuals.

While some writers have regarded rigid separation of H and L as an essential characteristic of diglossia, Ferguson (1991) acknowledges a continuum in his four defining languages and others have commented on a continuum in Arabic. There thus appears no good reason to exclude Indonesian solely because of the continuum; on the contrary, the concept of diglossia provides an excellent framework for description of the Indonesian sociolinguistic situation.

Hudson (1991:13) states that rigid compartmentalization is perhaps a necessary prerequisite for long-term maintenance of a diglossic situation. While research is needed into the origins of diglossia in Indonesian the present situation has been in existence since before independence, its origins certainly lying in part in the choice of High (Riau) Malay as the language of education in the nineteenth century.²¹ As to how long the present situation will last one can at present only speculate. In some domains where H was once the rule elements of L are becoming more common. Television dramas depicting real life are increasingly inclined to have characters use more natural informal language, as are some publications, particularly those for young readers.

²¹ Grijns (1991b) and Teeuw (1998) are valuable contributions to the study of the origins of modern Standard Indonesian but do not deal with the emergence of diglossia.

This trend has already been commented on by some researchers. Adelaar and Prentice (1996:678) write of colloquial Jakartan Indonesian:

It has become a prestige dialect among the urban elite in Jakarta and elsewhere in Indonesia. It often serves as an informal counterpart to the formal and stilted official Indonesian on which it exerts a strong influence. Jakartan Indonesian is increasingly being used in popular literature, comics, popular songs and performing arts.

Militating against the spread of L is the fact that all education is strictly in H. Nevertheless, this could change. As mentioned, large sections of the community feel alienated from the formal language. If authorities ever conclude that the national interest would be better served by a more positive attitude to colloquial speech, as advocated by Harimurti Kridalaksana, the result could well be a major shift away from H in the education system, which would have profound consequences for the future of the language.

Until the present there has been no study of the gradual intrusion of elements of L Indonesian into situations formerly closely associated with H nor of official attitudes to this development. Study of such phenomena could well give valuable insights into future trends.

Appendix

Conversations and interviews were selected in fairly random fashion from a larger number of texts. At the time of writing only two meetings had been recorded. Recorder/transcribers participated in conversations. They did not participate in meetings and their contribution in interviews has been excluded. Apart from the recorder several people participated in two conversations; no one participated in more than one interview or meeting.

Information on the texts below includes an identifying number, the approximate number of words, whether it was a conversation, interview or meeting, and basic information on the participants. Abbreviations: f = female, m = male, yo = years old.

- 002: 2,200 words – conversation between 5 male university staff (lecturers and administration staff), aged between 39 and 49
- 003: 10,600 words – conversation between 3 people in 20s (2 f, 1 m)
- 004: 8,100 words – conversation between 2 people in 20s (1 m, 1 f)
- 006: 10,600 words – conversation between 4 people in 20s (2 m, 2 f)
- 007: 12,500 words – conversation between 3 people in 20s (3 f)
- 010: 11,800 words – conversation between 4 people in early 20s (3 f, 1 m)
- 011: 8,500 words – conversation between 3 people in 20s (2 f, 1 m)

- 013: 8,000 words - interview with 26 yo female ballet instructor and dancer
- 015: 7,800 words - interview with 34 yo husband and 27 yo wife, boutique owners
- 016: 8,400 words - interview with 25 yo female teacher and travel agency worker
- 017: 7,500 words - interview with 22 yo female student
- 019: 12,400 words - interview with 47 yo female teacher
- 021: 6,100 words - interview with 20 yo female student
- 022: 7,900 words - meeting between 5 academics, 1 female in 20s, 4 males in 50s and 60s
- 023: 8,500 words - interview with 26 yo businessman
- 031: 7,600 words - meeting between 5 academics aged between late 30s and mid 40s (4 f, 1 m)
- 036: 4,400 words - conversation between 4 university staff (lecturers and administration staff) aged from early 40s to early 50s (3 f, 1 m)

Abbreviations

- TBB: *Tata bahasa baku bahasa Indonesia* 'Standard grammar of Indonesian'
- KB: *Kamus besar bahasa Indonesia* 'Comprehensive Indonesian dictionary'

References

- Adelaar, K. Alexander and D.J. Prentice
1996 'Malay, its history, role and spread', in: Stephen A. Wurm, Peter Mühlhäusler and Darrell T. Tryon (eds), *Atlas of languages of intercultural communication in the Pacific, Asia and the Americas*, pp. 673-93. Berlin: Mouton de Gruyter. [Trends in Linguistics, Documentation 13.]
- Alwi, Hasan, Soenjono Dardjowidjojo, Hans Lapoliwa, Anton M. Moeliono
1998 *Tata bahasa baku bahasa Indonesia*. Third edition. Jakarta: Balai Pustaka.
- Anderson, Benedict R.
1966 'The languages of Indonesian politics', *Indonesia* 1:89-116.
- Anwar, Khaidir
1980 *Indonesian; The development and use of a national language*. Yogyakarta: Gadjah Mada University Press.
- Atmosumarto, Sutanto
1995 *Colloquial Indonesian*. London: Routledge. [Colloquial Series.]
- Cartier, Alice
1980 'L'indonésien formel et le jakartanais: une approche socio-linguistique', *Journée d'Études* 3:15-36.

Chambert-Loir, Henri

1984 'Those who speak *Prokem*', *Indonesia* 37:105-17.

Collins, James T.

1998 *Malay, world language; A short history*. Kuala Lumpur: Dewan Bahasa dan Pustaka. [The History of Malay, Monograph Series.]

Dardjowidjojo, Soenjono

1998 'Strategies for a successful national language policy: the Indonesian case', *International Journal of the Sociology of Language* 130:35-47.

Eid, Mushira

1982 'The non-randomness of diglossic variation in Arabic', *Glossa* 16:54-84.

Englebretson, Robert

2003 *Searching for structure; The problem of complementation in colloquial Indonesian conversation*. Amsterdam: Benjamins. [Studies in Discourse and Grammar 13.]

Errington, J. Joseph

1986 'Continuity and change in Indonesian language development', *Journal of Asian Studies* 45:329-53.

Ewing, Michael C.

forthcoming 'Colloquial Indonesian', in: K. Alexander Adelaar and Nikolaus Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, London: Curzon.

Ferguson, C.A.

1959 'Diglossia', *Word* 15:325-40. [Reprinted in 1972 in Pier Paolo Giglioli (ed.), *Language and social context; Selected readings*, pp. 232-51. Harmondsworth: Penguin.]

1991 'Diglossia revisited', *Southwest Journal of Linguistics* 10:214-34.

Gil, David

1994 'The structure of Riau Indonesian', *Nordic Journal of Linguistics* 17:179-200.

Grijns, C.D.

1981 'Jakartan speech and Takdir Alishajbana's plea for the simple Indonesian word-form', in: Nigel Phillips and Khairid Anwar (eds), *Papers on Indonesian languages and literature*, pp. 1-34. London: Indonesian Etymological Project/Paris: Association Archipel. [Cahiers d'Archipel 13.]

1983 'Language use in three neighborhoods of Tanah Abang', in: James T. Collins (ed.), *Studies in Malay dialects Part 1*, pp. 37-51, Jakarta: Universitas Atma Jaya. [NUSA, Linguistic Studies of Indonesian and Other Languages in Indonesia 16.]

1991a *Jakarta Malay; A multidimensional approach to spatial variation*, Vol. 1. Leiden: KITLV Press. [Verhandelingen 149.]

1991b 'Bahasa Indonesia avant la lettre in the 1920s', in: H. Steinhauer (ed.), *Papers in Austronesian linguistics No. 1*, pp. 49-81. Canberra: Australian National University. [Pacific Linguistics, Series A, 81.]

- Grimes, Barabara Dix
 1991 'Exploring the sociolinguistics of Ambonese Malay', in: Ray Harlow (ed.), *VICAL 2; Papers from the Fifth International Conference on Austronesian Linguistics, Parts 1 & 2*. pp. 131-68. Auckland: Linguistic Society of New Zealand.
- Halim, Amran
 1981 *Intonation in relation to syntax in Indonesian*. Canberra: Australian National University. [Pacific Linguistics, Series D, 36.]
- Heryanto, Ariel
 1995 *Language of development and development of language; The case of Indonesian*. Canberra: Australian National University. [Pacific Linguistics, Series D, 86.]
- Hoffman, John
 1979 'A foreign investment; Indies Malay to 1901', *Indonesia* 27:65-92.
- Hudson, Alan
 1991 'Toward the systematic study of diglossia', *Southwest Journal of Linguistics* 10:1-22.
 1992 'Diglossia: a bibliographic review', *Language in Society* 21:611-74.
- Johns, Yohanni
 1996 *Bahasa Indonesia; Book Three; Langkah baru*. Canberra: Wilson.
- Kaswanti Purwo, Bambang
 1997 'Ihwal bahasa tak baku: bahasa yang baik dan tidak benar?', *Atma nan Jaya* 10-3:49-63.
- Kaye, A.S.
 2001 'Diglossia: the state of the art', *International Journal of the Sociology of Language* 152:117-29.
- Lauder, Multamia
 1998 'Indonesia's linguistic diversity', in: John H. McGlynn (ed.), *Indonesian heritage; Language and literature*, pp. 72-3. Singapore: Archipelago Press/Jakarta: Buku Antar Bangsa.
- Mahdi, Waruno
 1981 'Some problems in the phonology of metropolitan Indonesian', *Bijdragen tot de Taal-, Land- en Volkenkunde* 137:399-418.
- Maier, H.M.J.
 1993 'From heteroglossia to polyglossia; The creation of Malay and Dutch in the Indies', *Indonesia* 56:37-65.
- Moeliono, Anton M.
 1994a 'Contact-induced language change in present-day Indonesian', in: Tom Dutton and Darrell T. Tryon (eds), *Language contact and change in the Austronesian world*, pp. 377-88. Berlin: Mouton de Gruyter. [Trends in Linguistics 77, Studies and Monographs.]
 1994b 'Indonesian language development and cultivation', in: Abdullah Hassan (ed.), *Language planning in Southeast Asia*, pp. 195-213. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Oetomo, Dede
 1990 'The bahasa Indonesia of the middle class', *Prisma* 50:68-79.

- Peacock, James L.
1973 *Indonesia; An anthropological perspective*. Pacific Palisades, CA: Good-year. [Regional Anthropology Series.]
- Prentice, D.J.
1978 'The best chosen language', *Hemisphere* 22-3:18-23, 22-4:28-33.
- Sarumpaet, Jan Pieter
1980 *Modern usage in bahasa Indonesia*. Melbourne: Pitman.
- Sneddon, James Neil
1996 *Indonesian reference grammar*. Sydney: Allen and Unwin. [Published outside Australia as *Indonesian; A comprehensive grammar*, London: Routledge.]
2001 'Teaching informal Indonesian; Some factors for consideration', *Australian Review of Applied Linguistics* 24-2:81-95.
2002 'Variation in informal Jakartan Indonesian; A quantitative study'. Paper, Ninth International Conference on Austronesian Linguistics, Canberra.
2003 *The Indonesian language; Its history and role in modern society*. Sydney: UNSW Press.
- Steinhauer, Hein
1980 'On the history of Indonesian', in: *Studies in Slavic and general linguistics* Vol. 1, pp. 349-75. Amsterdam: Rodopi.
- Tanner, N.
1967 'Speech and society among the Indonesian elite; A case study of a multilingual community', *Anthropological Linguistics* 9-3:15-40.
- Teeuw, A.
1998 *De ontwikkeling van een woordenschat; Het Indonesisch 1945-1995*. Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen. [Mededelingen van de Afdeling Letterkunde (New Series) 61-5.]
- Tim Penyusun Kamus
1993 *Kamus besar bahasa Indonesia*. Second edition. Jakarta: Balai Pustaka. [First edition 1988.]
- Valdman, A.
1988 'Diglossia and language conflict in Haiti', *International Journal of the Sociology of Language* 71:67-80.
- Wallace, Stephen
1979 'The linguistic modernization of Jakarta', in: Gloria Davis (ed.), *What is modern Indonesian culture? Papers presented to the Conference on Indonesian Studies, July 29-August 1, 1976, Indonesian Studies Summer Institute, Madison, Wisconsin*, pp. 69-88. [Athens, OH]: Southeast Asia Program, Center for International Studies, Ohio University. [Papers in International Studies, Southeast Asia Series 52.]
- Wardhaugh, Ronald
1980 *An introduction to sociolinguistics*. Third edition. Oxford: Blackwell. [First edition 1986.]
- Wolff, John U., D. Oetomo and D. Fietkiewicz
1992 *Beginning Indonesian through self-instruction*. Third revised edition. Ithaca, NY: Southeast Asia Program, Cornell University. [First edition 1971.]

Wouk, Fay

1998

'Solidarity in Indonesian conversation; The discourse marker *kan*',
Multilingua 17:379-406.

1999

'Gender and the use of pragmatic particles in Indonesian', *Journal of Sociolinguistics* 3:194-219.

2