

# **DICTIONARY OF EASTERN PENAN**

INCORPORATING PRINCIPLES OF A LEXICOGRAPHIC MODEL  
KNOWN AS THE

EXPLANATORY-COMBINATORIAL DICTIONARY

**INCLUDING A GRAMMAR AND  
AN ENGLISH-PENAN INDEX**

© Ian Mackenzie, 2015  
[www.rimba.com](http://www.rimba.com)

© Ian Mackenzie  
[www.rimba.com](http://www.rimba.com)

# TABLE OF CONTENTS

FORWARD	1
FORMALISMS OF THE DICTIONARY	3
STRUCTURE OF AN ENTRY	6
OTHER FORMALISMS AND CONVENTIONS	14
SYMBOLS AND ABBREVIATIONS	17
PENAN PHONOLOGY AND ORTHOGRAPHY	18
PENAN GRAMMAR – A SKETCH	20
DICTIONARY OF EASTERN PENAN	35
ENGLISH-PENAN INDEX	217

© Ian Mackenzie  
[www.rimba.com](http://www.rimba.com)

# FORWARD

## 1. The Penan language and people

The Eastern Penan language belongs to the Bornean branch of the Austronesian language family. It is spoken in the eastern part of the Malaysian state of Sarawak on the island of Borneo. The number of its speakers is usually estimated to be ten thousand.

The data in this dictionary come primarily from two different regions -- the Upper Baram River, and the Tutoh River (the Tutoh is a major tributary of the Baram). The speech of these areas does not differ phonologically or morphologically, and only a relatively small number of vocabulary items distinguish them.

The language most closely related to Eastern Penan is Western Penan, spoken in the Belaga and Silat River drainages and elsewhere. While Eastern Penan is the name of a single language, "Western Penan" is an umbrella term covering a number of different speech communities, some of them so different from each other they are probably separate languages rather than dialects. Thanks to a collection of myths that was published in Belaga Penan,\* I have been able to compare one of these speech variants with Eastern Penan, and find myself unable to read it with comprehension, the reason being that something like 20% of the words differ between the two languages.

Unlike the majority of the indigenous peoples of Borneo, the Penans have no history of sedentary rice farming. They were the last nomadic hunter-gatherers to inhabit the vast virgin rainforest that covered northeastern Borneo until recently.

## 2. A new kind of dictionary

This is a new kind of dictionary: it is designed for *production*, not just interpretation. It is based on a model of lexicology known as the Explanatory-Combinatorial dictionary, or ECD. This will be explained in the next section.

It is relatively meaningless to ask how many "words" a dictionary contains. What is a word? In any case a "word" cannot be an idiom, yet idioms make up a significant percentage of any natural language. It is only meaningful to talk about the number of dictionary entries. In the ECD approach, each distinct word sense or idiom has its own entry.† The dictionary you see here contains something like eight thousand entries. At this point in my research I have no reason to suppose that Penan is less lexically rich than any of the world's other languages -- at least in their oral forms. Therefore, a complete dictionary of Penan would doubtless contain many tens of thousands of entries.

(The text you are viewing here dates from 2006. Since that time I have accumulated a vast amount of lexicographic data, increasing the number of entries by fifty percent. But these recent additions are still in rough form, and a year or two of editing would be necessary before I could include them in a published version.)

---

\* Langub, Jayl. *Suket: Penan Folk Stories*. University Malaysia Sarawak, 2001

† For example, in an ECD of English, there would be three entries for the noun "Yankee": 1. a person from New England. 2. a person on the Northern side in the war between the states. 3. a citizen of the United States. A more complex example is the form "spring". The following capitalized words or phrases must all be separate entries in a complete ECD (the definitions are abbreviated): SPRING<sup>1</sup> = 'source of water emerging from the ground'; SPRING<sup>2</sup> = 'season that follows winter'; SPRING<sup>3</sup> 1. = 'an elastic device'; SPRING<sup>3</sup> 2. = 'degree to which something is elastic' (e.g. *the spring of the board*); SPRING CHICKEN 1. = 'young chicken'; SPRING CHICKEN 2. = 'young person, as if that person were a "spring chicken 1"'. (Of course, the vocable SPRING covers many more entries than those listed.)

Not only do the number of entries in the current work fall far short of a complete dictionary, but so does their content. Definitions are often provisional, examples unrepresentative, and lists of collocations<sup>‡</sup> incomplete.

### **3. History and methodology of the project**

I first visited Sarawak in 1991, and have been returning ever since to conduct linguistic and ethnographic fieldwork, normally two or three times a year. As of 2015 my cumulative time in the field exceeds three years. The dictionary began with my first Penan word list, collected in 1993. I assembled it by questioning native speakers using the medium of Malay, a language spoken by most younger Penans. In 1995 I met Jokim, who became my primary language informant. He has a high school education, and although he speaks both Malay and English, our interactions have always been in Penan. Jokim is intelligent and patient, good at explaining meanings, and creative in constructing illustrative sentences. He is the author of a significant number of the examples in the present work.

In 2001 I met Galang Ayu, who became my primary ethnographic informant. He is a man with great fondness for Penan tradition, and has a formidable memory. Among all Penans I know, he alone recalls most details of traditional culture, and remembers the ancient religion. (Christian missionaries arrived on the scene beginning in the late 1940s.) Over the years he has dictated thousands of pages of ethnographic information, amounting to hundreds of thousands of words, and the resulting texts are a rich source of linguistic material. Citations from them make up a large percentage of the examples in this dictionary.

### **4. A work in progress**

The dictionary before you contains errors. Most of these will be invisible to the casual reader, or even to the beginning student of the language. However, it is to be expected that mistakes of a trivial nature also occur; it is a daunting task to proofread such a massive manuscript.

Over the years, a surprising range of people, including aid workers, filmmakers, social activists, researchers, even long-term "tourists", have asked for preliminary versions of this dictionary. In general I have acceded to their requests. As a result, there are a number of copies in circulation, including early, inferior versions, some of which have been photocopied (with or without my consent). It is in the interests of all that only the best available version be in circulation. It is for that reason, and to meet an ongoing need, that I have reluctantly decided to publish a work with so many shortcomings.

Ian Mackenzie  
Vancouver  
August 2015

---

<sup>‡</sup> A *collocation* is a group of two or more words that frequently co-occurs and thus becomes lexicalized, without being a full-fledged idiom. It may be unpredictable or idiosyncratic, but only to a degree. Here are some examples, all of which must be listed under some of the entries for "spring" mentioned in a previous footnote: "thermal spring", "artesian spring", "early spring", "false spring", "release a spring", "compression spring", "X is no spring chicken". (It is usually, but not always clear whether a given idiomatic expression should be treated as a collocation, or as a full-fledged idiom with an entry of its own.)

# FORMALISMS OF THE DICTIONARY

## 1. GENERAL

This is a new kind of dictionary, incorporating as it does principles of the Explanatory-Combinatorial lexicographic model.<sup>§</sup> Conventional dictionaries are *interpretation* dictionaries; they enable one to decipher texts in the language in question, but provide little or no information on how to create new texts. This dictionary aims to be a *production* dictionary; that is, in principle, even a person with little knowledge of the language should be able to use it (in conjunction with an appropriate grammar) to produce well-formed Penan texts.

A complete Explanatory-Combinatorial Dictionary (ECD) of Eastern Penan, along with a complete grammar, would, in principle, constitute a complete written record of the language.

Each entry of an ECD thoroughly describes a given unit of vocabulary, defined as either an individual word taken in one of its senses (a *lexeme*) or an idiom (a *phrase*). The meaning of each item is precisely specified, and so too are the particular rules that govern how it combines with other elements in a sentence -- rules that do not belong to the general grammar of the language.

In what follows, I explain the formalisms of this dictionary. It is important to note that I will not be describing a fully fledged ECD, but rather a simplified version adapted to the current task.

I have not sought to create a fully developed ECD for two reasons. First, the data is incomplete; the biggest obstacle is that I am not a native speaker of Penan, and therefore am unable to reliably fill in missing information from my own intuition. Second, even with complete data, the writing of even one complete ECD entry is a major task, and often takes days to accomplish; a typical entry is a page long, and some may extend for several pages.

The present dictionary, its simplified formalisms notwithstanding, strives to fulfill the promise of the two goals implied in the name "Explanatory-Combinatorial".

The "Explanatory" component of the name refers to the first major goal of the dictionary, namely defining the meaning of every lexical item in a rigorously correct and complete fashion. When providing a definition of any lexeme or idiom, most dictionaries make do with a list of partial synonyms. The ECD, on the other hand, uses a single expression consisting of lexemes that are less complex (=more primitive) than the lexeme being defined. By this approach we avoid a number of pitfalls of traditional lexicography, including circularity (e.g. giving "select" as the definition of "choose", and then "choose" as the definition of "select".)

In principle the definitions in an ECD of a given language are written in that language. For a number of reasons this principle has not been followed here: definitions and explanations are given in English. As a result, and because of my often provisional understanding of the meaning of a given lexeme, these definitions are often less formally rigorous than those in a fully developed ECD.

The "Combinatorial" component of the name refers to the second major goal of the dictionary, namely the description of the criteria governing all possible co-occurrences of each lexical item. That is, the dictionary aims to show the user how to combine a given lexical item with other items to produce well-formed utterances.

The two most striking features of the present dictionary correspond to the dual goal just enunciated. These are, first, the logical form of the definition, and second, the compilation of collocations.

---

<sup>§</sup> Mel'čuk, Igor A., and Alexander K. Zholkovsky, *Tolkovo-kombinatornyj slovar' sovremennogo russkogo jazyka*, Vienna, 1984 ('Explanatory combinatorial dictionary of modern Russian', which includes an introduction in English explaining its basic principles.) Also see Mel'čuk, Igor et al., *Dictionnaire explicatif et combinatoire du français contemporain*, Les Presses de l'Université de Montréal, 1984. (First of a series: three further volumes were published, in 1988, 1992, and 1999.) Also see Mel'čuk, Igor A. (1988). *Dependency syntax: Theory and practice*. Albany, NY: SUNY Press.

## 2. LOGICAL FORM OF THE DEFINITION

Most lexical items -- whether lexemes or phrasemes -- fall into two semantic classes: (1) those that name objects or classes of objects -- e.g. "sun, water, squirrel" -- and (2) those that label predicates, that is to say relations, properties, actions, states, events and so on -- e.g. "mother, employee, pride, height, war, decay". The former class of lexical items can only be defined by a simple reference to the object they label -- a picture for example -- while the latter class must be defined with reference to what we call their actants. The actants in the above examples are as follows: "mother of X, employee of Y in job Z, pride of X about Y, height of X above Y, war between X and Y over Z, X decays from Y into Z". Each of these expressions illustrates what we call the *propositional form* of the lexeme in question. In an ECD, the propositional form of a lexeme plays a central role in the latter's dictionary entry. Consider, for example, a portion of the entry for the Penan lexeme *bet*:

**X bet Y jin lem Z tai W** = 'X removes Y from Z and moves Y into <onto> W'

The upper-case letters "X", "Y" etc. are variables. We use the term "variable" in an algebraic sense, except that our variable symbols represent lexical items rather than numbers. In this case the letters denote all possible actants of the lexeme *bet*. (In the present dictionary, the letters "X, Y, and Z" are normally reserved for nouns, whereas the variable "V" is normally reserved for verbs -- see the next example. For more on which letters are used for which kind of variable, see the section "Other Formalisms and Conventions".) In the above example, one could for example substitute *redo* ('woman'), *Daud* ('David'), *pengeja'au* ('headman') for X; for Y one could substitute *ba* ('water'), *napun* ('sand'), *kekat éh maréng nelih* ('everything that was just bought'); for Z, anything that has a *lem* ('inside'), e.g. *luvang* ('hole'), *kerita* ('car'), and for W anything that can serve as a location -- e.g. *alut* ('boat'), or *jalan* ('road'). Thus this definition could be expanded into an almost infinite number of utterances -- e.g. *Redo bet napun jin lem luvang tai alut*. 'The woman removes sand from the hole and puts it into the boat.' or *Daud bet kekat éh maréng nelih jin lem kerita tai jalan*. 'Daud removes everything that was just bought from inside the car and puts it onto the road.'

Similarly, the propositional form

**payo X V** = 'it is to X 's liking to V'

permits an almost infinite number of substitutions: e.g. *Payo ké' lakau seminga'*. 'It is to my liking to take walks.' or *Payo lakei inah pakai keleput*. 'It is to that man's liking to use a blowpipe.'

## 3. COLLOCATIONS

In an ECD, most collocations are described using a kind of semantic metalanguage making use of the notion *lexical function*. In keeping with the simplified formalisms of the present work, we dispense with this metalanguage, and simply translate each collocation into English. I will therefore avoid a discussion of semantic theory, and simply say that many lexemes combine with others in idiomatic ways. For example, in English one says "throw a party", "deliver a lecture", and "launch an attack" -- but not "\*deliver a party", "\*launch a lecture", "\*throw an attack". Such usage is idiomatic, and must be learned when one learns the lexemes "party, lecture, attack". But unlike the expressions "kick the bucket" or "it's a piece of cake" (where neither a real bucket nor a real piece of cake are involved), the former expressions are not full-fledged idioms. We are, after all, dealing with a real 'party', a real 'lecture', an actual 'attack' -- not metaphorical ones -- and the idiomatic ways in which these lexemes combine with others must be listed in the entries for "party, lecture, attack" respectively. (Note that full-fledged idioms -- e.g. "kick the bucket", "with a high hand", "put one's best foot forward"-- are given separate entries in an ECD, since they are distinct lexical items.) Collocations are listed after the definition, and in the present dictionary each is prefixed by the + sign. Here are examples of some collocations, namely those belonging to the Penan lexeme *penyakit* 'disease that causes pain'.



+ **penyakit ja'au** 'serious illness' + **penyakit keta** <éh peketa> 'serious and painful illness' + **X maneu penyakit** 'X causes disease' + **penyakit X kabit** <pekabit> **tai Y jin Z** 'disease X spreads to Y from Z' + **penyakit éh lumang pekabit** 'contagious disease' + **X kabit penyakit Y** 'X catches disease Y' + **X keta neu penyakit Y** 'X suffers severely from disease Y' + **penyakit X tai vat vat** 'disease X spreads or gets worse' + **X matai neu penyakit Y** 'X dies from disease Y' + **X ngeretep neu penyakit Y** 'X manages to endure disease Y' + **penyakit X pegaha** 'X 's illness is getting better' + **X pawah jin penyakit Y** *or* + **X ma'o jin penyakit Y** 'X recovers from disease Y' + **X peposot penyakit Y** 'X relieves Y's disease' + **X ngema'o** <ngepawah> **penyakit Y** 'X cures Y's disease' + **penyakit X ma'o neu Y** 'X's disease gets better because of (treatment) Y' + **penga'o penyakit X** 'end of X 's illness' + **X nahan penyakit** 'remedy X alleviates disease for a while' + **X ngeju usah X jin penyakit** 'X protects X -self from disease'

# STRUCTURE OF AN ENTRY

We will now outline the structure of an entry and the formalisms used therein. Every entry consists of a series of as many as eleven components, but in practice most entries do not display values for all of these. We will label these components as follows:

1. Indexing term
2. Part of speech
3. Usage and regional variation
4. *Definiendum*
5. Semantic and other conditions
6. Definition
7. Informal note[s], including encyclopaedic information
8. Collocations
9. Examples
10. Cross references
11. English equivalent

## 1. INDEXING TERM

### (a) General

The indexing term is the expression, consisting of one or more words written in **boldface**, that immediately follows the " § " sign that stands at the head of each entry.

In the event that a given lexeme is homophonous with other lexical items (cf. English box 1 'six-sided container' with box 2 'square that is printed on a form that is meant to be checked'), the indexing term displays a numerical index to disambiguate it. There are four types of numerical indices; three of them are illustrated in the following four entries.

- § **tepun**<sup>1</sup> I.1. n. § -- **tepun** X = 'X 's grandparent'  
§ **tepun**<sup>1</sup> I.2. n. § -- **tepun** X = 'ancestor[s] or forebears of X'  
§ **tepun**<sup>1</sup> II. n. § -- **tepun** X = 'owner or master of X'  
§ **tepun**<sup>2</sup> n. § -- **tepun** = 'tiger'

As can be seen, when two or more homophones exist having no semantic relationship, they are distinguished by Arabic numeral superscripts. In this way **tepun** meaning 'tiger' is distinguished from the group of three **tepun** which all denote a person. This latter group of lexemes falls itself into two groups, and since the semantic relationship between these two groups is rather loose, they are labelled with Roman numerals. Thus, "I." applies to the two **tepun** which denote human forbears, and "II." applies to the **tepun** that denotes a person enjoying

authority or ownership. The last degree of differentiation is provided by normal Arabic numerals, which thus distinguish the lexeme meaning 'grandparent' from that denoting 'ancestor'.

If a finer degree of differentiation is needed, the letters "a, b, c" etc. are used. Thus:

§ **anak** 1.a. n. § -- **anak X** = 'baby or child of human X'

§ **anak** 1.b. n. § -- **anak X** = 'child, son, or daughter, of any age, of human X'

§ **anak** 2. n. § -- **anak X** = 'a person junior to X in age and authority and under the leadership of X, as if X were "anak 1" of X'

The purpose of the indexing term is not only to uniquely label each lexeme or idiom, but to position it in its appropriate place in the dictionary. This is no trivial issue in the case of Penan, since it is the initial phonemes of words, not their endings, that undergo morphological modification. In English, for example, *king*, *kingdom*, *kingly*, and *kingmaker* all naturally fall together on the same page of the dictionary, allowing the reader to easily compare the members of this family of lexemes. But by placing the following two Penan lexemes in alphabetical order, we would be widely separating them, and make it more difficult to grasp their semantic relationship. It would make the language learning process more difficult.

§ **olong** 1. n., adj. § -- **X olong Y** = 'animal X that is domesticated or kept alive by Y'

§ **molong** 1. v. § – **X molong Y** / **nolong**\* = 'X keeps animal Y as a pet or as a domesticated animal'

Instead, we choose to group them together under their common root, **olong**, by the expedient of re-writing the latter entry in the following fashion:

§ **olong - molong** 1. v. § – **X molong Y** / **nolong** = 'X keeps animal Y as a pet or as a domesticated animal'

Thus, whenever the indexing term consists of two expressions separated by a single dash " – ", it is the second of these expressions (in this case, **molong** 1.) that actually labels the lexeme being defined. The first expression (in this case, **olong**) serves purely to position the entry in the dictionary.

It behoves me to offer some further justification for this method of ordering the dictionary. It is true that departing from strict alphabetical order often makes looking for a word a roundabout process; even, as is the case, when the reader invariably finds a cross-reference to guide him on his way. But I am following the tradition of Malay and Indonesian lexicography. In those languages too the initial phoneme of a word frequently undergoes modification, and dictionary writers long ago decided that it was better to group together all words of a given family, rather than scattering them about. (For example, **mengadakan** 'bring into existence', **keadaan** 'situation', **berkeadaan** 'in the state of', **keberadaan** 'existence', are all grouped under their root word **ada** 'exist'.) In fact, Penan is considerably more extreme than Malay in its modification of word-initial phonemes; while Malay uses suffixes as well as prefixes, Penan never does; all morphological changes operate on the beginnings of words.<sup>†</sup>

In general, idioms are listed under the first letter of the first component of the idiom; thus, **lamin besara** 'court house' and **lamin gebala** 'pastor's house' occur in alphabetical order, right after the entry **lamin** 'house'.

---

\* **nolong** is the passive voice; more on this shortly

<sup>†</sup> One minor problem that arises while determining the appropriate order of entries relates to certain compounds. In order to avoid ambiguity, Penan speakers often prepose the generic terms **juhit** 'bird', **seluang** 'fish', **laka** 'vine' etc. in front of the names of certain animals and plants. For example, **ujung bolo** is the name of a fish, but it also quite literally means 'bamboo leaf', so one always preposes the word meaning 'fish' to prevent confusion – thus, **seluang ujung bolo**. Expressions of this type are listed alphabetically under the generic term – in the above example, after **seluang** rather than after **ujung**. Similarly, **kelit bolo**, which literally means 'bamboo bat', is listed under **kelit** 'bat', and not under **bolo** 'bamboo'. However, in the majority of cases botanical and zoological terms are quite unambiguous, and are therefore listed under their names proper, rather than after the generic term. For example, **benua** is listed under the letter **b**, because although it is always correct to say **kayeu benua** 'benua tree', one usually says just **benua**, because no confusion is possible.

## (b) Roots

One of the more awkward aspects of Penan lexicology is identifying the roots of certain verbs. This problem affects the ordering of items in the dictionary. In most cases it is obvious what the root is: for example, the verb *ngebu'un* 'to begin' is derived from the noun *bu'un* 'beginning'; thus the former must be placed immediately after the latter, by means of an entry having the following form:

§ **bu'un - ngebu'un** v. § – **X ngebu'un Y / kebu'un**

However, in a significant minority of cases the identity of the root is not obvious. Consider the following examples. The capitalized form is the proposed root, and a form in parentheses is a conceivable alternative root (the latter only in examples 10 - 13). The second word (in lower case) is the active form of the (corresponding) verb, the third word is the passive, and the fourth is the nonvolitional passive (an explanation of the latter is given in a future section):

ROOT	ACTIVE VOICE	PASSIVE VOICE	NONVOLITIONAL PASSIVE VOICE
1. <b>KEP</b> thing that blocks or screens off	mekep screen off	nekep be screened off	pekep get screened off

In this example, it is obvious that the root is *kep*.

2. <b>OTONG</b> a thing burned	motong burn	notong be burned	potong get burnt
3. <b>OLONG</b> something kept as a pet	molong keep as a pet	nolong be kept as a pet	polong get kept as a pet
4. <b>ANYAM</b> something woven	manyam weave	nanyam be woven	panyam get woven

These three examples are also clear-cut; each has only one possible root. They illustrate a somewhat predictable relationship which can be viewed as a process of prefixing *m* to nouns beginning with a vowel to form transitive verbs.

5. <b>TANEM</b> place of burial	nanem bury	tenanem be buried	tanem get buried
6. <b>TUTUP</b> get closed	nutup close	tenutup be closed	tutup get closed
7. <b>SOPÉ</b> get defeated	nyopé defeat (v)	senopé be defeated	sopé get defeated

In these cases, *tanem*, *tutup*, and *sopé* are the roots and not *nanem*, *nutup*, and *nyopé*. This is because the passives are created by inserting an infix *en* in the nonvolitional passive forms -- rather than by a transformation of the active forms. This is evidence for the nonvolitional passives being the most primitive forms, and therefore the roots.

8. <b>SUAI</b> create	suai create	senuai be created	pesuai get created
9. <b>PEI</b> put	pei put	nepei be put	pepei get put

In these two cases the roots do not correspond to the form of the nonvolitional passive. Rather, they are identical to the active forms of the verbs. Both the simple passive and nonvolitional passive are derived from them.

Note that there is no Penan verb where one could reasonably propose that the root has the same form as the passive voice. A verb's root always has the form of a semantically related noun (examples 1 - 5), of its nonvolitional passive voice (examples 5 - 7),\* or of its active voice (8 and 9).

10. MOHÉ	(OHÉ??)	(POHÉ)	mohé	nohé	pohé
clean	(clean state??)	(get cleaned)	clean	be cleaned	get cleaned
11. MUJEK	(UJEK??)	(PUJEK)	mujek	nujek	pujek
spill (v)	(spill (n)??)	(get spilled)	spill	be spilled	get spilled
12. MOLO	(OLO??)	(POLO)	molo	nolo	polo
drain liquid	(draining??)	(get drained)	drain liquid	be drained	get drained
13. MENAT		(PENAT)	menat	nenat	penat
pull		(get pulled)	pull	be pulled	get pulled
14. MENYAT			menyat	nenyat	
ask for			ask for	be asked for	

Cases 10 - 13 are problematic. According to the pattern of the three earlier examples *otong*, *olong*, and *anyam*, *ohé*, *ujek* and *olo* should be the respective roots of 10 - 12. Unfortunately I have never attested these, either in heard speech or in the million-odd words of Penan dictation that I have taken down. It is my guess that *ohé* and *olo* never occur, given what would be their marginal semantic usefulness and non-concrete meanings. The form *ujek* seems plausible -- the expression *ujek nyak* 'oil spill' strikes me as acceptable. But if any of these do exist, the forms *mujek*, *nujek*, *pujek*, *mohé*, *nohé*, *pohé*, and *molo*, *nolo*, *polo* surely occur hundreds or thousands of times more frequently. (I have not yet had a chance to test these three forms on native speakers.)

As for 13, a root \*ENAT is not theoretically possible; no Penan word begins with a schwa. (Other than a single strange exception -- see the section on phonology.) Nor does there exist a form \*NAT. There might conceivably exist a form *pengenat*, 'pulling', by analogy with *pengenyat*, which is derived from *menyat* (example 14). *Pengenyat* is an attested form meaning 'request', derived by prefixing *peng* to *menyat* 'ask for'.

Note that 14 appears to be defective; it apparently has no nonvolitional passive voice *penyat*. This is for semantic and pragmatic reasons; asking for something is quintessentially a volitional activity, and it would make no sense to have a verb form that meant 'get accidentally asked for'. (Although one could conceivably concoct a strange situation in which it would be appropriate.) By default, the active form *menyat* is the only possible root.

In the above array I suggest *pohé*, *pujek*, *polo*, and *penat* as possible roots, by analogy with examples 5 - 7 where the roots have the same form as the nonvolitional passive. There is in fact no logical reason why we should not assume such is also the case in 10 - 13. However, I have opted against this idea. *Mujek* and *mohé* appear roughly ten times more frequently in my corpus than do *pujek* and *pohé*; and *menat* is something like seventy times more frequent than *penat*. (*Molo* and *polo* are much rarer and therefore cannot be compared with statistical significance in a corpus

---

\* Note that example 5, *tanem*, has a root identical to both a noun and the nonvolitional passive.

"only" a million words long.) It would be perverse to force the reader to look up these words twice, with his first search ending at a cross-reference, e.g. **menat** see **penat - menat**.

(I am not enthusiastic about theoretical representations; languages are concrete, and people learn them as they hear them. The ordering of entries in a dictionary should reflect psychological reality rather than sterile theoretical or logical considerations. For example, it would be contrary to the goal of a practical dictionary to insist that *ujek* is the root rather than *mujek*, with the latter being ordered not as § **mujek** but rather as § **ujek - mujek**. Even if the form *ujek* should turn out to exist, *mujek* must still occur a thousand times more often. In the intuition of a native speaker it is therefore the latter that is the basic form, and although in some circumstances he might be inclined to derive *ujek* from it, he will do so only by analogy with other, much more common derivations.)

## 2. PART OF SPEECH

The next component of an entry is an indication of the formal category or part of speech to which the lexical item belongs – this is usually given in abbreviated form, and includes the following terms: n. 'noun', v. 'verb', adv. 'adverb', adj. 'adjective', pron. 'pronoun', idiom n. p. 'idiomatic noun phrase'. (A complete list is given under the section "Symbols and Abbreviations".)

## 3. NOTE ON USAGE AND REGIONAL VARIATION

Next may follow an expression, in parentheses, giving certain kinds of general information about the lexeme in question, including:

### (a) Regional Variation

As mentioned earlier, I have encountered different regional variants. The notation (Tutoh) indicates that the expression in question is used in the Tutoh River area, and not in the Upper Baram. Conversely, the notation (Upper Baram) implies that an expression is not used in the Tutoh River area. These are the two regional variants of Eastern Penan with which I have consistent contact. I have occasionally encountered a third population, the people who live in the Upper Limbang River area. My impression is that their language usage is closer to that of the Tutoh people than to the Upper Baram people.

### (b) Neologisms

Sometimes the notation (neol., from M.) appears. This should be read as 'neologism, from Malay'. This is used to indicate "marked" expressions, that is, Malay words that have very recently been borrowed and are not fully integrated into the language and acceptable to all speakers. Malay is by far the greatest source of borrowing, and has been so for some time; and it is therefore important to distinguish between those words of Malay origin that have been adopted by all speakers, versus those that are only current among younger, bilingual Penans. It should be emphasized, however, that there is a continuum between fully integrated Malay words and those that have only taken the first tentative steps on the path to assimilation. Generally speaking, I am more likely to use the designation (M) in those cases where Penan already has a perfectly adequate expression for the meaning in question. Fully assimilated borrowings from Malay – and there are many of these -- are not marked; if you know Malay, you will easily identify many of them.

There are also a few lexemes marked (neol., from E.). These are borrowed from English directly, or via Malay. Like the above-mentioned marked borrowings from Malay, they are in general not known to all speakers.

### (c) Archaisms

I sometime indicate if a lexical unit is archaic, obsolete, or rare. Another notation that occasionally appears is (dilem). *Dilem* is a Penan word with the primary meaning 'deep', and a secondary meaning that signifies any

linguistic expression which the speaker views as archaic, *recherché*, or unlikely to be understood by the uninitiated, including children. It is a term supplied by my informants, and which I include only in those cases when it is the only datum concerning usage that I have so far obtained. A future edition of the dictionary will replace all instances of this vague term with more precise designations.

#### 4. DEFINIENDUM

*Definiendum* is Latin for 'that which is to be defined'. It is the expression that immediately follows the second occurrence of the symbol " § ", which in turn is followed by a double dash " -- ". It is written in **boldface**, and consists of one or more words, and usually other symbols as well.

When the lexical item represents a predicate, that is, has actants, then this expression is written in what we call *propositional form*, that is its actants are represented by the algebraic variables **X**, **Y**, and so on. Here is one example, chosen at random from among thousands.

§ **payo** 4. v. § -- **X payo tong Y** = 'Y suits X'

The expression **X payo tong Y** is the *definiendum*, and since **payo** denotes a logical predicate, it is written in *propositional form*.

Thus, the first **payo** in this entry is part of the indexing term, while the second **payo** is part of the *definiendum*.

Before going further we will mention a systematic exception to the convention we have just expounded. A whole class of entries does not display a *definiendum* separate from the indexing term; instead, the latter does double duty, serving both as a label for the entry and its *definiendum* as well. This occurs only in the cases of those lexical items that (a) have no homophones, and therefore no need for a numerical index to disambiguate them; and (b) not naming predicates, have no government pattern. Examples of these entries are easy to find; here is just one.

§ **aseu** n. = 'dog'

Strictly speaking, this entry should be written

§ **aseu** n. § -- **aseu** = 'dog'

It is obvious why I opt for the abbreviated form: information is not lost, nor is confusion generated.

Occasionally the *definiendum* is presented as two distinct alternative propositional forms, which are followed by a single definition. For example,

-- **X tuai [tong] Y jin Z** || **X tuai jin Z tai Y** = 'X comes to Y from Z'

Thus, using the above propositional forms one could construct the two following synonymous texts:

*Balang tuai tong Long Kerong jin Long Leleng.* or *Balang tuai jin Long Leleng tai Long Kerong.* 'Balang comes to Long Kerong from Long Leleng.'

In the case of transitive verbs, the *definiendum* normally terminates with one or more elements introduced by a slash " / ". Each such slash introduces one of the voices of the verb. The form that is represented in the *definiendum* is always in the active voice; the form that follows a slash without any other notation is in the passive voice; the form that follows a slash and the acronym "nvp" is in the nonvolitional passive voice; and the form that follows a slash and the abbreviation "reciproc." is in the reciprocal voice. (For information on these latter voices, see the grammar.) Only a few verbs display all four voices; here is one example.

§ **muja'** 1. v. § -- **X muja' Y pala Z / nuja'** / nvp **puja'** / reciproc. **pepuja'** = 'X strikes [at] Y with the sharp end of object Z, thrusting Z longitudinally'

In this example, four different patterns are shown, one in full, and three by means of conventional abbreviations. These conventional abbreviations are always to be expanded in the same, predictable fashion. This is best explained by example.

The passive form, represented by / **nuja'**, should be read as follows:

**Y nuja' X pala Z** = 'Y is struck by X with the sharp end of Z, thrust longitudinally'

The nonvolitional passive, represented by / nvp **puja'**, should be read as follows:

**Y puja' neu Z** = 'Y gets struck by Z, which is thrust longitudinally'

The reciprocal voice, represented by / reciproc. **pepuja'**, should be read as follows:

**X ngan Y pepuja' pala Z** = 'X and Y strike each other with Z, which is <are> thrust longitudinally'

## 5. SEMANTIC AND OTHER CONDITIONS

Sometimes following the *definiendum* is an expression introduced by a colon " : ". This is a condition or conditions that apply to one or more elements of the *definiendum*. Here are just two of many possible examples:

§ **kemedut** adv. § -- **X V kemedut : V** denotes travel = 'X V -s backwards (=backing up)'

§ **kelaset - ngelaset**<sup>2</sup> v. § -- **X ngelaset N Y kelebé** : **N** denotes a number, **Y** denotes an interval of time = 'X goes hungry, intentionally or unintentionally, for N Y' (roughly speaking, 'fast')

## 6. DEFINITION

Then comes the definition proper, or *definiens* (Latin for 'that which is doing the defining'). This follows the equal sign " = " and is enclosed in single quotes.

In a full-fledged ECD considerable care is taken in the choice of the terms used in the definition. Among other things, these must always be more semantically "primitive" than the lexical item being defined. This condition is essential in the case of a definition written in the same language as the expression being defined. However, as mentioned earlier, my definitions are written in English, and this gives me a significant degree of freedom in how I write them. This is because there may be more than one English expression which is synonymous with a given Penan lexical item. (Note that, in strict ECD terms, these are not definitions at all, but simply English translations; but they do match what people normally understand as "definitions", namely expressions used in dictionaries to describe the meanings of words. At this point I should repeat a caveat I have already expressed: given the provisional nature of the data, many of these English translations are bound to be inaccurate to one degree or another.)

## 7. INFORMAL NOTES, INCLUDING ENCYCLOPAEDIC INFORMATION

Following the definition there is sometimes a note enclosed in parentheses. This is additional explanatory material, often further clarifying the meaning, and presented in an informal fashion. Sometimes this includes encyclopaedic information, which is non-linguistic information about that part of the real world connected with, or labelled by, the lexical item in question.



## 8. COLLOCATIONS

Next are the collocations, each prefixed by the symbol "+". Each collocation has its own *definiendum* and *definiens*. In a full-fledged ECD collocations are listed in a conventional order, and defined using a kind of semantic metalanguage (making use of the already mentioned theory of *lexical functions*.) Because of the provisional nature of this dictionary we follow no particular order, and go no further than translating them into simple English. (We have already seen an illustration of a collocation zone, the one pertaining to the entry for *penyakit*, 'disease'.)

## 9. EXAMPLES

The symbol | is used to introduce the examples. These are normal Penan texts that use the lexical item being defined. I gloss almost all of them in English.

## 10. CROSS REFERENCES

Many entries have cross references appended to them. These are introduced by such terms as "see", "cf", "syn", and "ant". (N.B. In a fully developed ECD, synonyms and antonyms are listed among the lexical functions. In this dictionary I present them as cross references.)

## 11. ENGLISH EQUIVALENT

Finally, there may be an English equivalent, indicated by the sign •. Note that perfect synonymy between lexical items of different languages is rather rare. This is particularly true when one is dealing with cultures as mutually alien as English and Penan. Only about a quarter of Penan lexical items can be translated, with reasonable accuracy, by a single English word. That is why full-fledged definitions, often consisting of many words, are so necessary. We supply an "English equivalent" to a particular lexical item in Penan only if the meaning of both the English and Penan terms are quite close. When this "English equivalent" differs at all from the definition, it is of course the definition that takes precedence.

# OTHER FORMALISMS AND CONVENTIONS

## I. BRACKETS

### 1. ANGLE BRACKETS < >

Consider the following example:

-- **X menyat Y ngan <jin> Z / nenyat** = 'X asks for Y from Z'

The angle brackets < > enclose an alternative element. That is, this definition can be read in two ways:

**X menyat Y ngan Z / nenyat**

**X menyat Y jin Z / nenyat**

When interpreting angle brackets (wherever these may be used in the Penan text, e.g. in the definition or example zones), the following conventions apply.

(a) When the angle brackets contain a single element, then this element is an alternative to the element immediately preceding the angle brackets. For example,

Pala ojo <tapong> ma'ung ba. 'Use one's hand <a hat> to scoop up water.'

(b) When the angle brackets contain more than one element, then this group of elements is an alternative to the single element immediately preceding the angle brackets. For example,

Pala ojo <ujung kayeu> ma'ung ba. 'Use one's hand <a leaf> to scoop up water.'

(c) When the angle brackets contain more than one element, *and* when the first of these elements is a repeat of an element in the text immediately preceding the angle brackets, then the group of elements in the angle brackets is an alternative to the group of elements in the preceding text beginning with the element that is repeated. The following examples make this clear:

Luten jin malem keto mahang <keto bé' pata'> uban kayeu meden. -- This should be read as either -- Luten jin malem keto mahang uban kayeu meden. *or* Luten jin malem keto bé' pata' uban kayeu meden. 'The fire from last night is still burning because the wood is very resistant to going out.'

Tana' mapé neu levam <neu besalé> <neu kereja batang>. 'The land was washed away by the flood <swept away in the landslide> <levelled by the logging>.'

### 2. SQUARE BRACKETS [ ]

Square brackets [ ] enclose an optional element. For example:

Anak éh [jam] ngaken. 'a clever child'

This text can be read in either one of two ways, namely

Anak éh ngaken.

Anak éh jam ngaken.

### 3. "SYMMETRICAL" INTERPRETATION OF BRACKETS

When either angle or square brackets are used in both the *definiendum* and *definiens*, they are to be interpreted in a "symmetrical" fashion. That is, if the *definiendum* is read as including the bracketed text, then the *definiens* must also be read in the same way. Conversely, if the *definiendum* is read as not including the bracketed text, then the *definiens* must be read likewise. Some examples will make this clear.

-- **X nekuja' jin Y tai [bau] Z** = 'X jumps up from Y onto [the top of] Z'

must be read as either

-- **X nekuja' jin Y tai Z** = 'X jumps up from Y onto Z'

or

-- **X nekuja' jin Y tai bau Z** = 'X jumps up from Y onto the top of Z'

## II. CHOICE OF LETTERS TO REPRESENT VARIABLES

In some entries we specify how a given variable must be interpreted, e.g.

**X V Y pekua' éh W Z** : **W** is the passive form of verb **V**, **Z** is its agent = 'X V -s Y in the same way as Y is W -ed by Z'

However, in those entries (the majority) where no such specification is given, the letters standing for variables are to be interpreted as follows:

To represent nouns and noun phrases, we use the following letters, in order of preference: **X**, **Y**, **Z**, **W**, and **U**. e.g.,

**X tengé** = 'X alone'

**X medai Y** = 'X is afraid of Y'

**X belah Y ngan Z** = 'X is in the space between Y and Z'

**X miso Y jin Z tai W / niso** = 'X moves Y from Z to W'

**X pala Y ma'ung Z jin W tai U / na'ung** = 'X uses Y to scoop up liquid Z from W and pour <throw> it onto U'

To represent verbs or adjectives, we use, in order of preference, **V** and **W**. e.g.,

**X juk V** = 'X will or shall V'

**nyeliko X V, Y W** : **V** and **W** are verbs = 'at the same time as, and to an extent proportional to X's V -ing, Y W -s'

Note the use of the English verb morphology markers, -ing and -s, that we have attached to the variables; these have no formal status in the definition and could be omitted, but we include them to improve readability.

To represent clauses, we use, in order of preference, **Q** and **P**. "Clause" is meant in the widest possible sense: as an expression containing, or acting as, a predicate -- and since in Penan not only verb complements but also the

subject can be omitted, **Q** and **P** may designate anything from a single verb or adjective to an expression that could stand on its own as an independent sentence.

**Q jak** = 'Q now, before anything else' (e.g., Jian ke' purung surat si'ik iteu jak. 'Kindly read this small book first.')

**sadat Q hun P** = 'Q (will be caused) if P' (e.g., Sadat ké' nebet bilun hun tuah tosok lebé. -- lit., reason my get-left-behind-by plane if we-two talk for-a-long-time -- 'I will miss the plane if we talk too long.')

To represent cardinal numerals, we use **N**.

**N X [éh] lepah** = 'N units of time X ago'

### III. PAKAI versus PALA

It will be noticed that many *definienda* make use of an instrumental phrase introduced by either **pakai** or **pala**. e.g.

-- **X nawan Y pakai Z / tenawan** = 'X treats Y with medicine Z'

-- **X muja' Y pala Z / nuja'** = 'X strikes [at] Y with the sharp end of object Z, thrusting Z longitudinally'

Both **pakai** and **pala** mean 'with, by means of, using', and I believe that they are synonymous. It would seem that **pala** is the original Penan word, and **pakai** is a Malay borrowing that has become so thoroughly implanted in the language that it is nowadays used more often than **pala**. The fact that some entries use **pala** while others use **pakai** is a result of happenstance -- some examples I collected contain one, some the other.

## SYMBOLS AND ABBREVIATIONS

<i>symbol</i>	<i>meaning</i>
adj.	adjective
adv.	adverb
ant	antonym
cf	"see" (an invitation to look at another related entry)
conj.	conjunction
dilem	a Penan word meaning 'profound, hard to understand' (cf. explanation above)
E	English
interj.	interjection
n.	noun
idiom n. p.	idiomatic noun phrase
idiom v. p.	idiomatic verbal phrase
idiom adj. p.	idiomatic adverbial phrase
idiom adv. p.	idiomatic adverbial phrase
LF	lexical function (in future editions I will not use this term, but rather "collocation".)
M	Malay
neol., from E.	neologism, from English
neol., from M.	neologism, from Malay
nvp	nonvolitional passive voice
poss. adj.	possessive adjective
pron.	pronoun
reciproc.	reciprocal voice
see	you must go to the entry indicated to find the information you need
syn	synonym
v.	verb
var	variant of

BIBLICAL CITATIONS (The Penan translation of the Bible is a useful source of texts)

B30:23	Genesis chapter 30 verse 23 (the form of a biblical quote)
P10:12	Exodus chapter 10 verse 12 (the form of a biblical quote)

## PENAN PHONOLOGY AND ORTHOGRAPHY

Penan has a phonemic inventory similar to that of Malay, and the orthography devised for it by missionaries in the 1970's is similar to that of Malay. The spelling system created at that time is adequate in principle, but has been applied in a flawed manner in Bible translations. More on this presently.

The consonants are as follows:

b	d		g
p	t		k
	s		h
v			
	j		
m	n	ny	ng
	l		
	r		

There are also the glides *y* and *w*.

Most of these letters are sounded as they are in IPA (the International Phonetic Alphabet). For those not familiar with the latter system, they may safely give these letters the value they have in Malay or Indonesian orthography. "v" is a voiced bilabial fricative (the upper lip is pressed against the lower lip, rather than, as in English, against the lower teeth). "ʔ" is a glottal stop. "ng" is a velar nasal, and "ny" is a palatal nasal -- both having the same value as in Malay. "r" is flapped, as in Malay, Spanish, or Italian. "j" is as in English or Malay.

All consonants may occur word initially. (But note that when the glottal stop is word initial, it is never phonemic (= significant or meaning bearing), so it is never written in that position. The situation here is very much like in English, where, in careful or emphatic speech, we often insert a glottal stop before a word-initial vowel.) Only seven consonants may occur word finally, viz. /p/, /t/, /k/, /h/, /ʔ/, /n/ and /ng/.

Penan has six vowels, represented by the letters *u*, *i*, *o*, *a*, *é*, and *e*. The first four of these letters have the same value as they have in Malay, Spanish or Italian. *é* is somewhere between the *e* in "bed" and the *é* in "café". *e* represents a schwa (the neutral vowel, present in many languages including English - the second vowel in each of the words *bucket*, *malice*, *pigeon* is a schwa). In Penan the schwa has a high functional load (= is used a great deal and often distinguishes meaning), and, unlike in English, a schwa can occur as the vowel of a stressed syllable.

Any of the vowels /a/, /é/, /i/, /o/, /u/, if occurring in a word-final syllable, can be either long or short. The vowel is pronounced long if the word-final syllable is open, that is if the vowel forms the last sound of the word - e.g. *pala*, *bé*, *kari*, *mero*, *ngelepu*. The vowel is short if the syllable is closed, i.e. if any consonant, including the glottal stop, is the last sound of the word - e.g. *tanaʔ*, *béʔ*, *suhat*, *térék*, *murip*. (Note that final long vowels are often, and optionally, pronounced with a glottal creak (i.e. a glottal semi-closure), which is never to be confused with the abrupt phonemic glottal stop /ʔ/ which can only follow a short vowel.)

The schwa may not appear in a word-final open syllable - i.e. it must always be followed by a consonant. (e.g. *merem*, *pet*, *keʔ*, but never *\*ke*, *\*te*, *\*me*, etc.) However, there is a single exception to this rule -- *e*, 'mummy' or

'mom' (This is also the single case where a schwa may occur as the first phoneme of a word!).

The following diphthongs occur: /ai/, /au/, /éu/, /oi/, /eu/, /ei/. These only occur word finally, and in open syllables (e.g. *barei*, *pesikeu*, *malai*). They are often terminated with a glottal closure, but this is by definition never phonemic -- to repeat, diphthongs are never followed by a consonant except across word boundaries. (And consequently diphthongs -- and the schwa for that matter -- do not participate in the aforementioned long-short vowel contrast).

Note that the letter combinations *ia*, *ie*, *iu* and *io* do not denote true diphthongs, even though they are written as such in words like *iah* 'he, she, it', *ieng* 'there is none', *keviu* 'whistle' and *tio* 'directly'. These represent, phonetically, /iyah/, /iyeng/, /keviyu/, and /tiyo/. Similarly, the letter combinations *ua*, *ue*, *ui* and *uo* -- in e.g. *lua*, *ueng*, *uit*, and *keruong*, -- also do not denote diphthongs. In fact they represent /uwa/, /uwe/, /uwi/, and /uwo/ respectively. However, since in both sets of cases, *y* and *w* are phonetic, not phonemic (= they are predictable in their occurrence, and not meaning-differentiating) we refrain from writing them. In those cases where *y* and *w* are written -- e.g. *awah*, *liwen*, *ayah*, *oyoi* -- they are full-fledged phonemes, and serve to distinguish meaning (cf. *liwen* vs. *lisen*, *ayah* vs. *anah*).

The only possible syllabic structures are V, CV, CVC, CVCV, CVCVC, etc. ("V" means 'vowel' and "C" means 'consonant'.) In other words, there are no consonant clusters in Penan, not even combinations of a nasal and following homorganic stop. For example, when Malay words like *sendiri* or *belanja* are borrowed into Penan, they become *sediri* and *belaja*. Consonant clusters can only occur across word boundaries, e.g. *kang kaput* 'Indian cuckoo', *jalan tekelet* 'shortcut', *sakit kenin* 'to feel emotionally hurt'.

All syllables are equally stressed.

The spelling used in the Penan Bible is deficient in respect of word-final syllables. The most frequently occurring defect of this orthography is its marking of a phonemic (=distinctive) glottal stop where it does not occur - e.g. *\*palaa'* instead of *pala*, *\*leboo'* instead of *lebo*, *\*avéé'* for *avé*, *\*akeu'* for *akeu*, *\*iteu'* for *iteu*.<sup>\*</sup> The second most important defect is the arbitrary way in which final vowels in open syllables are represented. Most are written with a double vowel letter, but some are not. Although any given word is spelled in a consistent way throughout the text, there is no logic as to which words are spelled how - thus although *pané* is always written correctly, *beté*, a word having the same long final vowel sound, is always spelled *\*betéé'*. Similarly, while *bé* is always rendered correctly, *sé* is written *\*séé'*.

With the agreement and encouragement of our principal linguistic informant, Jokim, we have chosen to correct these deficiencies and inconsistencies. Therefore all double vowel graphemes have been suppressed, and /' only marks glottal stops that are phonemic.

---

<sup>\*</sup>When linguists place an asterisk in front of a word or expression, they are identifying the latter as being ill-formed or ungrammatical. I believe there are two reasons for these errors made by the missionary. First, a vowel at the end of a word is pronounced long, and often with what linguists call a "glottal creak". The glottal creak is predictable and therefore not phonemic (=not meaning-bearing), but the missionary heard it and wanted to represent it, so she used the symbol normally reserved for the glottal stop. Second, diphthongs terminate with a true glottal stop, and in a "narrow" or phonetic transcription, as opposed to a phonemic one, it is not incorrect to spell e.g. *iteu'* rather than *iteu*. However, this glottal stop is not phonemic; it is entirely predictable and therefore should not be represented in a practical orthography.

# PENAN GRAMMAR – A SKETCH

What follows is a brief and very provisional outline of Penan grammar. Some areas of grammar are dealt with in much more detail than others; some important areas are left out entirely. I have decided to include it despite its obvious deficiencies, since it will be of considerable utility to those interested in learning the basics of the language.

## I. GENERAL

Nouns, adjectives, and intransitive verbs lack any form of inflection. In general, Penan lexemes do not undergo morphological changes which signify case, number, sex, gender, or tense.

## II. WORD ORDER AND THE BASIC SENTENCE

The word order of a sentence containing a transitive verb is Subject - Verb - Object.

*Labang na'at lamin.* 'Labang see house.' (Labang sees a house.)

Thus word order determines semantic relationship.

*Aseu nga'at anak.* 'Dog bite child.' (The dog bites the child.)

*Anak nga'at aseu.* 'Child bite dog.' (The child bites the dog.)

In Penan sentences it is quite normal for the object or other complements to be omitted, as well as the grammatical subject. This can result in a sentence that is quite ambiguous unless used in context.

*Bé' kelo.* 'No want.' "We don't want it" or perhaps "They don't want it" or perhaps "I don't want to do it", etc.

There is no verb "to be", so a complete sentence can be formed without an element that is obviously a verb.

*Lakei merek, redo besau.* 'Man angry, woman sad.' (The man is angry, the woman is sad.)

However, note that an expression such as

*Lakei merek*

can constitute either a sentence, as in the preceding example, or a noun phrase, as it appears in the following sentence:

*Lakei merek bé' moko siteu.* 'Man angry not stay here.' (The angry man is not staying here.)

Noun phrases may be constructed using the relativizing particle *éh*, meaning 'that, which, who'.

*Akeu teme'u ngan lakei éh merek.* 'I meet with man who angry.' (I met with the angry man.)

*lamin éh lepah potong.* 'House that already get-burned. (The house that has burned down.)

The normal order of nouns and their modifiers is noun first - modifier following.

*lamin maréng.* 'house new' (a new house)

*kerita peritah* 'car government' (a government car)

*redo iteu* 'woman this' (this woman)

However, sometimes adjectives are used in front of nouns – compare the two following examples.

*Lakei inah merek.* 'Man that angry.' (That man is angry.)



*Merek lakei inah.* 'Angry man that.' (That man is **angry**.)

It is more accurate to speak of the preposing of predicates rather than the preposing of adjectives. Consider the following pair, in which the second example contains a preposed verb:

*Akeu lakau.* 'I leave.' (I am leaving.)

*Lakau akeu.* 'Leave I. (I am **leaving**.)

Such preposing of a predicate serves to present it as the most important element in the sentence, one that denotes, for example, new or vital information.

It can also serve to create a declarative sentence, where the corresponding post-posed form would simply be a noun phrase. E.g.,

*ba pana* 'water hot' (hot water)

*Pana ba.* 'Hot water.' (The water is hot.)

The above cases should be distinguished from that of complex adjectives, which are phrases consisting of a noun and a preposed adjective. Some of these are idiomatic. For example,

*jian adet* 'well behaved'

*Anak iteu jian adet.* 'Child this well behaved.' (This child is well-behaved.)

*Jian adet anak iteu.* (This child is **well**-behaved.)

*sa'at akam* 'in bad health'

*Iah sa'at akam.* 'He bad health.' (He is in ill health.)

*besau kenin* 'sad, sorrowful'

*Redo eh besau kenin.* 'Woman who sad hearted.' (A woman who is sad.)

*Besau kenin redo inah.* 'Sad hearted woman that.' (That woman is **sad**.)

### III. DEMONSTRATIVES, DEFINITE VS INDEFINITE

The demonstrative adjectives are *iteu* 'this, these' and *inah* 'that, those', and *itai* 'yon (= far from both the speaker and the hearer)'. The first two of these are often abbreviated to *teu* and *nah*, respectively. They always follow the noun they modify.

*lakei iteu* 'this man <these men>'

*lamin inah* 'that house <those houses>'

*iteu*, *inah* and *itai* can also be pronouns, 'this one <these ones>' and 'that one <those ones>' respectively.

There is no indefinite article; nor is there an obligatory definite article. Thus,

*Babui kuman bua.* 'Pig eat fruit.'

is vague, and would most likely mean 'pigs eat fruit'. To give a definite referent to either of the two nouns, we can use demonstrative adjectives, such as *iteu* or *teu* ('this'), or *inah* or *nah* ('that'), which thus often play the role of definite article. Also commonly used is the element *ri* -- *X ri* means 'the X just mentioned'. Thus,

*Babui nah kuman bua.* 'That pig eats fruit.'

*Babui nah kuman bua ri.* 'That pig ate the fruit.'

*Babui kuman bua ri.* 'A pig (or pigs) ate the fruit.'

#### IV. VERBS

Penan verbs do not change in form or add affixes to denote number, mood, aspect, or tense. For example,

*Akeu suai alut.* 'I make boat.'

could be glossed as 'I made a boat <boats> or I am making a boat <boats> or I make boats.

Vague sentences like the above are not a problem as long as the context indicates when the action denoted by the verb occurs. However, one may add certain elements to the sentence to clearly indicate whether the past, the present or the future is meant, as outlined in the following section.

##### 1. TENSE MARKERS

There is no obligatory tense; that is, it is possible, even normal, to use verbs without reference to the time frame of the event they denote. However, both the past and the future can be denoted by the preposed auxiliary verbs *juk* 'will', *lepah* 'already', and *maréng* 'just'.

(a) *juk* 'shall, will' (denotes any future time)

*Akeu juk lakau.* 'I will travel.'

*Kitong juk tai beté sagam.* 'Kitong will go hunting tomorrow.'

(b) *lepah* 'already' (denotes any past time)

*Akeu lepah tai Marudi.* 'I already go Marudi.' (I went to Marudi.)

(c) *maréng* 'just, recently' (denotes a time in the very recent past)

*Lakei inah maréng ala babui.* 'Man that just get pig.' (That man just bagged a pig.)

Note that adverbs of time can often be used instead of these tense markers, and in some cases can be used along with them. Unlike the tense markers, when an adverb of time occurs immediately next to a verb, it follows it. Some examples:

(d) *malem* 'yesterday', *dau iteu* 'today', *sagam* 'tomorrow', *hun iteu* 'now', *sahau* 'before, long ago, in the old days'

*Iah peleka malem.* 'He leave yesterday.' (He left yesterday.)

*Lejeng lakau sagam.* or *Lejeng juk lakau sagam.* 'Lejeng (will) travel tomorrow.'

*Akeu suai alut hun iteu.* 'I make boat now.'

*Sahau iah moko tong tana.* 'In the old days she stay on land.' (She used to live in the forest.)

## 2. THE IMPERATIVE AND HORTATIVE

Penan verbs have two moods: indicative and imperative. Commands and exhortations are denoted by the normal active verb, and/or by certain lexical elements.

(a) An imperative can consist of an active verb without subject.

*Tuai!* 'Come!  
*Moko sitai!* 'Stay there!'

(b) A pronoun acting as subject can be used for additional clarity or emphasis.

*Lakau ketenah ke!* 'Travel ahead thou.' (**You** go straight ahead.)

(c) Very frequently the element *jian* 'good' is anteposed to a command. Some examples:

*Jian ke' mihin tinen ko!* 'Good thou bring mother thy.' (Bring your mother.)  
*Jian ka'au teu éh moko uban sagam pina kelunan juk tuai miting.* 'Good thou this who remain because tomorrow many person will come meeting.' (It would be best if it were you who stayed because tomorrow many people will come to the meeting.)  
*Hun néh kenat, jian lu' moko, amé pané.* 'If it thus, good we remain, we speak.' (If such is the case, let us stay, we will speak.)  
*Boh lakei inah seruh, "Hun mé' la'au ri' jian ku' put ka'an iteu na' péh éh bu'un néh malui."* 'So man that think, "If we hungry like-this, good I shoot animal this al though (two words) it beginning it transform.' (And the man thought, "Since we are hungry, let me just shoot this animal, even though it has started to transform.)  
*Jian tuah petat.* 'Good we move\_apart.' (Let us go separate ways.)  
*"Jian pa' éh avé da' mai awah ke' kelap.* 'Good even\_though it arrive for\_sure do\_not only thou flee.' (Even if he comes right up to you, do not run away.)

(d) The lexemes *kejera* or *turut*, both meaning 'let', can be used :

*Kejera jak redo lemanai inah moko ngan amé ngio avé jah polo dau.* 'Allow now woman unmarried that remain with us approximately until ten day.' (Let that unmarried woman stay with us for about ten days.)

*Turut ke' tuah tai kuman jak.* 'Let in\_order we go eat now.' (Let's go and eat now.)

(e) An extremely common way of expressing an imperative meaning is by means of the particle *da'*. *Da'* occurs at the end of a clause. A speaker appends *da'* to a clause in order to signal that the action or state denoted by that clause is prescribed, desirable, or necessary. As such it has a broad meaning, or perhaps a number of meanings, only one of which can be described as imperative. Here are some examples of its use in an imperative sense:

*Hun inah da' tuai keteleu.* 'When that IMPERATIVE\_PARTICLE come you\_three.' (Kindly come at that time.)  
*Kivu jalan iteu ke' da'.* 'Follow path this thou IMPERATIVE\_PARTICLE.' (Follow this path here.)  
*Jam jam ke' lakau da'.* 'Be\_good\_at be\_good\_at thou travel IMPERATIVE\_PARTICLE.' (Be careful while travelling -- lit., 'be very good at travelling')

(f) Negative commands are denoted by the anteposing of *mai* 'do not'.

*Mai kon!* 'Do\_not eat [it]!  
*Mai pemung ngan kelunan éh sa'at.* 'Do\_not associate with person who bad.' (Do not associate with bad people.)  
*Mai pekerek medok inah, dai néh nga'at ka'au.* 'Do\_not anger pig\_tailed\_macaque that, lest it bite thee.' (Do not make that monkey angry or it will bite you.)

### 3. SERIALIZATION OF VERBS

In Penan the same subject can govern two, three or even four verbs in series.

*Irah ngelangan nyelapang paka*. 'They aim-at shoot-at branch.' (They aimed and shot at the branch.)

*Irah pekelawan ngelangan nejau pakai inan tobo*. 'They mutually-fight aim-at throw-in-spear-fashion use stalk ginger.' (They fought each other by hurling stalks of ginger at each other. (presumably in fun!))

*Ka'au nekedéng petem nerokon maneu réh*. 'Thou stand be-immobile dig-in-heels assault them.' (You stand immobile with your feet dug in the ground and fight them.)

A clause containing serialized verbs refers to a single circumstance or activity which is characterized by a series of events or states. The order of the verbs corresponds to the chronological or logical order of those events or states.

Note that passive verbs can also be serialized:

*Ba Buto napit kelapah néh malem rai*. 'Tutoh River was-crossed was-passed-over by her last night.'

### 4. VOICE

Transitive verbs have four voices: active, passive, nonvolitional passive, and reciprocal.

#### (a) Active Voice

The verb is used in the active voice in sentences of the form X V Y, where X is the subject and denotes the first semantic actant\* of verb V, and Y is the direct object and denotes the second semantic actant of V. For example,

*Lakei put tahat*. 'Man shoot blowdart.' (A man shoots a blowdart, men shoot blowdarts, etc.)

#### (b) Passive Voice

Verbs use the passive about as frequently as the active voice. Almost every transitive verb has a passive form.

*Tahat neput lakei*. 'Blowdart shot-by man.' (A blowdart was shot by a men, blowdarts are shot by men, etc.)

#### (i) Syntax of the Passive

The passive form of the verb can simply be interchanged with the active form, and the only syntactic difference between an active construction and the corresponding passive construction is the reversal of the position of the nouns or pronouns that label the first and second actants of the verb. For example,

*Labang suai alut*. 'Labang built a boat.'

*Alut senuai Labang*. 'The boat was built by Labang.'

#### (ii) Function of the Passive

As in English, the passive is used to topicalize an element in the sentence by promoting it from direct object to grammatical subject. For example, note the subtle difference in meaning rendered by the translations of the above

---

\* *semantic actant* is a term used in Meaning-Text theory (which is the framework for the Explanatory-Combinatorial Dictionary). Roughly speaking, a semantic actant of a given lexeme is an argument of the predicate that that lexeme encodes. The lexeme *put* 'shoot' illustrated immediately below encodes a predicate with two arguments: the "shooter" and the "thing shot". These two semantic actants can be realized in the surface text in different ways. In the case of the active verb, the "shooter" is realized as the subject, and thus we label it as the first semantic actant.

examples. In the first sentence you are being told something about Labang, and in the second you are being told something about the boat. It is unlikely that you will want to topicalize "alut" unless you have a *particular* boat in mind, and that is why we choose to say "the boat" in the second translation rather than "a boat". Since Penan has no obligatory definite article, one of the uses of the passive is therefore to add the element of determinacy to the meaning of a noun or noun phrase by promoting it to subject.

### (iii) Morphological Form of the Passive.

A passive verb is formed from an active verb by the addition of a prefix, or an infix to the initial syllable, often along with certain other phonological changes. These rules governing these are almost exceptionless, that is, it is almost always possible to predict the passive form of a verb simply by looking at its active form. Below is a list of initial sounds of active verbs followed by the initial sounds of the corresponding passive verbs, and then examples of each. (V denotes any given vowel)

#### 1. mV --> nV

**menyat --> nenyat.** *Amé menyat penolong jin peritah.* 'We ask for help from the government.' *Penolong éh nenyat mé' jin peritah.* 'The help that is asked by us from the government.'

**mukap --> nukap.** *Akeu mukap tapé.* 'I open the door.' *Tapé nukap ké'.* 'The door is opened by me.'

#### 2. sV --> senV

**suai --> senuai.** *Labang suai lamin.* 'Labang builds a house.' *Lamin senuai Labang.* 'The house is built by Labang.'

**seva' --> seneva'.** *Anak tekep seva' tinen.* 'A child should respect a mother.' *Tinen tekep seneva' anak.* 'A mother should be respected by a child.'

**salin --> senalin.** *Guru salin surat.* 'The teacher translates the book.' *Surat senalin guru.* 'The book is translated by the teacher.'

#### 3.i. nyV --> senV

**nyurung --> senurung.** *Ayau nyurung lebo mé'.* 'The enemy attacks our land.' *Lebo mé' senurung ayau.* 'Our land is attacked by the enemy.'

**nyatek --> senatek.** *Akeu nyatek uai.* 'I cut through the rattan.' *Uai senatek ké'.* 'The rattan is cut through by me.'

3.ii. **nyV --> sV** NOTE: This rule applies only in the case of verbs derived from nouns that begin with /s/. A number of verbs are formed from such nouns by the expedient of replacing the initial /s/ with /ny/. For example, "surat" 'letter, book' has a corresponding verb, "nyurat", meaning 'create a letter or a book', i.e. 'write'. Similarly, "selapang" means 'shotgun', and "nyelapang" 'shoot with a shotgun.'

**nyurat --> surat.** *Iah nyurat surat iteu.* 'He wrote this letter.' *Surat iteu surat néh.* 'This letter was written by him.' (Note that the alternative form *senurat* is also possible -- *Surat iteu senurat néh.*)

**nyelapang --> selapang.** *Lakei inah nyelapang babui.* 'That man shot a pig.' *Babui selapang lakei inah.* 'A pig was shot by that man.'

#### 4. nV --> tV

**neranying --> teranying.** *Iah neranying kayeu.* 'He sharpened the wood to a point.' *Kayeu teranying néh.* 'The wood was sharpened to a point by him.'

**nanem --> tanem.** *Iah nanem patai.* 'She buried the corpse.' *Patai éh tanem néh.* 'The corpse was buried by her.'

**tV--> tenV**

**taket --> tenaket.** *Redo taket lamin.* 'The woman steps into the house.' *Lamin tenaket redo.* 'The house is stepped into by the woman.'

**tutup --> tenutup.** *Akeu tutup usit tapé.* 'I have closed the door.' *Usit tapé tenutup [ké].* 'The door has been closed [by me].'

**6. ngV --> kenV**

**ngasau --> kenasau.** *Buang éh ngasau lamin.* 'A bear that makes a mess of a house.' *Lamin éh kenasau buang.* 'A house that is messed up by a bear.'

**ngaléng --> kenaléng.** *Anak ngaléng bateu.* 'The child threw a rock.' *Bateu kenaléng anak.* 'A rock was thrown by a child.'

**7. ngV --> kV**

**ngejeku --> kejekeu.** *Redo ngejeku kerayung.* 'The woman folded the shirt.' *Kerayung inah kejekeu redo.* 'That shirt was folded by the woman.'

**ngelepek --> kelepek.** *Iah ngelepek kayeu maneu bo viheu.* 'He bent the stick to make a spring trap.' *Kayeu iteu kelepek néh.* 'This stick was bent by him.'

**8. kV --> kenV**

**kahang --> kenahang.** *Kahang luten létrik.* 'Light an electric lamp.' *luten éh kenahang.* 'a fire that has been lit.'

**kivu --> kenivu.** *Labang kivu jalan iteu.* 'Labang followed this trail.' *Jalan éh kenivu Labang tai sitai.* 'The trail followed by Labang goes there.'

**9. memV --> penV**

**memata' --> penata'** *Memata' titui létrik.* 'Turn off an electric light.' *Titui létrik penata'.* 'An electric light is turned off.'

**memutui --> penutui.** *Memutui tulang.* 'Break a bone.' *Tulang éh penutui.* 'A bone that has been broken.'

**10. pV --> penV**

**purung --> penurung.** *Anak purung surat.* 'The child reads the book.' *Surat penurung anak.* 'The book is read by the child.'

**pitah --> penitah.** *Pitah ka'an ngan ké.* 'Look for food for me.' *Ka'an éh penitah.* 'Food that was looked for.'

11. A few passives are formed in an irregular fashion.

For example, we would expect the passive of *nga'at* 'bite' to be *\*kena'at*. In fact, it is *ga'at*.

**nga'at --> ga'at.** *Aseu nga'at lakei.* 'The dog bit a man.' *Lakei ga'at aseu.* 'The man was bitten by a dog.'

**na'at --> ta'an.** *Sara' na'at lakei inah.* 'Sarah saw that man.' *Lakei inah ta'an Sara'.* 'That man was seen by Sarah.'

**jam --> kejam.** *Akeu jam ha' Penan.* 'I know the Penan language.' *Kekat ha' éh kejam ké.* 'All the languages known by me.'

**papit --> napit.** *Iah papit ba.* 'He crossed the river.' *Ba napit néh.* 'The river is crossed by him.'

12. A number of passives are formed according to the following rule:

**peC --> peC** (C is usually either *p* or *k*. Normally, this form of the passive only occurs if a verb does not already undergo one of the alterations listed above. For example, as already noted *purung* 'count' is passivized as *penurung*, and therefore does not take the form *\*pepurung*.)

As can be seen, passive verbs of this type have exactly the same morphological form as their active counterparts.

**peposot --> peposot.** *Akeu peposot anak.* 'I put the child to rest.' *Anak peposot ké'.* 'The child is put to rest by me.'

**pepesep --> pepesep.** *Iah pepesep anak.* 'She fed liquid to the child. *Anak pepesep néh.* 'The child was fed liquid by her.'

**pekelap --> pekelap.** *Iah pekelap ka'an.* 'He frightened the animals away.' *Bé' pina ka'an éh pekelap néh.* 'The animals frightened away by him were not many in number.'

It must be noted that passive verbs of this type are problematic when ambiguity with the active forms would result. For example, although *Sara' pepesep Balang.* could logically mean either 'Sarah fed liquid to Balang' or 'Sarah was fed liquid to by Balang', it would be interpreted as the former. Only by using the Class 3 pronouns *ké'* 'me', *ko'* 'thee', *néh* 'her, him, it', which are used to identify the agent of a passive verb, can these passive verbs be disambiguated from the homophonous active forms. Refer to the section on pronouns.

### (c) Nonvolitional Passive Voice

In addition to the active and the passive, many transitive verbs present a third voice, which I have chosen to call the nonvolitional passive. While the normal passive, described above, preserves the propositional meaning of the active form, the nonvolitional passive alters the meaning in a particular way. It is used to denote changes of state that are viewed as accidental, for example acts of nature, and therefore not brought about as a result of anyone's will or volition. I will explain by way of example.

Take the active verb *nasa'* and its passive form *tenasa'*. *Akeu nasa' lipan.* 'I destroyed the bulldozer' vs. *Lipan tenasa' ké'.* 'The bulldozer was destroyed by me.' But this verb also has a nonvolitional passive: *tasa'*. Thus, *Lipan tasa' neu keloho bateu.* 'The bulldozer was destroyed by a falling rock.'

The active verb *meseng* and its passive form *neseng* mean 'prevent from moving'. Thus, *Seradu meseng kerita.* means 'The soldiers prevented the car from travelling onwards.' The passive form would be *Kerita neseng seradu.* *Peseng* is the nonvolitional passive of this verb, and allows such sentences as *Ayau peseng neu jawin uai.* 'The enemy was immobilized by the (thorny) rattan vines (i.e. because he got caught up in them).'

As we see from these examples, the element *neu* (sometimes *naneu*) links the nonvolitional passive to the instrument of the action. However, in most cases the instrument is not mentioned; this should not be surprising, since the whole point of this verb form is to emphasize the fortuitous nature of the event, and in many cases the cause may not even be known.

Here are some more examples containing nonvolitional passives, followed by the active and passive forms of the relevant verbs in parentheses.

*Lamin néh potong.* 'His house got burned.' (*motong / notong* 'burn')  
*Kekat ujung kayeu purat [neu kepu].* 'All the leaves got scattered [by the wind].' (*murat / nurat* 'scatter')  
*Utang putui.* 'The staff got broken.' (*memutui / penutui* 'break')  
*Tajau bila'.* 'The urn got shattered.' (*memila' / benila'* 'shatter')

Note that sometimes either the regular passive or the nonvolitional passive can be used, even where one and the same event is being described. However, there is an important difference in meaning between the two forms. For example, you could say both *Lipan tenasa' néh* and *Lipan tasa' neu néh* -- but while the first means simply 'The bulldozer was destroyed by him', the second means 'The bulldozer got destroyed on account of him', that is, the event in question is viewed as the accidental result of his action. Similarly, *Lem pasa vam, ka'au adang juk senopé ké'* uses the regular passive, *senopé*, of the verb *nyopé*, 'defeat'; it means 'In the coming race you will surely be defeated by me.' However, *Kayan sitai sopé neu ayau réh*, which uses the nonvolitional form *sopé*, means 'The Kayans there got defeated by their enemies' -- in other words, in the latter case the defeat is depicted as an accident or misfortune, and their adversaries the mere instruments of this misfortune, rather than its conscious agents.

Here are some more examples of nonvolitional passives:

*mujek / nujek* 'spill' --> *pujek* 'get spilled'  
*mavut / navut* 'pull out' --> *pavut* 'get pulled out'  
*mikah / nikah* 'scatter' --> *pikah* 'get scattered'  
*migu' / nigu'* --> *pigu'* 'get banged against'  
*mohé / nohé* 'clean' --> *pohé* 'get cleaned'  
*mejé / nejé* 'tear' --> *pejé* 'get torn'  
*melen / nelen* 'roll up' --> *pelen* 'get rolled up'  
*mepa' / nepa'* 'squash' --> *pepa'* 'get squashed'  
*meseng / neseng* 'plug up' --> *peseng* 'get plugged up'  
*ngelepu / kelepu* 'set free' --> *pelepu* 'get free, escape'  
*ngaténg / kenaténg* 'tighten' --> *pekaténg* 'get taut'  
*nanem / tenanem* 'bury' --> *tanem* 'get buried'

#### (d) Reciprocal Voice

Many transitive verbs accept a prefix *pe-* (often accompanied by a change of the initial phoneme) and thereby denote a mutual or reciprocal action. This is best shown by example:

*tosok* 'talk' --> *petosok* 'talk to each other' Irah petosok. 'They talk to each other.'

*nuyang* 'have illicit sex with' --> *petuyang* 'Have illicit sex with each other.' Rawah petuyang. 'Those two have illicit sex with each other.'

*nga'at* 'bite' --> *pega'at* 'bite each other' Aseu pega'at. 'The dogs bite each other.'

*nesen* 'remember' --> *petesen* 'remember each other' Rawah keto petesen. 'The two of them still remember each other.'

These verbs accept the transformation  $X \text{ ngan } Y \text{ pe-}V \leftrightarrow X \text{ pe-}V \text{ ngan } Y$ . For example,

*Amé ngan irah éh tuai nepah petosok.* --> *Amé petosok ngan irah éh tuai nepah.* 'We and those who drop in to visit are talking' --> 'We talk with the people who drop in to visit'

*Redo inah ngan lakei iteu petuyang.* --> *Redo inah petuyang ngan lakei iteu.* 'That woman and this man have illicit sex with each other' --> 'That woman has illicit sex with this man.'

(Some adjectives and intransitive verbs also have reciprocal forms. For example,

*tengé [jin]* 'distinct, different [from]' --> *petengé* 'separate from each other' *Duah lamin inah petengé.* 'Those two houses are separated from each other.'

*pemung* 'connected or mixed with' --> *pepemung* 'mixed together' *Tulin jelai pepemung ngan tulin parai.* 'The maize grains are mixed in with the rice grains.'



## V. PRONOUNS

### 1. INVENTORY, MEANINGS, FORMS

Pronouns in Penan reflect number and person, but not sex. Below is a list of the semantic categories that Penan pronouns distinguish.

1st person singular (I)	akeu
2nd person singular (thou (=singular you))	ka'au
3rd person singular (he, she, it)	iah
1st person dual inclusive (I and thou)	tuah
1st person plural inclusive (we who are here, including you)	itam
1st person plural inclusive (we who are many, including you)	uleu
1st person dual exclusive (I and he <she>, but not you)	amo
1st person paucal exclusive (we few, but not you)	mételeu
1st person plural exclusive (we many, but not you)	amé
2nd person dual (you two)	kawah
2nd person paucal (you, being few in number)	keteleu
2nd person plural (you many)	ka'ah
3rd person dual (those two)	rawah
3rd person paucal (they, being few in number)*	rételeu
3rd person plural (they)	irah

Most of these pronouns have more than one morphological form, in a manner similar to English pronouns (cf he, him, his). Many have three distinct forms, and which of these is to be used in a given situation depends on the syntactic and/or stylistic role that the pronoun plays therein. It is thus convenient to distinguish three morphological classes of pronouns, and I have chosen to label these Class 1, Class 2, and Class 3. All pronouns listed above are of Class 1.

The following table lists their Class 2 and Class 3 counterparts.

CLASS 1	CLASS 2	CLASS 3
akeu	ku'	ké'
ka'au	ke'	ko'
iah	éh	néh
tuah	toh	toh
itam	tam	tam
uleu	lu'	lu'
amo	mo	mo
mételeu	mételeu	mételeu
amé	mé'	mé'

---

\* The pronouns *mételeu*, *keteleu*, and *rételeu* are trial in form, in that each incorporates the word *teleu* 'three'. However, nowadays they are commonly used to refer to a group of three, four or five, and sometimes even more. Probably they were originally semantically trial: when questioned on this matter, some informants say that these pronouns can only be used correctly to denote three persons, and cases where they denote a greater number are due to laziness or a lapsus on the part of the speaker.

kawah	koh	koh
keteleu	keteleu	keteleu
ka'ah	keh	keh
rawah	roh	roh
rételeu	rételeu	rételeu*
irah	réh	réh

As can be seen, only the singular pronouns have three distinct forms. The dual and plural pronouns have just two distinct forms, their Class 2 and Class 3 forms being identical. The paucal pronouns are morphologically invariable.

With rare exceptions only Class 1 forms are used when the pronoun is a subject which is the first element of a sentence. Class I forms may also be used when the pronoun is a subject in non-initial position, or when the pronoun is a direct object. They are also used in isolation, that is when the pronoun is the only element of an utterance.

*Akeu na'at iah.* 'I see him'  
*Iah na'at akeu.* 'He sees me.'  
*Boh akeu na'at iah.* 'Then I saw him.'  
*Boh iah na'at akeu.* 'Then he saw me.'  
*Sé na'at iah?* -- *Akeu.* 'Who saw him? -- Me.'

Class 2 forms may be used when the pronoun is a subject in non-initial position, or when the pronoun is a direct object. They may not be used in isolation or when the pronoun is a subject in initial position.

*Boh ku' na'at éh.* 'Then I saw him.'  
*Boh éh na'at ku'.* 'Then he saw me.'

(compare the impossible

\**Ku' na'at iah.* 'I see him'  
 \**Éh na'at akeu.* 'He sees me.')

Class 3 forms are used

(1) to represent the agent of a passive sentence

*Akeu ta'an néh.* 'I was seen by him.'  
*Iah ta'an ké.* 'He was seen by me.'

(2) as possessive adjectives

*Lamin ké.* 'My house.' *Lebo réh.* 'Their land.' *Aseu ko.* 'Your dog.'

(3) after certain conjunctions and prepositions.

*Uban néh bé' omok tuai,*... 'Because he cannot come,...'  
*Jalan réh pegen.* 'So that they can sleep.' (can also mean 'a place for them to sleep')  
*Ke-mé' poko tana.* 'So that we can keep our land.'  
*Pu'un rigit ngan ko?* 'Do you have money on you?'

(Class 2, and especially Class 3, pronouns display certain properties of clitics. They cannot be used independently, but are postposed to other elements, which can be called their "hosts". However, they are not clitic-like insofar as they maintain complete prosodic autonomy from their hosts; all syllables in any Penan sentence are pronounced with an even, equal stress. Furthermore, it is hard to see how they exhibit any tendency to become affixes; they are all postposed, but Penan is completely bereft of suffixes; it has only prefixes and infixes (the latter being inserted only in the first syllable – see the section on passive verbs). Furthermore, in the above examples of Class 2 and Class 3 pronouns, with two exceptions, none can be in any way phonetically conjoined with its host; such a circumstance would violate Penan phonology, which sticks to a purely CVCV

---

\* In very slow and careful speech, *rételeu* may be pronounced *réh teleu*.

regime. The first exception is seen in the case of *Boh éh na'at ku'*, in which the first two words could be conjoined without violating the CVCV condition. However, they are not; they are not pronounced [bohéh], but rather [boh 'éh], and very distinctly at that. The other exception is the penultimate string, which in standard orthography is written *Kemé' poko tana'*. But in this case it is *ke* that is a clitic – a proclitic to be precise (in fact, probably a full-fledged prefix) – and *mé'* which is its host.)

## 2. PRONOUNS USED AS COMITATIVES

Certain pronouns can be used to link two nouns where, in English, we would use *and* or *along with*. For example,

*Musa rawah Arun.* (lit., 'Moses those-two Aaron') 'Moses and Aaron.'

*O Ada kawah Sila jian koh menéng ha' ké'.* (lit., 'O Ada you-two Sila kindly you-two listen word my') 'O Ada and Sila, listen to my words.'

In most cases when a pronoun is used in this way, i.e. *comitatively*, only one noun is used, and the pronoun may either precede or follow it. How this works, and what is denoted, is best explained by example:

*Amé Kimu tuai malem.* 'We-many Kimu arrive last-night.' (Kimu and some others of us came last night.)

*Sagap amo kivu jalan inah.* 'Sagap we-two follow path that.' = *Amo Sagap kivu jalan inah.* 'We-two Sagap follow path that.' (Sagap and I took that trail.)

*Boh Musa rawah kivu.* 'So Musa the-two-of-them follow.' (So Musa and his companion followed.)

*Boh roh lakei éh singat tulat ka'an roh ri'.* 'Then the-two-of-them man who-is greedy divide-up meat belonging-to-the-two-of-them the.' (Then he and the greedy man divided up their meat.)

*Tovo Sagap rételeu kivu jalan inah...* 'When Sagap the-three-of-them follow path that.' = *Tovo rételeu Sagap kivu jalan inah...* (When Sagap and the others took that trail...)

*Irah Luti peleka malem.* 'Luti and the others left last night.'

The following syntactic rule emerges from the above examples:

P X or X P (where P is the pronoun, and X the noun) is interpreted as 'a person or persons denoted by P, along with the other person or persons denoted by P, which includes X' (e.g. *rételeu Musa* would be processed by this rule as "a person or persons denoted by *rételeu* (=those persons who are several in number), along with the other person or persons denoted by *rételeu*, which includes Musa." -- that is to say, 'those (several) people including Musa'.

## VI. COMPARATIVE AND SUPERLATIVE

The comparative is of the form

*X V jin Y*

where *X* and *Y* are the terms being compared, and *V* is an adjective.

Some examples:

*Surat inah ja'au jin surat iteu.* 'That book is bigger than this book.'

*Bok ké' kebit jin bok ko'.* 'My hair is longer than your hair.'

*Urip ko' si'ik jin urip ké'.* 'Life your small than life mine.' (You are younger than me.)

*Juhit marang pelapah bau jin paka' kayeu.* 'The bird flies much higher than the tree branches.'

To express the meaning 'X is the V –est of all X –s' we use the following construction.

*X V jin kekat Y*

*Surat inah ja'au jin kekat surat.* 'Book that big than all book.' (That book is the biggest of all.)

*Bakéh néh mu'un jin kekat bakéh néh.* 'Friend his true from all friend his.' (The best of all his friends.)

To express an absolute superlative we use *matek* 'the ...-est'.

*Éh mah matek jian?* 'Which most good?' (Which is the best?)  
*Tokong inah matek bau.* 'That hill is the highest.'

## VII. NUMERALS

The cardinal numerals are as follows:

1	jah
2	duah
3	teleu
4	pat
5	lemah
6	nem
7	tuju
8	ayah
9	pi'en
10	polo <i>or</i> jah polo
11	jah polo jah
12	jah polo duah
13	jah polo teleu
20	duah polo
21	duah polo jah
30	teleu polo
100	jah ato
200	duah ato
1000	jah ibeu
2000	duah ibeu

Ordinal numbers (except for the first one) are derived by the prefix *ke-*

*X bu'un* 'the first X'; *keruah X* 'the second X'; *keteleu X* 'the third X'; *kepat X* 'the fourth X'; *kelemah X* 'the fifth X'; and so on.

## VIII. DERIVATIONAL MORPHOLOGY

Derivational morphology is the process by which lexemes are formed from other lexemes by the addition or alteration of phonemes. In Penan, morphological alterations operate only on the beginning of the word, through consonant alteration, prefixing, or a combination of the two.

Here are some examples of common kinds of word formation in Penan.

### 1. VERB FORMATION

#### (a) Transitive verbs formed from intransitive verbs

There is a process of verb formation whereby an intransitive verb with a propositional form Y V is transformed into a transitive verb with a propositional form X V Y, the latter having the meaning 'X causes that Y V'. This is done by prefixing *pe-*, sometimes with modification of the initial phoneme[s] as well.

*kelap* 'flee' --> *pekelap* / *pekelap* 'scare away' *Mai pekelap babui*. 'Don't scare away the wild pig.'  
*sala* 'wrong' --> *pesala* / *pesala* 'blame' *Iah pesala padé néh*. 'He blamed his brother.'  
*pemung* 'connected or mixed with' --> *pepemung* / *pepemung* 'combine with' *Redo tulin inah pepemung tulin parai ngan tulin jelai*. 'That crazy woman mixed rice grains in with corn kernels.'  
*to'o* 'dry' --> *peto'o* / *teno'o* 'dry, make dry' *Iah peto'o kerayung tong peté*. 'She dried the shirt in the sun.'  
*lena* 'clear, lucid' --> *pekelena* / *pekelena* 'explain' *Akeu pekelena tosok*. 'I explained the story.'  
*vun vun* 'be piled up' --> *pepevun* / *pepevun* 'pile up' *Akeu pepevun kayeu*. 'I pile up wood.'  
*nesen* 'be conscious of' --> *petesen* / *petesen* 'remind' *Ko' petesen anak namung inah tong tinen néh*. 'You be sure to remind that adopted child about its mother.'

#### (b) Verbs formed from nouns or adjectives

Some verbs are formed from a noun or adjective by prefixing *ng* or *nge* to the latter:

*ngejian* 'to improve' (from *jian*, 'good'); *ngejajan* 'to chant' (from *jajan*, 'chant'); *ngavet* 'to wear a loincloth' (from *avet*, 'loincloth'); *ngeradau* 'to shout' (from *radau*, 'shout'); *ngeripen* 'to enslave' (from *ripen*, 'slave').

### 2. NOUN FORMATION

#### (a) The prefix *peng-*

Many nouns are formed from other words by adding the prefix *peng*. These fall into the following types:

(i) *peng-* prefixed to a word beginning with a vowel:

*urip* 'life' --> *pengurip* 'age'  
*ada* 'light' --> *pengada* 'glass'

(ii) *peng-* prefixed to a word beginning with a consonant, without phonological alteration of the latter, but with obligatory insertion of *e* (to retain the CVCV syllable structure):

*kaya* 'rich' --> *pengekaya* 'wealth'  
*jam* 'know' --> *pengejam* 'knowledge'  
*keta* 'be suffering' --> *pengeketa* 'affliction'  
*lakau* 'travel' --> *pengelakau* 'trip'  
*ga* 'end, final' --> *pengega* 'end, termination'  
*bé* 'all' --> *pengebé* 'totality'

*mutau* 'tired' --> *pengemutau* 'fatigue'

(iii) *peng* prefixed to a word beginning with a consonant, resulting in replacement of the latter with the *ng* of the prefix.

*mena* 'give' --> *pengena* 'gift'

*maneu* 'do' --> *penganeu* 'deed'

*ngelan* 'hope' --> *pengelan* 'hope, faith'

*kuman* 'eat' --> *penguman* 'food'

*kelo* 'want' --> *pengelo* 'desire'

(iv) *peng* prefixed to a word beginning with *s*, resulting in the suppression of both this *s*, and the *ng* of the prefix, and their replacement by the phoneme *ny*:

*suai* 'create' --> *penyuai* 'creation'

*seruh* 'think' --> *penyeruh* 'thought'

*sala* 'wrong' --> *penyala* 'evil deed'

*sakit* 'be hurting' --> *penyakit* 'illness'

The semantic alteration resulting from noun formation is not always regular or predictable. For example, one might incorrectly guess that *pengeja'au* means 'size'.

*ja'au* 'big' --> *pengeja'au* 'important person or chief'<sup>\*</sup>

#### **(b) The prefix *ke-***

Some nouns are derived by prefixing *ke-* to a verb or adjective. This process creates abstract nouns in much the same way as the suffix *-ness* does in English (e.g. fit --> fitness, pale --> paleness)

*ja'au* 'big' --> *keja'au* 'size'

*gahang* 'strong' --> *kegahang* 'strength'

*mé* 'tasty' --> *kemé* 'tastiness'

*menya* 'ashamed' --> *kemenya* 'shame'

---

<sup>\*</sup> *pengeja'au* can also mean 'the greater part of'.