Lexicography Considerations for Tok Pisin

Karl J. Franklin

Contents

Abstract
0. Introduction
   1. Basic notions
   2. Words have histories
   3. Acquisition, analysis and evaluation of data
   4. Some lexical categories of meaning
   5. Some applications
   6. Definitions and the dictionary
   7. Speech acts and definitions
   8. Body parts as initiators
   9. Some conclusions

Selected glossary
Notes
References

Abstract

This article examines a number of issues that are involved in dictionary compilation, including terminology, elicitation, kinds of meanings, lexical relationships, componential analysis and definitions. Definitions may be both folk and technical and the methodology suggested here profits from the semantic framework and theory of Wierzbicka (1985, 1986, 1987), as well as from other sources. Illustrations are given in Tok Pisin as well as in English. The article concludes with a number of practical suggestions which are relevant in compiling dictionaries in the vernacular languages of Papua New Guinea.

0. Introduction
This is a brief overview of Lexicography, the compiling and editing of dictionaries as applied to Tok Pisin. Lexicography is a complex undertaking involving observations of linguistics, translation, anthropology, and sociolinguistics and it requires numerous decisions, many of which might be unanticipated by one just beginning a dictionary compilation. Though focusing on practical lexicography, I mention important issues and point to recommended sources for those wishing to research them in the abundant helpful research available. Without going into detailed definitions I will also identify some basic terminology of lexicography and apply principles of elicitation and discovery techniques to some Tok Pisin and English words.

1. Basic notions

Some authors use the word lexicon instead of dictionary, but in this survey, as generally used elsewhere, lexicon is a technical term used to refer to lexemes, the set of all words and idioms of a language. Some authors use lexicon to refer to the vocabulary of a language as used by an individual speaker, but the term is not so used in this article.

At the heart of lexicology is the relationship between the meanings of lexemes. The task of the lexicologist is to make specific statements about the form, meaning, and usage of the lexemes that appear as entries in the dictionary.

Lexemes are the smallest units of a theoretical lexicon but may also occur as a phrase, a compound word, or in special combinations. Each lexeme is differentiated from other lexemes in the meaning system of a language. A lexeme may be abstract or may occur in different forms, such as bearing different inflections. The dictionary information on a lexeme as a dictionary entry generally includes, as the minimum, its pronunciation, part of speech, inflected forms, and various meanings, generally grouped according to its senses and subsenses. Specialized dictionaries may also include particular collocations (as in Benson et. al 1986).

Though more limited than syntactic or grammatical semantics, the lexical semantics included in a dictionary may describe the distinctive semantic features that characterize lexemes and their semantic fields (Lehrer 1974, Grandy 1987). The semantic field is often organized such that a set of terms represents a particular domain which refers to a speaker's area of experience. The domain can be described according to various organizing principles and semantic relationships.

Various terms are used to refer to meaning, such as: definitional meaning (how the semantic structures of lexemes, as represented in a dictionary, are outlined), referential meaning (resulting in the lexicalization of extralinguistic concepts and objects), conceptual meaning (how the meanings are related as far as the mental image of the speaker is concerned), and collocational meaning (the relationship of words and their meanings in combination or sequence). Concepts in turn are classified as events, things or attributes and are related to one another by case roles (Larson 1984:26, 199-222).

Closely related to semantics is pragmatics, which deals with the speech situations or social usage of utterances (Healey and Healey 1992:185). Within this sub-field are the notions of speech acts (Austin 1962, Searle 1969), with utterances classified according to their locutionary and illocutionary meanings.
2. Words have histories

To gain some idea of the extent of the task of lexicography consider how variable over time a lexicon is. For 1969 alone the *New World Dictionary* publisher examined 250,000 new words. Of these some 1,000 needed to be updated or included with a dictionary definition. In 1983 there were 53,000 book titles published in the U.S., plus newspapers and magazines. All of these contain some new words which are, potentially, included later in dictionaries.

New words are formed in various ways. In examining 6,000 new words for inclusion in the revision of the 1961 *Webster's* (some 12 years after the 1961 version was published) most of the additions were due either to affixation (2/5 of them) or compounding (some 3/5 of them). In other instances writers simply coin new words. Linguists are no exception: see Crystal (1993) and Richards (1985) for terminology which has been recorded in linguistics recently.

Lexicographers are also interested in the history of words. In the vernacular languages of parts of Papua New Guinea (henceforth PNG), an increasing number of words have been borrowed from Tok Pisin (henceforth TP) and can be fairly easily recognized. The borrowings often show reworking of the phonemes, occur as lexical and syntactic blends, or have new meanings. In much the same way that the vernaculars borrow from TP, TP has in turn borrowed from English. These new words quite naturally include technical terms from a variety of trades and disciplines.

Words also have an internal history: they are borrowed or develop locally in response to factors like word taboo or social identity. The words or expressions can be identified with various dialects, such that lexicographers can plot occurrences in terms of "dialect geography" (Anderson 1987, Kirk et al. 1985, Stevens 1978). It is possible to document how non-native speakers use English quite differently; their differences often reflect the pronunciation and intonation of their mother tongues or may correlate with the location and social position of the speakers of a language.

For a dictionary to be useful the compiler must decide how to incorporate many of these sociolinguistic factors because, properly conceived and developed, a dictionary is a code to the whole language. Therefore, lexicographers are guided not only by what they find in formal parameters of grammar; they go beyond these areas into, semantics and pragmatics as well.

3. Acquisition, analysis and evaluation of data

Elicitation is but one aspect of the process of discovering words or expressions and their meanings. The compiler must arrange the data systematically according to particular cultural domains to better determine contextual constraints on any word or phrase. Observations are verified by checking the data with other native speakers and experts.

Data acquisition for dictionaries can be aided by ethnosemantics, Hammel 1965, Nader 1965, and Romney and D'Andrade 1964, are early examples. Various procedures, such as question and answer dialogues, employing the use of lists and experts, constructing non-trivial questions for every recorded observation answer, and systematically studying the components of any given
lexeme, are helpful to understand how meanings are used in a culture. These procedures can be supplemented with others, including steps recommended for componential analysis (Nida 1975), constructing folk taxonomies (Spradley 1979, 1980), conceiving of cultural activities in terms of scripts and plans (Schank and Abelson 1977), and so on. Regardless of the methodology some evaluation of the description is needed.

The implications for the dictionary compiler are that first-hand observation and cultural participation are necessary to hypothesize about the meanings of words. Long-term field workers who are concerned with cultural knowledge and usage have an advantage in gaining the necessary insights because one must go beyond mere glossing in a second language; one must participate in the culture in order to use the forms and expressions properly. A broad interest in the culture is imperative.

Dictionary compilation is an experimental process that can be verified. By handing a TP dictionary to language assistants and having them fill out the vernacular equivalents, a field worker is conducting an experiment that can be replicated. But this is not an exploration of the semantic features of the lexemes, nor would it constitute a rigorous control of the definitions.

Strictly speaking, providing good definitions is the rationale of monolingual dictionaries, i.e. the lexicographer must explain meanings in the language using its own metavocabulary. In PNG languages the existing dictionaries are generally bilingual or multilingual and the so-called definitions are the best translation equivalents that speakers and compilers can manage at the time. By way of contrast, literacy books for advanced readers in vernacular languages (e.g., K. Franklin and Y. Kirapeasi 1975) assume an intuitive understanding of lexemes.

Proper elicitation techniques are an initial step in dictionary making because they help field workers discover key semantic components in various cultural domains.

Some years ago Nida (1964) classified words into four semantic sets: objects, events, qualities, and relationships. More recently (Nida and Louw 1992:17) these are called entities (objects and participants), activities (events and happenings), characteristics (quantities and qualities) and relations. Probably anyone involved in translation is acquainted with the restructuring of semantic relationships which employs these basic classes. In a previous work Nida (1975) began with the fundamental division of objects v. events, then subdivided objects into those which are count and those which are not, i.e. mass objects. Countable objects included both animate and inanimate while non-countable objects incorporated liquids, gas, and dry substances. Animate objects were divided as plant, animal, and persons, while inanimates were manufactured, natural, and mobile (natural, movable, and not movable). Events were partitioned into participants on the one hand (communicative, physiological, physical, position, change of state, change of position, and making) and abstracts on the other hand (psychological, sensory, and meteorological).

These basic and general divisions of the real world have varying conceptual and cognitive representations in cultures.

4. Some lexical categories of meaning
4.1 Denotations

A denotative meaning relates the lexeme to the real world. For example, the Merriam Webster dictionary (1993) reports that a cassowary is

"any of a genus (Casuaris) of large ratite birds chiefly of New Guinea and northern Australia that have a horny casque on the head and closely related to the emu."

Both the cassowary and the emu are considered to be real-world objects. The cassowary is given taxonomic or referential status by comparing it to the emu which is, in turn, defined (in Merriam Webster 1993) as "a swift-running Australian bird (Domiceius novae-hollandiae) with underdeveloped wings and is related to and smaller than the ostrich".

Turning to ostrich, we read that it is a

"swift-footed 2-toed flightless ratite bird that has valuable wing and tail plumes, and is the largest of existing birds, often weighing 300 pounds" (Merriam Webster 1993).

But does a dictionary user need to know what ratite means to know how to identify an ostrich, emu, or cassowary? No, because the term ratite is part of the scientific metalexicon used to place birds within a scientific taxonomy. The term is used in the superorder of Ratitae to refer to birds with a flat breastbone, including the kiwi and moa as well. The term is helpful to scientists, but unnecessary in practical, folk-lexicography.

Denotative meaning is viewed as the central or core meaning of a lexical item and is generally equated with the referential (i.e. cognitive or conceptual) meaning. The average dictionary user is not concerned with the fact that a cassowary is a ratite. In calling the bird a ratite the dictionary compilers were attempting to place not only the cassowary, but all flora and fauna within a particular taxonomy, one that is known and generally accepted by the scientific community.

Dictionary entries can achieve classification in two main ways:

1. by providing the reader with a dictionary description of how the items are placed in the folk taxonomy;
2. in the case of flora and fauna, by marking the "scientific" name and classification in some special way.

Taylor (1989:68) refers to this dual classification as a combination of folk categories and expert categories. For example, according to research in parts of PNG (Majnep and Bulmer 1977) expert categories identify those technical features which place certain birds in a taxonomy having at the most general level the contrasting features of [+FLIGHT] or [-FLIGHT]. Birds such as the hawk or sparrow fly but others like the cassowary or emu do not.
For many people, folk categories are more useful. It is doubtful that average English speakers could develop the taxonomy of a cassowary further and supply the scientific taxonomic label for this group, or that their knowledge of ratite would serve in any defining process. Further, they are unlikely to list a kiwi in the same group as an ostrich.

The expression a cassowary is a muruk (TP, compare Mihalic 1971 and Mihalic and Verhaar, forthcoming) contains corresponding denotative synonyms for the two words. Both refer to the same object in the real world and there is a direct equivalence between the object of the real world and the word used in the translation. I can call the thing a cassowary or a muruk, depending on my language, because each word stands for the same cognitive object, refers to something which is [-flight], and which may, depending on the culture, belong to the class of all birds. Folk taxonomies help us to know which features to include in the defining process.

4.2. Connotative considerations

Connotations are supplementary meanings which extend beyond the central linguistic one. They represent people's emotions or attitudes about, say, a cassowary. For example, in PNG the cassowary is often associated with cultural beliefs, which show up in legends, idioms, comments or observations on its food, strength, habitat, and usefulness.

For example, connotative meanings would be that a cassowary is prized for its plumes, or that (in some countries) a dog is "a household animal which some people love so much that they allow them to sleep on their beds". Connotative meanings are open-ended according to the culture, the historical period and the individual experience.

A further type of connotative meaning is stylistic, i.e. it describes or implies the social circumstances that relate to the use of the lexeme. A word may be used in a particular location, be written rather than spoken, be colloquial, slang, or used in jokes and riddles.

Consider, for example, the word yana in Kewa. It is the most widespread form for "dog" and has cognates as far away as the Enga Province. However, dog has many taboo lexemes as well, especially near the areas where wild dogs are featured in stories. Therefore forms such as irikai, pudiala, ponape, iya, riale, usapu, iriga, and yapa tala are used as well. These are stylistic forms because factors such as intonation accompany the connotative or associational meanings. The lexeme igira refers to a hopeless dog, one which will not hunt. The word is intended as an insult or curse, thus communicating the feelings and attitudes of the speaker. Likewise the term yapa tala "it is hitting/killing the possum" can be abusive because a semantic reversal is intended whereby the hearer understands that the dog is not, in fact, a good hunter.[5]

4.3. Encyclopedic meaning

There is a difference between the kinds of meanings which would be found in an encyclopedia as opposed to a dictionary because categories in the latter are more narrowly defined.

A dictionary is
"a reference book containing words usually alphabetically arranged with information about their forms, pronunciations, functions, etymologies, meanings, and syntactical and idiomatic uses (Merriam Webster 1993)."

An encyclopedia, on the other hand, is

"a work that contains information on all branches of knowledge or treats comprehensively a particular branch of knowledge, usually arranged alphabetically and often by subject" (Merriam Webster 1993).

For example, in an older edition of the Encyclopedia Britannica (1963) there are seven main sections in the entry on dogs, including their origin and history, behavior, reproduction, genetics and breeding, care and training, breeders' associations, dog shows, as well as classifications and breeds. Classification is subdivided into English and European, as well as American, and uses the components of sporting, hound, working, Terrier, Toy, non-sporting, and other varieties (like those around a village or town). There are almost 9 pages about dogs plus 8 plates with 55 pictures. There are 24 sporting varieties, 19 kinds of hounds, 28 working (or sled) dogs, 20 kinds of terriers, 16 toy varieties and 10 which are simply non-sporting.

But even an encyclopedia does not presume to tell the reader everything. The variety of dog called Lhasa Apso is not described, although the reader is told (in a dictionary) that it comes from Tibet. The reader would have to consult a book which specializes in dogs to learn more.

Some linguists (Haiman 1980) do not believe that there can be any theoretical division between an encyclopedia and a dictionary. Practically, however, space, time and cost limit what a compiler can include in a dictionary. There are theoretical problems as well. Taylor (1989:81) asks how one can draw the line between what a speaker knows from the language and what the speaker knows from acquaintance with the world. Taylor believes that if the latter is taken into consideration then features are added which provide an indefinitely long definition in a particular category.

4.4. Relationships between lexemes

There are many ways to study the relationships which lexemes hold one to another. Cruse (1986:86ff.) speaks of congruence relations and variants (noting class identity, inclusion, overlap, and disjunction), cognitive synonymy, hyponymy, as well as lexical compatibility and incompatibility. Often lexical relations are not clear cut and in such instances Cruse talks of partial, quasi-, pseudo-, and para- relations.

Relationships can also be defined by means of paradigmatic and syntagmatic substitutions. These help to determine collocations (e.g. `she died' v. `she passed away', which fit into the same semantic frame); idioms (`she kicked the bucket' fits into the same frame as `she died', but has a different effect); synonymy (`die' and `not alive'): antonymy (`die' v. `live'); relational opposites (`bury' v. `resurrect'); and semantic fields [DEATH]. All of these help to classify meanings of particular lexemes.
Other kinds of relationships include **hierarchical**, such as taxonomies, and **part-whole**, which Cruse (1986:157) calls meronomies; for example, the `paw' of a dog is part of its `leg'. These may show **class inclusion**, as opposed to **class membership**; for example, a `dog' is a kind of `animal'.

**Class meaning** results from morphological or syntactic analysis (Nida and Louw 1992:62). Here the meaning is generalized for a given class such as a grammatical category, like `tense' for present, past, future, or other relevant sets.

The semantic reading of a lexeme includes a list of those **markers** which distinguish the lexeme's meaning from that of something else in the same general class or category, e.g. proper vs. count, or count vs. mass. According to Lakoff, **markedness** is a feature of asymmetry within categories and represents a kind of **prototype** effect (1987:60).

Greenberg has written extensively on **markedness** (1963, 1966). He shows that the unmarked category occurs with the higher frequency in languages. It is considered to be a more basic or implied feature. The more complex the marking of a lexeme, the less that lexeme is used. In grammar, for example, the third person is the least marked, the second is the most marked, and the first person is intermediate. In adjectives the positive is unmarked in English, but the comparative and superlative are marked.

Markedness (Andrews 1990) is thought to be a universal property of all languages. Greenberg suggests ten criteria that relate to markedness in the grammar or lexicon. Some of these are applicable to a language like **TP**:

1. If there is a marked category there is always an unmarked. This is the universal implicational statement. For example, if there is a plural (marked) there will also be a singular.
2. The unmarked may have zero expression, e.g. *prince* v. *princess*.
3. There may be a generic term for an entire category as the opposite of the marked, e.g. man vs. woman. This is what is called the **par excellence** expression.
4. The marked as overt may be optional, such as author to mean either sex v. authoress to be marked as female.
5. Certain kinds of marking may be avoided due to neutralization or syncretization, e.g. *he*, *she*, or *it* in the singular but only *they* in the plural.
6. There may be contextual neutralization such that the unmarked appears only in certain environments.
7. The unmarked may contain more morphological alternation.
8. The unmarked may represent more categories.
9. The unmarked may be dominant and stand for the whole category, e.g. in Spanish *los padres* can mean `the parents, the father and mother', but *las madres* means only `the mothers'.
10. The unmarked is more frequent in texts.

**5. Some applications**
Various techniques can be used to determine related meanings, which in turn assist in establishing the most general definition of a lexeme. To give some idea of the problem, as well as methodology, I first examine the lexeme *house* in English, followed by *haus* and some of its related forms in TP.

5.1 House in English

In the *Macmillan dictionary for children* (1982), the definition for *house* is: "a building in which people live".

Assumed as already understood in the definition are the words *building*, *people*, and *live*. However, the definition of *building* is that it is "something *built".* The definition of *build* (built is a variant listed) is "to make by putting *parts* or materials together."

Consider further the definitions of other lexemes which were used to explain what a house is: *people* are "men, women and children or persons". *Persons* are in turn a "man, woman or child". A *human being* is defined as "having to do with a *person* or *persons"*, or alternatively as "being a *person* or *persons".

The most difficult part of the original definition has to do with *live*, which, in turn, is defined extensionally as:

(1) to be *alive*, to have *life*, (2) to stay *alive*, (3) to *support* oneself, (4) to *feed*, and (5) to make one's *home*.

However, this is not end of the lexeme *live*. Other supplementary components add the subsenses of:

(1) having *life* or *living*, (2) *burning*, (3) carrying an *electric current*, (4) seen while *actually happening*, (5) not on *tape*.

Related to *live* is *alive*, which refers to something

(1) having *life*, *living*, (2) having *power*; *active*.

A further relationship of *alive* is with the word *life*, which is defined in a number of ways:

(1) *quality* of life that only *plants* and *animals* have; *quality* of life which makes it possible for them to *reproduce*, (2) a *living being*, a *person*, (3) the *period* from *birth* to *death*, (4) the *period* during which something *lasts* or *works*, (5) a *way* of living, (6) a *story* of a person's life, (7) *energy*, *spirit*.

This is as far as I shall go to make the point: it is impossible to understand definitions without certain basic cultural reference points. Other lexemes, such as the verb *make* in the first definition (which can also be deverbalized), were discounted and its relationships to *put* and the latter's association to *together* were not pursued. The various senses of *parts* and *material*, or
how man, woman, children (or child), and person are related but were not examined either. This is because there is no natural end-point in the semantic relationships which exist in often circular dictionary definitions (Wierzbicka 1987).

5.2 Haus in TP

Mihalic (1971:96-7) gives the following definition for haus, with nearly 50 subentries subsequent to it:

    haus, (E) house, home, dwelling, hut, shed,

where the "definition" of haus is a near translation into English of a family of synonyms, i.e. house, home, etc. As Nida and Louw (1992:43) demonstrate, a list of glosses does not resolve the differences of meanings in specific contexts, nor do the glosses help to show how sets of particular meanings of a lexeme may cluster. Leaving aside for the moment the problem of what a monolingual definition of haus would be, let us outline some methodology that can be used to examine the meanings of it.

First of all, compare only particular meanings of the subentries (Nida 1975:55ff.) and not the meanings of the set of all 'houses'. Meanings from the same semantic domain are examined, e.g. haus ka 'garage' and haus balus 'hangar', but neither of these compare with haus blut 'menstrual hut' or haus mani 'bank'.

To group the kinds of houses in some satisfactory (i.e. consistent and replicable) way, I first attempt to distinguish those components which define the set. Certain of the components are diagnostic, i.e. essential for the contrast, while others will be supplementary. They will be important for an extensive definition but are not treated as essential components. No more components than necessary are used to establish differences between the lexemes. Sometimes it is helpful to represent features or components with plus and minus symbols, but I do not consider features as strictly binary. Some sets are complex and incorporate functional features as well.

In the following discussion I am using the term domain in a general, even abstract, sense. Taylor (1989:85) speaks of certain basic domains that are not reducible to more primitive cognitive structures. He interprets Lakoff's models (Lakoff 1987:74ff.) as domains. Crystal (1993:112), on the other hand, equates semantic domains with semantic fields.

The following two frames are useful in establishing contrasts within a domain that employs the lexeme haus:

1. X lives in a______________. [HABITAT]
2. Y is placed in a______________. [STORAGE]

These frames establish that certain types of objects are placed, or events take place, in certain kinds of houses. Inanimate objects are kept in houses, while humans and animals live in houses. For example, a ka, balus, or kopra (garage, hangar, and copra shed) are all [+STORAGE] kinds
of houses, while man, kiap, boi, sista, and marit (men's dormitory, government officer's dwelling, servant's quarters, nun's quarters, and married quarters) are [+HUMAN, +HABITAT] and spaida, kakaruk, pik, pisin, bulamakau (spider web, chicken pen, pig pen, bird nest, barn) are [-HUMAN, +HABITAT].

By considering other instances of haus additional components can be suggested from this frame:

3. Z takes place in a ___________. [ACTIVITY]

Here certain other features are established: [TASK] for houses which are described secondarily as kat, skul, holi, kamda, kuk, tambaran, or pepa (surgery, classroom, church, workshop, kitchen, spirit house, and office); [PERSONAL] or [EXPERIENTIAL] for those which include blut, pekpek, slip, or sir (menstrual hut, toilet, bedroom hospital); the component [DISPOSAL] for those which are marked as marasin, ais, lait, mani, dring, or kaikai (dispensary, freezer, power plant, bank, hotel, restaurant).

There are various classification problems which emerge, based on such an analysis. For example, are all the 'houses' which contain people of the same type? What further components need to be added? Does, for example, the fact that a haus kiap was classically "a temporary dwelling for officers who travel to an area" suggest additional components?

5.3 Haus in syntax

Definitions are also understood in terms of their syntactic context. Consider the following examples of haus from the TP New Testament:

(a) N + N (functioning as ADJ) = kind of house, where the following serve as submembers of the set and define certain functions:

- haus holi (holy) = temple, sanctuary, Most Holy Place (house holy)
- haus kaikai (food) = barn(s) haus bilong holim kaikai (house food)
- haus lotu (worship) = synagogue (house worship)
- haus sel /lotu sel (canvas or cloth) = shelter, tabernacle, tent, shrine (house canvas)
- haus takis (tax) = booth (house taxes)
- haus matmat (grave) = tomb (house grave)
- haus pasindia (guest) = Inn (house guest)
- haus kalabus (enclose) = jail
- haus skul (school) = lecture hall
- haus tambu (ancestor) = shrine (house ancestor)
- haus tambaran (spirit) = idol(s) (house spirit)

(b) N + bilong + NP (where NP includes haus), and is the pivotal member in the description. All of these are understood metaphorically and are candidates for the compounding syntax that occurs in (a) above:
manmeri bilong dispela haus ( > haus famili) = people from this house > house family > family
papa bilong haus ( > haus papa) = father of the house > owner
bun bilong dispela haus ( > haus bun) = bone of this house > cornerstone
pos bilong haus ( > haus pos) = posts of the house > posts; foundation

(c) haus bilong NP, where haus refers to a specific location. Several of these are candidates for compounding as well:

haus bilong ol king ( > haus king) = house for all kings > palace
haus bilong en stre ( > ?haus stre) = his very own house > hometown
haus bilong prea ( > haus prea) = house for prayer > prayer house
haus bilong yupela ( >haus Juda) = your own place > Jerusalem
haus bilong Devit (i pundaun) = house of David's > tent
haus bilong mi yet = my very own house
haus bilong Holi Spirit ( > haus Holi Spirit) = house of the Holy Spirit > spiritual house

There are other sets which are descriptive phrases and are not newly or potentially derived instances:

(d) LOCATION + haus or haus + LOC:
antap long haus = on top of the house > housetop
stap long haus oltaim = to live in the house always > permanent place
dua autsait long haus = the door which is outside of the house > entrance
haus antap long pos = house which is on top of the posts > foundation post (cf. pos bilong haus or haus pos, above)

(e) ADJ + haus (cf. N + N):
longpela haus = a long or high house > tower
Tripela haus pasindia = three guest houses > Three Taverns
dispela haus hia = this house here > (our) present body
bikpela haus king = a large house of the king > palace of the King
wanpela haus i bruk long tupela lain = a single house that has been divided into two groups > Kingdom which is divided

(f) haus + VP/ VP + haus, in which there is a complement:
haus i kamap doti = house which has become dirty > defiled
haus God i givim mipela = house which God has given us > earthly tent
haus man i no bin wokim long han = house which man has not made by hand > eternal
pisin i gat haus = bird which has a house > nest
yupela i kamap haus = you all have become a house > dwelling

(g) Other instances:
The compounding system in English, which TP often follows (Mühlhäusler 1979), is complex. Consider, for example, some of the derivations of the domain house which are typically found in an English dictionary: house-boat, house-boy, house-break, house-clean, house-coat, house-fly, house-guest, house-hold, house-keep, house-maid, house-mother, house-party, house-physician, house-plant, house-seat, house-warming, house-wife, house-work, house-of-cards, house-of-correction, house-of Lords, house-of representatives, housing development, housing estate, housing project, and so on.

This does not even consider idioms such as on the house, the House of God, house of ill fame, a full house, bring the house down, eat someone out of house and home, a half-way house, house arrest, put one's house in order, and like a house on fire.

Few of these English compounds have been transferred into TP and those that have similar formations mean something different, e.g.:

E house-boat `a boat which serves as a house'
TP haus bot `a storage shed for a boat'

E house-boy `a servant who works indoors'
TP haus boi `quarters for labourers'

To establish the semantic components the procedure is again to deal with families of lexemes and to group those of like meanings with others:

1. house boat is assigned to a group which includes liner, cruiser, barge, sailboat, or anything which sails on the water and holds people;
2. house-boy groups with servant, maid, butler;
3. house-train groups with potty-train;
4. house-clean with spring-clean;
5. house-coat with garment, and so on.

Rather than common components there may be links or chains of components that are common to a subset. Further, as they are examined the componential connections will show dialectal changes or that a word or expression is in the process of change.

Certain components will be unordered while others will show well-defined and ordered chains or constellations of meanings. The analyst first establishes the primary components, then any unordered meanings, then attempts to arrange these in a logical order of dependency. This process will assist in the definition of the meanings.

Componential analysis is helpful in delimiting the cultural components. In Kewa, for example, trees have been defined according to their
1. location, size, fruit (which may be plenteous or sparse, edible or not);
2. for human consumption or not;
3. difficult to obtain, have a certain colour, etc.;
4. function, such that the skin is used for ropes, bilum; leaves for decoration, mumu, tangles; or more generally for houses, fences, sugar cane poles, firewood, etc.);
5. colour (of both the body of the tree and the leaves);
6. varieties (i.e. relationships, similarities);
7. history of introduction to the area;
8. texture (of the bark, leaves, e.g. twisted, scaly, fuzzy, smooth, prickly, or nondescript);
9. strength; and even the
10. nature of the sap.

It is possible to divide hundreds of trees simply on the basis of these diagnostic components.

6. Definitions and the dictionary

Definitions cannot be compiled or understood simply by examining the components of words in a dictionary. I once saw a definition of "lifeboat" on a community school blackboard in the Southern Highlands which read "enjoyable boats filled with happiness." It highlights some of the problems of using the dictionary to form new definitions. Generally the task of defining is one of the more neglected aspects of dictionary making.

Translating monolingual definitions for a dictionary, rather than simply giving cross-language equivalents, helps the analyst to understand the cultural function of a lexeme. One of the duties of the lexicographer is to edit and formulate definitions in the dictionary in such a way that the reader can easily find the materials wanted. Dictionary users will want a variety of information, including pronunciation, spelling, grammatical class, synonyms (or antonyms), or etymology, which traces the origin of words. But basically the user is concerned with meanings. The word must therefore be defined clearly enough that ambiguities are resolved and the word can be placed in its proper context. Ilson (1987:71) writes that a "definition is the de-lexicalization of a lexical unit into semantic and syntactic components which are then presented in a single phrase...". The phrase is then said to characterize the lexeme, both semantically and syntactically. There are a number of principles which are used when lexicographers consider definitions (Ayto 1983, Wierzbicka 1987):

1. Avoid circularity, where the reader is led on a wild goose chase through the dictionary.
2. Define every word that is used in the definition; however, in practice this is rarely achieved. Certain words are the primitive members of the definition.
3. Define the word rather than simply talking about it, illustrating it, or telling how it is used.

To achieve these goals, definitions often will show the priority of essence, where the essential meanings are given first; they will be simple and brief, avoiding ambiguity; their grammatical function will be shown and the semantic substitutability of the items will be demonstrated.

6.1 Haus in TP
Returning to the TP word haus: how would this be described in a monolingual dictionary to show the components which were listed earlier? At the most generic (inclusive) level:

**haus**: kain ples we ol pipel ol i ken sindaun na kaikai o slip long en. "A kind of place where all the people can sit and eat or sleep in it."

Each of the subentries is then "marked" with supplementary semantic information, such as:

**haus kapa**: wane pela kain haus ol i wokim long kapa antap long ruf bilong en na bai i stap longtaim. "A kind of house which they have made with an iron roof so that it will last a long time."

**haus kunai**: wane pela kain haus ol i wokim long kunai antap long ruf bilong en na bai i stap 4 yia samting. "A kind of house which has been made with a grass roof and which will last about four years."

**haus sel**: wane pela kain haus ol i putim na rausim kwiktaim long wokim long laplap samting. "A kind of house which can be put up and taken down quickly which is made out of canvas-like material."

**haus ais**: wane pela liklik haus o rum bilong kolim kaikai o samting na bai i no inap sting kwiktaim. "A small house or room which cools the food and things so that they will not spoil quickly."

**haus balus**: wane pela bikpela haus bilong holim balus o bilong ol mekanik i ken fiksim balus. "A large house which can contain an airplane or where mechanics can fix airplanes."

**haus pisin**: wane pela liklik kain haus ol pisin i wokim long diwai stik o lip samting. "A small house which is built by birds out of sticks or leaves."

**haus spaida**: wane pela liklik haus ol spaida i wokim long liklik rop bilong em yet. "A small house which spiders make out of their own thread."

Are these definitions sufficient? Do they account for the fact that, for example, a haus sel is used "taim ol kiap i wokabout sampela ol i yusim dispela kain haus na ami tu ol i putim haus olsem" = "when the government officers hike around some of them use this kind of house and the army also puts up houses of this type"? In the same way, have the features of the hangar, cooling room, and so on been distinguished? When can the lexicographer judge that haus and its derivations have sufficient definitions?

In the case of **haus** there are two senses, which incorporate features noted earlier (i.e. habitat and storage):
1. wanpela kain ples we manmeri i ken stap long en. I gat ruf na windua na dua. "A kind of place where people can live. It has a roof, windows and doors."

2. wanpela kain ples we manmeri i ken putim sampela samting bilong ol. "A kind of place where people can put things that belong to them."

This would account for living and interaction on the one hand and shelter or cover for protection on the other hand. It introduces a new item to be defined however, viz. **ples**. In TP "**ples**" is defined as place, village, region, area, or town. It groups, diagnostically, with ples ____, where one can have **balus, bung, daun, kanaka, kik, kol, kunai, pilai, singsing, waswas, or nating**, (hangar, meeting, depression, native, playing field, freezer, open field, playing area, dance grounds, washing area, non-descript area) but not occurring with **kapa** (iron).

### 6.2. Folk definitions

It is helpful to discover some definitions on the basis of the semantic relationships which native speakers give in their classifications of language items and actions (cf. [Casagrande and Hale 1967](#)). For example, note the following relationships:

1. **Attributive**: X is defined with respect to one or more distinctive or characteristic attributes Y. Haus i gat ruf na windua. "A house has a roof and windows."

2. **Contingency**: X is defined with relation to a usual or necessary antecedent or concomitant Y. Sapos manmeri i gat sik bai ol i go long haus sik. "If people are sick they go to the hospital."

3. **Function**: X is defined as the means of effecting Y. Sampela manmeri ol i slip long haus pasinda. "Some people sleep in the guest house."

4. **Spatial**: X is oriented spatially with respect to Y. Haus tambaran i stap klostu long ples singsing. "The ancestral house is located near the dance grounds."

5. **Operational**: X is defined with respect to an action Y of which it is a characteristic goal or recipient. Haus lotu em i ples bilong bungim ol manmeri long lotu. "Church is where people gather to worship."

6. **Comparison**: X is defined in terms of similarity with Y. Haus win i olsem haus kunai tasol i no gat blain samting long arede. "An outdoor covered house is like a temporary house but it doesn't have sides on it."

7. **Exemplification**: X is defined by citing an appropriate co-occurrent, Y. Haus ov Asembli em i wanpela bikpela haus gavman. "The house of Assembly is a very big government house."

8. **Class Inclusion**: X is defined with respect to its membership in a hierarchical class Y. Haus kat i wanpela rum i stap insait long haus sik. "The surgery is a room that is inside a hospital."

9. **Synonymy**: X is defined as being equivalent to Y. Haus lotu ol i save kolim sios. "A church is also called a church."

10. **Antonym**: X is defined as the negation of Y. Haus marit em i no ples bilong ol singel man o meri. "Married quarters are not for single men or women."

11. **Provenience**: X is defined with respect to its placement in a series or spectrum that also includes Y. Dispela kain haus i olsem haus bilong pisin, haus bilong spaida, haus bilong bulamakau. "This type of house is like that of the birds, spiders, or cows."
12. **Circularity**: X is defined as X. Em wapela kain haus. "It is a kind of house."

_Casagrande and Hale (1967)_ ask to what extent these semantic relationships are universals of language behaviour and would naturally show up in folk definitions. Are they routinely used by professional lexicographers? There may also be certain relationships that reflect different cognitive styles or worldviews. Certain relationships may associate with certain grammatical classes, e.g. antonym with adjectives, contingency with verbs, attributive and class inclusion with plants and animals. The semantic relationships delimited by Casagrande and Hale provide a useful heuristic for analyzing native-speaker definitions.

_Laycock (1977:177)_ gives examples of several TP definitions in earlier dictionaries. For example the definition of anis 'ant' includes the features of

1. liklik samting "something small" [small object];
2. stap long graun or antap long diwai "on the ground" or "in the trees" [habitat],
3. gat lek na han, tit i strong, "has legs and hands, the teeth are powerful" and ret, blak, smolpela, "red, black, small" and bikpela "large" [physical characteristics].

### 6.3 Testing and compiling folk definitions

_Wierzbicka (1985:146ff)_ has described a method which can help elicit and test folk definitions. She asks native speakers two questions (when considering objects) (1) Imagining an X, what would people say about it?, and (2) Imagining an X, what could people say about it?

Following Wierzbicka:

1. Imagining a **hand**, what **would** people say about it:
   
   - It is attached to the wrist.
   - It has (4/5) fingers and a thumb.
   - There are two of them.
   - There are other parts of it: nails, knuckles, etc.
   - They are used to grasp objects.
   - They are able to feel things.
   - etc.

2. Imagining a **hand**, what **could** people say about it:
   
   - It is larger than a finger.
   - It is used for making objects.
   - People wear rings/ gloves on them.
   - People wave goodbye with them.
   - etc.

What dimensions/ components enter into the folk definitions:
Now consider a dictionary definition, which states (in the primary or first sense) that a **hand** is:

"the terminal part of the vertebrate forelimb when modified (as in humans) as a grasping organ" (Merriam Webster 1993).

However, for a fieldworker studying and learning the English language, questions immediately raised are:

1. What is a **vertebrate**, a **forelimb**, or an **organ**?
2. What happens when something is **modified** or **grasped**?

### 6.4 Derivations and definitions

Merriam Webster (1993) subdivides the definition of **hand** into 12 senses and multiple subsenses for the noun alone, plus additional senses for the transitive verb.

Following the senses for the noun there are numerous prepositional idioms, such as: at **hand**, at the **hands** of, by **hand**, in **hand**, off one's **hands**, on all **hands** or on every **hand**, on one's **hands**, out of **hand**, to **hand**, and with a heavy **hand**.

There then follows a listing of various idioms and compounds, for example: **hand** and foot, **hand** axe, **hand** bag, **hand** ball, **hand** bill, **hand** book, **hand** cart, **hand** cheese, **hand** clasp, **hand** craft, **hand** cuff (v. **hand** cuff), **hand** down, **handed**, **handedness**, **hand** fast, **hand** feed, **hand** ful, **hand** glass, **hand** grip, **hand** gun, **hand** hold, **hand** icap, **hand** icapper, **hand** icer, **hand** ily, **hand** in glove, **hand** in **hand**, **hand** iwork, **hand** kerchief, **hand** le (several senses), **hand** le breadth, **hand** le down, **hand** le off, **hand** le out, **hand** le over, **hand** le over fist, **hand** pick, **hand** press, **hand** print, **hand** puppet, **hand** rail, **hand** saw, **hand** sbreadth, **hands** -down, **hand** set, **hand** shake, **hands** some, **hand** spring, **hand** stand, **hand** stand, **hand** to- **hand**, **hand** to- **mouth**, **hand** work, **hand** woven, **hand** write, **hand** writing, **hand** wrought, **handy**, and **handy** man.

Taking now only those examples of **hand** which occur as a unitary lexeme, how can it not only be differentiated from other senses, but also how can the derivations of it be generalized? English dictionaries do not generally include right-branching derivatives in the same alphabetical listing, but these must also be accounted for as well, e.g. fore**hand**, back**hand**, glad-**hand**, sleight-of-**hand**, open-**handed**, empty-**handed**, short-**handed**, free **hand**, old **hand**, high-**handed**, heavy
handed, left-handed, first hand, second hand, hand in hand, man handle, and so on. The prepositional forms are found by searching for the appropriate form, although some may be included in essays under the discussion of hand, e.g. at hand, off hand, out-of-hand, under hand(ed) hand over, hand off, hand to mouth, show of hands, with an even hand, be in good hands, and so on.

In general, the derivations are somehow semantically related to the primary lexeme hand. The problem in lexical semantics is to discover how and then to describe the semantic relationships. The problem is similar in other languages. Note, for example, TP, where han takes on different meanings in compounds:

hanbom: grenade (i.e. a bomb which can be held in the hand);
han pensil: fingerprint (i.e. a substitute for a pencil).

The relationship of hand to other parts of the body may be culture-specific, as shown in Kewa, where ki and àà refer to the whole extremity (upper and lower), as well as to certain parts of them (see Franklin 1963).

7. Speech acts and definitions

Wierzbicka (1987) has suggested a procedure for analyzing and defining speech act verbs (SAVs). She outlines a set of primitive metalexicon items from which all of the definitions are framed. There are about 200 words in the set, which differentiate 37 different groups of speech act verbs, each group comprising from four to ten verbs. For example, the ORDER group consists of: command, demand, tell, direct, instruct, require, and prescribe.

Wierzbicka uses what is called a Natural Semantic Metalanguage (NSM) to represent the meanings of the SAVs. The NSM currently consists of 28 elements which are used to provide a semantic description of the verb in question (Goddard 1990).

Some of these occur in TP as SAVs, and because one of the tests of the validity of the definition is its translatability (Wierzbicka 1986, Goddard 1990) I have attempted to frame some of the definitions in TP. I do this because speech acts (SAs) in English and TP can be compared to see what underlying universal semantic components are involved. In the following section the comparable meta-components are given in English in square brackets to represent equivalent semantic units between English and TP.

7.1 Order: givim toksave [UNDERSTAND, EXPECT, WANT, CAUSE, FUTURE]

1. Mi ting olsem yu ken save long wanem samting yu mas mekim. [UNDERSTAND] "I think that you can know what you must do."
2. Mi ting olsem bai yu bihainim tok bilong mi. [EXPECT; FUTURE] "I think that you will obey what I say."
3. **Mi tok olsem, mi laik yu mekim wanpela samting.** [WANT] "I say that I want you to do something."

4. **Tok bilong mi inap mekim yu bihainim dispela.** [CAUSE] "What I say will be able to make you obey this."

### 7.2 Command: mekim save [WANT, EXPECT, CAUSE, OFFICIAL, AUTOMATIC, PRESENT]

1. **Mi tok olsem mi laik yu mekim wanpela samting.** [WANT] "I say that I want you to do something."

2. **Mi ting olsem bai yu bihainim tok bilong mi.** [EXPECT] "I think that you will obey what I say."

3. **Tok bilong mi inap mekim yu bihainim dispela.** [CAUSE] "What I say is able to make you obey this."

4. **Mi tok olsem na mi gat strong long mekim tok olsem.** [OFFICIAL] "I say this and I have the power to say it."

5. **Mi tok olsem na tok bilong mi tasol inap mekim yu bihainim dispela.** [AUTOMATIC] "I say this and what I say is able to make you obey it."

6. **Mi tok olsem na yu mas bihainim laik bilong mi long nau tasol.** [PRESENT] "I say this and you must obey what I want right now."

### 7.3 Demand: tok strong long; mas mekim/ wokim [WANT, CAUSE, RELUCTANCE, CONVINCE, RIGHT]

1. **Mi tok olsem, mi laik yu mekim wanpela samting.** [WANT] "I say this and I want you to do something."

2. **Tok bilong mi inap mekim yu bihainim dispela.** [CAUSE] "What I say is able to make you obey it."

3. **Tingting bilong mi, yu no laik bihainim dispela.** [RELUCTANCE] "I think that you do not want to obey it."

4. **Mi ting olsem i gat gutpela as long mekim dispela.** [CONVINCE] "I think that there is a good reason for you to do it."

5. **dispela tok i orait long mi autim.** [RIGHT] "It is proper for me to say this."

### 7.4 Tell: tokim [WANT, EXPECT, CAUSE, INDEF RESULT]

1. **Mi tok olsem, mi laik yu mekim wanpela samting.** [WANT] "I say that I want you to do something."

2. **Tingting bilong mi bai yu bihainim tok bilong mi.** [EXPECT] "I think that you will obey my words."

3. **Tok bilong mi inap mekim yu bihainim dispela.** [CAUSE] "What I say is enough to make you obey it."

4. **Mi ting olsem bai yu save mekim tasol mi no save tru.** [INDEFINITE RESULT] "I think that you will do it but I am not sure."
7.5 Direct: tok long mekim [WANT, CAUSE, COOPERATE, OFFICIAL, IMPERSONAL]

1. *Mi tok olsem mi laik yu mekim wapela samting.* [WANT] "I say that I want you to do something."
2. *Tok bilong mi inap mekim yu bihainim dispela.* [CAUSE] "What I say is able to make you obey it.
3. *Tingting bilong mi olsem yu laik bihainim dispela.* [COOPERATE] "I think that you want to obey it."
4. *Mi tok olsem na mi gat strong long mekim tok olsem.* [OFFICIAL] "I say that and I have the power to say it."
5. *Tok bilong mi em inap harim.* [IMPERSONAL] "What I say he can obey."

7.6 Instruct: skulim [WANT, EXPECT, UNDERSTAND, AUTHORITY, DEF RESULT]

1. *Mi tok olsem mi laik yu mekim wapela samting.* [WANT] "I say that I want you to do something."
2. *Tingting bilong mi bai yu bihainim tok bilong mi.* [EXPECT] "I think that you will obey what I say."
3. *Tingting bilong mi yu ken save long wanem samting yu mas mekim.* [UNDERSTAND] "I think that you can understand what you must do."
4. *Tok bilong mi mi gat strong long mekim.* [AUTHORITY] "What I say I have the power to say."
5. *Mi ting olsem bai yu wokim tru laik bilong mi.* [DEF RESULT] "I think that you will really do what I want."

7.7 Prescribe: toksave long wokim [UNDERSTAND, CAUSE, WRITE, AUTHORITY, CONVINCE]

1. *Tingting bilong mi yu ken save long wanem samting yu mas mekim.* [UNDERSTAND] "I think that you will know what you must do."
2. *Tok bilong mi inap mekim yu bihainim dispela.* [CAUSE] "What I say is able to make you do it."
3. *Mi ting olsem mi mas autim dispela long pepa long mekim em bihainim.* [WRITE] "I think that I should write this so that you will do it."
4. *Tok bilong mi mi gat strong long mekim.* [AUTHORITY] "What I say I have the power to say."
5. *Mi ting olsem i gat gutpela as long mekim dispela.* [CONVINCE] "I think that there is a good reason to do this."

The above translations into *TP* explicate the semantic components contained in their English equivalents. By using speech act verbs it is possible to test the assumption about translatability: have the equivalent components been transferred? What components are carried over into *TP* from English? Are they appropriate and sufficient? In this first attempt about 46 *TP* words are used as primitives in the definitions. Goddard's (1990:258) shorter list now follows, compared
with that of Wierzbicka (1987:13), whose primitives appear here in bold italic and, if additional, placed in parentheses:

- **I, you, someone, something** [substantives]
- **place, world, become**
- **this, the same (other), two, all** [determiners, quantifiers]
- **good, bad** [evaluative]
- **be a kind of, be a part of** [taxonomy, partonymy]
- **be like** [attributive]
- **think (of), say, know** [mental predicates]
- **do, happen to** [action, patient]
- **want, (not want), no!, would, imagine, can** [mood, irrealis, modality]
- **when, after**, where [time, place]
- **because** [causation]
- **very** [intensifier]

## 8. Body parts as initiators

I suggest that a dictionary compiler can use body parts as a beginning natural and useful semantic domain. There are a number of reasons why this is the case:

1. As Mark Johnson's title *The Body in the Mind* (1987) suggests, meaning, imagination, and reason are closely tied to our personal experience and have a bodily basis.
2. Body parts are commonly categorized in ways that are reflected grammatically. By examining sets of body parts, notions of possession, instrumentality, and nominal concretions (Longacre 1976) are encountered.
3. They become highly metaphorical and illustrate metonymy and other figurative extensions (Nida 1975).
4. Although body parts are most easily elicited as nouns, they readily derive verbs, adjectives and compounds.
5. Other semantic domains are easily associated with body parts, for example, parts of animals and plants, as well as parts of automobiles (Blasso 1967).
6. Emotions spring from body avenues, as do many expressions concerning emotions (Lakoff 1987).
7. Verbal actions also arise naturally from body parts, for example hand > to handle, leg > to kick, or tongue > to taste.
8. Further extensions are readily built upon a schema involving body parts: go < by foot, carry < on the shoulder, sleep < with the head < on the pillow.

For these reasons alone (which may apply to other domains to a lesser degree) a dictionary compiler should begin with body parts and see how far the semantic extensions and relations will go.

### 8.1 Folk Anatomy
Ethnolinguistic procedures that are well known and established allow replicable elicitation. The first principle is simply to elicit all of the body parts that are commonly and collectively known by the community. This can be done in a number of ways:

1. Simply record all of the parts of the body that a language speaker and his/her companions give. If English speakers were asked "What are the names of the parts of the body?", they would give words like head, eye, mouth, ear, leg, and so on. They would not commonly give deltooid muscle (or even muscle), lymph glands, or artery, even though most English speakers would recognize the words. Further anatomical specializations would be even less common.

2. Further test the body parts elicited to see which ones can easily be used in figures of speech (McElhanon and Sácnemac 1975). The most generic and well known body parts will be found in idioms, but it is unlikely that there will be an idiom involving one's thoracic nerve.

3. Deal with lexemes, i.e. compounds, idioms, figures of speech and so on. In the initial stages the compiler is not worried about the eventual separation of the parts into various semantic domains. The English lexeme the mouth of the river will eventually be a part of the domain of bodies of water. Initially, word and conceptual associations are given full play.

4. Construct some sort of taxonomy for the body parts (Franklin 1963) which will then reflect suggested conceptual relationships.

8.2 Start with the Head

This procedural account for TP begins with het (head), because the lexeme is easily recognizable, and yet not as extensive in its semantic scope as its counterpart in English. However, from time to time I compare the lexeme with the English cognate head.

In TP het is a part of the bodi (body). If a person asks for body parts in TP words like ai, ia (yau in classical TP), nus, maus, pes, sol, bel, han, lek, (eye, jaw, nose, mouth, face, shoulder, stomach, hand, leg) and so on are commonly heard (cf. also Mühlhäusler 1979:224). A working taxonomy would look something like this (where ... represents an unlisted string of forms):
By noting particular syntactic frames it is possible to determine a number of meanings which involve the lexeme het:

1. **het bilong [X]:** \{diwai, kanu, pik, maunten, wara\} (The head or top of a tree, canoe, pig, mountain, water)
2. **[X] bilong het:** \{gras, bun, paspas\} (hair, bone, of the head)
3. **[X]het:** \{bik, gutpela, go\} (big head, good head, to go ahead)
4. **het i [X]:** \{pen, bruk, pas\} (the head pains, cut open, clogged)
5. **het[X]:** \{tok, pen, kela, kota, man, meri, win\} (main point, head pain, bald headed, headquarters, chief, responsible woman, head wind)
6. **bikheth-im** (talk back); bruk-im het (split the head open); daun-im het (bow the head); and in soccer, het-im bol (head the ball).

As expected, the nominal **het** may be derived as a compound (bikheth, hetman), as a verb (**hetim** bol), as part of a stylized expression (**daunim het** = bow the head), as an idiom (**het** bilong pik = pig headed), as a figurative extension (het bilong diwai = wooden headed), or as the topic in a complement phrase (**het i pen** = the head is paining).

Body parts therefore seem to be a particularly productive area for lexicographic study. This is for a number of reasons (cf. also **Lakoff 1987**):

1. **Ontological.** We perceive our own existence and explain the world around us accordingly, so that we see, hear, feel, etc. using body parts.
2. **Experiential.** Pain and emotions radiate from within us or from actual impairment to the body, hence the heart, liver, brain, etc. are used to describe these feelings.
3. **Metaphorical.** Literally we have hands, but when we hand-over something we go a step further and lexicalize the function in terms of the body-part.
4. **Spatial.** Our face is in front of us so when we face someone we describe this in reference to our position. These are also given Spatial metaphors, such as face-off, backslide, and so on.
5. **Symbolic.** Although this if a type of metaphor, it is actualized by means of a physical process: we shake hands on it, nod our head that it is OK, wave with our hand, and so on.
6. **Poetic.** The human body is a base point of departure and a source of meaning, serving as an icon, model or reference point for imaging and spacial categories (**Friedrich** 1979:499).
7. **Extensional.** Each body part leads to another.

In regard to the point consider, for example, the dictionary definition of face in English, and what it entails:

**Face:** the front of the human **head**, including the **chin, mouth, nose, cheeks, eyes** and usually the **forehead**.
To understand this definition a reader must also, to some degree, interpret the following definitions:

**Head**: the upper or anterior division of the body, containing the brain, the chief sense organs, and the mouth.

**Chin**: the lower portion of the face, lying between the lower lip and the lower jaw.

**Mouth**: the cavity externally bounded by the lips, internally by the pharynx that encloses the tongue, gums, and teeth.

**Cheeks**: the fleshy side of the face below the eye and above and to the side of the mouth.

**Eye**: the organ of sight, nearly spherical and hollow lined with sensitive retina and lodged in a bony orbit in the skull and is normally paired.

**Forehead**: the part of the face above the eyes.

Each of the items in bold face is either an actual part of the face, or it could be construed as such a part by someone who is learning English, e.g. *orbit* and *organ* are not parts of the head but could be read that way because of the definitions.

Without a good knowledge of spatial orientations (side, above, lower, below, to the side of, etc.) it would be difficult to read the definitions. In addition conceptions of descriptive words like smooth, shaped like, narrow, broad, round, fleshy, spherical, hollow, bony, paired are also necessary.

Even this much detail does not finish what is semantically related to the lexeme face. What about whiskers and what is the difference between them and a beard or a mustache? And what about a kind of mustache called a handlebar or beards called Vandykes or goatees? What are side-burns and mutton-chops, etc.?

### 8.3 A body part schemata

A schemata for body parts can begin with the literal human primary parts and then extend to secondary, but still literal, parts of flora and fauna. In figure 2 below the primary or main entry is of course hand. The three semantic categories given in square brackets at the top of the figure follow Nida (1964) although I have not included what he calls ABSTRACTS. In the OBJECT category a fuller explication would include the set of all those objects associated with hand, exemplified here with spoon, ring, glove, and gun. Either hand itself or the OBJECTS which correlate with it may further portray EVENTS, such as touch and feel or eat and feed (with or without a spoon). The RELATIONS generally involve a syntactic frame which includes a preposition. The prototype or schemata suggested is elementary and compares only to parallel extremities of flora and fauna.
By extending meaning structure from a prototype it is possible to elaborate a schema (Langacker 1987, Taylor 1989:65). Taylor, who follows Langacker in this regard, explains that a prototype is a typical instance of a category with other elements allowed membership on the basis of degree of similarity. In the case of a schema there is an abstract categorization that is compatible with all the members of the category which it defines. It embodies the commonality of its members. Taylor considers categorization by prototype and schema as aspects of the same phenomena.

A beginning schema for the prototypical extremities of humans, certain animals and trees would incorporate the following:

![Figure 2: Extensions of extremities](image)

Of course, the extensions are not identical between people, flora, and animals:

![Figure 3: Comparisons in the extensions](image)
Other body parts come into play for animals: tail, for example may parallel leg and arm, but it has no further parts. Leg, arm and tail do enter into various contrasting idioms: He legged it home, but not *He tailed/ armed it home. His voice tailed off, but not His voice legged/ armed off. He armed himself, but not *He legged/ tailed himself. The applies to other lexemes throughout figure 7, e.g. He clawed his way up the hill, but not He nailed his way up the hill; He pawed the ground, but not *He handed the ground. He is a budding (*nailing; ?clawing) linguist.

Although both hands (or fingers) and branches will wave, one by means of a human instrument and the other by the wind, a leg does not wave. Further, although hand is an extension of arm, once it applies to flora it becomes specific (a hand of bananas) and to many fauna it does not apply at all because a different lexical item is introduced (such as paw). Certain analogies terminate: there are toenails and fingernails but only a paw has claws. The verbal extensions converge in some areas: as mentioned, hands, fingers, branches, and twigs can all wave, but only the floral counterparts can sway (although one's body can sway). People can dig without fingers or a dog can dig with its claws, but a tree or branch does not dig with anything.

This is only a sample of how a body part can be used as an initializer to generate further lexical entries; other compounds, verbs, prepositional phrases, idioms, and metaphors arise out of a simple body-part entry. They in turn prompt new extensions and collocations, all extending from the body.

9. Some conclusions

I now summarize a number of practical suggestions with attendant hurdles in dictionary compilation. Many of the suggestions are self-evident; persistent lexicographers attempt to solve the hurdles.

1. There are numerous categories of information to include in a dictionary, including phonological, orthographic, morphophonemic, grammatical, syntactical, sociolinguistic and cultural information.
2. The compiler must therefore be very familiar with the language and the culture. Observation and participation are necessary to test the meanings of lexical items.
3. The dictionary is a tool for the compiler, the translator, the language learner, and to some extent, the native speaker of the language. Each of these audiences must be kept in mind.
4. Various meanings must be tested vis à vis semantic relationships, including taxonomies and domain analysis.
5. National assistants who are native-tongue speakers can be trained to write folk definitions. These definitions provide valuable basic information on lexical fields.
6. Fieldworkers need to understand and use the various eliciting and ethnosemantic techniques which are available.
7. Definitions should be attempted by using the language alone. This will entail establishing a metalexicon for further definition purposes (as done by Wierzbicka, passim).
8. Brevity and conciseness are not the foremost considerations in a first definition. Rather establishing sets of semantic components which distinguish related sets of words is of primary importance.
9. Defining is tedious and requires thinking, testing and attention to detail. However, only in this way can key terms or idioms (as in Galland and Hinds-Howell 1986) be adequately understood.

10. Idioms, metaphors, and poetic materials are part of the data for a dictionary.

11. Encyclopedic information should be coded. It follows that a reliable and consistent reference system should be used.

**Selected glossary**

**antonymy**
"the relationship of 'oppositeness of meaning'. antonyms are often thought of in the same breath as synonyms, but they are in fact very different. There may be no true synonyms, but there are several kinds of antonyms... -gradable antonyms, such as big/small, good/bad, which permit the expression of degrees (very big, quite small, etc.); - nongradable antonyms (also called complementary terms), which do not permit degrees of contrast, such as single/married... and -converse terms: two-way contrasts that re interdependent, such as buy/sell or parent/child..." Crystal 1993:105).

**class meaning**
the feature or features that all members of a set have in common, such as each of the parts of speech, a noun serving as a referential item (among other) and a verb identifying a state or action.

**componential analysis**
"...analysing lexemes into a series of semantic features, or components. Man, for example, could be analysed as ADULT, HUMAN, and MALE. The approach was originally devised by anthropologists as a means of comparing vocabulary from different cultures, and it has been developed by semanticists as a general framework for the analysis of meaning." (Crystal 1993:107).

**concepts**
meanings classified as events, things or attributes and are related to one another by case roles (Larson 1984:26, 199-222).

**connotations**
supplementary meanings which extend beyond the central linguistic one. They represent people's emotions or attitudes about, say, a cassowary.

**ethnosemantics**
a branch of semantics which studies meaning in its cultural context.

**folk definitions**
definitions on the basis of the semantic relationships which native speakers give in their classifications of language items and actions.

**folk taxonomy**
the way a people classifies things by the words they use to refer to the world or interpret the world.

**hyponymy**

"the notion of `inclusion', whereby we can say that `an X is a kind of Y'. For example, rose is a hyponym of flower, car of vehicle" (Crystal 1993:105).

**illocutionary act**

the complete speech act made in a typical utterance, that consists of the delivery of the propositional content of the utterance, including the references made and the predicate, with a particular illocutionary force, whereby the speaker asserts, suggests, demands, promises, or vows, etc.

**illocutionary force**

the intent of a speaker in making an utterance, and the particular presuppositions and attitudes that must accompany that intent, together comprising, on one analysis, the features of the illocutionary point, the strength of the illocutionary point, the preparatory conditions, the propositional content conditions, the mode of achievement, the sincerity conditions, and the strength of the sincerity conditions (as the illocutionary forces that distinguish acts of asserting, promising, excommunicating, exclaiming in pain, inquiring, or ordering).

**illocutionary verb**

a verb that as part of its meaning names at least one illocutionary force or a value of some component of illocutionary force (as "assert"; or "advise", which names more than one force, having either the assertive or directive illocutionary point; or "announce", which does not name a particular force, but appears to refer in part to a high degree of strength of sincerity conditions; or "insinuate", which cannot be used as a performative verb, but is nevertheless an illocutionary verb useful in reporting illocutionary acts second-hand).

**lexeme**

the smallest unit that is distinguished from other similar units in the meaning system of a language. It may be abstract and occur in different forms, i.e. in different inflections. For example walk may occur as walks, walking, walked, or in idioms such as cakewalk (in American English this refers to something which is easily done or achieved). A lexeme may also occur as a phrase, a compound word, or in special combinations. The dictionary information on a lexeme or lexical entry generally includes, as the minimum, its pronunciation, part of speech, inflected forms, and various meanings, generally grouped according to its senses and subsenses. Specialized dictionaries may also include particular collocations (as in Benson et. al 1986).

**lexical semantics**

information included in a dictionary that may describe the distinctive semantic features that characterize lexemes and their semantic fields (Lehrer 1974, Grandy 1987). The semantic field is often organized such that a set of terms represents a particular domain.
which refers to a speaker's area of experience. The domain can be described according to various organizing principles and semantic relationships.

**lexicology**
that branch of linguistics which is concerned with the relationship of the meanings of lexemes. A lexicologist makes specific statements about the form, meaning, and usage of individual lexical items which may be included as entries in the dictionary.

**lexicon**
sometimes used instead of "dictionary" though generally lexicon is a technical term which refers to the set of all words and idioms of the language in question; sometimes writers use lexicon to refer to the vocabulary of a language as used by an individual speaker.

**markedness**
the characteristics of members of a universal linguistic categories such that the least frequently found and least predictable is the most marked, and vice versa.

**meaning**
various terms are used to refer to meaning, such as **definitional** meaning (how the semantic structures of lexemes, as represented in a dictionary, are outlined), **referential** meaning (resulting in the lexicalization of extralinguistic concepts and objects), **conceptual** meaning (how the meanings are related as far as the mental image of the speaker is concerned), and **collocational** meaning (the relationship of words and their meanings in combination or sequence).[^4]

**paradigmatic**
"the way in which lexemes an substitute for each other" (Crystal 1993: 105).

**pragmatics**
the study of the aspects of meaning and language use that are dependent on the speaker, addressee, and other features of the context of utterance; including the study, on one hand, of the effect that context of utterance, generally observed principles of communication, and the goals of the speaker have on the choice of means of expression, and on the other hand, the effect such factors have on the interpretation made of an utterance, by the addressee; and centering programmatically on such concerns as the treatment of given versus new information (including presupposition), deixis, speech acts (especially illocutionary acts), implicature, and the relations of meaning or function between portions of discourse or turns of conversation.

**relational opposites**
the relation holding between a pair of lexemes such as father-son, such that if one person is father to another, that other is son to the first one.

**semantic field**
an organization of lexemes that interrelate such that one can be seen to define another in some way, like the lexemes that make up color, one complementing another.

**Semantics**

the study of meaning in linguistic expressions, apart from consideration of the effect that pragmatic factors such as features of the context, conventions of language use, and the goals of the speaker have on the meaning of language in use; also, the study of meaning in linguistic expressions, generally.

**Speech Acts**

utterances classified according to their *locutionary* and *illocutionary* meanings. Certain utterances have an illocutionary force about them and employ performative verbs. Speech acts have been classified as commissive, declarative, directive, expressive, and representative.

**Synonymy**

"the relationship of 'sameness' of meaning, e.g. kingly/royal/regal, pavement/sidewalk, youth,youngster" (*Crystal 1993*:105).

**Syntagmatic**

"the way lexemes occur in sequences" (*Crystal 1993*:105).

**Notes**

1 An early and much more expanded version of this paper formed the basis for several seminars on lexicography presented at Ukarumpa for the Summer Institute of Linguistics in 1990 and 1991. I have combined some of that material with a later version which I presented at the Linguistic Society of Papua New Guinea in Madang in September, 1992. I am grateful to a number of colleagues for comments on previous drafts including, in particular, David Snyder, Cindi Farr, and Janet Ezard. Snyder has since (*1991*) worked on the problem of a defining language for *TP* dictionaries. There are numerous collections of articles dealing with lexicography and dictionaries, such as Hartmann, ed. (*1983*) and Ilson, ed. (*1985*, *1986*, *1987*). On a more practical or applied level see Bartholomew and Schoenhals (*1983*), Newell (*1986*), and Coward and Grimes (*1995*).

2 *Zgusta (1971)* presents a definitive summary of all that is involved in lexicography. His extensive comments on the whole lexicography process are of particular interest to anyone involved in compiling a dictionary. A simple and very practical way to see what to include in a dictionary is to examine various kinds of dictionaries and analyze their range of audiences. *R. Merkin (1983)* traces the history of academic dictionaries and *A.P. Cowrie (1983)* considers what English dictionaries should include for the foreign learner.

3 *Lyons (1977*:250ff.) outlines the general theory of semantic fields, including the difficulty and differences in the terminology that is used by various authors, such as lexical field, conceptual field, and lexical structure. He uses the example of the color continuum as a conceptual field and applies this model to diachronic semantics. *Palmer (1981*:68ff.) also discusses semantic fields
and color systems, including the well-known work of Berlin and Kay (1969) who claimed that there is a universal inventory from which all languages derive eleven or fewer basic and ordered terms.

4Leech (1974:10ff.) discusses seven types of meaning: 1) conceptual or the sense; 2) connotative, which refers to what is communicated by the language; 3) stylistic; 4) affective, or the feelings and attitudes of the speaker; 5) reflected; 6) collocative; and 7) thematic, or the organization of the message in terms of ordering, forms and emphasis. Meanings 2-6 are considered by Leech to depict associative meaning, that is the mental connections based upon the contiguities of experience. Nida (1975:26) describes meaning in terms of the intersection of two sets of factors: the cognitive and emotive on the one hand and the extralinguistic (referential) and intralinguistic (grammatical) on the other hand.

5These observations come from studies on the Kewa language (e.g. Franklin 1971, Franklin and Kirapeasi, eds. 1972, 1975, Franklin, Franklin and Kirapeasi, eds. 1974, Franklin and Franklin, assisted by Kirapeasi, 1978).

6Semantic fields like DEATH are highly metaphorical and figure into special saying, poems and other figurative literature. See, for example, the work by Lakoff and Turner (1989).

7Authors describe sense relationships using a variety of terms. See, for example, the classic study on folk taxonomies by Conklin (1962). Lyons (1977) treats, among other things, opposition and contrast, including non-binary contrasts, as well as hierarchy, hyponymy, lexical gaps, part-whole relations and componential analysis. Nida (1975) is much more extensive in his treatment of componential analysis. For other terminology see also Cruse (1986) on congruence and partial relations and Parker-Rhodes (1978) on the theory of inferential semantics.

8Lyons (1977:305ff.) reviews the history and use of the words marked and unmarked. He discusses formal marking (such as host v. hostess), where one member of the pair of opposites is marked and the other is not. Usually one member is distributed more widely than the other or one is more specific in sense than the other (lioness v. lion). Semantic marking takes place in a wide variety of relationships including hyponymy (nurse implies female), part-whole (arm:body), and collectives (cattle, furniture). He refers also to the distinction of distinguishers and markers by Katz and Fodor (1963). Systematic meaning was presented by the markers, the residue by distinguishers.

References


Franklin, Karl J. and Yapua Kirapeasi, eds. 1975. Pisini agaapara adaa agaa laapo i buku. Ukarumpa: Summer Institute of Linguistics. [A common-usage dictionary translating TP words and expressions into West Kewa.]


Mihalic, F. and John Verhaar, eds. Forthcoming. TP Dictionary.


Date created: 25-Mar-1998
Questions/Comments: SILEWP@sil.org

[SILEWP 1998 Contents | SILEWP Home | SIL Home]