

THE NATURE OF A LEARNER'S DICTIONARY*

Leonard E. Newell
Canada Institute of Linguistics

1. WHAT IS A LEARNER'S DICTIONARY?

A learner's dictionary, simply put, attempts to provide the user with tools not only to be able to understand what a text means, but also to be able to compose natural texts in the language being described. Language being of the nature that it is, the task of compiling such a dictionary is extremely complex and, in the long run, probably unattainable with any degree of completeness. We still do not have the analytical tools available to adequately unravel the semantic intricacies of a language. Sinclair states "...it is not yet understood how meanings are distributed among forms of a lemma, and a new branch of study is looming — the interrelationships of a lemma and its forms" (1991:41). He goes on to say (p.45), "Research will, in due course, offer guide-lines which will gradually improve the choice of texts, sampling methods, processing of evidence, and application of the results to lexicography. Until then, we must use the evidence with care, but we must use it."

Our obligation, then, is to employ whatever tools we have and attempt, no matter how imperfectly, to compile dictionaries to aid beginning learners to effectively understand and compose texts.

2. KINDS OF LEARNERS' DICTIONARIES

Basically there are two kinds of learners' dictionaries — one in which both the source and target languages are the same; a monolingual dictionary, and one in which the source language (metalanguage) is different from the target language, a bilingual dictionary (or a multilingual dictionary if more than one source language is involved).

2.1. A Monolingual Learner's Dictionary

A monolingual learner's dictionary is intended mainly for advanced language learners who, in addition to using the dictionary for looking up spellings or the meaning of uncommon words, will need to consult it for matters such as grammatical distribution of lexemes as well as various semantic relationships among them. It might also include words not generally used in the core vocabulary of average speakers as well as etymological, archaic, botanical, and other information. Examples of this kind of dictionary include *The Oxford Advanced Learner's Dictionary of Current English*, *Collins Cobuild English Language Dictionary*, and *Chambers Universal Learner's Dictionary*.

2.2. A Bilingual Learner's Dictionary

A bilingual (or multilingual) learner's dictionary is specifically designed for those with little knowledge of the target language. It focuses on the common-core vocabulary and must concern itself with a good deal of information necessary for the formulation of discourse

which is generally considered not necessary in a standard or advanced learner's dictionary. This includes a somewhat detailed grammatical distribution of lexemes and detailed descriptions of various semantic relationships among them. It has little, if any, etymological, archaic, botanical, or other information.

3. SOME FEATURES OF A LEARNER'S DICTIONARY

3.1. Language Naturalness

A learner's dictionary seeks to present the language, as much as possible, as it is naturally spoken and written. This is accomplished primarily by working from a large textual corpus. This has a number of advantages. One is that it helps minimize unnaturalness in lexical description. It is all too easy to be unduly influenced by the patterns of the metalanguage in compiling a bilingual dictionary. The result is sometimes to describe grammatical and semantic patterns that perhaps could but do not actually exist in the writing and speech of the target language as it is naturally formulated. And for those compiling a monolingual dictionary, it helps avoid the description of word forms and usage that are often included in traditional dictionaries but are not actually found in natural speech or writing. Based on frequency of occurrence using a large corpus, hard decisions about what to include and what to leave out are confidently made.

3.2. Use of Full Sentences in Definitions and Lexical Glosses

A good learner's dictionary, beyond explaining the designative meaning of a word, attempts to illustrate usage in terms of semantic and grammatical contexts.

3.2.1. Lexical Context

Most traditional dictionaries are relatively adequate in describing the *taxonomic* (vertical, substitutionary) relationships of lexemes, but fail to indicate the *tactic* (horizontal) relationships. Note how one dictionary defines the word **whimsical**:

whimsical - full of, actuated by, or exhibiting whims

whim - a capricious or eccentric and often sudden idea or turn of the mind

But there is no indication of what (or who) is full of whim, is actuated by whim, or exhibits whim. We are left to conjecture that the following phrases are well-formed: a whimsical story, whimsical sense of humor, whimsical ways, whimsical smile, or a whimsical person. This information would need to be verified by consulting an extensive textual database, and would need to be described as generically as possible in a lexical description of **whimsical**.

To accomplish this, some dictionaries use full sentences in describing words. *Cobuild*, for example, casts all lexical descriptions in full sentences:

immunize...immunizes, immunizing, immunized... If people or animals are **immunized** they are made immune to a particular disease, often by being given an injection. EG *Everyone who is going abroad will need to be immunized against typhoid.*

In addition to the information that this word means "to be made immune" which is commonly given in standard dictionaries, it also includes the following information:

- 1) The use of this verb is restricted (exclusively, or almost exclusively) to its passive form.

- 2) The patient of the action is a person or animal.
- 3) The instrument is often an injection.
- 4) The negative beneficiary of the action is a particular disease which is expressed following the proposition *against*. (The use of *against* occurs only in the illustrative sentence.)

Unless context for possessed nouns is specified, users in many cases are either dependent on their own knowledge of the language to provide this information, or are unable to use possessed nouns accurately within sentences. This is commonly the case when the entry lexeme benefits a co-occurring substantive in some way or is a noun (thing or nominalized event) of which a co-occurring noun is source.

Ifugao example:

adaw possd n. someone's **small gift**, which he or she accepts without formal obligation to repay, as betel nut, tobacco, a little rice, an item of clothing

Hindi example:

dān n. someone's **gift** which he or she gives to someone

English example:

Cobuild: A **gift** is something that you give someone as a present.

But perhaps one can say:

"This is my gift that you gave me (? as Ifugao)."

3.2.2. Syntactic Variation

Full sentences can also be used to express alternate ways a particular word is used, as illustrated in the Romblomanon bilingual dictionary (in preparation):

bāgu 1 tmp conj. ⇒ **Before** something happens or a state exists something else happens; or something happens **before** something else happens or a state exists. *Bāgu ka mag'ūli' sa Rumlun tudlu'i 'ānay 'ang 'imu manga ka'ibāhan.* **Before** you go back to Romblon please teach your companions. *Didtu na kami nagkatuyug bāgu bumālik naman pagkaMyirkulis.* We slept there again **before** returning on Wednesday.

3.3. Importance of Gloss Accuracy

Lexical glosses should, as much as possible, follow the exact pattern of usage so that, by following it, the user will be able to accurately formulate the context of the lexeme.

sūkuy v. ⇒ An object is **measured** by someone to determine its length or width. (not: "The length or width of an object is **measured**...")

Ginsūkuy ku `ang lamīsa kung ma'unu kalāpad. I measured the table
[to determine] how wide [it is].

3.4. Word Forms in a Learner's Dictionary

A chronic problem in compiling dictionaries for agglutinative languages is the matter of describing the various inflected forms of nouns and verbs.

For English and other Indo-European languages this matter does not present major problems. The uninflected form of the verb in English is usually chosen as the entry form (lemma). Other inflected forms are few and are generally handled immediately following the part of speech at the beginning of the entry. For regular verbs, e.g. **hunt**, *Cobuild* lists the third person singular form **hunts**, present participle **hunting** and past **hunted**. *Cobuild*, however, does not distinguish between homonyms or homographs. As a result, only shared forms are listed following the entry form; e.g. **ring...**, **rings**, **ringing**. Differences (**rang** *pt.* and **rung** *pp.* for one homonym and **ringed** both *pt.* and *pp.* for the other) are listed in the text that immediately follows. For other irregular verbs, e.g. **sing** *Cobuild* lists four forms following the lemma: **sings**, **singing**, **sang**, and **sung**.

Oxford and *Webster* separate homonyms and homographs into separate entries, so they do not have this complication. Forms of regular verbs are not listed at all. For irregular forms, the past, e.g. **rang** and past participle **rung** are cited. In addition, *Webster's* includes the present participle **ringing**.

Compound tenses of verbs in English, including features of aspect, mood, etc., are expressed as phrases and are not cited as lexical forms. They occur only in example sentences.

In contrast, features of tense, aspect, mood, case, manner, etc. for agglutinative languages are commonly indicated by verb affixation. The result is a large number of forms for any given verb. Before our ability to compile exhaustive lists of such forms by computer we were, at least to some extent, unaware of how many actually occur in natural, everyday speech and writing for any given verb. In any case, to compile this information manually would be virtually impossible, since it requires surveying several hundreds (if not thousands) of pages of texts for each verb. The computer has solved this problem for us (but has created others). For the Rombomanon form **bakay** which has to do with 'buying,' for example, 83 distinct forms were recorded from a limited textual database, along with the sentence contexts in which they were found. The list of affixed forms and affix combinations is displayed below.

ANALYSIS OF BAKAY: 1243 INSTANCES FOUND. FORMS: 83	
111 instances of nagabakay (naga-)	30 instances of ginabakya (gina-a)
109 instances of mabakay (ma-)	30 instances of makabakay (maka-)
91 instances of magbakay (mag-)	28 instances of gingbakay (giN-)
89 instances of ginabakay (gina-)	27 instances of nabakay (na-)
67 instances of pagbakay (pag-)	25 instances of bakay
65 instances of bakyun (-un)	23 instances of nakabakay (naka-)
55 instances of nagbakay (nag-)	20 instances of nagapamakay (paN-naga-)
55 instances of `ibakay ('i-)	17 instances of binakyan (-in-an)
39 instances of pangbakay (paN-)	16 instances of bakyan (-an)

NATURE OF A LEARNER'S DICTIONARY

16 instances of `ikabakay (ka-`i-)	4 instances of pagkabakay (ka-pag-)
15 instances of ginbakay (giN-)	3 instances of bakya (-a)
15 instances of nakakabakay (ka-naka-)	3 instances of baki (-i)
14 instances of mababakay (CV-ma-)	3 instances of magpabakay (pa-mag-)
14 instances of nababakay (CV-na-)	3 instances of `ipabakay (pa-`i-)
13 instances of balakyānun (-al-anun)	3 instances of `iyugpamākay (paN-`iyug-)
13 instances of mamakay (maN-)	2 instances of bakyānun (-anun)
12 instances of bumakay (-um-)	2 instances of byakyun (-Vy--un)
12 instances of ginabakyan (gina-an)	2 instances of gapabakay (pa-ga-)
11 instances of nagpamakay (paN-nag-)	2 instances of ginapamakay (paN-gina-)
9 instances of gabakay (ga-)	2 instances of gingpamakay (paN-giN-)
9 instances of ginbakyan (giN-an)	2 instances of kabākay (CV-ka-)
9 instances of makakabakay (ka-maka-)	2 instances of magpakākay (ka-pag-mag-)
9 instances of miyugpamākay (paN-miyug-)	2 instances of minugbākay (minug-)
9 instances of pinamakay (paN-in-)	2 instances of nabakyan (na--an)
8 instances of nagpabakay (pa-nag-)	2 instances of nagabinākay (-in-naga-)
8 instances of `iyugbākay (`iyug-)	2 instances of pagbakyan (pag--an)
7 instances of ginapabakay (pa-gina-)	2 instances of pagbakyun (pag--un)
7 instances of gingbakyan (giN--an)	2 instances of pagpabakay (pa-pag-)
7 instances of kabākay (ka-)	2 instances of `inugbākay (`inug-)
7 instances of pagpamakay (paN-pag-)	1 instance of bayakyan (ay--an)
7 instances of `ikakabakay (ka-CV-`i-)	1 instance of bayakyun (-ay--un)
6 instances of mabakyan (ma--an)	1 instance of biyakyun (-iy--un)
6 instances of nagapabakay (pa-naga-)	1 instance of gapamakay (paN-ga-)
5 instances of ginpabakay (pa-giN-)	1 instance of ginabaki (gina-i)
5 instances of pamakay (paN-)	1 instance of gingpabakay (pa-giN-)
5 instances of pambākay (paN-)	1 instance of ginpangbakay (paN-giN-)
4 instances of binakay (-in-)	1 instance of magabakay (maga-)
4 instances of byakyānun (-ay--anun)	1 instance of magbinākay (in-mag-)
4 instances of miyugbākay (miyug-)	1 instance of magpamakay (paN-mag-)
4 instances of pabakya (pa--a)	1 instance of magpinabākay (pa--in-mag-)
4 instances of pagbinākay (-in-pag-)	1 instance of makapamakay (paN-maka-)
	1 instance of mapabakay (pa-ma-)

But the ability to compile lists of forms along with illustrative sentences in which they occur has presented us with at least one even bigger problem — how to handle these multitudinous forms in a dictionary. For a standard dictionary, used mainly by native speakers to look up spellings or the meanings of relatively rare words, it might be argued that most of these various forms can be ignored or at least minimized. Perhaps only a representative sample would be included. And it might also be argued that even for non-native users, a representative sample would suffice if it were accompanied by a good grammar sketch which would include a description of paradigms of various morphological forms. Though this approach might be marginally adequate for native users, my experience indicates that it is quite inadequate to rely on a grammatical sketch which, for the most part, is able only to handle regularities, to care for most of these forms. Without knowing much more both grammatically and semantically about a given verb than does a beginner in the language, it is simply not possible with any degree of accuracy to predict which affixes naturally occur on a given verb (except, perhaps, for some of the more simple affixes) and, when they do, what the combined meaning is. And to complicate the situation, some inflected forms of many verbs are restricted to specific senses of the lexical form. This requires variation in the function of a given affix when two or more verb senses are involved. In addition to this, some verb affixes have multiple functions when combined with some verbs.

It is necessary, then, to include many of these forms in a learner's dictionary. Since they are forms of single verb lexemes, they need to be grouped together in a systematic way. Forms irregular in shape should have separate minor entries referring the user to the entry where they are described. The following excerpt is taken from the Romblomanon dictionary.

bakay 1 v. ⇒ Someone **buys** an item for sale from someone else, a particular place.

nagbakay, bumakay 'Adu `ang sa kay Pransis ning tambur hay si Gilbert `adtu `ang **nagbakay**. That tambourine that Francis has is what that Gilbert **bought**.

nagapamakay, gapamakay May mutur nga Sibuhānu nga **nagapamakay** sīhi'. There is a Cebuano pumpboat that **is buying** seashells.

mabakay, magabakay *Mabakay* na lang kita ning kagamitan nga human sa kawāyan. We **will just buy** furniture that is made of bamboo.

mamakay Ma `igma' lang `aku kag **mamakay** `aku nang manga gāmit nāmun sa bayay. I will just eat my lunch and **will go buying** things for our house.

magbakay [Follows a negative] 'Ini `ang kandīla' puydi `ini nga `iwag sa bayay kung gab `i pāra nga `indi' ka na **magbakay** ning pitrulyu. It is possible for this candle to be a light at night in the house so that you **will not buy** kerosene.

magbakay, bumakay [Indicates a request or command.] 'Adu `ang dagku' nga tināpay, sa panadirya ka gid **magbakay!** Those large buns, **buy** them at the bakery! **Bumakay** ka na lang ning dāyum kag bunang kag tahi'un `ang `imu manga gisi' nga manga bāru'! You just **buy** a needle and thread and sew your worn out dresses!

⇒ Money is **used** by someone to **buy** an item for sale.

ginbakay, binakay 'Indi' mu gāni' masakar `ang kwarta nga `imu **ginbakay** ning pūkut. You will not get back the money which you **used to buy** the fishnet.

ginbakyan [Follows a negative.] *Waya' da gali' ni Pidru ginbakyan 'ang kwarta kay gin'utang* .Pedro did not use the money to buy because [he] lent [it].

pinamakay *Pinamakay ni 'Ana 'ang kwarta ning baru' para sa 'iya manga 'apu. Ana used the money to frequently buy dresses for her grandchildren. [Prefix pinaN- indicates multiple things bought. Focus is either on that used to buy, illustrated here, or that bought, illustrated below.]*

ginabakay *'Ang narimidyu ku hay ginabakay ku naman ning gatas sa mga 'unga'. My wages are what I use to buy milk for my children.*

'ibakay *'Ibakay ninda ning pagka'un sa 'inda pamilya. They will use [the money] to buy food for their family.*

ma'ibakay [This form follows an existential and *ning*.] *'Igwa kami ning ma'ibakay ning 'asin, puspuru, 'asukar kag habun. We have [money] which [we] will use to buy salt, matches, sugar and soap.*

pangbakay [This form follows a determiner:] *'Kung sinda hay makabaligya' ning 'isda', 'igwa na sinda ning pangbakay ning sigarilyu. If they will be able to sell fish, they will then have [money] to be used for buying cigarettes.*

⇒ An item for sale is bought by someone from someone else, from a particular place.

ginbakay, binakay *'Igwa kami papilis nga pinirmahan; ina' nga duta' hay binakay ku na kay Pinu. We have signed papers; that is the land which was already bought by me from Pino.*

bakya, ginbakya [These forms follow a negative.] *'Ang 'iya 'ini balayan hay waya' ini ginbakya. The shell of this [shell animal] was not bought.*

'bakya, ginbakay, ginbakya [These forms follow the interrogative pronoun *di'in*.] *Di'in niya bakya 'ang sim? Where was the galvanized iron bought by him?*

pinamakay [Prefix *pinaN-* indicates multiple things bought. Focus is either on that bought, illustrated here, or that used to buy, illustrated above.] *'Ang 'inda pinamakay nga pabisti hay kumplitu. The wedding wardrobe which was bought by them is complete.*

ginpangbakay, pinangbakay [These forms follow *nga* and indicate multiple things bought.] *'Ang manga gamit nga ginpangbakay ku hay 'ipadaya ku na lang. As for the many things which were bought, I will let [someone] bring [them].*

ginabakay *Kung 'imu malukat hay 'igwa 'akuy pagasayluha 'ina' nga kwarta kay may ginabakay 'aku nga duta'. If you will redeem [your mortgage] there will be money that I will transfer, because there is land that I am buying.*

ginabakya [This form follows a negative.] *Waya' 'ini ginabakya sa banwa nga baskit. This basket is not being bought in town.*

bakyun *Kung 'ikaw hay nagabakay ning tila 'ina' hay dipindi sa 'imu kung 'anu klasi nga tila 'ang 'imu bakyun. If you buy cloth, that depends on you [as to] what kind of cloth will be bought by you.*

pagbakyun [This form follows a negative.] *Kung waya' da 'ini ning sabat hay 'indi' da pagbakyun nang miyugpamakay ning banig. If there is no design, [that mat] will not be bought by the commercial mat buyers.*

bakyun, bakya [These forms are used in command or request sentences.] *Bakyun mu na lang 'ang plantya ku kag 'adtung pinggan kag t̄asa!* My iron and those plates and cups **should be bought** by you! *Bakya, Idna, 'ang 'isda' kay 'aku mamalantya nang 'ākun ning b̄aru'!* Edna, the fish **should be bought** [by you] because I will iron my clothes!

⇒ **For** someone or something, an item for sale **is bought** by someone else.

ginbakyan, binakyan *Ginbakyan si Birtu ning 'alambri, p̄ara humanun nga tabūnan.* For Berto, chicken wire **was bought**, in order to make a fish trap.

ginabakyan 'Ang 'iya manghud nga dalāga hay **ginabakyan** si Linda ning bag'u nga b̄aru'. As for her younger sister, a new dress **is being bought** for Linda [by her, i.e. by her younger sister].

ginabakyi [This form follows a negative.] *Kung waya' ning sagnat 'u sakit 'ang 'unga' waya' da ginabakyi ning pipsi 'u mirinda 'ūrins.* If a child has no fever or sickness, **for** [him], Pepsi or Mirinda Orange **is not bought**.

bakyan Bakyan ku kamu ning b̄aru' kag sapātus. **For** you, I **will buy** a dress and shoes.

⇒ **It is time** that an item for sale **be bought for** someone or something by someone else.

bayakyan [This present-tense form indicates a focus on both the time and one who benefits by something being bought.] 'Ang 'isa nga 'alādin nāmun hay waya' ning ḡasa, ḡāni' **bayakyan** pa 'adtu ning n̄ipul. As for one of our pressure lamps, there is no mantle, so **it is now time** that a nipple **be bought for** [the generator].

⇒ a person or place **from whom** or **from which** someone **buys** an item for sale

mabakyan *Kung ma'āyu da 'ang tag'īya nga 'īmu mabakyan ning buli hay ginataw'an ka da ning pa'āman.* If the owner **from whom** you **will buy** the buri palm fronds is kind, you will be given an extra frond. 'Ang 'isa ka bilug nga kawāyan hay baynti 'ang ma'intik; waya' ka pa ning **mabakyan**. As for one piece of bamboo, a small one is twenty [pesos]; there is no place **from which** you **will buy**.

bakay, pagbakay part ⇒ **buying** something *Kung 'indi' ma'ūbus bakay 'ang 'utanun nang 'īmu manga kalāpit, 'īmu na 'ini 'ilibud sa 'iban.* If your neighbors do not finish **buying** the vegetables, take them from place to place to others. 'Ang kwarta nga 'āmun ginggāmit sa **pagbakay** ning pukut, hay ging'ambag'ambagan nāmun. The money which we used for **buying** a fishnet, we each contributed.

2 bakay, pagbakay n. ⇒ the **buying price** of something *Sa dayanun kadāmu' 'ang nagapangutana kung pila kunu 'ang 'ākun bakay nang tiyu'.* Along the road, many are asking how much my **buying price** of the puppy [might be]. *Kung nilāta naman 'ang pagbakay nang kamūti hay kwarinta 'ang 'isa ka lāta.* If buying is by the utility can, the **buying price** of sweet potatoes is forty [pesos] for one sixteen-liter measure.

3 'iyugpamākay n. ⇒ **resources for buying** 'Ang 'īmu manga paninda hay 'īmu lang 'ini 'ipa'ūtang sa 'īmu manga kalāpit, waya' ka naman ning '**iyugpamākay** sa banwa.

As for your items for sale, if you let your neighbors get [them] on credit, you will not have resources for buying in town.

4 *miyugpamākay, miyugbākay, minugbākay, minugpamākay* *n.* ⇒ a commercial buyer of items made for sale *Kung waya' da 'ini ning sabat hay 'indi' da pagbakyun nang miyugpamākay ning banig.* If there is no design, [the mats] will not be bought by commercial mat buyers.

5 *kwarta nga 'inugbākay* *n.* ⇒ spending money *Nalū'uy gid sa 'iya 'ang 'iya maninuy, gāni' gintaw'an siya nga dā'an ning kwarta nga 'inugbākay.* His godfather really pitied him, so he was freely given spending money.

3.5. Use of “Natural” Sentence Illustrations

Many dictionaries, and even some learners' dictionaries, have not provided users with natural sentence illustrations. The result is often illustrations that do not convince the discriminating user of the validity of the usage described. *Webster's Collegiate Dictionary*, tenth edition (not a learner's dictionary), gives as one definition of *top* “the part that is nearest in space or time to the source or beginning”. No illustrative sentence is provided. By allowing this definition, one might expect to be able to say “*We arrived at the top of the river by evening*” meaning its headwaters, or “*During the top of Professor Delgado's lecture....*,” meaning the introductory part. Neither of these would probably be naturally formulated by a native English speaker.

There is, however, a tradition among a few dictionary compilers for using natural sentence illustrations. These are sentences produced when speakers or writers are not concentrating on the actual words or forms of sentences. The *Oxford English Dictionary* beginning as early as 1878 has generally followed this policy. This has also been followed more recently by the *Cobuild Dictionary of the English Language*. In describing a second meaning of the word **legless**, for example, *Cobuild* gives the following example: *He was more or less legless already and we were only on our third pint.* From this example sentence alone we can deduce that this refers to someone who is drunk and unable to walk.

It seems clear that illustrative sentences, taken from natural text materials and used to illustrate usage, are necessary to provide a helpful guide in composing sentences.

3.6. Description of Multiple Senses

With the use of a large corpus of texts as a database for compiling a dictionary, there is little chance that a major usage of any given word will be missed. In fact, so much valuable information is assembled for the compiler that a main concern is what to include and what to leave out. But it seems clear that, for one who wishes to use the dictionary as a guide, not only for reading and understanding the object language, but also for speaking and writing it, at least all major sense distinctions must be described and illustrated.

Cobuild has partially, but not completely, followed the principle of placing similar senses and sub-senses close together regardless of parts of speech. Under the lemma **head**, for example, sense 4 is marked as a noun with the following explanation: “If you are the **head** of something such as a company or organization, you are in charge of it and in charge of the people in it. EG ...*the head of the English department.* ⇒ used as an adjective. EG ...*the head gardener...*”

In this case, the noun and adjective are described under the same sense which indicates that, although there is a *formal* distinction between them (signaled by syntax), there is not a *functional* distinction.

However, not until sense 17 does the following appear, marked 'verb': "If you **head** a company, organization, etc. you are the person in charge of it. EG *The firm is headed by John Murray.*" (Note that there is little semantic difference between this sentence and the following: "*John Murray is head of the firm.*") The verb is, in fact, the same sense as the noun and adjective and, in my opinion, should be described with them.

Oxford makes the initial division between parts of speech. The noun is entered under the lemma **head**¹ and the word "head" in the phrase "*head waiter*" is treated as an attributive under this sense. Parts of speech are handled under separate lemmas in the dictionary, recognizing parts of speech as having a primary function and semantics as having a secondary function. Thus the verb form of "head" in the phrase "*head a rebellion, government, delegation*" is found 1 1/2 pages away, under **head**².

Oxford recognizes a word like spinning **top** as semantically unrelated to the various senses of **top** which have to do with physical or figurative orientation, and treats them as homonyms. These are indicated as **top**² and **top**¹, respectively. *Cobuild* does not distinguish between senses of a lexical form and homonyms and lists them serially under the same lexical entry. This avoids decisions about semantic relatedness. However, it is difficult to explain putting "playing top" as sense 9 between "a piece of clothing that a woman wears on the upper half of her body" (sense 8) and "the highest point on something such as a scale" (sense 10).

Romblomanon orders senses according to meaning without regard to parts of speech. Thus under sense 1 of the lemma **hāpun**, for example, the parts of speech occur in the order *noun, verb, noun, noun, adverb, and interjection* respectively for each sub-sense. Senses 2-4 are nouns, and sense 5 is a verb. This is because a dictionary should be primarily a description of meaning, and only secondarily of form.

hāpun 1a n. ⇒ afternoon, i.e., any period from about one p.m. to five p.m. *May yāra' nga nagapanamkun nga ginahagkut kāda hāpun.* There are those who have pregnancy cravings, who feel cold every **afternoon**. (Periods within an afternoon have a possessed, i.e. part-whole, relationship with 'afternoon', beginning about one p.m. 'ala 'ūna nang hāpun to about five p.m. 'alas singku nang hapun.) *Naka'abut 'aku diri sa bayay 'iksaktu nga 'alas tris na nang hāpun.* I was able to reach here at the house **at exactly three o'clock in (lit. of) the afternoon.**

1b v. ⇒ [it] becomes afternoon

maghāpun v. *Ka'ūna 'adtu tanan bāgu-maghāpun.* Eat all of it before [it] becomes afternoon!

1c sa hāpun n ⇒ **this coming afternoon** *Hambay ku hay sīgi, makara' gāling nga nākun kami sa hāpun.* I said, go ahead for we will all come **this afternoon**. See Appendix 32.

1d hāpun na n. ⇒ **late afternoon** *Waya' da si Nīna nakabakay ning balinghuy kay hāpun na.* Nena was unable to buy cassava because it was **late afternoon**.

1e hay pagkahāpun, nadverb, ang pagkahāpun, hay nang pagkahāpun adv. ⇒ **When it was afternoon**, something happened or a state existed. *Hay pagkahāpun*

naman 'ina', Martis na nang hāpun, nagkadtu kami kana Manang Zenaida kay 'imaw 'ang 'āmun 'inirgūhan. **When it was afternoon** again, that Tuesday afternoon we went to older sister Zenaida's place because we told our stories to each other there.

1f ma'āyu nga hāpun, ma'āyung hāpun inj. ⇒ **Good afternoon!** *'Ang hambay ku, "Ma'āyu nga hāpun"; tapus nagsabat da si Nīna sa 'ākun nga "Dāyun".* I said, **"Good afternoon"**; then Nena answered me, "Come in".

2 kahāpun time n. ⇒ **yesterday** *Nakapamāti' kunta' kita ning mīting kahāpun kung waya' mag'uyan.* We would have been able to listen to the meeting **yesterday** if it had not rained.

3 maghāpun n. ⇒ **the whole day** *Manga tuyu ka 'adlaw buyad, 'uga nga dā'an 'ang niyug basta tudutūdu 'ang sīlak maghāpun.* [After] about three days drying, the coconut fruit will be dry then if the sunlight is very strong the **whole day**.

4 kasanghāpun (from hāpun + ka- + 'isa + nga) time n. ⇒ **the day before yesterday**

5 'ihapun v. ⇒ **Someone eats supper.**

nag'ihapun Nag'ihapun kami kag kumatūyug na. We **ate supper** and then slept.

In many cases nouns and verbs have no sense differences at all.

bātuk n. ⇒ An object has **stripes** on its surface. *'Ang pakpak nang pispis nga māyang Tagālug may bātuk nga kapi, puti' kag 'itum.* The wings of tree sparrow birds have dark yellow, white and black **stripes**.

v. An object is **striped**. *'Ang pakpak nang pispis nga māyang Tagālug hay batūkan ning kapi, puti' kag 'itum.* The wings of tree sparrow birds **are striped** with yellow, white and black.

3.7. Idioms and Set Expressions

The following Romblomanon set expressions are cited under the entry *'isa* which, in some contexts, is the cardinal number one:

kung ka 'isa adv. ⇒ **Sometimes** an event occurs or a state exists.

'isa nga buwas time n. ⇒ **the day after tomorrow**

This illustrates the fact that many phrases are best treated as lexical units rather than strings of meaning units.

It is especially common among languages for small words which occur frequently within texts to combine to create new and often unexpected meanings. The word *'put* in English, for example, combines with many other words to derive new meanings, as in *put across, put off, put on, put down, put over, put through, put up, put up with, put up to, put up a fight, put to death, put aside, put away, put back, put upon, put together, put down for, put down to, put forth, put forward, put in, put into (harbor), hard put, stay put, put through (their) paces*. To treat them as strings of lexical units, each word with its own meaning, would, in some cases, involve major semantic difficulty in description and sometimes could not be done at all. The solution is to consider them set expressions with meanings that are not identified with the constituent parts of the phrases.

NEWELL

Phrases that occur frequently, however, involve *degrees* of “setness”, from those in which the meaning of the constituents more or less equals the meaning of the phrase to those in which the meaning of at least one of the constituents bears little or no resemblance to the meaning of the phrase.

3.7.1. The meaning of the constituents equals the meaning of the phrase.

Romblomanon:

bayay *n.* a house

ʼibābaw nang bayay *n.* the upstairs of a house

ʼidāyum nang bayay *n.* the area under a house

luwas nang bayay *n.* the outside ground area of a house

sākup nang bayay, suyud nang bayay *n.* the area inside a house

bayaybāyay *n.* a shed

bayaybāyay nang bābuy *n.* a pig shed

bayaybāyay nga ʼinūman *n.* a drinking shed

bayaybāyay nang manuk *n.* a chicken house

(**bagʼu** *adj.* new, **tūʼig** *n.* a year, **-ng** [var. of *nga*] modifying ligature)

bagʼung tūʼig *n.* the new year

In each case, the phrase meaning is predictable by identifying the meaning of the words involved. Nevertheless, these phrases are included in the dictionary because they have a high frequency of occurrence and would be helpful, especially in the early stages of language learning.

3.7.2. The meaning of the constituents almost equals the meaning of the phrase.

The combined meaning of constituents of a large number of recurring phrases in many languages almost, but does not quite, equal the meaning of the phrase.

Romblomanon:

(**tāwu** person)

bagʼung tāwu *n.* a newborn baby (lit. a new person)

ʼinaʼ *dem pron.* that near; **niyan** *adv.* now

ʼinaʼ niyan *adv.* soon

3.7.3. The meaning of the parts bears *some* resemblance to the meaning of the phrase.

Romblomanon:

(*kasay n. marriage, wedding day, wedding ceremony*)

bag'ung kasay n. a newlywed couple (not: a new marriage)

The word *kung*, at the beginning of dependent clauses, has a conditional meaning 'if' or 'when'. However, in combination with interrogative adverbs and pronouns, it serves to change the interrogative to a relative adverb or relative pronoun.

Bāsi' hay gabinākay siya ning 'itlug didtu sa magūyang?, Why does she keep on buying eggs from that old woman?'

'Ang kahugā'an sa pagpangabuhi' hay 'isa gid nga dahilan nga kung bāsi' 'ang 'isa nga tāwu hay naga'isip nga magkadtu sa 'iban nga lugar. Hardship in life is one reason why some people consider going to other places.

Note the change in meaning (still a set expression) when the order is reversed:

bāsi' kung adv. perhaps, it might be that...

Makadtu na 'aku kay 'alas dus na; bāsi' kung ma'ulihī 'aku sa padāwi. I will go now because it is already twelve o'clock; perhaps I will be late for the time fish bite.

But notice that with *kay* 'because', there are two possibilities:

1) *Kay* functions alone with its own meaning, *bāsi' kung* functions as a set expression:

kay bāsi' kung = because perhaps

Nagdyāgan 'aku kag sīgi 'ang panglikīta' ku sa 'ibābaw kay bāsi' kung may mahūyug nga nyug. I ran and I continued looking above because perhaps there would be a coconut which would fall down.

2) *Kay bāsi'* functions as a set expression, *kung* functions alone:

kay bāsi' kung = 'so that if'

Ginhambayan ku 'ang manga ka'ibāhan ku sa bayay nga maglūtu' nga dā'an mas timprānu, kay bāsi' kung mag'abut hay priparādu na. I told my companions at home to cook right away so that if he comes [we will be] prepared already.

Dī'in ka mabu'uy nang 'imu kaka'ūnun masunud nga 'adlaw? Where will you get that which you will eat the next day?

Wayā' ni Nīna bati' bafī'a kung dī'in na 'ang 'iya nanay. Nena did not hear from any source where her mother is now.

3.8. Additional Semantic, Cultural, and Grammatical Information

To aid in language use, especially a non-native speaker who wishes to speak and write in a natural manner, a learner's dictionary commonly needs to give information in addition to what is found in the definition or lexical gloss and illustrative sentences. This includes general semantic, cultural, and grammatical information.

NEWELL

3.8.1. Semantic information

Romblomanon examples.

Under the entry **bayay** 'house':

A phrase consisting of a determiner and *bayay* without further expansion has the meaning 'our house'. (Note: English 'home' denotes a similar meaning.) *Naka'isi 'aku nga magbāhuy lang gid 'aku hay babayā'an ku gid kamu diri sa bayay* I thought that when I became grown up, I would surely leave you here at **our house**.

Under the entry '*isa* 'one':

Romblomanon numbers from one '*isa* to ten *napūyu* are used in numbering a measurement or object, esp. if it is designated by a Romblomanon word, e.g. '*isa ka dāngaw*, 'one handspan'. Numbers of Spanish words are usually designated by Spanish numbers, e.g. *dus kīlu* 'two kilos'. All numbers from eleven '*unsi* up to ninety-nine *nubinta 'i nuwībi* are Spanish.

Under the entry '*ūnu*:

Spanish numbers, '*ūnu, dus, tris, kwatru*, etc. are used 1) for abstract counting, and numbering things if items are not mentioned; 2) for measuring the time of day (Appendix 22); 3) to indicate money values, both pesos and centavos (Appendix 28); 4) usually for measuring a thing or the length, size, or weight of a thing if it is designated by a Spanish word (Appendix 26); all numbers from eleven '*unsi* up to ninety-nine *nubinta 'i nuwībi* are Spanish.)

Under entry '*ūnsi* and above:

Numbers eleven '*unsi* up to ninety-nine *nubinta 'i nuwībi* Spanish. (For a description of the use of Spanish numbers, see entry '*ūnu*.)

Under entry *ka*²:

Romblomanon numbers from one to ten are linked with a thing or a measure of something by the ligature *ka*; eleven and higher are not usually linked by a ligature: '*unsi 'adlaw* 'eleven days'. Numbers of Spanish origin are not linked by a number ligature: *dyis gantas* 'ten gantas'.

3.8.2. Cultural information

Romblomanon examples:

bayaybāyay nga 'inūman 'a drinking shed' (under *bayaybāyay*)

A shelter next to a convenience store *tyindāhan* where people congregate to socialize and drink alcoholic beverages.

Under **kalag** 'a departed soul, i.e., the soul of a dead person'

A departed soul wanders around and sometimes inflicts sickness *lū'uy* 2 on the living.

Under **tuba** 'coconut wine'

Coconut wine is measured for sale by the glass *bāsu*, four liter measuring bottle *galun*, or one liter measuring bottle *plastik* 2. It is sold in stores *tindāhan*, and markets *mirkādu*.

3.8.3. Grammatical information

Under *bilug*¹ 2b 'the whole area, the whole thing':

An area or item referred to in its entirety commonly follows *nga*. *Bilang bāyus kay Huwan, naghambay 'ang prinsisa nga mapapakadtu kunu sa 'iya 'ang bilug nga ginhari'an*. As a reward to John, the princess said that the whole kingdom area would surely be given to him. *Masakit 'ang 'akun bilug nga lāwas*. My entire body aches.

kung pila 'how much', 'how many,' (sense 2 under entry *pila*):

Followed by the number ligature *ka* and a measure noun. *'Ari nyan ya' ku pa da masasayuri kung pila ka 'iktarya 'ang 'akun ning dūta*. As of now, I really don't yet know how many hectares my land [amounts to].

tanūman 'a planting area,' (sense 1c under *tanum*):

This form is distinguished from the verb form *tanuman* (no length) which, when nominalized by a preceding determiner, has the meaning 'an area which will be planted by someone'.

3.9. Lexical Comparison and Contrast

A fundamental principle of lexicography is contrast. The semantics of a word described without considering other related words often results in a non-native dictionary user's inability to use the word accurately or effectively.

The *Oxford Dictionary* does occasionally cross-reference to other related words. It has a category of words following *Cf* which includes synonyms, antonyms, and generics (with no indication of which category is cited). In addition, it includes many pictures which illustrate kinds-of, part-whole, lexical-field, and chronological relationships, as well as illustrations of synonyms, sense discriminations, and words difficult to describe. But many words are not cross referenced to other words. Under *head*¹, for example, *Oxford* has, as sense 14a, "chief person of a group or organization, etc...[attrib] *head waiter*". And under *top*¹ *Oxford* lists "⇒ adj [usu attrib] highest in position, rank or degree:...one of Britain's top scientists..." But, although definitions are similar, since *head* and *top* are not compared, we do not know if *top waiter* and *head scientist* are well formed.

Cobuild indicates some synonyms (following =), antonyms (following ≠) and generics (following ↑). But in some cases citations of related words are not given, and in other cases citations are not reliable. For example, it has, under the lemma *chief*, "A chief is 1.1 a person who has authority over a group or an organization... EG ...*Jean Ducret, chief of the Presidential Security Corps*..." 'Head' is not given as a synonym of 'chief' under *chief* ... 1.1 which leads one to conclude that *head of the Presidential Security Corps* is not natural English (though it probably is). Under *head* ...4, *Cobuild* has "If you are the head of something such as a company or organization, you are in charge of it and in charge of the

NEWELL

people in it EG ...*I must telephone head office in London.*" The word 'chief' is here cited as a synonym of the adjective 'head'. It should then be acceptable to say *I must telephone chief in London*, which is probably not natural in any English dialect. Cross referencing must be accurate and as complete as possible to serve as a reliable guide for composing sentences.

Cross referencing is displayed by indicating synonyms, antonyms, related words, and by the use of pictures, tables, and appendices.

4. CONCLUSION

If we are to do more than provide information about spellings and the meaning of uncommon words in a dictionary, the task becomes increasingly complex. And if we are to give good information about how beginning users are to use the words found in the dictionary to speak and write in a natural way, the task becomes extremely complex indeed.

We live in an ever-shrinking world that impacts not only international communication, but also communication among various speech communities within nations. We must communicate efficiently in order to understand each other if we are to achieve any degree of success in promoting cultural and social acceptance of each other. A learner's dictionary is a fundamental step in building a bridge between and among cultural communities to accomplish this. The task is large, it is complicated, and we work with imperfect tools. Large amounts of resources are needed in terms of finances and expert personnel well trained in the mechanics and art of dictionary making. But formidable as this task is, we cannot be content with less than our best efforts in accomplishing what must be done. Good learners' dictionaries attest to the finest efforts of communities desiring interlingual and intercultural communion.

NOTES

*Paper presented as the 1998 Andrew Gonzalez, FSC Professorial Chair Lecture in Linguistics, De La Salle University, Manila, Philippines, July 10, 1998.

¹English dictionaries in general describe words rather than lexemes. The present participle actually involves the lexemes *be* and an enclitic *-ing* suffixed to the verb root. Auxiliaries are *have-en* (past participle), *be-ing* (present participle), and *be-en* (passive). Compound tenses are formed by combinations of these three.

REFERENCES

- COLLINS COBUILD ENGLISH LANGUAGE DICTIONARY. 1987. London: Collins ELT.
- MERRIAM-WEBSTER'S COLLEGIATE DICTIONARY. Tenth Edition. 1993.
- NEWELL, LEONARD E. 1995. Handbook on lexicography for Philippine and other Languages. Manila: Linguistic Society of the Philippines.
- OXFORD ADVANCED LEARNER'S DICTIONARY OF CURRENT ENGLISH. Fourth Edition. 1981. Oxford: Oxford University Press.
- SINCLAIR, J.M. 1991. Corpus, concordance, collocation. Oxford: Oxford University Press.