The lexicon of a language is the sum of the vocabularies of its speakers. If the language is alive, its lexicon will consist of a core known to all its adult (and mentally unchallenged) speakers. More peripheral are words that are rarely used, or known only to some speakers. These may be words used on special occasions only, words referring to infrequent phenomena, or words that are stylistically marked or have become archaic. Perhaps even more peripheral are words used or known only by speakers of a nonstandard dialect, for instance, members of subcultures (e.g., religions), specialists, in-crowds, elites, and outcasts. Peripheral in a semantic sense are words that are meant to insult and offend.

Under the heading Macrostructure, below, I first discuss the choices the authors of this dictionary made in selecting their lexical units: lemmas, sublemmas, and subsublemmas; and monomorphemic words, derivations, and compounds. In the subsequent section, called Microstructure, I deal with the content and structure of the dictionary articles. The review closes with a section titled Other Matters and Final Appreciation.

Macrostructure

A lexicographer has to decide which and how many of those “peripheral” words should be included in her or his dictionary, apart from the “core” vocabulary. Comprehensiveness is an ideal but, in practice, it’s a relative notion: lexicographers rarely have either the time or funds (or the market) on their sides. The long-awaited Indonesian-English dictionary by Stevens and Schmidgall-Tellings (hereafter referred to as CIED) is in many respects exceptional. For one thing, the adjective in the title is no indication of chutzpah: none of the existing Indonesian dictionaries, whether bilingual or monolingual, equal CIED in comprehensiveness.

One of the special problems that the Indonesian-X lexicographer has to solve is related to the definition of the source language, Indonesian. In the last decades of Dutch colonial rule, a standardized form of Malay (the predecessor of Indonesian) had begun to develop, but its status was far from official, let alone national, and its relation to other varieties of Malay was one of “pure” versus “corrupted” or, at best, “high” versus “low.” The corrupted (or low) varieties could be neglected or treated as curiosa. In colonial times, therefore, Malay could be defined on the basis of a limited corpus of Malay data (texts mainly) that were considered sufficiently “pure.” Since 1945, however, the autonym of the language, bahasa Indonesia (“the language of Indonesia”), has become official, with the consequence that its grammar had to be standardized, and its lexicon “developed.” Its new status also changed the status of the regional and social varieties of Malay: these could now be considered dialects and registers of the standard language. But which ones, and to what extent? Indonesia has 500 or more regional languages, many of which are only distantly or not at all genetically related to Indonesian. For that matter, regional Malay vernaculars may be considered languages.
of their own (instead of dialects), not to be accounted for in an Indonesian-X dictionary.

For practical reasons, most Indonesian dictionaries, whether monolingual or bilingual, confine themselves to the official standard language. *CIED* takes a more liberal approach. It fairly systematically includes Jakarta Malay words as well as morphological forms that differ from their standard equivalents. Typical examples are forms such as *nabrak*, *ketabrak*, and *nabrakin* (marked somewhat inconsistently as *coq* [colloquial], *J* [Jakarta], and *J coq*, respectively), which are given with the standard Indonesian sublemmas *tertablak*, *menablak*, and *menablakkan*. Frequent and typical words from other varieties of Malay are also included as lemmas of their own, for instance, words from Banjarese, Minahassa Malay, Ambon (or “East Indonesian” Malay), Malaysian Malay, and Singapore Malay. Many lemmas are marked as originating from a non-Malay regional language, such as Javanese, Sundanese, and Batak. Some of these are commonly known. In other cases, the etymological label presumably indicates that the word represented by the lemma has appeared in one or more Indonesian texts, without necessarily having become common and known to Indonesians with another home language. This also holds for borrowings from foreign languages—in particular, words from Arabic and from pre-independence Dutch. All lemmas that are known to be borrowings are provided with an etymological label, even words belonging to the basic vocabulary (such as *kepala*, meaning “head,” which is dutifully marked as being of Sanskrit origin).

In line with its liberal approach toward the definition of Indonesian, *CIED* contains many examples taken from a large variety of sources. In the introduction the authors mention newspapers, books, magazines, personal and government documents, legal and court documents, tape-recorded conversations, street signs, graffiti, restaurant menus, testimony given in the US Immigration Court and at US civil and criminal trials, Internet sources, information from numerous Indonesians and observers of Indonesian, and from secondary sources—sixty-seven of which are listed with their author(s) and titles (they include general and special dictionaries, glossaries, and terminology lists). This variety of sources covers not only the official standard language but also texts and constructions of a more substandard or nonofficial nature. This is reflected in the sociolinguistic and stylistic labels used for some of the lemmas: classical, colloquial, derogatory, obsolete, poetical, vulgar, young people’s disguised language (prokem), slang, and—specifically—teen slang (*bahasa gaul*).

The status of Indonesian as the official and national language of the country required the adoption and creation of new terms and expressions for all aspects of modern urban life and for all current scientific concepts. As a result, modern Indonesian vocabulary contains thousands of words recognizable as international vocabulary. In *CIED* these are marked as “D” (Dutch), “E” (English), or “D/E” (having either Dutch or English as the immediate source). Sometimes Dutch is given too much credit: a word like *kondomisasi*, “spreading/popularizing the use of condoms,” has no Dutch original, but is the product of Indonesian linguistic creativity using the assimilated suffix *(n)isasi* (also with original roots). Likewise, *CIED* has *kuningisasi*, “using the color yellow (*kuning*) in public places (e.g., painting walls yellow) as a sign of support for Golkar” and *kantongisasi*, in the construction *kantongisasi kotoran kuda* (in Klaten), “putting a gunny sack (*kantong*) on the shaft of a carriage or cart to catch
the horse’s manure.” Gems such as these, by the way, are exemplary of the lexicographers’ keen eye for the creative vitality of Indonesian.

In special domains—especially in euphemisms and terminology developed in the 1980s and early 1990s for concepts in the military and entrepreneurial spheres—the Indonesian language “engineers” often fell back on Sanskrit patterns and roots. CIED has quite a collection of those terms, labelled as “Skt neo” (Sanskrit-based neologism), e.g., wisudapurnawira, or “ceremony at one’s retirement from the army”; wirapuspaniaga, “salespromotor, canvasser”; and tunasusilawan, “male prostitute.”

Compared with the two most-voluminous Indonesian dictionaries of the last two decades, the monolingual Kamus Besar Bahasa Indonesia (KBBI)\(^1\) and the Indonesian–Russian dictionary (KBBIR)\(^2\), CIED has by far the largest number and variety of lemmas. The five last lemmas discussed above, for instance, are not found in either KBBI or KBBIR. Only for specialized modern (scientific) terminology (the English equivalents of which are practically the same anyway) is KBBI more complete. Words such as éntomofili, fémtovolt, and resorsinol are explained in KBBI, but are absent from CIED (and KBBIR).

Apart from creating and adopting regular words, Indonesian has enriched its vocabulary by the creation of acronyms and initialisms, both of which can in principle be the basis for further derivations. A sequence of the same letter in initialisms may even be replaced by a letter + cipher (pronounced as such!). Some examples in CIED are:

- **klompencapir** (kelompok pendengar, pembaca, dan pemirsa), “listeners, readers and viewers”;
- **tongpés** (kantong kempes), “broke, penniless” (literally: “flat pocket”);
- **rudal** (peluru kendali), guided missile, with the derived verb **merudal**, 1. “to launch a missile attack on,” 2. “to shoot down with a missile”;
- **PHK** (pemutusan hubungan kerja), “termination of employment,” with the derivations **mem-PHK-kan**, “to lay off (a worker)”; and **PHK-wan**, “dismissed labourer, someone laid off”;
- **WNI** (warganegara Indonesia), “Indonesian citizen,” with the derivations **ke-WNI-an**, “Indonesian citizenship” and **me-WNI-kan**, “to naturalize as an Indonesian citizen”; and
- **B3** (bahan berbahaya dan beracun), “toxic materials” (literally: “dangerous and toxic materials,” pronounced as [betiga]).

The number of short-lived acronyms and initialisms is so large that no two dictionaries will make the same selection. KBBIR has comparable entries to those just quoted, but often different ones, whereas KBBI provides an appendix (pp. 1322–1340) in which no derivations are given. CIED has quite a number of such “words,” albeit

---


not all of them as lemmas of their own, but as alternatives for full-word examples, such as with a number of the collocations with badan, in the meaning of “board, agency,” e.g. Pembinaan Pendidikan Pelaksanaan Pedoman Penghayatan dan Pemgamaran Pancasila” (BP7), “Board for Developing Education and the Implementation of Guidelines for Instilling and Applying Pancasila” and Koordinasi Masalah Kenakalan Remaja Penyalagunaan Narkotika (BAKORLANTIK), “Coordinating Board for Combating Juvenile Delinquency and Drug Abuse.”

The convention in Malay/Indonesian-X lexicography is to order polymorphemic lemmas as sublemmas under the lemma for their root. Contrary to the polymorphemic lemmas, the root lemma does not necessarily occur as an independent word. All lemmas that occur as words and all sublemmas are lexemes, i.e., each lexeme represents an inflectional paradigm of one to several words. Neither KBBI, KBBIR, nor any Indonesian-X dictionary I am aware of makes these paradigms explicit. CIED is no exception. However, whereas other dictionaries take many productive derivations for granted, CIED is—again—admirably comprehensive. Indeed, in some cases CIED is perhaps over-comprehensive. For gradable adjectives, for instance, prefixation with se- is completely predictable, “se-[adjective]” meaning “as [adjective] as.” Yet the form with the prefix is (often but not consistently) presented as a sublemma. For example, selébar, ”as wide as”; seluas, ”as wide as”; and selucu, ”as funny as” are given as entries, but selangsing (bintang sinétron), “as slender as (a soap star)”; selamban (kura-kura), “as slow as a turtle”; and selezat (bubur ayam), “as tasty as (chicken porridge)” are not. More important, however, is that with regard to less-productive derivations, CIED is more elaborate and complete than are other dictionaries.

Indonesian is rich in compounds, which (with very few exceptions) are left-headed (i.e., consisting of a head followed by a specification) and written according to the standard orthography as separate words, unless they are prefixed and suffixed at the same time: cf. tanggung jawab, “responsible”; and bertanggung jawab, “to be responsible (for)”; but pertanggungjawaban, “responsibility.” Compounds written as separate words are treated in dictionaries as examples of usage of their first component. If these are subject to further derivational affixation, they may be treated as sublemmas of their own and the derivations as subsublemmas. CIED treats all “first-order compounds” as examples of usage and further derivations as sublemmas. Under CIED’s entry lemma anak, for instance, one of the many collocations given is the compound anak tiri, “stepchild.” From this compound the verb menganak-tirikan is derived, meaning “to treat someone like a stepchild, neglect.”

A consequence of the fact that compounds that are not prefixed and suffixed cannot be distinguished from syntactic constructions, either in writing or in structure (at least in a majority of cases), is that in lexicographical practice (when Indonesian is the source language) compounds are treated with restraint. That is, they may be deemed transparent, but even if they are not, they don’t look sufficiently like words not to be considered second-rate and less important. Among the existing dictionaries of Indonesian, KBBI and KBBIR stand out by their number of compounds. As it turns out, by comparison CIED is far more inclusive. A comparison of four randomly chosen lemmas (which are the first component of many compounds) is illustrative. The following table compares the number of compounds beginning with kata, batu, burung, and membawa that CIED (1), KBBIR (2), and KBBI (3) have and do not have in common:
The large number of compounds beginning with *kata* in *KBBIR* is due to the inclusion of alternative (and less standard) terminology for word-classes and subclasses. On the whole, *CIED* is by far the most complete as regards compounds and collocations.

As the number of compounds with *burung* suggests (see the table above), *CIED* is comprehensive in its account of Indonesian flora and fauna. Many hyponyms in these lexical domains are compounds in Indonesian, either beginning with a hyperonym like *burung*, “bird”; *ikan*, “fish”; *bunga*, “flower”; *pohon*, “tree”; and *buah*, “fruit.” Also included are those terms beginning with an intermediate hyponym such as *elang*, “various species of predatory birds, such as buzzards, eagles, hawks, kestrels, kites, falcons, ospreys, and goshawks,” which appears as the first component of over twenty-five compounds for different species (all with their scientific equivalent), as against nine in *KBBI* and none in *KBBIR*.

### Microstructure

With the exception of abbreviations, acronyms, and initialisms, a dictionary article in *CIED* is headed by a monomorphemic root, which usually occurs as an independent word. If not, it is immediately followed by an affixed or reduplicated form, or an indication of the compound of which it is a component. For example:

**pingit berpingit**, “to be locked up, isolated, secluded (of a marriageable girl at home)”

**pinak anak** – a) “children and grandchildren”; b) “descendants”

Bear in mind that, in Indonesian lexicographical tradition, lemmas for transitive verbs always have the active prefix *me(N)*- and the first form in an entry should be interpreted as the imperative. For instance, *racuk meracuk*, 1. “to try to pick up (a girl)”; 2. “to bother, annoy, tease.”

As indicated above, main lemmas are followed by etymological, dialectal, or stylistic information. A typical example of a lengthy article is the one for the inherited word *batu*. This lemma is followed by a semantic profile that is divided into nine different meanings, which are numbered and ordered more or less according to frequency. Some of these meanings are immediately followed by an example of usage, e.g. “4. flashlight battery. *lampu séntér dua*—a flashlight with two batteries...
8. renal/urinary calculus, kidney stone…. *alat pemecah—ginjal ultrasonic léwat kulit* percutaneous ultrasonic lithotripter." This semantic block is followed by a number of compounds/collocations in which *batu* is the second component: *air batu* (ice), *anak batu* (slate pencil), *arang batu* (coal), *damar batu* (hard resin), *gula batu* (rock candy), *hujan batu* (hail), *rumah batu* (brick house), and *tukang batu* (bricklayer). The inclusion of such items is an important deviation from current lexicographical practice with regard to Indonesian. After these compounds, or collocations, the dictionary offers a list of proverbs and expressions, such as *Itu baru kena -nya!* (That hits the nail right on the head!) and *bulat tak bersanding* (a brave man fears nobody). These are followed finally by 179 compounds with *batu* as its first component, ranging from *batu air*, "river stone, river boulder" through *batu kali*, "river stone," and its derivation *membatu kali*, "to remain silent," to *batu usus*, "intestinal calculus."

Many words can only be properly translated in context. *CIED*, which is based on an enormous variety of texts, presents many examples of usage faithfully taken from those texts, which are illustrative of the different senses the word in question may have. Some of these verbatim examples are useful for more than just word meaning. They also present insights into Indonesian sentence structure, life, and stylistics. The subentry *berpuas*, for instance, has the following example, typical of intellectual journalism, for the collocation *berpuas diri*, "to feel satisfied/pleased with oneself, complacent": "*Pemerintah jangan cepat berpuas diri dengan adanya organisasi-organisasi kemasyarakatan yang berpacu untuk mencantumkan Pancasila sebagai satu-satunya asas dalam anggaran dasarnya* (the government should not content itself quickly with social organizations that move swiftly to insert *Pancasila* as the one and only principle in their statutes)." Another example is found with the lemma *belépotan*, "messy, greasy, muddy, dirty, filthy, soiled, smeared," the syntactic valence of which would be sufficiently illustrated by *belépotan lumpur*, "smeared with mud" and *belépotan dengan abu*, "filthy with ash." Instead, the authors adduce the following three colorful sentences:

**Hidungnya belépotan sambal:** “Your nose is smeared with sambal.”

**Akibat hujan abu itu, kendaraan-kendaraan bermotor angkutan umum jurusan Yogyakarta-Semarang dan Solo-Semarang, Jumat pagi ~ dengan abu kendati tipis:** “Due to the ash rain the public transport motor vehicles on the Yogyakarta–Semarang and Solo–Semarang run were covered Friday morning with an ash layer, though it was very thin.”

**Pokoknya, dengan uang lima ratus saja bisa ngibing sampai ~ keringat:** “The main thing is, for only five hundred (rupiah), you can dance with a dogér till you are soaked with sweat.”

These examples may be lengthy, but they do contribute to the attractiveness of the dictionary by their “couleur local,” if only because the English translations cannot be fully understood without an additional search for the meaning of *dogér*, explained as 2. “a female dancer,” apparently involved in what is given as the first sense: 1. *tari dogér* “k.o. street dance performed by women” and possibly of *sambal*: “hot spicy sauce/paste made from ground red chilli peppers, salt, etc. and served along with cooked rice.” A problem with all examples is that they in their turn may require contextual information, too. In the first example (inset, above), the third-person-
possessive suffix, *nya*, is translated as “your,” which can only be a contextual interpretation of the proper meaning “his,” “her,” or “its” (i.e., belonging to the thing/situation, and therefore translatable as “the,” and used to avoid an expression for the second person which might be perceived as inappropriate).

Many lemmas in the dictionary are followed by two or more English equivalents, due to the lack of a single one with more or less exactly the comparable meaning, as well as to the rich English lexicon. *Hantu*, meaning “phantom, ghost, demon, spectre, apparition, or evil spirit” is an example. *Belépotan*, quoted above, is another. A speaker of English in these cases may be able to make the right choice given the context. An Indonesian, on the other hand, may need additional contextual explanation to avoid translations like “muddy with sweat” for *belépotan keringat*, or “uninsulated female” for *perempuan telanjang* (naked woman). A malicious reviewer may find more examples. But the fact is that it is practically impossible to produce a dictionary that is both receptive and productive, i.e., equally geared to the needs of speakers of the target language and those of the source language. This dictionary is in the first place a dictionary for speakers of English. This is also apparent from the many examples of encyclopaedic explanations for items pertaining to Indonesian culture that lack proper English equivalents. Examples such as *sambal* and *kantongisasi* are abundant. More are noted here (I could fill a whole volume of this journal with such examples):

**Baharanjang** KA – [Keréta api batubara rangkaian panjang] coal-transport train consisting of 32 vans and one administration car pulled by two locomotives, which runs from Tanjungenim to Tarahan in Lampung (South Sumatra).

**mantri** 1. a title for a number of low-level positions; 2. (in older documents and stories) minister or counsellor of the sovereign (subordinate to the *patih*), first official of the *paséban*.

**wayang I** 1. shadow-play performance with puppets (made of buffalo hide or carved wood) manipulated by the puppeteer (*dalang*) and which cast their shadows against a cloth screen (*kelir*), illuminated by an oil lamp. The *dalang* is accompanied by a *gamelan* and by a female singer or singers [= *pesind(h)én*]; 2. the shadow puppets themselves.

**phio** illegal paper currency in the form of a piece of paper on which Chinese characters and certain codes are written; widely circulated for business transactions and gambling in Chinese circles in Tanjungpinang, Riau archipelago.

**serimpi** a classical female dance performed by a group of four principal dancers representing the heroines of the Menak romance.

**kuntilanak** the spirit of a woman who has a hole in her back, who died during pregnancy or confinement and who now wants to take possession of a woman who has just given birth, or to kidnap a baby in order to experience the joys of motherhood.

**pelesit I** … 2. a spirit controlled by a sorcerer to suck the blood of a person or eat a child’s corpse; it often appears in the shape of a blood-sucking cricket.

None of the English expressions for creatures from the other world would be an adequate translation equivalent for the last two items. Examples like these illustrate the
difficulties of cross-cultural lexicography and the illuminating way the authors of CIED have approached these.

Other Matters and Final Appreciation

No review of a dictionary is complete without some remarks on printing errors and wrong translations. To quote the authors themselves: “no work of this size and complexity can be entirely free of...errors or omissions.” And yes, there are some. The instructions for finding the lexical root, for instance (which in Malay/Indonesian lexicographical tradition is the lemma heading all derivational lexemes), are rather succinct, and for those users of the dictionary who lack sufficient insight into the morphology of the language, the lists of prefixes and suffixes on page x should be supplemented by ku- and kau- and -ku, -mu, -nya, -kah, and -lah. They may be considered clitics, but the orthography treats them as affixes, so they contribute in concealing the root. Also, the table illustrating the nonstandard forms of the meN-prefix (page x) could be improved upon by changing nc- and nj- into n-, which it should be before stems beginning with c- and j-, respectively, and by treating stems beginning with v- in the same way as those with f-. The nonexistent prefix kersi-, finally, should be read as bersi-.

The spelling used in the Indonesian lemmas, sublemmas, and examples is the common Indonesian–Malaysian orthography of 1972, with two exceptions. The minor one is that hyphens are used in derivations from compounds such as bertanggung-jawab and menganak-tirikan, whereas the standard spelling would have bertanggung jawab and menganaktirikan. A major (and for the non-Indonesian user very welcome) difference is that the phonemic opposition [ɛ] versus [e] is reflected in the spelling: CIED renders them as é and e, respectively, whereas both phonemes are written as e in the standard orthography. In this connection it should be mentioned that the Indonesian rules for splitting words over lines are not always followed. Words should be broken at a syllable boundary, but neither syllable onsets nor codas may be separated from the morpheme they belong to. Yet one finds forms such as penga-ngkut, pemili-han, and ketidakjuju-ran instead of peng-angkut, pemilih-an, and ketidakjujur-an.

Printing errors are rare, but do occur. They are of the nature pramusisiwi instead of pramusiwi for “babysitter,” and—a bit more serious—of an unjustified hyphen (the first one) representing the lemma in one of the proverbs with batu: hujan - emas di negeri orang hujan - di negeri sendiri, baik juga di negeri sendiri, “east, west, home’s best; be it ever so humble there is no place like home” (literally: “rain of gold in a foreign country, rain of stones in one’s own, it’s nonetheless good in one’s own”). I have not succeeded, however, in pinpointing serious lexicographical mistakes such as wrong translations or blatant gaps. One may find cases of the opposite: under batu, for instance, the expression seperti batu jatuh ke lubuk occurs twice: once with the translation “vanished into thin air; spirited away” and, again, after some other expressions, with the explanation “refers to someone who has left his kampung and never comes back.”

The dictionary does not give information on word class, which is defendable since word class divisions in Indonesian are not undisputed and difficult to define. Insofar as word class labels would be helpful to predict the inflected forms of the paradigm
represented by the lemma or sublemma in question, the English equivalents and the examples of usage are, as far as I could ascertain, always sufficient.

Incidentally, non-standard example sentences appear as an illustration of a standard lemma: under the lemma justru, for instance, one finds this sentence: sekali-kali tidak pernah saya mencaci dia, justru saya pujinya, “I’ve never insulted him, on the contrary I’ve praised him” (where the standard grammar would require menujinya). But cases like these only reflect the faithfulness of the authors to their sources. Real life examples, such as the lengthy quotations above (with berpuas diri and belépotan), or more colloquial ones, like Bréngsék! Mau tau urusan orang!, which is translated freely as “Damn it! You’re always sticking your nose into someone else’s business!” turn CIED into more than a practical aid for reading other texts: it is a joy to read the dictionary itself. In short, the dictionary is a must for anyone who is interested in Indonesia beyond the shallow level of the culturally autistic expatriate.

Glamour has never been a by-product of lexicography. When a new dictionary is published, lexicographers have to face the easy criticism of reviewers and of users who may have unrealistically high expectations. Rare eureka-like moments may brighten the many years of painstaking labor and cutting of knots, but if the dictionary is reliable and the best of its kind, the names of its authors will live on for decades. The indefinite article in the title of A Comprehensive Indonesian-English Dictionary is too modest for this volume that will serve well for many years to come. Stevens and Schmidgall-Tellings: those are names to remember.