Bilingual Dictionaries in Language Cultivation

Deny Arnos Kwary, Centre for Lexicography, Aarhus University, Aarhus, Denmark (deak@asb.dk)

Abstract

Dictionaries have often been mentioned as being an indispensible tool for language cultivation. However, in many countries, the attention only focuses on monolingual dictionaries. Bilingual dictionaries are rarely, if at all, used as one of the tools for language cultivation. Given the prevalence of English in Indonesia, it is more common for Indonesian people to have English-Indonesian dictionaries than the Indonesian Comprehensive Dictionary (*Kamus Besar Bahasa Indonesia*). Consequently, it is recommended that the National Center for Indonesian Languages should produce an English-Indonesian dictionary to strengthen language cultivation. This dictionary production should be based on modern lexicographical theories in order to ensure that the main function of the dictionary can be fulfilled satisfactorily.

  The discussion in this paper is divided into three parts. The first part examines the most popular printed and online English-Indonesian dictionaries among Indonesian people, and shows how those dictionaries may have a negative effect on language cultivation efforts. The second part explains the problems associated with the descriptive and prescriptive approaches, and proposes a new approach which allows both the fulfilment of the dictionary function and the inclusion of an empirical basis. The third part demonstrates how to implement this new approach, along with the modern theory of lexicographical function to create a better English-Indonesian dictionary for Indonesian people, which can be used as one of the tools for language cultivation.

Keywords: Language Cultivation, Dictionary, Descriptive, Prescriptive, Proscriptive, Lexicographical Functions, User Needs, Text Reception, Text Production, Translation.

1. Introduction

As a relatively new language, Indonesian needs various tools to support its cultivation.

In this paper, I propose the creation of a new English-Indonesian dictionary to be used

as one of these tools. I am encouraged to propose this tool after reading the article

written by Moeliono in Kompas newspaper, republished on the *Pusat Bahasa* website

on 17 November 2009. In this article, Moeliono questioned the use of the English word *summit* in the title *National Summit 2009*, instead of the Indonesian word *rembuk*. In addition, Moeliono expressed his concern about the fact that most government officials do not have, and never refer to, the Indonesian Comprehensive Dictionary (*Kamus Besar Bahasa Indonesia—KBBI*). As a lecturer at an Indonesian university, I find the same case among Indonesian students. It is unusual to find students who have *KBBI*, except those who major in the Indonesian language.

Law No. 24/2009 has prescribed the use of the Indonesian language in national forums. However, it is not wise to simply issue such Law without providing the relevant tools to support its implementation. *KBBI*, which has been used as one of the main tools in language cultivation, can also be used to support this implementation. The problem is that *KBBI* is not popular among Indonesian people, simply because they do not find any need to use it. In the above example, the organisers in charge of determining the topic of the forum seemed to be quite familiar with the English word *summit*, but unfamiliar with the Indonesian word *rembuk*. If they were aware of the Laws, they might have tried to look for the Indonesian equivalent of the English word *summit* in an English-Indonesian dictionary. If they used the website http://kamus.net, which claims to be the largest and most popular Indonesian dictionary, they would only find the equivalent *puncak* ('peak'), which was definitely inappropriate for that context. Alternatively, if they searched for the equivalent using the English-Indonesian glossary website (http://pusatbahasa.depdiknas.go.id/glosarium) created by the National Center for Indonesian Languages (*Pusat Bahasa*), they would not find any result. Since they could not find the appropriate Indonesian equivalent for the English word *summit*, it was reasonable that they would just use the English word.

Dictionaries have often been pointed out as being one of the most valuable information tools for language cultivation. Lewis and Trudell (2010: 269) state that the production of dictionaries is one of the linguistic processes in language cultivation. Coulmas (1999: 127) calls dictionaries an indispensable tool in language cultivation. Most experts, however, only associate monolingual dictionaries with language cultivation. Bergenholtz and Gouws (2006: 29) are some of the first experts who explicitly mention bilingual dictionaries as an information tool in language policy which is closely related to language cultivation. In this paper, it is recommended that a bilingual dictionary, or an English-Indonesian dictionary, is utilized as one of the tools to support language cultivation efforts.

Nowadays, as English is becoming more prevalent in Indonesia, it is more common to find Indonesian people who have English-Indonesian dictionaries than those who have *KBBI*. Nevertheless, as shown above, the current English-Indonesian dictionaries have not been able to play an eminent role in Indonesian language cultivation. In some cases, they can even damage the language cultivation efforts. Further explanations of the negative influences of the current printed and online English-Indonesian dictionaries on Indonesian language cultivation are presented in the following section.

2. Problems with the Current English-Indonesian Dictionaries

English-Indonesian dictionaries are more widely used by Indonesian people than *KBBI*. Unfortunately, the use of the current English-Indonesian dictionaries as reference and information tools may actually weaken the language cultivation efforts. In this paper, two popular English-Indonesian dictionaries are analyzed. The first is *Kamus Inggris-*

Indonesia (Echols and Shadily, 1975/2007), which was already in its 29th printing in 2007, clear evidence that it is ubiquitous in Indonesia. The second is the online English-Indonesian dictionary that has been quite popular among Indonesians and is available for free at http://kamus.net. At least three problems can be identified in those dictionaries that make them a factor in the deficiency of Indonesian language cultivation. These three problems are outdated equivalents, misspelled equivalents, and inappropriate equivalents. Each of these is discussed in the following paragraphs.

Language develops over time, and changes may occur either in orthography or in the meanings of words. Such changes are occuring particularly quickly in the Indonesian language, which is a relatively new language. Unfortunately, these changes have not been incorporated in the bilingual dictionaries. One example of an outdated equivalent is the English word *bachelor*. Both *Kamus Inggris-Indonesia* and *Kamus.net* still mention *sarjana muda* as one of the equivalents for the English noun *bachelor*. In the past, the education system in Indonesia differentiated between the *sarjana muda* level, with the degree of *Bachelor of Arts*, and the *sarjana* level, with the degree of *Doktorandus* or *Doktoranda*. Since 1993, as a response to the decree of the Minister of Education and Culture No. 036/U/1993, the degree *Doktorandus* or *Doktoranda* has no longer been used in Indonesia. With the current education system in Indonesia, the English word *bachelor* should be translated as *sarjana strata satu*. Unfortunately, this new equivalent has not been included in the bilingual dictionaries. Therefore, some people, including translators, may be trapped into using the outdated equivalent for this English word.

The second problem with the bilingual dictionaries is the misspelled equivalents. With the influence of local languages, it is quite common to find misspelled Indonesian

equivalents in English-Indonesian dictionaries. Some examples of words which are misspelled are *pilem* instead of *film,* and *kopor* instead of *koper*. Misspelled equivalents may also be due to the poor understanding of bilingual lexicographers about the standard spelling rules in the Indonesian language. One example is the English noun *object*. Both *Kamus Inggris-Indonesia* and *Kamus.net* translate *object* as *obyek*, which is a non-standard spelling, instead of its standard spelling *objek*. Indonesian people who consult these bilingual dictionaries will tend to use these misspelled equivalents, and thus weaken the language cultivation efforts.

The third problem with the bilingual dictionaries is the provision of inappropriate equivalents. A typical example of this type of problem is the Indonesian equivalent for the English word *summit*, mentioned in Section 1. A search for the equivalent of this word in the website http://kamus.net on 4 August 2010 only showed one result, i.e. *puncak*, which is related to the peak of a mountain. If the dictionary user consults *Kamus Inggris-Indonesia*, he will find the following dictionary article:

summit /ˈsʌmit/ *kb*. puncak. *s. conference* konperénsi tingkat tinggi.

Since the national summit is related to a conference, this dictionary may lead its users to translate it as *konperensi tingkat tinggi*, which is not only imprecise but also misspelled. This is the equivalent that Moeliono (2009) mentions as the favourite of the press media, but with the letter *-f-* (*konferensi*) instead of the non-standard spelling with the letter *-p-* (*konperensi*) given in the dictionary. Unfortunately, the non-standard equivalent *konperensi tingkat tinggi* is still used in some official documents. Based on an advanced search on http://google.co.id for the exact phrase and only in websites in the Indonesian language, there are still hundreds of results for *konperensi tingkat tinggi*. One of them is in the press release No. 84/PR/XII/2005 of the Ministry of Foreign

Affairs of 6 December 2005. Another is in the Presidential Decree No. 20/2001 issued on 5 February 2001.

To address the problems mentioned above, it is necessary to create an English-Indonesian dictionary which does not have those problems and can be used as a tool for language cultivation. However, before discussing the proposed steps to create the dictionary, it is necessary to evaluate the two approaches which have often been used in lexicography, in order to identify the problems associated with those approaches and to rationalize the use of the new approach called proscriptive lexicography.

3. Selecting between Prescriptive, Descriptive, and Proscriptive Lexicography

The two approaches which are commonly used in dictionary-making are prescriptive lexicography and descriptive lexicography. Hartmann and James (2002: 111) define prescriptive lexicography as an 'approach to dictionary-making which is based on normative attitudes as to how a language or language variety should be used rather than the facts observed about its usage.' If this approach is followed in making an English-Indonesian dictionary, an English entry word like *influence* (verb), will only be accompanied by one Indonesian equivalent, i.e. *memengaruhi*. Most Indonesian people may find this equivalent strange, because the spelling does not conform to the facts observed about its usage. It is more common to find the word *mempengaruhi* than the word *memengaruhi* in Indonesian documents. As a comparison, an advanced search on 5 August 2010 using http://google.co.id for all documents written in the Indonesian language showed 233,000 hits for *memengaruhi* and 1,050,000 hits for *mempengaruhi*.

Considering the fact that *mempengaruhi* is used a lot more often than *memengaruhi*, I have never found any English-Indonesian dictionaries that provide the

equivalent *memengaruhi* for the English word *influence* (verb). This may show that the current English-Indonesian dictionaries use the descriptive approach, instead of the prescriptive one. According to Hartmann and James (2002: 37), descriptive lexicography is an 'approach to dictionary-making which is based on the observed facts about a language or language variety rather than attitudes on how it should be used.' With this approach, it is reasonable that the bilingual dictionaries tend to show the equivalents which are commonly used, rather than the standard equivalents suggested by *Pusat Bahasa*. This condition certainly has a negative effect on the Indonesian language cultivation efforts.

A further implementation of the descriptive approach, which relies heavily on corpus evidence, shows even more problems, not only for language cultivation efforts, but also for the dictionary users. *The Cambridge Guide to English Usage* (CGEU) written at Macquarie University's Dictionary Research Centre by Peters in 2004, is one of the examples of a work using the descriptive approach and based mainly on corpus evidence. In the Preface of CGEU, Peters (2004: vii) mentions that CGEU allows the readers to choose their own style. Unfortunately, dictionary users do not usually know which one to choose. Consider the following dictionary article from CGEU (2004: 20).

adviser or advisor

Both these spellings are in current use, though **adviser** is the dominant spelling in both the US and the UK. The ratio in American data from CCAE is 20:1 and in British data from the BNC it's 6:1. Curiously, **advisor** is sometimes said to be "the American spelling." The *Oxford Dictionary* (1989) notes the frequency of the **-or** spelling in the titles of persons who give advice "especially in the US," and this quasi-official usage has no doubt helped to make people aware of it. Yet the *Oxford* lists **advisor** only as a variant of **adviser**, with no independent headword even for cross-referencing. The spelling **adviser** is consistent with the majority of agent words formed in English (see **-er/-or**), and it goes back to C17, according to *Oxford* citations, whereas **advisor** is first recorded just before 1900. Whether it is simply a respelling or **adviser** or a backformation from **advisory** is a matter of debate. But whatever its past, **advisor** is registered alongside **adviser** in major British, American, Canadian and Australian dictionaries.

After reading this lengthy dictionary article, I doubt that the dictionary users know which spelling—*adviser* or *advisor*—they should use, especially after reading the last sentence. This is supported by Allen (2009: 358) who points out that 'reporting findings rather than giving rulings is likely to leave many users unsure how to act.' A similar case will occur if we apply the descriptive approach and corpus evidence to an English-Indonesian dictionary. If we only list all the possible equivalents in the dictionary, it is unlikely that the users will be able to choose the right one.

As explained above, using either prescriptive or descriptive lexicography is not the best approach in dictionary-making. In this case, a new approach is required in order to provide better lexicographical solutions for the dictionary users. Bergenholtz (2003) proposes an approach called proscriptive lexicography. It should be noted here that the term proscriptive is not related to the English word which means 'forbidden.' Bergenholtz derived the term proscriptive from the Latin word *proscribere* which means to 'announce' or 'make public'.

Based on the proscriptive approach, a lexicographer still uses the results of corpus evidence as in the case of the descriptive approach. However, the lexicographer does not only inform the users about the language use, but he also tells the users which variant they should choose if there is more than one possibility (Bergenholtz 2003: 77). The implementation of the proscriptive lexicography can be seen in the following dictionary articles for the Danish words *vejlede* ('instruct') and *kraftvarmeværk* ('combined power and heating plant').

> **vejlede** verb ⟨-r, vejledte, -t⟩
> The Danish Language Council allows only preterite and past tense *vejledede* and *vejledt*, inflections which are not very frequent.

kraftvarmeværk noun ⟨et; -et, -er, -erne⟩

Other spellings with hyphens are possible: *kraft-varmeværk* or *kraft-varme-værk*. They are not recommended. They are quite rare in language use, e.g. by special field experts. The Danish Language Council allows only the spelling with two hyphens: *kraft-varme-værk.*

<div align="right">(Source: Bergenholtz, 2003: 78)</div>

In the above examples, we can see that all variants are presented, and the options recommended by the Danish Language Council are also stated, no matter whether they are frequent or not in the corpus. If we use this proscriptive lexicography approach for an English-Indonesian dictionary, we can have a dictionary article like the following for the English word *influence* (V).

influence verb ⟨-s, -d, -d, -ing⟩

memengaruhi, *mempengaruhi*. The form *mempengaruhi* is the more frequent one. Pusat Bahasa allows only *memengaruhi*. Click HERE to see the explanation of Indonesian word formation rules.

The above dictionary article is suitable for an electronic dictionary. If it is used in a printed dictionary, the last sentence in the example above should be changed to "See the explanation of Indonesian word formation rules in the Appendix." This additional information is useful for dictionary users who want to know why the form *memengaruhi* is selected as the standard one, and why the form *mempengaruhi* is not acceptable. Since not all dictionary users need this information, it does not have to be put in the dictionary article. It is better to place it in the Appendix which functions as a guide to the Indonesian word formation rules.

As we can see in the dictionary article, both the standard and the non-standard variants are presented, along with the explanation based on the corpus evidence and the recommendation from the language council. By using this proscriptive lexicography approach, the English-Indonesian dictionary will show the users all the possible forms

found in the corpus and support the language cultivation efforts. A step-by-step guide to create the dictionary based on a theoretical foundation is presented in the following section.

4. Implementing the Lexicographical Function Theory in Creating the Dictionary

Dictionaries should be created based on a sound theoretical foundation. In this paper, the theory selected to be applied is the modern theory of lexicographical functions. There are two main reasons for choosing this theory. The first is its transformative nature, which means that it does not only lay a solid theoretical foundation for reviews and criticism of dictionaries that do not perform their role optimally, but also strives to provide new and better solutions to lexicographical problems (Tarp 2008: 84). The second is its focus on dictionary users. The significance of focusing on the users has been stated by several experts. Bogaards (1999: 118) urged lexicographers to adopt more often and more systematically the viewpoint of the learners or the dictionary users, whereas Atkins and Rundell (2008) dedicate a special section on tailoring the entry to the user who needs it.

The modern theory of lexicographical functions originates from the Centre for Lexicography of Aarhus School of Business, Denmark, and has been developed since the early 1990s, cf. Tarp (1992, 2008), Bergenholtz and Tarp (1995, 2003), Nielsen (1994a), and Bergenholtz and Nielsen (2006). This theory has been called one of the most important and more recent contributions in the field of theoretical lexicography (Gouws 2007: 66), since it has enabled the creation of innovative concepts and been successfully implemented in several outstanding dictionaries, such as the English-Danish Dictionary, Danish-German Dictionary, and Danish-Spanish Dictionary.

4. 1. Determining the Dictionary Function and the User Profile

As the name of the theory suggests, this lexicographical function theory requires lexicographers to determine the function of the dictionary at the initial stage in creating the dictionary. Tarp (2008: 81) defines a lexicographical function as 'the satisfaction of the specific types of lexicographically relevant need that arise in a specific type of potential user in a specific type of extra-lexicographical situation.' For communication-oriented functions, Bergenholtz and Tarp (2003: 176) state that the most important functions are to assist the users in solving problems related to:

1. text reception in the native language
2. text production in the native language
3. text reception in a foreign language
4. text production in a foreign language
5. translation of texts from the native language into a foreign language
6. translation of texts from a foreign language into the native language.

From the six communication-oriented functions above, the one chosen in this study is the translation of texts from a foreign language into the native language because it is the most important function in connection with an English-Indonesian dictionary to support language cultivation efforts. It is also related to the problem identified in Section 1, in connection with the need to translate the English word *summit* into Indonesian. In particular, the dictionary is expected to be able to assist Indonesian people when they want to find the Indonesian equivalents of English words.

After selecting the main function of the dictionary, the next step is to determine the user profile in order to ascertain what they need in a particular situation, and the assistance that the dictionary can provide to meet the user needs. The user profile of the dictionary should be drawn up meticulously when creating a dictionary. Without

establishing its specific type of users, a dictionary will only become a display without anyone benefiting from it. A user profile is a set of characteristics of the intended user group of a dictionary. Landau (2001: 345) points out that the identification of the market or dictionary users is the first consideration in dictionary publishing, but does not explain how to determine the profile of potential dictionary users. Atkins and Varantola (2008: 340) provide a clearer technique to determine the profile of dictionary users, i.e. by using the following five questions:

- Question 1 asks for information about the dictionary users' mother tongue.

- Question 2 asks for the users' own assessment of their level of competence (beginner, intermediate, advanced) in the L2 of their choice.

- Question 3 relates to the actual translation task being performed, whether it is L1 to L2 or L2 to L1.

- Question 4 asks for information about the language of the translation passage selected (easy, intermediate, or difficult).

- Question 5 asks for information about all the dictionaries that were used in the course of the translation work being recorded.

Those five questions, however, are not transformative in nature. They mostly ask about the past experience of the users in using dictionaries and may show some inconsistencies. The answers to Questions 2 and 4, which are based on the personal opinion of the users, are examples of such inconsistencies. For Question 2, a user may identify himself as advanced, while his competence is probably the same as another user who identifies himself as intermediate. The same applies to Question 4, where a user may consider the selected text easy while another user may consider it difficult, although both of them identify themselves as intermediate in Question 2.

Question 5 is not very useful either for transformative research, because it only identifies which dictionaries have been used by the users. Besides, the answers may not reflect the actual conditions for at least the following two situations. The first situation is a user who has two dictionaries but s/he uses one of them much more often than the other one. The second situation is a user who has several dictionaries and uses all of them but s/he does not remember the titles and the publishers. Those two situations will certainly distort the results. Therefore, Question 5 should not be used in the analysis of the user profile.

Consequently, it is only useful to consider Questions 1 and 3 when creating better concepts for future dictionaries. Question 1 is necessary because by knowing the mother tongue of the users, lexicographers will be able to ensure some of the dictionary features focus on a particular language and culture; for example, by providing specific cultural notes to avoid any misunderstanding. Question 3 is also necessary to determine the source language which will be used as the first framework for selecting the headwords for the dictionary. However, it is not necessary to distribute questionnaires to find the answers to these two questions. Lexicographers will be able to answer these questions easily after they determine the function of the dictionary.

As stated previously, the main function selected for this proposed dictionary is the translation of texts from a foreign language into the native language. If we relate this to Question 3 in Atkins and Varantola (2008: 340), the actual translation task being performed is from L2 to L1, or from English to Indonesian. This also implies the answer to Question 1, i.e. that the dictionary users' mother tongue is Indonesian. Some people may argue that there are still some Indonesian people whose mother tongue is not the Indonesian language, but their local language. Such an argument is usually related to the

differences in the level of Indonesian language proficiency of most Indonesian people. To cope with this concern, the following user characteristics proposed by Bergenholtz and Nielsen (2006: 258) in determining the user profile can be considered.

1. Which language is their native language?
2. At what level do they master their native language?
3. At what level do they master a foreign language?
4. How extensive is their experience in translating between the languages in question?
5. What is the level of their general cultural and encyclopaedic knowledge?
6. At what level do they master the special subject field in question?
7. At what level do they master the corresponding LSP in their native language?
8. At what level do they master the corresponding LSP in the foreign language?

The eight questions listed above are actually intended to be used to draw up the profile of an LSP dictionary user group, but with several adjustments, these questions can also be applied to a general dictionary user group. As we can see from the list, Questions 6, 7, and 8 are particularly appropriate for LSP, so they can be omitted when determining the user profile of the proposed dictionary in this paper. The other five questions, however, can be used to determine the user profile of a general dictionary like the dictionary proposed in this paper.

In connection with the problem identified in Section 1 and the main function selected for the proposed dictionary, the main users are Indonesian government officials. The answer to Question 1 is that the native language of these users is Indonesian. The previous argument about the position of Indonesian as the native language of these users can be addressed by using Question 2. For this question, the users can be considered to have a mastery of the Indonesian language at an intermediate level. They have used the Indonesian language in many domains and have studied it from elementary school until at least senior high school. However, due to the influences

from their local languages, their mastery of the Indonesian language is not at an advanced level.

Next, as the answer for Question 3, the users are at an intermediate level in their mastery of the English language because they have studied it for at least six years, and with the widespread use of English in Indonesia, these users have been in contact with English every day. For Question 4, the experience of these users of translating from English to Indonesian is not extensive because they are not professional translators. Lastly, for Question 5, their general cultural and encyclopaedic knowledge is at an intermediate level because they are adults, and most of them have studied at university.

The users need to translate English words into standard Indonesian words. Therefore, the dictionary should not only provide the Indonesian equivalents which enable the users to understand what the English words mean, but it should also show the standard and the most appropriate Indonesian equivalents. In connection with this dictionary assistance, the user group of this dictionary may actually include a wider group, such as translators, students and Indonesian people in general, who need help in translating English words into Indonesian. After determining the dictionary function and the profile of the intended user group, the next step is to select the headwords for the dictionary, which is discussed in the following section.

4. 2. Selecting the Headwords for the Dictionary

The headwords for a dictionary are usually selected from a corpus which has been built according to the dictionary function and user profile. Since English is an international language which has drawn the interest of many language researchers, it is possible to find good resources from which the headwords for the dictionary can be selected. In this paper, two resources that are considered particularly suitable to be referred to when

selecting the headwords for the proposed dictionary are OALD (Oxford Advanced Learner's Dictionary) and DANTE (Database of Analysed Texts of English).

If cooperation with Oxford University Press can be established, the headwords available in OALD can be used to select the headwords for the proposed dictionary. The headwords in OALD must have been selected meticulously from the large and balanced corpus, the British National Corpus, which consists of 100 million words, and the Academic Wordlist. In addition, OALD is the most popular monolingual English dictionary in Indonesia, so it has been the main reference work which is consulted for definitions of English words. Unlike the OALD in several countries, such as China and Malaysia, the OALD in Indonesia has not been bilingualised. Therefore, there is still room for a project to bilingualize this popular dictionary, or use its headwords to create an English-Indonesian dictionary for the text translation function.

The other resource which can be used to select the headwords is DANTE (Database of Analysed Texts of English). The corpus of DANTE includes the 100 million word British National Corpus, the 25 million word Hiberno-English corpus, the 100 million word corpus of American English, and the 1.5 billion word ukWAC corpus. The front page of the website http://webdante.net states that DANTE is a database which can be used to make new bilingual dictionaries with English as the source language. DANTE allows the automatic insertion of translation fields that can hold target-language equivalents at the appropriate points. Consequently, the work to create the bilingual dictionary will be easier and can be completed faster.

One of the problems of using the headwords from OALD or DANTE is the exclusion of other local English language varieties. Indonesia is located between

Australia and Singapore which have local English vocabularies that may not be included in the OALD or DANTE. This means that if the headwords of the proposed dictionary are only selected from the OALD and DANTE, there will be some situations where the main users of the proposed dictionary cannot find the solutions to their lexicographical problems.

In order to ensure that all the English words needed by the users are included as headwords in the dictionary, it is necessary to build a corpus collected from the English texts which include the local English varieties. With the recent advances in computer technology, the process to collect the corpus and extract the keywords will not take too much time. One example of a tool which can be used to build the corpus is the WebBootCat tool available at http://sketchengine.co.uk. WebBootCat can construct a corpus instantly by crawling particular Internet domains, such as: .au (Australia), .sg (Singapore), and .id (Indonesia), that have web pages in the English language. The tool can also produce a wordlist instantly, from which the lexicographers can select the headwords for the dictionary. After the headwords are selected, the next step is to provide the equivalents and other relevant data for the dictionary.

4. 3. Determining the Equivalents and Other Relevant Data

There are various types of data that can be addressed to the headwords and provided in a dictionary. Bergenholtz and Tarp (1995) mention the following types of data: word combinations, synonyms, antonyms, labels, pronunciations, examples, notes, encyclopaedic information, illustrations, etc. Since the main function of the proposed dictionary is translation, the main data that have to be addressed to the English headwords are the Indonesian equivalents, notes and encyclopaedic information.

The Indonesian equivalents of English words can be determined by asking a group of professional translators to translate the English words which have been selected as the headwords. In addition, a parallel corpus is sometimes useful to assist the translation process. Atkins and Rundell (2008: 476) distinguish between two types of parallel corpus: a translation corpus and a comparable corpus. A translation corpus consists of the translation of the source language texts. This means that the quality of the equivalents depends on the quality of the translation. If the lexicographers are in doubt about the quality of the translation, it is better to use a comparable corpus than a translation corpus. A comparable corpus 'is made up of two individual language corpora, selected on the basis of at least one shared parameter, usually the subject matter, together with possibly other properties shared by the texts, such as the date and/or the medium' (Atkins and Rundell 2008: 479). An English-Indonesian comparable corpus may consist of the news taken from quality newspapers in the UK and in Indonesia, reporting the same events in English and in Indonesian, respectively. This corpus is an excellent resource for creating a bilingual dictionary because the texts are original and no translation is involved.

For the dictionary proposed in this paper, the use of both professional translators and a comparable corpus is recommended. Professional translators will provide the translations in a relatively short amount of time. The comparable corpus can be used to check the translations. However, the Indonesian texts in the comparable corpus must consist of two sub-corpora, labelled colloquial and standard. The colloquial corpus may contain Indonesian texts from newspapers and websites. WebBootCat can be used to construct this corpus instantly by crawling the Indonesian Internet web pages that are in the Indonesian language. The other sub-corpus, the standard one, must be constructed

only from texts which use standard Indonesian language, such as KBBI and the publications of *Pusat Bahasa*.

The use of two sub-corpora ensures that both colloquial and standard Indonesian words are taken into account in the dictionary. This means that the dictionary includes both the standard and local variants. However, based on the proscriptive lexicography approach explained in the previous section, the lexicographers should state which form is recommended. Since the dictionary is intended to support the language cultivation efforts, the recommendation is always based on the standards determined by *Pusat Bahasa*.

In addition to the equivalents and the notes regarding the recommended forms, the rationalization for selecting particular forms to be the standard may also be included in the dictionary. This is usually called the encyclopaedic section and can be placed in the front matter of the dictionary. Nielsen (1994: 92) suggests that the encyclopaedic section be divided into separately numbered paragraphs to facilitate cross-referencing. The numbered paragraphs will make it easier for the users to find the data they need. Consider the previous example for the English word *influence* (V). If the encyclopaedic section is not made into numbered paragraphs, the dictionary article will look like the following:

influence verb ⟨-s, -d, -d, -ing⟩

memengaruhi, *mempengaruhi*. The form *mempengaruhi* is the more frequent one. Pusat Bahasa allows only *memengaruhi*. See the explanation of Indonesian word formation rules in Appendix 1.

The Appendix may consist of the word formation rules of all affixes in the Indonesian language, which will make it quite lengthy. When the users open the Appendix and find that it consists of several pages, they might be reluctant to look for

the paragraph that provides the explanation they need. Let us compare it with the following dictionary article when the encyclopaedic section has been made into numbered paragraphs.

influence verb 〈-s, -d, -d, -ing〉

memengaruhi, *mempengaruhi*. The form *mempengaruhi* is the more frequent one. Pusat Bahasa allows only *memengaruhi*. See the explanation on Page 10, Paragraph 20.

The cross-reference to a specific paragraph on a particular page enables the users to find the data easily. To save space in the printed form of the dictionary, the sentence "See the explanation on Page 10, Paragraph 20" can be simplified into "cf. Page 10, #20". If the dictionary is in an electronic form, the paragraph number in the dictionary article should be embedded with a direct link to the specific paragraph. Therefore, the users only need to click on the paragraph number and they will be taken directly to the page where the explanation is located.

In some cases, it is not necessary to place the explanation for selecting a particular form or for refusing a variant in the encyclopaedic section. If the colloquial variants are due to the influence of a specific local language or pronunciation, the explanation may be placed directly in the dictionary article. For example, the dictionary article for the English word *conference* may look like the following:

conference noun 〈-s〉

konferensi, konperensi. Some Indonesian people find it difficult to pronounce the sound /f/ and substitute it with the sound /p/. This is not recommended in writing. Pusat Bahasa allows only the spelling with /f/.

From the discussion above, it is clear that there are three main groups of data which should be addressed to the headwords of the proposed dictionary: equivalents, notes, and encyclopaedic information. The equivalents include both the standard and the

colloquial variants. Since the dictionary will be used as one of the tools to support Indonesian language cultivation efforts, the standard equivalent is always put in first, followed by the non-standard equivalent. To ensure that the users choose the standard ones, notes are given in the dictionary articles based on the proscriptive lexicography approach. Finally, further information or explanations may be provided either within the article or in the encyclopaedic section of the dictionary.

5. Conclusion

The discussion in this paper shows that it is necessary to rely not only on the Indonesian Comprehensive Dictionary (*Kamus Besar Bahasa Indonesia*) but also on an English-Indonesian dictionary for Indonesian language cultivation. However, based on the analysis of the current English-Indonesian dictionaries, we can conclude that they are not suitable for language cultivation. In some cases, they can even have a negative effect on language cultivation efforts. Therefore, there is an urgent need to create a new English-Indonesian dictionary which can be used as one of the tools for language cultivation.

The principles to create this dictionary have been explained in this paper. The prescriptive and descriptive approaches are deemed to be insufficient. The new approach, called proscriptive lexicography, has been demonstrated to be the most appropriate one for the proposed dictionary. In addition, it is necessary to follow the principles stated in the lexicographical function theory in order to ensure that the dictionary will provide solutions that both satisfy the needs of the users and support language cultivation efforts.

References

Allen, R., 2009. Dictionaries of usage. In A. P. Cowie, ed. *The Oxford History of English Lexicography Volume II*. Oxford: Oxford University Press, pp. 339–360.

Atkins, B. T. S. & Rundell M., 2008. *The Oxford Guide for Practical Lexicography.* Oxford: Oxford University Press.

Atkins, B. T. S. & Varantola, K., 2008. Monitoring dictionary use. In Th. Fontenelle, ed. *Practical Lexicography: A Reader*. Oxford: Oxford University Press, pp. 337–375.

Bergenholtz, H. & Gouws, R., 2006. How to do policy with dictionaries. *Lexikos*, 16, pp. 13–45.

Bergenholtz, H., 2003. User-oriented understanding of descriptive, proscriptive and prescriptive lexicography. *Lexikos*, 13, pp. 65–80.

Bergenholtz, H. & Nielsen, S., 2006. Subject-field components as integrated parts of LSP dictionaries. *Terminology*, 12(2), pp. 281–303.

Bergenholtz, H., & Tarp, S. eds., 1995. *Manual of Specialised Lexicography: The Preparation of Specialised Dictionaries.* Amsterdam: John Benjamins.

Bergenholtz, H., & Tarp, S., 2003. Two opposing theories: on H. E. Wiegand's recent discovery of lexicographic functions. *Hermes*, 31, pp. 171-196.

Bogaards, P. 1999. Access structures of learner's dictionaries. In T. Herbst and K. Popp, eds. *The perfect Learners' dictionary (?)*. Tübingen: Max Niemeyer Verlag, pp. 113–130.

Cabinet Secretariat, Republic of Indonesia, Presidential Decree *No. 20/2001*, on 5 February 2001.

Coulmas, F. ed., 1999. *The Blackwell Encyclopedia of Writing Systems*. Oxford: Blackwell.

Echols, J. M. & Shadily, H., 1975/2007. *Kamus Inggris-Indonesia*. Jakarta: Gramedia Pustaka Utama.

Gouws, R. H., 2007. Sublemmata or main lemmata. In H. Gottlieb and J. E. Mogensen, eds. *Dictionary Visions, Research, and Practice*. Amsterdam: John Benjamins, pp. 55–69.

Hartmann, R. R. K. & James, G., 2002. *Dictionary of Lexicography.* 2nd ed. London: Routledge.

http://google.co.id. [Accessed 5 August 2010].

http://kamus.net. [Accessed 4 August 2010].

http://pusatbahasa.depdiknas.go.id/glosarium. [Accessed 4 August 2010].

http://sketchengine.co.uk. [Accessed 6 August 2010].

http://webdante.net. [Accessed 6 August 2010].

Landau, S. I., 2001. *Dictionaries. The art and craft of lexicography*. 2nd ed. Cambridge: Cambridge University Press.

Lewis, M. P. & Trudell, B., 2010. Language cultivation in contexts of multiple community languages. In B. Spolsky and F. M. Hult, eds. *The Handbook of Educational Linguistics*. West Sussex: Wiley-Blackwell, pp. 266–279.

Ministry of Education and Culture, Republic of Indonesia, Decree *No. 036/U/1993*, on 9 February 1993.

Ministry of Foreign Affairs, Republic of Indonesia, Press release *No. 84/PR/XII/2005*, on 6 December 2005.

Ministry of Laws and Human Rights, Republic of Indonesia, Laws *No. 24/2009*, on 9 July 2009.

Moeliono, A. M., 2009. National Summit?. *Kompas*, 17 Nov.

Nielsen, S., 1994a. *The Bilingual LSP Dictionary. Principles and Practice for Legal Language*. Tübingen: Gunter Narr Verlag.

Nielsen, S., 1994b. Encyclopedic information in the front matter: bilingual law dictionaries. *Netværk LSP Nyhedsbrev*, 9, pp. 85–93.

Peters, P., 2004. *The Cambridge Guide to English Usage*. Cambridge: Cambridge University Press.

Sugono, D. ed., 2008. *Kamus Besar Bahasa Indonesia.* 4th ed. Jakarta: Gramedia Pustaka Utama.

Tarp, S., 1992. *Prolegomena til Teknisk Ordbog*. Ph.D. Dissertation. Aarhus: Spanish Institut, Aarhus School of Business.

Tarp, S., 2008. *Lexicography in the Borderland between Knowledge and Non-knowledge: General Lexicographical Theory with Particular Focus on Learner's Lexicography.* Niemeyer: Tübingen.

Turnbull, J. ed., 2010. *Oxford Advanced Learner's Dictionary.* 8th ed. Oxford: Oxford University Press.