# Contents

# Chapter 1

# Lojban As We Mangle It In Lojbanistan: About This Book

*coi lojban.* *coi rodo*

## 1.1 What is Lojban?

Lojban (pronounced *LOZH-bahn*) is a constructed language. Previous versions of the language were called *Loglan* by Dr. James Cooke Brown, who founded the Loglan Project and started the development of the language in 1955. The goals for the language were first described in the open literature in the article *Loglan*, published in *Scientific American*, June, 1960. Made well-known by that article and by occasional references in science fiction (most notably in Robert Heinlein's novel *The Moon Is A Harsh Mistress* and computer publications, Loglan and Lojban have been built over four decades by dozens of workers and hundreds of supporters, led since 1987 by The Logical Language Group (who are the publishers of this book).

There are thousands of artificial languages (of which Esperanto is the best-known), but Loglan/Lojban has been engineered to make it unique in several ways. The follow-

ing are the main features of Lojban:

- Lojban is designed to be used by people in communication with each other, and possibly in the future with computers.
- Lojban is designed to be neutral between cultures.
- Lojban grammar is based on the principles of predicate logic.
- Lojban has an unambiguous yet flexible grammar.
- Lojban has phonetic spelling, and unambiguously resolves its sounds into words.
- Lojban is simple compared to natural languages; it is easy to learn.
- Lojban's 1300 root words can be easily combined to form a vocabulary of millions of words.
- Lojban is regular; the rules of the language are without exceptions.
- Lojban attempts to remove restrictions on creative and clear thought and communication.
- Lojban has a variety of uses, ranging from the creative to the scientific, from the theoretical to the practical.
- Lojban has been demonstrated in translation and in original works of prose and poetry.

## 1.2 What is this book?

This book is what is called a *reference grammar*. It attempts to expound the whole Lojban language, or at least as much of it as is understood at present. Lojban is a rich language with many features, and an attempt has been made to discover the functions of those features. The word *discover* is used advisedly; Lojban was not *invented* by any one person or committee. Often, grammatical features were introduced into the language long before their usage was fully understood. Sometimes they were introduced for one reason, only to prove more useful for other reasons not recognized at the time.

By intention, this book is complete in description but not in explanation. For every rule in the formal Lojban grammar (given in Chapter 21), there is a bit of explanation and an example somewhere in the book, and often a great deal more than a bit. In essence, Chapter 2 gives a brief overview of the language, Chapter 21 gives the formal structure of the language, and the chapters in between put semantic flest on those formal bones. I hope that eventually more grammatical material founded on (or even correcting) the explanations in this book will become available.

Nevertheless, the publication of this book is, in one sense, the completion of a long period of language evolution. With the exception of a possible revision of the language that will not even be considered until five years from publication date, and any revisions of this book needed to correct outright errors, the language described in this book will not be changing by deliberate act of its creators any more. Instead, language change will take place in the form of new vocabulary — Lojban does not yet have nearly the

vocabulary it needs to be a fully usable language of the modern world, as Chapter 12 explains — and through the irregular natural processes of drift and (who knows?) native-speaker evolution. (Teach your children Lojban!) You can learn the language described here with assurance that (unlike previous versions of Lojban and Loglan, as well as most other artificial languages) it will not be subject to further fiddling by language-meisters.

It is probably worth mentioning that this book was written somewhat piecemeal. Each chapter began life as an explication of a specific Lojban topic; only later did these begin to clump together into a larger structure of words and ideas. Therefore, there are perhaps not as many cross-references as there should be. However, I have attempted to make the index as comprehensive as possible.

Each chapter has a descriptive title, often involving some play on words; this is an attempt to make the chapters more memorable. The title of Chapter 1 (which you are now reading), for example, is an allusion to the book *English As We Speak It In Ireland*, by P. W. Joyce, which is sort of informal reference work about Hiberno-English. *Lojbanistan* is both an imaginary country where Lojban is the native language, and a term for the actual community of Lojban-speakers, scattered over the world. Why *mangle*? As yet, nobody in the real Lojbanistan speaks the language at all well, by the standards of the imaginary Lojbanistan; that is one of the circumstances this book is meant to help remedy.

## 1.3 What are the typographical conventions of this book?

Each chapter is broken into numbered sections; each section contains a mixture of expository text, numbered examples, and possibly tables.

The reader will notice a certain similarity in the examples used throughout the book. One chapter after another rings the changes on the self-same sentences:

**Example 1.3.1**
mi klama le zarci
I go-to that-which-I-describe-as-a store.
I go to the store.

will become wearisomely familiar before Chapter 21 is reached. This method is deliberate; I have tried to use simple and (eventually) familiar examples wherever possible, to avoid obscuring new grammatical points with new vocabulary. Of course, this is not the method of a textbook, but this book is not a textbook (although people have learned Lojban from it and its predecessors). Rather, it is intended both for self-learning (of course, at present would-be Lojban teachers must be self-learners) and to serve as a reference in the usual sense, for looking up obscure points about the language.

It is useful to talk further about Example 1.3.1 for what it illustrates about examples in this book. Examples usually occupy three lines. The first of these is in Lojban, the second in a word-by-word literal translation of the Lojban into English, and the third in colloquial English. The second and third lines are sometimes called the *literal translation* and the *colloquial translation* respectively. Sometimes, when clarity is not sacrificed thereby, one or both are omitted. If there is more than one Lojban sentence, it generally means that they have the same meaning.

Words are sometimes surrounded by angle brackets. In Lojban texts, these enclose optional grammatical particles that may (in the context of the particular example) be

either omitted or included. In literal translations, they enclose words that are used as conventional translations of specific Lojban words, but don't have exactly the meanings or uses that the English word would suggest. In Chapter 3, forward slashes surround phonetic representations in the International Phonetic Alphabet.

Many of the tables, especially those placed at the head of various sections, are in three columns. The first column contains Lojban words discussed in that section; the second column contains the grammatical category (represented by an UPPER CASE Lojban word) to which the word belongs, and the third column contains a brief English gloss, not necessarily or typically a full explanation. Other tables are explained in context.

A few Lojban words are used in this book as technical terms. All of these are explained in Chapter 2, except for a few used only in single chapters, which are explained in the introductory sections of those chapters.

## 1.4 Disclaimers

It is necessary to add, alas, that the examples used in this book do not refer to any existing person, place, or institution, and that any such resemblance is entirely coincidental and unintentional, and not intended to give offense.

When definitions and place structures of gismu, and especially of lujvo, are given in this book, they may differ from those given in the English-Lojban dictionary (which, as of this writing, is not yet published). If so, the information given in the dictionary supersedes whatever is given here.

## 1.5 Acknowledgements and Credits

Although the bulk of this book was written for the Logical Language Group (LLG) by John Cowan, who is represented by the occasional authorial *I*, certain chapters were first written by others and then heavily edited by me to fit into this book.

In particular: Chapter 2 is a fusion of originally separate documents, one by Athelstan, and one by Nora Tansky LeChevalier and Bob LeChevalier; Chapter 3 and Chapter 4 were originally written by Bob LeChevalier with contributions by Chuck Barton; Chapter 12 was originally written (in much longer form) by Nick Nicholas; the dialogue near the end of Chapter 13 was contributed by Nora Tansky LeChevalier; Chapter 15 and parts of Chapter 16 were originally by Bob LeChevalier; and the YACC grammar in Chapter 21 is the work of several hands, but is primarily by Bob LeChevalier and Jeff Taylor. The BNF grammar, which is also in Chapter 21, was originally written by me, then rewritten by Clark Nelson, and finally touched up by me again.

The research into natural languages from which parts of Chapter 5 draw their material was performed by Ivan Derzhanski. LLG acknowledges his kind permission to use the fruits of his research.

The pictures in this book were drawn by Nora Tansky LeChevalier, except for the picture appearing in Chapter 4, which is by Sylvia Rutiser.

I would like to thank the following people for their detailed reviews, suggestions, comments, and early detection of my embarrassing errors in Lojban, logic, English, and cross-references: Nick Nicholas, Mark Shoulson, Veijo Vilva, Colin Fine, And Rosta, Jorge Llambias, Iain Alexander, Paulo S. L. M. Barreto, Robert J. Chassell,

Ivan Derzhanski, Jim Carter, Irene Gates, Bob LeChevalier, John Parks-Clifford (also known as *pc*), and Nora Tansky LeChevalier.

Nick Nicholas (NSN) would like to thank the following Lojbanists: Mark Shoulson, Veijo Vilva, Colin Fine, And Rosta, and Iain Alexander for their suggestions and comments; John Cowan, for his extensive comments, his exemplary trailblazing of Lojban grammar, and for solving the *manskapi* dilemma for NSN; Jorge Llambias, for his even more extensive comments, and for forcing NSN to think more than he was inclined to; Bob LeChevalier, for his skeptical overview of the issue, his encouragement, and for scouring all Lojban text his computer has been burdened with for lujvo; Nora Tansky LeChevalier, for writing the program converting old rafsi text to new rafsi text, and sparing NSN from embarrassing errors; and Jim Carter, for his dogged persistence in analyzing lujvo algorithmically, which inspired this research, and for first identifying the three lujvo classes.

Of course, the entire Loglan Project owes a considerable debt to James Cooke Brown as the language inventor, and also to several earlier contributors to the development of the language. Especially noteworthy are Doug Landauer, Jeff Prothero, Scott Layson, Jeff Taylor, and Bob McIvor. Final responsibility for the remaining errors and infelicities is solely mine.

## 1.6 Informal Bibliography

The founding document for the Loglan Project, of which this book is one of the products, is *Loglan 1: A Logical Language* by James Cooke Brown (4th ed. 1989, The Loglan Institute, Gainesville, Florida, U.S.A.) The language described therein is not Lojban, but is very close to it and may be considered an ancestral version. It is regrettably necessary to state that nothing in this book has been approved by Dr. Brown, and that the very existence of Lojban is disapproved of by him.

The logic of Lojban, such as it is, owes a good deal to the American philosopher W. v.O. Quine, especially *Word and Object* (1960, M.I.T. Press). Much of Quine's philosophical writings, especially on observation sentences, reads like a literal translation from Lojban.

The theory of negation expounded in Chapter 15 is derived from a reading of Larry Horn's work *The Natural History of Negation*.

Of course, neither Brown nor Quine nor Horn is in any way responsible for the uses or misuses I have made of their works.

Depending on just when you are reading this book, there may be three other books about Lojban available: a textbook, a Lojban/English dictionary, and a book containing general information about Lojban. You can probably get these books, if they have been published, from the same place where you got this book. In addition, other books not yet foreseen may also exist.

## 1.7 Captions to Pictures

The following examples list the Lojban caption, with a translation, for the picture at the head of each chapter. If a chapter's picture has no caption, *(none)* is specified instead.

**Example 1.7.1**
coi lojban. coi rodo
Greetings, O Lojban! Greetings, all-of you

**Example 1.7.2**
(none)

**Example 1.7.3**
.i .ai .i .ai .o
[untranslatable]

**Example 1.7.4**
jbobliku
Lojbanic-blocks

**Example 1.7.5**
(none)

**Example 1.7.6**
lei re nanmu cu bevri le re nanmu
The-mass-of two men carry the two men
Two men (jointly) carry two men (both of them).

**Example 1.7.7**
ma drani danfu
.i di'e
.i di'u
.i dei
.i ri
.i do'i
[What sumti] is-the-correct type-of-answer?
The-next-sentence.
The-previous-sentence.
This-sentence.
The-previous-sentence.
An-unspecified-utterance.

**Example 1.7.8**
ko viska re prenu poi bruna la santas.
[You!] see two persons who-are brothers-of Santa.

**Example 1.7.9**
(none)

**Example 1.7.10**
za'o klama
[superfective] come/go
Something goes (or comes) for too long.

**Example 1.7.11**
le si'o kunti
The concept-of emptiness

**Example 1.7.12**
(none)

**Example 1.7.13**
.oi ro'i ro'a ro'e
[Pain!] [emotional] [social] [mental]

**Example 1.7.14**
(none)

**Example 1.7.15**
mi na'e lumci le karce
I other-than wash the car
I didn't wash the car.

**Example 1.7.16**
drata mupli pe'u .djan.
another example [please] John
Another example, John, please!

**Example 1.7.17**
zai xanlerfu by. ly. .obu .jy by. .abu ny.
[Shift] hand-letters l o j b a n
"Lojban" in the manual alphabet

**Example 1.7.18**
no no
0 0

**Example 1.7.19**
(none)

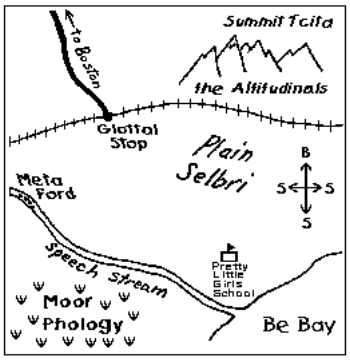
**Example 1.7.20**
(none)

**Example 1.7.21**
(none)

## 1.8 Boring Legalities

# Chapter 2

# A Quick Tour of Lojban Grammar, With Diagrams




## 2.1 The concept of the bridi

This chapter gives diagrammed examples of basic Lojban sentence structures. The most general pattern is covered first, followed by successive variations on the basic components of the Lojban sentence. There are many more capabilities not covered in this chapter, but covered in detail in later chapters, so this chapter is a *quick tour* of the material later covered more slowly throughout the book. It also introduces most of the Lojban words used to discuss Lojban grammar.

Let us consider John and Sam and three statements about them:

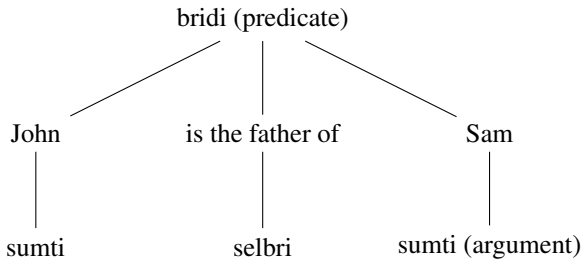
**Example 2.1.1**
John is the father of Sam.

**Example 2.1.2**
John hits Sam.

**Example 2.1.3**
John is taller than Sam.

These examples all describe relationships between John and Sam. However, in English, we use the noun *father* to describe a static relationship in Example 2.1.1, the verb *hits* to describe an active relationship in Example 2.1.2, and the adjective *taller* to describe an attributive relationship in Example 2.1.3. In Lojban we make no such grammatical distinctions; these three sentences, when expressed in Lojban, are structurally identical. The same part of speech is used to represent the relationship. In formal logic this whole structure is called a *predication*; in Lojban it is called a *bridi*, and the central part of speech is the *selbri*. Logicians refer to the things thus related as arguments, while Lojbanists call them *sumti*. These Lojban terms will be used for the rest of the book.

**Diagram 2.1**




In a relationship, there are a definite number of things being related. In English, for example, *give* has three places: the donor, the recipient and the gift. For example:

**Example 2.1.4**
John gives Sam the book.

and

**Example 2.1.5**
Sam gives John the book.

mean two different things because the relative positions of *John* and *Sam* have been switched. Further,

**Example 2.1.6**
The book gives John Sam.

seems strange to us merely because the places are being filled by unorthodox arguments. The relationship expressed by *give* has not changed.

In Lojban, each selbri has a specified number and type of arguments, known collectively as its *place structure*. The simplest kind of selbri consists of a single root word, called a *gismu*, and the definition in a dictionary gives the place structure explicitly.

The primary task of constructing a Lojban sentence, after choosing the relationship itself, is deciding what you will use to fill in the sumti places.

This book uses the Lojban terms *bridi*, *sumti*, and *selbri*, because it is best to come to understand them independently of the English associations of the corresponding words, which are only roughly similar in meaning anyhow.

The Lojban examples in this chapter (but not in the rest of the book) use a blue background behind each sumti, and a green background behind each selbri, to help you to tell them apart.

## 2.2 Pronunciation

Detailed pronunciation and spelling rules are given in Chapter 3, but what follows will keep the reader from going too far astray while digesting this chapter.

Lojban has six recognized vowels: *a*, *e*, *i*, *o*, *u* and *y*. The first five are roughly pronounced as *a* as in *father*, *e* as in *let*, *i* as in *machine*, *o* as in *dome* and *u* as in *flute*. *y* is pronounced as the sound called *schwa*, that is, as the unstressed *a* as in *about* or *around*.

Twelve consonants in Lojban are pronounced more or less as their counterparts are in English: *b*, *d*, *f*, *k*, *l*, *m*, *n*, *p*, *r*, *t*, *v* and *z*. The letter *c*, on the other hand is pronounced as the *sh* in *hush*, while *j* is its voiced counterpart, the sound of the *s* in *pleasure*. *g* is always pronounced as it is in *gift*, never as in *giant*. *s* is as in *sell*, never as in *rose*. The sound of *x* is not found in English in normal words. It is found as *ch* in Scottish *loch*, as *j* in Spanish *junta*, and as *ch* in German *Bach*; it also appears in the English interjection *yecchh!*. It gets easier to say as you practice it. The letter *r* can be trilled, but doesn't have to be.

The Lojban diphthongs *ai*, *ei*, *oi*, and *au* are pronounced much as in the English words *sigh*, *say*, *boy*, and *how*. Other Lojban diphthongs begin with an *i* pronounced like English *y* (for example, *io* is pronounced *yo*) or else with a *u* pronounced like English *w* (for example, *ua* is pronounced *wa*).

Lojban also has three *semi-letters*: the period, the comma and the apostrophe. The period represents a glottal stop or a pause; it is a required stoppage of the flow of air in the speech stream. The apostrophe sounds just like the English letter *h*. Unlike a regular consonant, it is not found at the beginning or end of a word, nor is it found adjacent to a consonant; it is only found between two vowels. The comma has no sound associated with it, and is used to separate syllables that might ordinarily run together. It is not used in this chapter.

Stress falls on the next to the last syllable of all words, unless that vowel is *y*, which is never stressed; in such words the third-to-last syllable is stressed. If a word only has one syllable, then that syllable is not stressed.

All Lojban words are pronounced as they are spelled: there are no silent letters.

## 2.3 Words that can act as sumti

Here is a short list of single words used as sumti. This list provides examples only, not the entire set of such words, which may be found in Chapter 7.

**Definition 2.1**

| | |
|---|---|
| **mi** | I/me, we/us |
| **do** | you |
| **ti** | this, these |
| **ta** | that, those |
| **tu** | that far away, those far away |
| **zo'e** | unspecified value (used when a sumti is unimportant or obvious.) |

Lojban sumti are not specific as to number (singular or plural), nor gender (masculine/feminine/neutral). Such distinctions can be optionally added by methods that are beyond the scope of this chapter.

The cmavo *ti*, *ta*, and *tu* refer to whatever the speaker is pointing at, and should not be used to refer to things that cannot in principle be pointed at.

Names may also be used as sumti, provided they are preceded with the word *la*:

**Definition 2.2**

| | |
|---|---|
| **la meris.** | the one/ones named Mary |
| **la djan.** | the one/ones named John |

Other Lojban spelling versions are possible for names from other languages, and there are restrictions on which letters may appear in Lojban names: see Chapter 6 for more information.

## 2.4 Some words used to indicate selbri relations

Here is a short list of some words used as Lojban selbri in this chapter:

**Definition 2.3**

| | |
|---|---|
| **vecnu** | $x_1$ (seller) sells $x_2$ (goods) to $x_3$ (buyer) for $x_4$ (price) |
| **tavla** | $x_1$ (talker) talks to $x_2$ (audience) about $x_3$ (topic) in language $x_4$ |
| **sutra** | $x_1$ (agent) is fast at doing $x_2$ (action) |
| **blari'o** | $x_1$ (object/light source) is blue-green |
| **melbi** | $x_1$ (object/idea) is beautiful to $x_2$ (observer) by standard $x_3$ |
| **cutci** | $x_1$ is a shoe/boot for $x_2$ (foot) made of $x_3$ (material) |
| **bajra** | $x_1$ runs on $x_2$ (surface) using $x_3$ (limbs) in manner $x_4$ (gait) |
| **klama** | $x_1$ goes/comes to $x_2$ (destination) from $x_3$ (origin point) via $x_4$ (route) using $x_5$ (means of transportation) |
| **pluka** | $x_1$ pleases/is pleasing to $x_2$ (experiencer) under conditions $x_3$ |
| **gerku** | $x_1$ is a dog of breed $x_2$ |
| **kurji** | $x_1$ takes care of $x_2$ |
| **kanro** | $x_1$ is healthy by standard $x_2$ |
| **stali** | $x_1$ stays/remains with $x_2$ |
| **zarci** | $x_1$ is a market/store/shop selling $x_2$ (products) operated by $x_3$ (storekeeper) |

Each selbri (relation) has a specific rule that defines the role of each sumti in the bridi, based on its position. In the table above, that order was expressed by labeling the sumti positions as $x_1$, $x_2$, $x_3$, and $x_4$. Like the table in Section 2.3, this table is far

from complete: in fact, no complete table can exist, because Lojban allows new words to be created (in specified ways) whenever a speaker or writer finds the existing supply of words inadequate. This notion is a basic difference between Lojban (and some other languages such as German and Chinese) and English; in English, most people are very leery of using words that "aren't in the dictionary". Lojbanists are encouraged to invent new words; doing so is a major way of participating in the development of the language. Chapter 4 explains how to make new words, and Chapter 12 explains how to give them appropriate meanings.

## 2.5 Some simple Lojban bridi

Let's look at a simple Lojban bridi. The place structure of the gismu *tavla* is

**Example 2.5.1**
$x_1$ talks to $x_2$ about $x_3$ in language $x_4$

where the *x*es with following numbers represent the various arguments that could be inserted at the given positions in the English sentence. For example:

**Example 2.5.2**
John talks to Sam about engineering in Lojban.

has *John* in the $x_1$ place, *Sam* in the $x_2$ place, *engineering* in the $x_3$ place, and *Lojban* in the $x_4$ place, and could be paraphrased:

**Example 2.5.3**
Talking is going on,
with speaker John
and listener Sam
and subject matter engineering
and language Lojban.

The Lojban bridi corresponding to Example 2.5.1 will have the form

**Example 2.5.4**
$x_1$ ⟨*cu*⟩ tavla $x_2$ $x_3$ $x_4$

The word *cu* serves as a separator between any preceding sumti and the selbri. It can often be omitted, as in the following examples.

**Example 2.5.5**
mi tavla do zo'e zo'e
I talk to you about something in some language.

**Example 2.5.6**
do tavla mi ta zo'e
You talk to me about that thing in a language.

**Example 2.5.7**
mi tavla zo'e tu ti
I talk to someone about that thing yonder in this language.

(Example 2.5.7 is a bit unusual, as there is no easy way to point to a language; one might point to a copy of this book, and hope the meaning gets across!) When there are one or more occurrences of the cmavo *zo'e* at the end of a bridi, they may be omitted, a process called *ellipsis*. Example 2.5.5 and Example 2.5.6 may be expressed thus:

**Example 2.5.8**
mi tavla do
I talk to you (about something in some language).

**Example 2.5.9**
do tavla mi ta
You talk to me about that thing (in some language).

Note that Example 2.5.7 is not subject to ellipsis by this direct method, as the *zo'e* in it is not at the end of the bridi.

## 2.6 Variant bridi structure

Consider the sentence

**Example 2.6.1**
mi ⟨*cu*⟩ vecnu ti ta zo'e
seller-$x_1$ sells goods-sold-$x_2$ buyer-$x_3$ price-$x_4$
I sell this to that for some price.
I sell this-thing/these-things to that-buyer/those-buyers.
(The price is obvious or unimportant.)

Example 2.6.1 has one sumti (the $x_1$) before the selbri. It is also possible to put more than one sumti before the selbri, without changing the order of sumti:

**Example 2.6.2**
mi ti ⟨*cu*⟩ vecnu ta
seller-$x_1$ goods-sold-$x_2$ sells buyer-$x_3$
I this sell to that.
(translates as stilted or poetic English)
I this thing do sell to that buyer.

**Example 2.6.3**
mi ti ta ⟨*cu*⟩ vecnu
seller-$x_1$ goods-sold-$x_2$ buyer-$x_3$ sells
I this to-that sell.
(translates as stilted or poetic English)
I this thing to that buyer do sell.

Example 2.6.1 through Example 2.6.3 mean the same thing. Usually, placing more than one sumti before the selbri is done for style or for emphasis on the sumti that

are out-of-place from their normal position. (Native speakers of languages other than English may prefer such orders.) If there are no sumti before the selbri, then it is understood that the $x_1$ sumti value is equivalent to *zo'e*; i.e. unimportant or obvious, and therefore not given. Any sumti after the selbri start counting from $x_2$.

**Example 2.6.4**
ta ⟨*cu*⟩ melbi
object/idea-$x_1$ is-beautiful (to someone by some standard)
That/those is/are beautiful.
That is beautiful.
Those are beautiful.

when the $x_1$ is omitted, becomes:

**Example 2.6.5**
melbi
unspecified-$x_1$ is-beautiful (to someone by some standard)
Beautiful!
It's beautiful!

Omitting the $x_1$ adds emphasis to the selbri relation, which has become first in the sentence. This kind of sentence is termed an observative, because it is often used when someone first observes or takes note of the relationship, and wishes to quickly communicate it to someone else. Commonly understood English observatives include *Smoke!* upon seeing smoke or smelling the odor, or *Car!* to a person crossing the street who might be in danger. Any Lojban selbri can be used as an observative if no sumti appear before the selbri.

The word *cu* does not occur in an observative; *cu* is a separator, and there must be a sumti before the selbri that needs to be kept separate for *cu* to be used. With no sumti preceding the selbri, *cu* is not permitted. Short words like *cu* which serve grammatical functions are called *cmavo* in Lojban.

## 2.7 Varying the order of sumti

For one reason or another you may want to change the order, placing one particular sumti at the front of the bridi. The cmavo *se*, when placed before the last word of the selbri, will switch the meanings of the first and second sumti places. So

**Example 2.7.1**
mi tavla do ti
I talk to you about this.

has the same meaning as

**Example 2.7.2**
do se tavla mi ti
You are talked to by me about this.

The cmavo *te*, when used in the same location, switches the meanings of the first and the third sumti places.

**Example 2.7.3**
mi tavla do ti
I talk to you about this.

has the same meaning as

**Example 2.7.4**
ti te tavla do mi
This is talked about to you by me.

Note that only the first and third sumti have switched places; the second sumti has remained in the second place.

The cmavo *ve* and *xe* switch the first and fourth sumti places, and the first and fifth sumti places, respectively. These changes in the order of places are known as *conversions*, and the *se*, *te*, *ve*, and *xe* cmavo are said to convert the selbri.

More than one of these operators may be used on a given selbri at one time, and in such a case they are evaluated from left to right. However, in practice they are used one at a time, as there are better tools for complex manipulation of the sumti places. See Chapter 5 for details.

The effect is similar to what in English is called the *passive voice*. In Lojban, the converted selbri has a new place structure that is renumbered to reflect the place reversal, thus having effects when such a conversion is used in combination with other constructs such as *le selbri* ⟨*ku*⟩ (see Section 2.10).

## 2.8 The basic structure of longer utterances

People don't always say just one sentence. Lojban has a specific structure for talk or writing that is longer than one sentence. The entirety of a given speech event or written text is called an utterance. The sentences (usually, but not always, bridi) in an utterance are separated by the cmavo *ni'o* and *.i*. These correspond to a brief pause (or nothing at all) in spoken English, and the various punctuation marks like period, question mark, and exclamation mark in written English. These separators prevent the sumti at the beginning of the next sentence from being mistaken for a trailing sumti of the previous sentence.

The cmavo *ni'o* separates paragraphs (covering different topics of discussion). In a long text or utterance, the topical structure of the text may be indicated by multiple *ni'o*s, with perhaps *ni'oni'oni'o* used to indicate a chapter, *ni'oni'o* to indicate a section, and a single *ni'o* to indicate a subtopic corresponding to a single English paragraph.

The cmavo *.i* separates sentences. It is sometimes compounded with words that modify the exact meaning (the semantics) of the sentence in the context of the utterance. (The cmavo *xu*, discussed in Section 2.15, is one such word — it turns the sentence from a statement to a question about truth.) When more than one person is talking, a new speaker will usually omit the *.i* even though she/he may be continuing on the same topic.

It is still O.K. for a new speaker to say the *.i* before continuing; indeed, it is encouraged for maximum clarity (since it is possible that the second speaker might merely be adding words onto the end of the first speaker's sentence). A good translation for *.i* is

the *and* used in run-on sentences when people are talking informally: “I did this, and then I did that, and . . . , and . . . ”.

## 2.9 tanru

When two gismu are adjacent, the first one modifies the second, and the selbri takes its place structure from the rightmost word. Such combinations of gismu are called *tanru*. For example,

**Example 2.9.1**
sutra tavla

has the place structure

**Example 2.9.2**
$x_1$ is a fast type-of talker to $x_2$ about $x_3$ in language $x_4$
$x_1$ talks fast to $x_2$ about $x_3$ in language $x_4$

When three or more gismu are in a row, the first modifies the second, and that combined meaning modifies the third, and that combined meaning modifies the fourth, and so on. For example

**Example 2.9.3**
sutra tavla cutci

has the place structure

**Example 2.9.4**
$s_1$ is a fast-talker type of shoe worn by $s_2$ of material $s_3$

That is, it is a shoe that is worn by a fast talker rather than a shoe that is fast and is also worn by a talker.

Note especially the use of *type-of* as a mechanism for connecting the English translations of the two or more gismu; this convention helps the learner understand each tanru in its context. Creative interpretations are also possible, however:

**Example 2.9.5**
bajra cutci
runner shoe

most probably refers to shoes suitable for runners, but might be interpreted in some imaginative instances as “shoes that run (by themselves?)”. In general, however, the meaning of a tanru is determined by the literal meaning of its components, and not by any connotations or figurative meanings. Thus

**Example 2.9.6**
sutra tavla
fast-talker

would not necessarily imply any trickery or deception, unlike the English idiom, and a

**Example 2.9.7**
jikca toldi
social butterfly

must always be an insect with large brightly-colored wings, of the family *Lepidoptera*. The place structure of a tanru is always that of the final component of the tanru. Thus, the following has the place structure of *klama*:

**Example 2.9.8**
mi ⟨*cu*⟩ sutra klama la meris.
I quickly-go to Mary.

With the conversion *se klama* as the final component of the tanru, the place structure of the entire selbri is that of "se klama": the $x_1$ place is the destination, and the $x_2$ place is the one who goes:

**Example 2.9.9**
mi ⟨*cu*⟩ sutra se klama la meris.
I quickly am-gone-to by Mary.

The following example shows that there is more to conversion than merely switching places, though:

**Example 2.9.10**
la tam. ⟨*cu*⟩ melbi tavla la meris.
Tom beautifully-talks to Mary.
Tom is a beautiful-talker to Mary.

has the place structure of *tavla*, but note the two distinct interpretations. Now, using conversion, we can modify the place structure order:

**Example 2.9.11**
la meris. ⟨*cu*⟩ melbi se tavla la tam.
Mary is beautifully-talked-to by Tom.
Mary is a beautiful-audience for Tom.

and we see that the modification has been changed so as to focus on Mary's role in the bridi relationship, leading to a different set of possible interpretations. Note that there is no place structure change if the modifying term is converted, and so less drastic variation in possible meanings:

**Example 2.9.12**
la tam. ⟨*cu*⟩ tavla melbi la meris.
Tom is talkerly-beautiful to Mary.

**Example 2.9.13**
la tam. ⟨*cu*⟩ se tavla melbi la meris.
Tom is audiencely-beautiful to Mary.

and we see that the manner in which Tom is seen as beautiful by Mary changes, but Tom is still the one perceived as beautiful, and Mary, the observer of beauty.

## 2.10 Description sumti

Often we wish to talk about things other than the speaker, the listener and things we can point to. Let's say I want to talk about a talker other than *mi*. What I want to talk about would naturally fit into the first place of *tavla*. Lojban, it turns out, has an operator that pulls this first place out of a selbri and converts it to a sumti called a *description sumti*. The description sumti *le tavla ku* means "the talker", and may be used wherever any sumti may be used.

For example,

**Example 2.10.1**
mi tavla do le tavla ku

means the same as

**Example 2.10.2**
I talk to you about the talker

where *the talker* is presumably someone other than me, though not necessarily. Similarly *le sutra tavla ku* is "the fast talker", and *le sutra te tavla ku* is "the fast subject of talk" or "the subject of fast talk". Which of these related meanings is understood will depend on the context in which the expression is used. The most plausible interpretation within the context will generally be assumed by a listener to be the intended one.

In many cases the word *ku* may be omitted. In particular, it is never necessary in a description at the end of a sentence, so:

**Example 2.10.3**
mi tavla do le tavla
I talk-to you about-the talker

means exactly the same thing as Example 2.10.1. There is a problem when we want to say "The fast one is talking." The *obvious* translation *le sutra tavla* turns out to mean *the fast talker*, and has no selbri at all. To solve this problem we can use the word *cu*, which so far has always been optional, in front of the selbri.

The word *cu* has no meaning, and exists only to mark the beginning of the selbri within the bridi, separating it from a previous sumti. It comes before any other part of the selbri, including other cmavo like *se* or *te*. Thus:

**Example 2.10.4**
le sutra tavla
The fast talker

**Example 2.10.5**
le sutra cu tavla
The fast one is talking.

**Example 2.10.6**
le sutra se tavla
The fast talked-to one

**Example 2.10.7**
le sutra cu se tavla
The fast one is talked to.

Consider the following more complex example, with two description sumti.

**Example 2.10.8**
mi ⟨*cu*⟩ tavla le vecnu ⟨*ku*⟩ le blari'o ⟨*ku*⟩
I talk-to the seller about the blue-green-thing.

The sumti *le vecnu* contains the selbri *vecnu*, which has the *seller* in the $x_1$ place, and uses it in this sentence to describe a particular *seller* that the speaker has in mind (one that he or she probably expects the listener will also know about). Similarly, the speaker has a particular blue-green thing in mind, which is described using *le* to mark *blari'o*, a selbri whose first sumti is something blue-green.

It is safe to omit both occurrences of *ku* in Example 2.10.8, and it is also safe to omit the *cu*.

## 2.11 Examples of brivla

The simplest form of selbri is an individual word. A word which may by itself express a selbri relation is called a *brivla*. The three types of brivla are gismu (root words), lujvo (compounds), and fu'ivla (borrowings from other languages). All have identical grammatical uses. So far, most of our selbri have been gismu or tanru built from gismu.

**Example 2.11.1**
**gismu**:
mi ⟨*cu*⟩ klama ti zo'e zo'e ta
Go-er goes destination origin route means.
I go here (to this) using that means (from somewhere via some route).

**Example 2.11.2**
**lujvo**:
ta ⟨*cu*⟩ blari'o
That is-blue-green.

**Example 2.11.3**
**fu'ivla**:
ti ⟨*cu*⟩ djarspageti
This is-spaghetti.

Some cmavo may also serve as selbri, acting as variables that stand for another selbri. The most commonly used of these is *go'i*, which represents the main bridi of the previous Lojban sentence, with any new sumti or other sentence features being expressed replacing the previously expressed ones. Thus, in this context:

**Example 2.11.4**
ta ⟨*cu*⟩ go'i
That too/same-as-last selbri.
That (is spaghetti), too.

## 2.12 The sumti *di'u* and *la'e di'u*

In English, I might say "The dog is beautiful", and you might reply "This pleases me." How do you know what *this* refers to? Lojban uses different expressions to convey the possible meanings of the English:

**Example 2.12.1**
le gerku ⟨*ku*⟩ cu melbi
The dog is beautiful.

The following three sentences all might translate as "This pleases me."

**Example 2.12.2**
ti ⟨*cu*⟩ pluka mi
This (the dog) pleases me.

**Example 2.12.3**
di'u ⟨*cu*⟩ pluka mi
This (the last sentence) pleases me (perhaps because it is grammatical or sounds nice).

**Example 2.12.4**
la'e di'u ⟨*cu*⟩ pluka mi
This (the meaning of the last sentence; i.e. that the dog is beautiful) pleases me.

Example 2.12.4 uses one sumti to point to or refer to another by inference. It is common to write *la'edi'u* as a single word; it is used more often than *di'u* by itself.

## 2.13 Possession

*Possession* refers to the concept of specifying an object by saying who it belongs to (or with). A full explanation of Lojban possession is given in Chapter 8. A simple means of expressing possession, however, is to place a sumti representing the possessor of an object within the description sumti that refers to the object: specifically, between the *le* and the selbri of the description:

**Example 2.13.1**
le mi gerku cu sutra
The of-me dog is fast.
My dog is fast.

In Lojban, possession doesn't necessarily mean ownership: one may *possess* a chair simply by sitting on it, even though it actually belongs to someone else. English uses possession casually in the same way, but also uses it to refer to actual ownership or even more intimate relationships: "my arm" doesn't mean "some arm I own" but rather "the arm that is part of my body". Lojban has methods of specifying all these different kinds of possession precisely and easily.

## 2.14 Vocatives and commands

You may call someone's attention to the fact that you are addressing them by using *doi* followed by their name. The sentence

**Example 2.14.1**
doi djan.

means "Oh, John, I'm talking to you". It also has the effect of setting the value of *do*; *do* now refers to *John* until it is changed in some way in the conversation. Note that Example 2.14.1 is not a bridi, but it is a legitimate Lojban sentence nevertheless; it is known as a *vocative phrase*. Other cmavo can be used instead of *doi* in a vocative phrase, with a different significance. For example, the cmavo *coi* means "hello" and *co'o* means "good-bye". Either word may stand alone, they may follow one another, or either may be followed by a pause and a name. (Vocative phrases with *doi* do not need a pause before the name.)

**Example 2.14.2**
coi. djan.
Hello, John.

**Example 2.14.3**
co'o. djan.
Good-bye, John.

Commands are expressed in Lojban by a simple variation of the main bridi structure. If you say

**Example 2.14.4**
do tavla
You are-talking.

you are simply making a statement of fact. In order to issue a command in Lojban, substitute the word *ko* for *do*. The bridi

**Example 2.14.5**
ko tavla

instructs the listener to do whatever is necessary to make Example 2.14.4 true; it means "Talk!" Other examples:

**Example 2.14.6**
ko sutra
Be fast!

The *ko* need not be in the $x_1$ place, but rather can occur anywhere a sumti is allowed, leading to possible Lojban commands that are very unlike English commands:

**Example 2.14.7**
mi tavla ko
Be talked to by me
Let me talk to you.

The cmavo *ko* can fill any appropriate sumti place, and can be used as often as is appropriate for the selbri:

**Example 2.14.8**
ko kurji ko

and

**Example 2.14.9**
ko ko kurji

both mean "You take care of you" and "Be taken care of by you", or to put it colloquially, "Take care of yourself".

## 2.15 Questions

There are many kinds of questions in Lojban: full explanations appear in Chapter 19 and in various other chapters throughout the book. In this chapter, we will introduce three kinds: sumti questions, selbri questions, and yes/no questions.

The cmavo *ma* is used to create a sumti question: it indicates that the speaker wishes to know the sumti which should be placed at the location of the *ma* to make the bridi true. It can be translated as *Who?* or *What?* in most cases, but also serves for *When?*, *Where?*, and *Why?* when used in sumti places that express time, location, or cause. For example:

**Example 2.15.1**
ma tavla do mi
Who? talks to-you about-me.
Who is talking to you about me?

The listener can reply by simply stating a sumti:

**Example 2.15.2**
la djan.
John (is talking to you about me).

Like *ko*, *ma* can occur in any position where a sumti is allowed, not just in the first position:

**Example 2.15.3**
do ⟨*cu*⟩ tavla ma
You talk to what/whom?

A *ma* can also appear in multiple sumti positions in one sentence, in effect asking several questions at once.

**Example 2.15.4**
ma ⟨*cu*⟩ tavla ma
What/Who talks to what/whom?

The two separate *ma* positions ask two separate questions, and can therefore be answered with different values in each sumti place. The cmavo *mo* is the selbri analogue of *ma*. It asks the respondent to provide a selbri that would be a true relation if inserted in place of the *mo*:

**Example 2.15.5**
do ⟨*cu*⟩ mo
You are-what/do-what?

A *mo* may be used anywhere a brivla or other selbri might. Keep this in mind for later examples. Unfortunately, by itself, *mo* is a very non-specific question. The response to the question in Example 2.15.5 could be:

**Example 2.15.6**
mi ⟨*cu*⟩ melbi
I am beautiful.

or:

**Example 2.15.7**
mi ⟨*cu*⟩ tavla
I talk.

Clearly, *mo* requires some cooperation between the speaker and the respondent to ensure that the right question is being answered. If context doesn't make the question specific enough, the speaker must ask the question more specifically using a more complex construction such as a tanru (see Section 2.9). It is perfectly permissible for the respondent to fill in other unspecified places in responding to a *mo* question. Thus, the respondent in Example 2.15.7 could have also specified an audience, a topic, and/or a language in the response.

Finally, we must consider questions that can be answered *Yes* or *No*, such as

**Example 2.15.8**
Are you talking to me?

Like all yes-or-no questions in English, Example 2.15.8 may be reformulated as

**Example 2.15.9**
Is it true that you are talking to me?

In Lojban we have a word that asks precisely that question in precisely the same way. The cmavo *xu*, when placed in front of a bridi, asks whether that bridi is true as stated. So

**Example 2.15.10**
xu do tavla mi
Is-it-true-that you are-talking to-me?

is the Lojban translation of Example 2.15.8. The answer *Yes* may be given by simply restating the bridi without the *xu* question word. Lojban has a shorthand for doing this with the word *go'i*, mentioned in Section 2.11. Instead of a negative answer, the bridi may be restated in such a way as to make it true. If this can be done by substituting sumti, it may be done with *go'i* as well. For example:

**Example 2.15.11**
xu do kanro
Are you healthy?

can be answered with

**Example 2.15.12**
mi kanro
I am healthy.

or

**Example 2.15.13**
go'i
I am healthy.

or (Note that *do* to the questioner is *mi* to the respondent.)

**Example 2.15.14**
le tavla cu kanro
The talker is healthy.

or

**Example 2.15.15**
le tavla cu go'i
The talker is healthy.

A general negative answer may be given by *na go'i*. *na* may be placed before any selbri (but after the *cu*). It is equivalent to stating *It is not true that ...* before the bridi. It does not imply that anything else is true or untrue, only that that specific bridi is not true. More details on negative statements are available in Chapter 15.

## 2.16 Indicators

Different cultures express emotions and attitudes with a variety of intonations and gestures that are not usually included in written language. Some of these are available in some languages as interjections (i.e. Aha!, Oh no!, Ouch!, Aahh!, etc.), but they vary greatly from culture to culture.

Lojban has a group of cmavo known as "attitudinal indicators" which specifically covers this type of commentary on spoken statements. They are both written and spoken, but require no specific intonation or gestures. Grammatically they are very simple: one or more attitudinals at the beginning of a bridi apply to the entire bridi; anywhere else in the bridi they apply to the word immediately to the left. For example:

**Example 2.16.1**
.ie mi ⟨*cu*⟩ klama
Agreement! I go.
Yep! I'll go.

**Example 2.16.2**
.ei mi ⟨*cu*⟩ klama
Obligation! I go.
I should go.

**Example 2.16.3**
mi ⟨*cu*⟩ klama le melbi .ui ⟨*ku*⟩
I go to the beautiful-thing (and I am happy because it is the beautiful thing I'm going to).

Not all indicators indicate attitudes. Discursives, another group of cmavo with the same grammatical rules as attitudinal indicators, allow free expression of certain kinds of commentary about the main utterances. Using discursives allows a clear separation of these so-called *metalinguistic* features from the underlying statements and logical structure. By comparison, the English words *but* and *also*, which discursively indicate contrast or an added weight of example, are logically equivalent to *and*, which does not have a discursive content. The average English-speaker does not think about, and may not even realize, the paradoxical idea that *but* basically means *and*.

**Example 2.16.4**
mi ⟨*cu*⟩ klama .i do ⟨*cu*⟩ stali
I go. You stay.

**Example 2.16.5**
mi ⟨*cu*⟩ klama .i ji'a do ⟨*cu*⟩ stali
I go. In addition, you stay. (added weight)

**Example 2.16.6**
mi ⟨*cu*⟩ klama .i ku'i do ⟨*cu*⟩ stali
I go. However, you stay. (contrast)

Another group of indicators are called *evidentials*. Evidentials show the speaker's relationship to the statement, specifically how the speaker came to make the statement. These include *za'a* (I directly observe the relationship), *pe'i* (I believe that the relationship holds), *ru'a* (I postulate the relationship), and others. Many American Indian languages use this kind of words.

**Example 2.16.7**
pe'i do ⟨*cu*⟩ melbi
I opine! You are beautiful.

**Example 2.16.8**
za'a do ⟨*cu*⟩ melbi
I directly observe! You are beautiful.

## 2.17 Tenses

In English, every verb is tagged for the grammatical category called tense: past, present, or future. The sentence

**Example 2.17.1**
John went to the store

necessarily happens at some time in the past, whereas

**Example 2.17.2**
John is going to the store

is necessarily happening right now. The Lojban sentence

**Example 2.17.3**
la djan. ⟨*cu*⟩ klama le zarci
John goes/went/will go to-the store

serves as a translation of either Example 2.17.1 or Example 2.17.2, and of many other possible English sentences as well. It is not marked for tense, and can refer to an event in the past, the present or the future. This rule does not mean that Lojban has no way of representing the time of an event. A close translation of Example 2.17.1 would be:

**Example 2.17.4**
la djan. pu klama le zarci
John ⟨*past*⟩ goes to-the store

where the tag *pu* forces the sentence to refer to a time in the past. Similarly,

**Example 2.17.5**
la djan. ca klama le zarci
John ⟨*present*⟩ goes to-the store

necessarily refers to the present, because of the tag *ca*. Tags used in this way always appear at the very beginning of the selbri, just after the *cu*, and they may make a *cu* unnecessary, since tags cannot be absorbed into tanru. Such tags serve as an equivalent to English tenses and adverbs. In Lojban, tense information is completely optional. If unspecified, the appropriate tense is picked up from context. Lojban also extends the notion of *tense* to refer not only to time but to space. The following example uses the tag *vu* to specify that the event it describes happens far away from the speaker:

**Example 2.17.6**
do ⟨*cu*⟩ vu vecnu zo'e
You yonder sell something-unspecified.

In addition, tense tags (either for time or space) can be prefixed to the selbri of a description, producing a tensed sumti:

**Example 2.17.7**
le pu bajra ⟨*ku*⟩ cu tavla
The earlier/former/past runner talked/talks.

(Since Lojban tense is optional, we don't know when he or she talks.) Tensed sumti with space tags correspond roughly to the English use of *this* or *that* as adjectives, as in the following example, which uses the tag *vi* meaning *nearby*:

**Example 2.17.8**
le vi bajra ⟨*ku*⟩ cu tavla
The nearby runner talks.
This runner talks.

Do not confuse the use of *vi* in Example 2.17.8 with the cmavo *ti*, which also means *this*, but in the sense of *this thing*.

Furthermore, a tense tag can appear both on the selbri and within a description, as in the following example (where *ba* is the tag for future time):

**Example 2.17.9**
le vi tavla ⟨*ku*⟩ cu ba klama
The here talker [future] goes.
The talker who is here will go
This talker will go.

## 2.18 Lojban grammatical terms

Here is a review of the Lojban grammatical terms used in this chapter, plus some others used throughout this book. Only terms that are themselves Lojban words are included: there are of course many expressions like *indicator* in Chapter 16 that are not explained here. See the Index for further help with these.

**Definition 2.4**

**bridi** predication; the basic unit of Lojban expression; the main kind of Lojban sentence; a claim that some objects stand in some relationship, or that some single object has some property.

**sumti** argument; words identifying something which stands in a specified relationship to something else, or which has a specified property. See Chapter 6.

**selbri** logical predicate; the core of a bridi; the word or words specifying the relationship between the objects referred to by the sumti. See Chapter 5.

**cmavo** one of the Lojban parts of speech; a short word; a structural word; a word used for its grammatical function.

**brivla** one of the Lojban parts of speech; a content word; a predicate word; can function as a selbri; is a gismu, a lujvo, or a fu'ivla. See Chapter 4.

**gismu** a root word; a kind of brivla; has associated rafsi. See Chapter 4.

**lujvo** a compound word; a kind of brivla; may or may not appear in a dictionary; does not have associated rafsi. See Chapter 4and Chapter 12.

**fu'ivla** a borrowed word; a kind of brivla; may or may not appear in a dictionary; copied in a modified form from some non-Lojban language; usually refers to some aspect of culture or the natural world; does not have associated rafsi. See Chapter 4.

**rafsi** a word fragment; one or more is associated with each gismu; can be assembled according to rules in order to make lujvo; not a valid word by itself. See Chapter 4.

**tanru** a group of two or more brivla, possibly with associated cmavo, that form a selbri; always divisible into two parts, with the first part modifying the meaning of the second part (which is taken to be basic). See Chapter 5.

**selma'o** a group of cmavo that have the same grammatical use (can appear interchangeably in sentences, as far as the grammar is concerned) but differ in meaning or other usage. See Chapter 20.

# Chapter 3

# The Hills Are Alive With the Sounds of Lojban

*.i .ai .i .ai .o*

## 3.1 Orthography

Lojban is designed so that any properly spoken Lojban utterance can be uniquely transcribed in writing, and any properly written Lojban can be spoken so as to be uniquely reproduced by another person. As a consequence, the standard Lojban orthography must assign to each distinct sound, or phoneme, a unique letter or symbol. Each letter or symbol has only one sound or, more accurately, a limited range of sounds that are permitted pronunciations for that phoneme. Some symbols indicate stress (speech emphasis) and pause, which are also essential to Lojban word recognition. In addition, everything that is represented in other languages by punctuation (when written) or by tone of voice (when spoken) is represented in Lojban by words. These two properties together are known technically as *audio-visual isomorphism*.

Lojban uses a variant of the Latin (Roman) alphabet, consisting of the following letters and symbols:

**Definition 3.1**
**Alphabet** ’ , . a b c d e f g i j k l m n o p r s t u v x y z

omitting the letters *h*, *q*, and *w*.

The alphabetic order given above is that of the ASCII coded character set, widely used in computers. By making Lojban alphabetical order the same as ASCII, computerized sorting and searching of Lojban text is facilitated.

Capital letters are used only to represent non-standard stress, which can appear only in the representation of Lojbanized names. Thus the English name *Josephine*, as normally pronounced, is Lojbanized as *DJOsefin.*, pronounced /ˈdʒo sɛ finʔ/. (See Section 3.2 for an explanation of the symbols within square brackets.) Technically, it is sufficient to capitalize the vowel letter, in this case *O*, but it is easier on the reader to capitalize the whole syllable.

Without the capitalization, the ordinary rules of Lojban stress would cause the *se* syllable to be stressed. Lojbanized names are meant to represent the pronunciation of names from other languages with as little distortion as may be; as such, they are exempt from many of the regular rules of Lojban phonology, as will appear in the rest of this chapter.

## 3.2 Basic Phonetics

Lojban pronunciations are defined using the International Phonetic Alphabet, or IPA, a standard method of transcribing pronunciations. By convention, IPA transcriptions are always within square brackets: for example, the word *cat* is pronounced (in General American pronunciation) /kæt/. Section 3.10 contains a brief explanation of the IPA characters used in this chapter, with their nearest analogues in English, and will be especially useful to those not familiar with the technical terms used in describing speech sounds.

The standard pronunciations and permitted variants of the Lojban letters are listed in the table below. The descriptions have deliberately been made a bit ambiguous to cover variations in pronunciation by speakers of different native languages and dialects. In all cases except *r* the first IPA symbol shown represents the preferred pronunciation; for *r*, all of the variations (and any other rhotic sound) are equally acceptable.

| Letter | IPA | Description |
|---|---|---|
| ’ | /h/ | an unvoiced glottal spirant |
| , | /-/ | the syllable separator |
| . | /ʔ/ | a glottal stop or a pause |
| a | /a/, /ɑ/ | an open vowel |
| b | /b/ | a voiced bilabial stop |
| c | /ʃ/, /ʂ/ | an unvoiced postalveolar fricative |
| d | /d/ | a voiced dental/alveolar stop |
| e | /ɛ/, /e/ | a front mid vowel |
| f | /f/, /ɸ/ | an unvoiced labial fricative |

| | | |
|---|---|---|
| g | /g/ | a voiced velar stop |
| i | /i/ | a front close vowel |
| j | /ʒ/, /z̬/ | a voiced postalveolar fricative |
| k | /k/ | an unvoiced velar stop |
| l | /l/, /l̩/ | a voiced lateral approximant (may be syllabic) |
| m | /m/, /m̩/ | a voiced bilabial nasal (may be syllabic) |
| n | /n/, /n̩/, /ŋ/, /ŋ̩/ | a voiced dental or velar nasal (may be syllabic) |
| o | /o/, /ɔ/ | a back mid vowel |
| p | /p/ | an unvoiced bilabial stop |
| r | /r/, /ɹ/, /ɾ/, /ʀ/, /r̩/, /ɹ̩/, /ɾ̩/, /ʀ̩/ | a rhotic sound |
| s | /s/ | an unvoiced alveolar sibilant |
| t | /t/ | an unvoiced dental/alveolar stop |
| u | /u/ | a back close vowel |
| v | /v/, /β/ | a voiced labial fricative |
| x | /x/ | an unvoiced velar fricative |
| y | /ə/ | a central mid vowel |
| z | /z/ | a voiced alveolar sibilant |

The Lojban sounds must be clearly pronounced so that they are not mistaken for each other. Voicing and placement of the tongue are the key factors in correct pronunciation, but other subtle differences will develop between consonants in a Lojban-speaking community. At this point these are the only mandatory rules on the range of sounds.

Note in particular that Lojban vowels can be pronounced with either rounded or unrounded lips; typically *o* and *u* are rounded and the others are not, as in English, but this is not a requirement; some people round *y* as well. Lojban consonants can be aspirated or unaspirated. Palatalizing of consonants, as found in Russian and other languages, is not generally acceptable in pronunciation, though a following *i* may cause it.

The sounds represented by the letters *c*, *g*, *j*, *s*, and *x* require special attention for speakers of English, either because they are ambiguous in the orthography of English (*c*, *g*, *s*), or because they are strikingly different in Lojban (*c*, *j*, *x*). The English *c* represents three different sounds, /k/ in *cat* and /s/ in *cent*, as well as the /ʃ/ of *ocean*. Similarly, English *g* can represent /g/ as in *go*, /dʒ/ as in *gentle*, and /ʒ/ as in *garage* (in some pronunciations). English *s* can be either /s/ as in *cats*, /z/ as in *cards*, /ʃ/ as in *tension*, or /ʒ/ as in *measure*. The sound of Lojban *x* doesn't appear in most English dialects at all.

There are two common English sounds that are found in Lojban but are not not Lojban consonants: the *ch* of *church* and the *j* of *judge*. In Lojban, these are considered two consonant sounds spoken together without an intervening vowel sound, and so are represented in Lojban by the two separate consonants: *tc* (IPA /tʃ/) and *dj* (IPA /dʒ/). In general, whether a complex sound is considered one sound or two depends on the language: Russian views *ts* as a single sound, whereas English, French, and Lojban consider it to be a consonant cluster.

## 3.3 The Special Lojban Characters

The apostrophe, period, and comma need special attention. They are all used as indicators of a division between syllables, but each has a different pronunciation, and each is used for different reasons:

The apostrophe represents a phoneme similar to a short, breathy English *h*, (IPA /h/). The letter *h* is not used to represent this sound for two reasons: primarily in order to simplify explanations of the morphology, but also because the sound is very common, and the apostrophe is a visually lightweight representation of it. The apostrophe sound is a consonant in nature, but is not treated as either a consonant or a vowel for purposes of Lojban morphology (word-formation), which is explained in Chapter 4. In addition, the apostrophe visually parallels the comma and the period, which are also used (in different ways) to separate syllables.

The apostrophe is included in Lojban only to enable a smooth separation between vowels, while joining the vowels within a single word. In fact, one way to think of the apostrophe is as representing an unvoiced vowel glide.

As a permitted variant, any unvoiced fricative other than those already used in Lojban may be used to render the apostrophe: IPA /T/ is one possibility. The convenience of the listener should be regarded as paramount in deciding to use a substitute for /h/.

The period represents a mandatory pause, with no specified length; a glottal stop (IPA /ʔ/) is considered a pause of shortest length. A pause (or glottal stop) may appear between any two words, and in certain cases – explained in detail in Chapter 4 — must occur. In particular, a word beginning with a vowel is always preceded by a pause, and a word ending in a consonant is always followed by a pause.

Technically, the period is an optional reminder to the reader of a mandatory pause that is dictated by the rules of the language; because these rules are unambiguous, a missing period can be inferred from otherwise correct text. Periods are included only as an aid to the reader.

A period also may be found apparently embedded in a word. When this occurs, such a written string is not one word but two, written together to indicate that the writer intends a unitary meaning for the compound. It is not really necessary to use a space between words if a period appears.

The comma is used to indicate a syllable break within a word, generally one that is not obvious to the reader. Such a comma is written to separate syllables, but indicates that there must be no pause between them, in contrast to the period. Between two vowels, a comma indicates that some type of glide may be necessary to avoid a pause that would split the two syllables into separate words. It is always legal to use the apostrophe (IPA /h/) sound in pronouncing a comma. However, a comma cannot be pronounced as a pause or glottal stop between the two letters separated by the comma, because that pronunciation would split the word into two words.

Otherwise, a comma is usually only used to clarify the presence of syllabic *l*, *m*, *n*, or *r* (discussed later). Commas are never required: no two Lojban words differ solely because of the presence or placement of a comma.

Here is a somewhat artificial example of the difference in pronunciation between periods, commas and apostrophes. In the English song about Old MacDonald's Farm, the vowel string which is pronounced *ee-i-ee-i-o* in English could be Lojbanized with periods as:

**Example 3.3.1**
.i.ai.i.ai.o
/ʔi ʔaj ʔi ʔaj ʔo/
Ee! Eye! Ee! Eye! Oh!

However, this would sound clipped, staccato, and unmusical compared to the English. Furthermore, although Example 3.3.1 is a string of meaningful Lojban words, as a sentence it makes very little sense. (Note the use of periods embedded within the written word.)

If commas were used instead of periods, we could represent the English string as a Lojbanized name, ending in a consonant:

**Example 3.3.2**
.i,ai,i,ai,on.
/ʔi jaj ji jaj jonʔ/

The commas represent new syllable breaks, but prohibit the use of pauses or glottal stop. The pronunciation shown is just one possibility, but closely parallels the intended English pronunciation.

However, the use of commas in this way is risky to unambiguous interpretation, since the glides might be heard by some listeners as diphthongs, producing something like

**Example 3.3.3**
.i,iai,ii,iai,ion.

which is technically a different Lojban name. Since the intent with Lojbanized names is to allow them to be pronounced more like their native counterparts, the comma is allowed to represent vowel glides or some non-Lojbanic sound. Such an exception affects only spelling accuracy and the ability of a reader to replicate the desired pronunciation exactly; it will not affect the recognition of word boundaries.

Still, it is better if Lojbanized names are always distinct. Therefore, the apostrophe is preferred in regular Lojbanized names that are not attempting to simulate a non-Lojban pronunciation perfectly. (Perfection, in any event, is not really achievable, because some sounds simply lack reasonable Lojbanic counterparts.)

If apostrophes were used instead of commas in Example 3.3.2, it would appear as:

**Example 3.3.4**
.i'ai'i'ai'on.
/ʔi hai hi hai honʔ/

which preserves the rhythm and length, if not the exact sounds, of the original English.

## 3.4 Diphthongs and Syllabic Consonants

There exist 16 diphthongs in the Lojban language. A diphthong is a vowel sound that consists of two elements, a short vowel sound and a glide, either a labial (IPA /w/) or palatal (IPA /j/) glide, that either precedes (an on-glide) or follows (an off-glide) the main vowel. Diphthongs always constitute a single syllable.

For Lojban purposes, a vowel sound is a relatively long speech-sound that forms the nucleus of a syllable. Consonant sounds are relatively brief and normally require an accompanying vowel sound in order to be audible. Consonants may occur at the beginning or end of a syllable, around the vowel, and there may be several consonants in a cluster in either position. Each separate vowel sound constitutes a distinct syllable; consonant sounds do not affect the determination of syllables.

The six Lojban vowels are *a*, *e*, *i*, *o*, *u*, and *y*. The first five vowels appear freely in all kinds of Lojban words. The vowel *y* has a limited distribution: it appears only in Lojbanized names, in the Lojban names of the letters of the alphabet, as a glue vowel in compound words, and standing alone as a space-filler word (like English *uh* or *er*).

The Lojban diphthongs are shown in the table below. (Variant pronunciations have been omitted, but are much as one would expect based on the variant pronunciations of the separate vowel letters: *ai* may be pronounced /ɑj/, for example.)

| Letters | IPA | Description |
|---|---|---|
| ai | /aj/ | an open vowel with palatal off-glide |
| ei | /ɛj/ | a front mid vowel with palatal off-glide |
| oi | /oj/ | a back mid vowel with palatal off-glide |
| au | /aw/ | an open vowel with labial off-glide |
| ia | /ja/ | an open vowel with palatal on-glide |
| ie | /jɛ/ | a front mid vowel with palatal on-glide |
| ii | /ji/ | a front close vowel with palatal on-glide |
| io | /jo/ | a back mid vowel with palatal on-glide |
| iu | /ju/ | a back close vowel with palatal on-glide |
| ua | /wa/ | an open vowel with labial on-glide |
| ue | /wɛ/ | a front mid vowel with labial on-glide |
| ui | /wi/ | a front close vowel with labial on-glide |
| uo | /wo/ | a back mid vowel with labial on-glide |
| uu | /wu/ | a back close vowel with labial on-glide |
| iy | /jæ/ | a central mid vowel with palatal on-glide |
| uy | /wæ/ | a central mid vowel with labial on-glide |

(Approximate English equivalents of most of these diphthongs exist: see Section 3.11 for examples.)

The first four diphthongs above (*ai*, *ei*, *oi*, and *au*, the ones with off-glides) are freely used in most types of Lojban words; the ten following ones are used only as stand-alone words and in Lojbanized names and borrowings; and the last two (*iy* and *uy*) are used only in Lojbanized names.

The syllabic consonants of Lojban, /l̩/, /m̩/, /n̩/, and /r̩/, are variants of the non-syllabic /l/, /m/, /n/, and /r/ respectively. They normally have only a limited distribution, appearing in Lojban names and borrowings, although in principle any *l*, *m*, *n*, or *r* may be pronounced syllabically. If a syllabic consonant appears next to a *l*, *m*, *n*, or *r* that is not syllabic, it may not be clear which is which:

**Example 3.4.1**
brlgan.
/br̩l gan/
or /brl̩ gan/

is a hypothetical Lojbanized name with more than one valid pronunciation; however it is pronounced, it remains the same word.

Syllabic consonants are treated as consonants rather than vowels from the standpoint of Lojban morphology. Thus Lojbanized names, which are generally required to end in a consonant, are allowed to end with a syllabic consonant. An example is *rl.*, which is an approximation of the English name *Earl*, and has two syllabic consonants.

Syllables with syllabic consonants and no vowel are never stressed or counted when determining which syllables to stress (see Section 3.9).

## 3.5 Vowel Pairs

Lojban vowels also occur in pairs, where each vowel sound is in a separate syllable. These two vowel sounds are connected (and separated) by an apostrophe. Lojban vowel pairs should be pronounced continuously with the /h/ sound between (and not by a glottal stop or pause, which would split the two vowels into separate words).

All vowel combinations are permitted in two-syllable pairs with the apostrophe separating them; this includes those which constitute diphthongs when the apostrophe is not included.

The Lojban vowel pairs are:

| | | | | | |
|---|---|---|---|---|---|
| a'a | a'e | a'i | a'o | a'u | a'y |
| e'a | e'e | e'i | e'o | e'u | e'y |
| i'a | i'e | i'i | i'o | i'u | i'y |
| o'a | o'e | o'i | o'o | o'u | o'y |
| u'a | u'e | u'i | u'o | u'u | u'y |
| y'a | y'e | y'i | y'o | y'u | y'y |

Vowel pairs involving *y* appear only in Lojbanized names. They could appear in cmavo (structure words), but only *.y'y.* is so used — it is the Lojban name of the apostrophe letter (see Chapter 17).

When more than two vowels occur together in Lojban, the normal pronunciation pairs vowels from the left into syllables, as in the Lojbanized name:

**Example 3.5.1**
meiin.
mei,in.

Example 3.5.1 contains the diphthong *ei* followed by the vowel *i*. In order to indicate a different grouping, the comma must always be used, leading to:

**Example 3.5.2**
me,iin.

which contains the vowel *e* followed by the diphthong *ii*. In rough English representation, Example 3.5.1 is *May Een*, whereas Example 3.5.2 is *Meh Yeen*.

## 3.6 Consonant Clusters

A consonant sound is a relatively brief speech-sound that precedes or follows a vowel sound in a syllable; its presence either preceding or following does not add to the count

of syllables, nor is a consonant required in either position for any syllable. Lojban has seventeen consonants: for the purposes of this section, the apostrophe is not counted as a consonant.

An important distinction dividing Lojban consonants is that of voicing. The following table shows the unvoiced consonants and the corresponding voiced ones:

| **Unvoiced** | **Voiced** |
|---|---|
| p | b |
| t | d |
| k | g |
| f | v |
| c | j |
| s | z |
| x | - |

The consonant *x* has no voiced counterpart in Lojban. The remaining consonants, *l*, *m*, *n*, and *r*, are typically pronounced with voice, but can be pronounced unvoiced.

Consonant sounds occur in languages as single consonants, or as doubled, or as clustered combinations. Single consonant sounds are isolated by word boundaries or by intervening vowel sounds from other consonant sounds. Doubled consonant sounds are either lengthened like /s/ in English *hiss*, or repeated like /k/ in English *backcourt*. Consonant clusters consist of two or more single or doubled consonant sounds in a group, each of which is different from its immediate neighbor. In Lojban, doubled consonants are excluded altogether, and clusters are limited to two or three members, except in Lojbanized names.

Consonants can occur in three positions in words: initial (at the beginning), medial (in the middle), and final (at the end). In many languages, the sound of a consonant varies depending upon its position in the word. In Lojban, as much as possible, the sound of a consonant is unrelated to its position. In particular, the common American English trait of changing a *t* between vowels into a *d* or even a flap (IPA /ɾ/) is unacceptable in Lojban.

Lojban imposes no restrictions on the appearance of single consonants in any valid consonant position; however, no consonant (including syllabic consonants) occurs final in a word except in Lojbanized names.

Pairs of consonants can also appear freely, with the following restrictions:

1. It is forbidden for both consonants to be the same, as this would violate the rule against double consonants.

2. It is forbidden for one consonant to be voiced and the other unvoiced. The consonants *l*, *m*, *n*, and *r* are exempt from this restriction. As a result, *bf* is forbidden, and so is *sd*, but both *fl* and *vl*, and both *ls* and *lz*, are permitted.

3. It is forbidden for both consonants to be drawn from the set *c*, *j*, *s*, *z*.

4. The specific pairs *cx*, *kx*, *xc*, *xk*, and *mz* are forbidden.

These rules apply to all kinds of words, even Lojbanized names. If a name would normally contain a forbidden consonant pair, a *y* can be inserted to break up the pair:

**Example 3.6.1**
djeimyz.
/dʒɛj məzʔ/
James

The regular English pronunciation of *James*, which is /dʒɛjmz/, would Lojbanize as *djeimz.*, which contains a forbidden consonant pair.

## 3.7 Initial Consonant Pairs

The set of consonant pairs that may appear at the beginning of a word (excluding Lojbanized names) is far more restricted than the fairly large group of permissible consonant pairs described in Section 3.6. Even so, it is more than English allows, although hopefully not more than English-speakers (and others) can learn to pronounce.

There are just 48 such permissible initial consonant pairs, as follows:

**Definition 3.2**
**Pairs** bl br cf ck cl cm cn cp cr ct dj dr dz fl fr gl gr jb jd jg jm jv kl kr ml mr pl pr sf sk sl sm sn sp sr st tc tr ts vl vr xl xr zb zd zg zm zv

Lest this list seem almost random, a pairing of voiced and unvoiced equivalent vowels will show significant patterns which may help in learning:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| pl | pr | | | fl | fr | | |
| bl | br | | | vl | vr | | |
| | | | | | | | |
| cp | cf | ct | ck | cm | cn | cl | cr |
| jb | jv | jd | jg | jm | | | |
| sp | sf | st | sk | sm | sn | sl | sr |
| zb | zv | zd | zg | zm | | | |
| | | | | | | | |
| tc | tr | ts | | kl | kr | | |
| dj | dr | dz | | gl | gr | | |
| | | | | | | | |
| ml | mr | | | xl | xr | | |

Note that if both consonants of an initial pair are voiced, the unvoiced equivalent is also permissible, and the voiced pair can be pronounced simply by voicing the unvoiced pair. (The converse is not true: *cn* is a permissible initial pair, but *jn* is not.)

Consonant triples can occur medially in Lojban words. They are subject to the following rules:

1. The first two consonants must constitute a permissible consonant pair;
2. The last two consonants must constitute a permissible initial consonant pair;
3. The triples *ndj*, *ndz*, *ntc*, and *nts* are forbidden.

Lojbanized names can begin or end with any permissible consonant pair, not just the 48 initial consonant pairs listed above, and can have consonant triples in any location,

as long as the pairs making up those triples are permissible. In addition, names can contain consonant clusters with more than three consonants, again requiring that each pair within the cluster is valid.

## 3.8 Buffering of Consonant Clusters

Many languages do not have consonant clusters at all, and even those languages that do have them often allow only a subset of the full Lojban set. As a result, the Lojban design allows the use of a buffer sound between consonant combinations which a speaker finds unpronounceable. This sound may be any non-Lojbanic vowel which is clearly separable by the listener from the Lojban vowels. Some possibilities are IPA /I/, /ö/, /U/, or even /Y/, but there probably is no universally acceptable buffer sound. When using a consonant buffer, the sound should be made as short as possible. Two examples showing such buffering (we will use /I/ in this chapter) are:

**Example 3.8.1**
vrusi
/ˈvru si/
or /vI ˈru si/

**Example 3.8.2**
.AMsterdam.
/ʔam ster damʔ/
or /ˈʔa mI sI tɛ rI da mIʔ/

When a buffer vowel is used, it splits each buffered consonant into its own syllable. However, the buffering syllables are never stressed, and are not counted in determining stress. They are, in effect, not really syllables to a Lojban listener, and thus their impact is ignored.

Here are more examples of unbuffered and buffered pronunciations:

**Example 3.8.3**
klama
/ˈkla ma/
/kI ˈla ma/

**Example 3.8.4**
xapcke
/ˈxap ʃkɛ/
/ˈxa pI ʃkɛ/
/ˈxa pI ʃI kɛ/

In Example 3.8.4, we see that buffering vowels can be used in just some, rather than all, of the possible places: the second pronunciation buffers the *pc* consonant pair but not the *ck*. The third pronunciation buffers both.

**Example 3.8.5**
ponyni'u
/po næ ˈni hu/

Example 3.8.5 cannot contain any buffering vowel. It is important not to confuse the vowel *y*, which is pronounced /ə/, with the buffer, which has a variety of possible pronunciations and is never written. Consider the contrast between

**Example 3.8.6**
bongynanba
/boŋ gæ ˈnan ba/

an unlikely Lojban compound word meaning *bone bread* (note the use of /ŋ/ as a representative of *n* before *g*) and

**Example 3.8.7**
bongnanba
/boŋ ˈgnan ba/

a possible borrowing from another language (Lojban borrowings can only take a limited form). If Example 3.8.7 were pronounced with buffering, as

**Example 3.8.8**
/boŋ gI ˈnan ba/

it would be very similar to Example 3.8.6. Only a clear distinction between *y* and any buffering vowel would keep the two words distinct.

Since buffering is done for the benefit of the speaker in order to aid pronounceability, there is no guarantee that the listener will not mistake a buffer vowel for one of the six regular Lojban vowels. The buffer vowel should be as laxly pronounced as possible, as central as possible, and as short as possible. Furthermore, it is worthwhile for speakers who use buffers to pronounce their regular vowels a bit longer than usual, to avoid confusion with buffer vowels. The speakers of many languages will have trouble correctly hearing any of the suggested buffer vowels otherwise. By this guideline, Example 3.8.8 would be pronounced

**Example 3.8.9**
/boːŋ gI ˈnaːn baː/

with lengthened vowels.

## 3.9 Syllabication and Stress

A Lojban word has one syllable for each of its vowels, diphthongs, and syllabic consonants (referred to simply as *vowels* for the purposes of this section.) Syllabication rules determine which of the consonants separating two vowels belong to the preceding vowel and which to the following vowel. These rules are conventional only; the phonetic facts of the matter about how utterances are syllabified in any language are always very complex.

A single consonant always belongs to the following vowel. A consonant pair is normally divided between the two vowels; however, if the pair constitute a valid initial consonant pair, they are normally both assigned to the following vowel. A consonant triple is divided between the first and second consonants. Apostrophes and commas,

of course, also represent syllable breaks. Syllabic consonants usually appear alone in their syllables.

It is permissible to vary from these rules in Lojbanized names. For example, there are no definitive rules for the syllabication of names with consonant clusters longer than three consonants. The comma is used to indicate variant syllabication or to explicitly mark normal syllabication.

Here are some examples of Lojban syllabication:

**Example 3.9.1**
pujenaicajeba
pu,je,nai,ca,je,ba

This word has no consonant pairs and is therefore syllabified before each medial consonant.

**Example 3.9.2**
ninmu
nin,mu

This word is split at a consonant pair.

**Example 3.9.3**
fitpri
fit,pri

This word is split at a consonant triple, between the first two consonants of the triple.

**Example 3.9.4**
sairgoi
sair,goi
sai,r,goi

This word contains the consonant pair *rg*; the *r* may be pronounced syllabically or not.

**Example 3.9.5**
klezba
klez,ba
kle,zba

This word contains the permissible initial pair *zb*, and so may be syllabicated either between *z* and *b* or before *zb*.

Stress is a relatively louder pronunciation of one syllable in a word or group of words. Since every syllable has a vowel sound (or diphthong or syllabic consonant) as its nucleus, and the stress is on the vowel sound itself, the terms "stressed syllable" and *stressed vowel* are largely interchangeable concepts.

Most Lojban words are stressed on the next-to-the-last, or penultimate, syllable. In counting syllables, however, syllables whose vowel is *y* or which contain a syllabic consonant (*l*, *m*, *n*, or *r*) are never counted. (The Lojban term for penultimate stress is "da'amoi terbasna".) Similarly, syllables created solely by adding a buffer vowel, such as /I/, are not counted.

There are actually three levels of stress — primary, secondary, and weak. Weak stress is the lowest level, so it really means no stress at all. Weak stress is required for syllables containing *y*, a syllabic consonant, or a buffer vowel.

Primary stress is required on the penultimate syllable of Lojban content words (called *brivla*). Lojbanized names may be stressed on any syllable, but if a syllable other than the penultimate is stressed, the syllable (or at least its vowel) must be capitalized in writing. Lojban structural words (called *cmavo*) may be stressed on any syllable or none at all. However, primary stress may not be used in a syllable just preceding a brivla, unless a pause divides them; otherwise, the two words may run together.

Secondary stress is the optional and non-distinctive emphasis used for other syllables besides those required to have either weak or primary stress. There are few rules governing secondary stress, which typically will follow a speaker's native language habits or preferences. Secondary stress can be used for contrast, or for emphasis of a point. Secondary stress can be emphasized at any level up to primary stress, although the speaker must not allow a false primary stress in brivla, since errors in word resolution could result.

The following are Lojban words with stress explicitly shown:

**Example 3.9.6**
dikyjvo
DI,ky,ʒvo

(In a fully-buffered dialect, the pronunciation would be: /ˈdi kæ ʒI vo/.) Note that the syllable *ky* is not counted in determining stress. The vowel *y* is never stressed in a normal Lojban context.

**Example 3.9.7**
.armstrong.
.ARM,strong.

This is a Lojbanized version of the name *Armstrong*. The final *g* must be explicitly pronounced. With full buffering, the name would be pronounced:

**Example 3.9.8**
/ˈʔa rI mI sI tI ro nI gIʔ/

However, there is no need to insert a buffer in every possible place just because it is inserted in one place: partial buffering is also acceptable. In every case, however, the stress remains in the same place: on the first syllable.

The English pronunciation of *Armstrong*, as spelled in English, is not correct by Lojban standards; the letters *ng* in English represent a velar nasal (IPA /ŋ/) which is a single consonant. In Lojban, *ng* represents two separate consonants that must both be pronounced; you may not use /ŋ/ to pronounce Lojban *ng*, although /ŋg/ is acceptable. English speakers are likely to have to pronounce the ending with a buffer, as one of the following:

**Example 3.9.9**
/ˈʔarm stron gIʔ/
or /ˈʔarm stroŋ gIʔ/
or even /ˈʔarm stro nIgʔ/

The normal English pronunciation of the name *Armstrong* could be Lojbanized as:

**Example 3.9.10**
.ARMstron.

since Lojban *n* is allowed to be pronounced as the velar nasal /ŋ/.

Here is another example showing the use of *y*:

**Example 3.9.11**
bisydja
BI,sy,dja
BI,syd,ja

This word is a compound word, or lujvo, built from the two affixes *bis* and *dja*. When they are joined, an impermissible consonant pair results: *sd*. In accordance with the algorithm for making lujvo, explained in Chapter 4, a *y* is inserted to separate the impermissible consonant pair; the *y* is not counted as a syllable for purposes of stress determination.

**Example 3.9.12**
da'udja
da'UD,ja
da'U,dja

These two syllabications sound the same to a Lojban listener — the association of unbuffered consonants in syllables is of no import in recognizing the word.

**Example 3.9.13**
e'u bridi
e'u BRI,di
E'u BRI,di
e'U.BRI,di

In Example 3.9.13, *e'u* is a cmavo and *bridi* is a brivla. Either of the first two pronunciations is permitted: no primary stress on either syllable of *e'u*, or primary stress on the first syllable. The third pronunciation, which places primary stress on the second syllable of the cmavo, requires that — since the following word is a brivla — the two words must be separated by a pause. Consider the following two cases:

**Example 3.9.14**
le re nobli prenu
le re NObli PREnu

**Example 3.9.15**
le re no bliprenu
le re no bliPREnu

If the cmavo *no* in Example 3.9.15 were to be stressed, the phrase would sound exactly like the given pronunciation of Example 3.9.14, which is unacceptable in Lojban: a single pronunciation cannot represent both.

## 3.10 IPA For English Speakers

There are many dialects of English, thus making it difficult to define the standardized symbols of the IPA in terms useful to every reader. All the symbols used in this chapter are repeated here, in more or less alphabetical order, with examples drawn from General American. In addition, some attention is given to the Received Pronunciation of (British) English. These two dialects are referred to as GA and RP respectively. Speakers of other dialects should consult a book on phonetics or their local television sets.

**Definition 3.3**

**/ˈ/** An IPA indicator of primary stress; the syllable which follows /ˈ/ receives primary stress.

**/ʔ/** An allowed variant of Lojban .. This sound is not usually considered part of English. It is the catch in your throat that sometimes occurs prior to the beginning of a word (and sometimes a syllable) which starts with a vowel. In some dialects, like Cockney and some kinds of American English, it is used between vowels instead of *t*: *bottle* /boʔl̩/. The English interjection *uh-oh!* almost always has it between the syllables.

**/ː/** A symbol indicating that the previous vowel is to be spoken for a longer time than usual. Lojban vowels can be pronounced long in order to make a greater contrast with buffer vowels.

**/a/** The preferred pronunciation of Lojban *a*. This sound doesn't occur in GA, but sounds somewhat like the *ar* of *park*, as spoken in RP or New England American. It is pronounced further forward in the mouth than /ɑ/.

**/ɑ/** An allowed variant of Lojban *a*. The *a* of GA *father*. The sound /a/ is preferred because GA speakers often relax an unstressed /ɑ/ into a schwa /ə/, as in the usual pronunciations of *about* and *sofa*. Because schwa is a distinct vowel in Lojban, English speakers must either learn to avoid this shift or to use /a/ instead: the Lojban word for *sofa* is *sfofa*, pronounced /sfofa/ or /sfofɑ/ but never /sfofə/ which would be the non-word *sfofy*.

**/æ/** Not a Lojban sound. The *a* of English *cat*.

**/b/** The preferred pronunciation of Lojban *b*. As in English *boy*, *sober*, or *job*.

**/β/** An allowed variant of Lojban *v*. Not an English sound; the Spanish *b* or *v* between vowels. This sound should not be used for Lojban *b*.

**/d/** The preferred pronunciation of Lojban *d*. As in English *dog*, *soda*, or *mad*.

**/ɛ/** The preferred pronunciation of Lojban *e*. The *e* of English *met*.

**/e/** An allowed variant of Lojban *e*. This sound is not found in English, but is the Spanish *e*, or the tense *e* of Italian. The vowel of English *say* is similar except for the off-glide: you can learn to make this sound by holding your tongue steady while saying the first part of the English vowel.

**/ə/** The preferred pronunciation of Lojban *y*. As in the *a* of English *sofa* or *about*. Schwa is generally unstressed in Lojban, as it is in English.

It is a totally relaxed sound made with the tongue in the middle of the mouth.

**/f/** The preferred pronunciation of Lojban *f*. As in *fee*, *loafer*, or *chef*.

**/ɸ/** An allowed variant of Lojban *f*. Not an English sound; the Japanese *f* sound.

**/g/** The preferred pronunciation of Lojban *g*. As in English *go*, *eagle*, or *dog*.

**/h/** The preferred pronunciation of the Lojban apostrophe sound. As in English *aha* or *oh, hello*.

**/i/** The preferred pronunciation of Lojban *i*. Essentially like the English vowel of *pizza* or *machine*, although the English vowel is sometimes pronounced with an off-glide, which should not be present in Lojban.

**/ɪ/** A possible Lojban buffer vowel. The *i* of English *bit*.

**/ö/** A possible Lojban buffer vowel. The *u* of *just* in some varieties of GA, those which make the word sound more or less like *jist*. Also Russian *y* as in *byt*' (to be); like a schwa /ə/, but higher in the mouth.

**/j/** Used in Lojban diphthongs beginning or ending with *i*. Like the *y* in English *yard* or *say*.

**/k/** The preferred pronunciation of Lojban *k*. As in English *kill*, *token*, or *flak*.

**/l/** The preferred pronunciation of Lojban *l*. As in English *low*, *nylon*, or *excel*.

**/l̩/** The syllabic version of Lojban *l*, as in English *bottle* or *middle*.

**/m/** The preferred pronunciation of Lojban *m*. As in English *me*, *humor*, or *ham*.

**/m̩/** The syllabic version of Lojban *m*. As in English *catch 'em* or *bottom*.

**/n/** The preferred pronunciation of Lojban *n*. As in English *no*, *honor*, or *son*.

**/n̩/** The syllabic version of Lojban *n*. As in English *button*.

**/ŋ/** An allowed variant of Lojban *n*, especially in Lojbanized names and before *g* or *k*. As in English *sing* or *singer* (but not *finger* or *danger*).

**/ŋ̩/** An allowed variant of Lojban syllabic *n*, especially in Lojbanized names.

**/o/** The preferred pronunciation of Lojban *o*. As in the French *haute (cuisine)* or Spanish *como*. There is no exact English equivalent of this sound. The nearest GA equivalent is the *o* of *dough* or *joke*, but it is essential that the off-glide (a /w/-like sound) at the end of the vowel is not pronounced when speaking Lojban. The RP sound in these words is /ə/ in IPA terms, and has no /o/ in it at all; unless you can speak with a Scots, Irish, or American accent, you may have trouble with this sound.

**/ʔ/** An allowed variant of Lojban *o*, especially before *r*. This sound is a shortened form of the *aw* in GA *dawn* (for those people who don't pronounce *dawn* and *Don* alike; if you do, you may have trouble with this sound). In RP, but not GA, it is the *o* of *hot*.

**/p/** The preferred pronunciation of Lojban *p*. As in English *pay*, *super*, or *up*.

**/r/** One version of Lojban *r*. Not an English sound. The Spanish *rr* and

the Scots *r*, a tongue-tip trill.

**/ɹ/** One version of Lojban *r*. As in GA *right*, *baron*, or *car*. Not found in RP.

**/ɾ/** One version of Lojban *r*. In GA, appears as a variant of *t* or *d* in the words *metal* and *medal* respectively. A tongue-tip flap. /ʁ/ One version of Lojban *r*. Not an English sound. The French or German *r* in *reine* or *rot* respectively. A uvular trill.

**/r̩/, /ɹ̩/, /ɾ̩/, /ʁ̩/** These are syllabic versions of the above. /ɹ̩/ appears in the GA (but not RP) pronunciation of *bird*.

**/s/** The preferred pronunciation of Lojban *s*. As in English *so*, *basin*, or *yes*.

**/ʃ/** The preferred pronunciation of Lojban *c*. The *sh* of English *ship*, *ashen*, or *dish*.

**/ʂ/** An allowed variant of Lojban *s*. Not an English sound. The Hindi retroflex *s* with underdot, or Klingon *S*.

**/t/** The preferred pronunciation of Lojban *t*. As in English *tea*, *later*, or *not*. It is important to avoid the GA habit of pronouncing the *t* between vowels as /d/ or /ɾ/.

**/T/** Not normally a Lojban sound, but a possible variant of Lojban '. The *th* of English *thin* (but not *then*).

**/v/** The preferred pronunciation of Lojban *v*. As in English *voice*, *savor*, or *live*.

**/w/** Used in Lojban diphthongs beginning or ending with *u*. Like the *w* in English *wet* /wɛt/ or *cow* /kɑw/.

**/x/** The preferred pronunciation of Lojban *x*. Not normally an English sound, but used in some pronunciations of *loch* and *Bach*; *gh* in Scots *might* and *night*. The German *Ach-Laut*. To pronounce /x/, force air through your throat without vibrating your vocal chords; there should be lots of scrape.

**/Y/** A possible Lojban buffer vowel. Not an English sound: the *ü' of German "hüsch*.

**/z/** The preferred pronunciation of Lojban *z*. As in English *zoo*, *hazard*, or *fizz*.

**/ʒ/** The preferred pronunciation of Lojban *j*. The *si* of English *vision*, or the consonant at the end of GA *garage*.

**/z̬/** An allowed variant of Lojban *z*. Not an English sound. The voiced version of /ʂ/.

## 3.11 English Analogues For Lojban Diphthongs

Here is a list of English words that contain diphthongs that are similar to the Lojban diphthongs. This list does not constitute an official pronunciation guide; it is intended as a help to English-speakers.

| Lojban | English |
|---|---|
| ai | *pie* |

| | |
|---|---|
| ei | *pay* |
| oi | *boy* |
| au | *cow* |
| ia | *yard* |
| ie | *yes* |
| ii | *ye* |
| io | *yodel* (in GA only) |
| iu | *unicorn* or *few* |
| ua | *suave* |
| ue | *wet* |
| ui | *we* |
| uo | *woe* (in GA only) |
| uu | *woo* |
| iy | *million* (the *io* part, that is) |
| uy | *was* (when unstressed) |

## 3.12 Oddball Orthographies

The following notes describe ways in which Lojban has been written or could be written that differ from the standard orthography explained in the rest of this chapter. Nobody needs to read this section except people with an interest in the obscure. Technicalities are used without explanation or further apology.

There exists an alternative orthography for Lojban, which is designed to be as compatible as possible (but no more so) with the authority used in pre-Lojban versions of Loglan. The consonants undergo no change, except that *x* is replaced by *h*. The individual vowels likewise remain unchanged. However, the vowel pairs and diphthongs are changed as follows:

- *ai*, *ei*, *oi*, *au* become *ai*, *ei*, *oi*, *ao*.
- *ia* through *iu* and *ua* through *uu* remain unchanged.
- *a'i*, *e'i*, *o'i* and *a'o* become *a,i*, *e,i*, *o,i* and *a,o*.
- *i'a* through *i'u* and *u'a* through *u'u* are changed to *ia* through *iu* and *ua* through *uu* in lujvo and cmavo other than attitudinals, but become *i,a* through *i,u* and *u,a* through *u,u* in names, fu'ivla, and attitudinal cmavo.
- All other vowel pairs simply drop the apostrophe.

The result of these rules is to eliminate the apostrophe altogether, replacing it with comma where necessary, and otherwise with nothing. In addition, names and the cmavo *.i* are capitalized, and irregular stress is marked with an apostrophe (now no longer used for a sound) following the stressed syllable.

Three points must be emphasized about this alternative orthography:

- It is not standard, and has not been used.

- It does not represent any changes to the standard Lojban phonology; it is simply a representation of the same phonology using a different written form.
- It was designed to aid in a planned rapprochement between the Logical Language Group and The Loglan Institute, a group headed by James Cooke Brown. The rapprochement never took place.

There also exists a Cyrillic orthography for Lojban which was designed when the introductory Lojban brochure was translated into Russian. It uses the letters *a*, *be*, *ve*, *ge*, *de*, *e*, *zhe*, *ze*, *i*, *ka*, *el*, *em*, *en*, *o*, *pe*, *er*, *es*, *te*, *u*, *ef*, *kha*, and *sha* in the obvious ways. The Latin letter *y* is mapped onto the hard sign, as in Bulgarian. The apostrophe, comma, and period are unchanged. Diphthongs are written as vowel pairs, as in the Roman representation.

Finally, an orthography using the Tengwar of Féanor, a fictional orthography invented by J. R. R. Tolkien and described in the Appendixes to *The Lord Of The Rings*, has been devised for Lojban. The following mapping, which closely resembles that used for Westron, will be meaningful only to those who have read those appendixes. In brief, the tincotéma and parmatéma are used in the conventional ways; the calmatéma represents palatal consonants, and the quessetéma represents velar consonants.

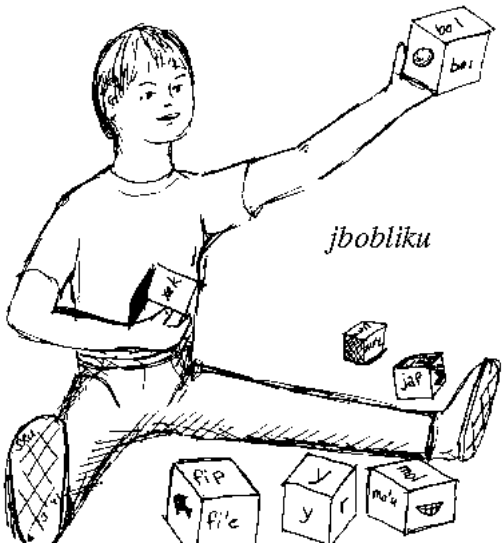
| | | | |
|---|---|---|---|
| t | tinco | p | parma |
| - | calma | k | quesse |
| d | ando | b | umbar |
| - | anga | g | ungwe |
| - | thule | f | formen |
| c | harma | x | hwesta |
| - | anto | v | ampa |
| j | anca | - | unque |
| n | numen | m | malta |
| - | noldo | - | nwalme |
| r | ore | u | vala |
| i | anna | - | vilya |

The letters *vala* and *anna* are used for *u* and *i* only when those letters are used to represent glides. Of the additional letters, *r*, *l*, *s*, and *z* are written with *rómen*, *lambe*, *silme*, and *áre/esse* respectively; the inverted forms are used as free variants.

Lojban, like Quenya, is a vowel-last language, so tehtar are read as following the tengwar on which they are placed. The conventional tehtar are used for the five regular vowels, and the under-dot for *y*. The Lojban apostrophe is represented by *halla*. There is no equivalent of the Lojban comma or period.

# Chapter 4

# The Shape of Words to Come: Lojban Morphology




## 4.1 Introductory

Morphology is the part of grammar that deals with the form of words. Lojban's morphology is fairly simple compared to that of many languages, because Lojban words don't change form depending on how they are used. English has only a small number of such changes compared to languages like Russian, but we do have changes like *boys* as the plural of *boy*, or *walked* as the past-tense form of *walk*. To make plurals or past tenses in Lojban, you add separate words to the sentence that express the number of boys, or the time when the walking was going on.

However, Lojban does have what is called *derivational morphology*: the capability of building new words from old words. In addition, the form of words tells us something about their grammatical uses, and sometimes about the means by which they entered the language. Lojban has very orderly rules for the formation of words

of various types, both the words that already exist and new words yet to be created by speakers and writers.

A stream of Lojban sounds can be uniquely broken up into its component words according to specific rules. These so-called *morphology rules* are summarized in this chapter. (However, a detailed algorithm for breaking sounds into words has not yet been fully debugged, and so is not presented in this book.) First, here are some conventions used to talk about groups of Lojban letters, including vowels and consonants.

1. V represents any single Lojban vowel except *y*; that is, it represents *a*, *e*, *i*, *o*, or *u*.

2. VV represents either a dipthong: *ai*, *ei*, *oi*, *au*; or a two-syllable vowel pair with an apostrophe separating the vowels: *a'a*, *a'e*, *a'i*, *a'o*, *a'u*, *e'a*, *e'e*, *e'i*, *e'o*, *e'u*, *i'a*, *i'e*, *i'i*, *i'o*, *i'u*, *o'a*, *o'e*, *o'i*, *o'o*, *o'u*, *u'a*, *u'e*, *u'i*, *u'o*, *u'u*.

3. C represents a single Lojban consonant, not including the apostrophe, one of *b*, *c*, *d*, *f*, *g*, *j*, *k*, *l*, *m*, *n*, *p*, *r*, *s*, *t*, *v*, *x*, or *z*. Syllabic *l*, *m*, *n*, and *r* always count as consonants for the purposes of this chapter.

4. CC represents two adjacent consonants of type C which constitute one of the 48 permissible initial consonant pairs: *bl*, *br*, *cf*, *ck*, *cl*, *cm*, *cn*, *cp*, *cr*, *ct*, *dj*, *dr*, *dz*, *fl*, *fr*, *gl*, *gr*, *jb*, *jd*, *jg*, *jm*, *jv*, *kl*, *kr*, *ml*, *mr*, *pl*, *pr*, *sf*, *sk*, *sl*, *sm*, *sn*, *sp*, *sr*, *st*, *tc*, *tr*, *ts*, *vl*, *vr*, *xl*, *xr*, *zb*, *zd*, *zg*, *zm*, *zv*.

5. C/C represents two adjacent consonants which constitute one of the permissible consonant pairs (not necessarily a permissible initial consonant pair). The permissible consonant pairs are explained in Chapter 3. In brief, any consonant pair is permissible unless it contains: two identical letters, both a voiced (excluding *r*, *l*, *m*, *n*) and and an unvoiced consonant, or is one of certain specified forbidden pairs.

6. C/CC represents a consonant triple. The first two consonants must constitute a permissible consonant pair; the last two consonants must constitute a permissible initial consonant pair.

Lojban has three basic word classes — parts of speech — in contrast to the eight that are traditional in English. These three classes are called cmavo, brivla, and cmene. Each of these classes has uniquely identifying properties — an arrangement of letters that allows the word to be uniquely and unambiguously recognized as a separate word in a string of Lojban, upon either reading or hearing, and as belonging to a specific word-class.

They are also functionally different: cmavo are the structure words, corresponding to English words like *and*, *if*, *the* and *to*; brivla are the content words, corresponding to English words like *come*, *red*, *doctor*, and *freely*; cmene are proper names, corresponding to English *James*, *Afghanistan*, and *Pope John Paul II*.

## 4.2 cmavo

The first group of Lojban words discussed in this chapter are the cmavo. They are the structure words that hold the Lojban language together. They often have no semantic

meaning in themselves, though they may affect the semantics of brivla to which they are attached. The cmavo include the equivalent of English articles, conjunctions, prepositions, numbers, and punctuation marks. There are over a hundred subcategories of cmavo, known as *selma'o*, each having a specifically defined grammatical usage. The various selma'o are discussed throughout Chapter 5 to Chapter 19 and summarized in Chapter 20.

Standard cmavo occur in four forms defined by their word structure. Here are some examples of the various forms:

| | | | | | |
|---|---|---|---|---|---|
| V-form | .a | .e | .i | .o | .u |
| CV-form | ba | ce | di | fo | gu |
| VV-form | .au | .ei | .ia | .o'u | .u'e |
| CVV-form | ki'a | pei | mi'o | coi | cu'u |

In addition, there is the cmavo *.y.* (remember that *y* is not a V), which must have pauses before and after it.

A simple cmavo thus has the property of having only one or two vowels, or of having a single consonant followed by one or two vowels. Words consisting of three or more vowels in a row, or a single consonant followed by three or more vowels, are also of cmavo form, but are reserved for experimental use: a few examples are *ku'a'e*, *sau'e*, and *bai'ai*. All CVV cmavo beginning with the letter *x* are also reserved for experimental use. In general, though, the form of a cmavo tells you little or nothing about its grammatical use.

*Experimental use* means that the language designers will not assign any standard meaning or usage to these words, and words and usages coined by Lojban speakers will not appear in official dictionaries for the indefinite future. Experimental-use words provide an escape hatch for adding grammatical mechanisms (as opposed to semantic concepts) the need for which was not foreseen.

The cmavo of VV-form include not only the diphthongs and vowel pairs listed in Section 4.1, but also the following ten additional diphthongs:

| | | | | |
|---|---|---|---|---|
| .ia | .ie | .ii | .io | .iu |
| .ua | .ue | .ui | .uo | .uu |

In addition, cmavo can have the form Cy, a consonant followed by the letter *y*. These cmavo represent letters of the Lojban alphabet, and are discussed in detail in Chapter 17.

Compound cmavo are sequences of cmavo attached together to form a single written word. A compound cmavo is always identical in meaning and in grammatical use to the separated sequence of simple cmavo from which it is composed. These words are written in compound form merely to save visual space, and to ease the reader's burden in identifying when the component cmavo are acting together.

Compound cmavo, while not visually short like their components, can be readily identified by two characteristics:

1. They have no consonant pairs or clusters, and
2. They end in a vowel.

For example:

**Example 4.2.1**
.iseci'i
.i se ci'i

**Example 4.2.2**
punaijecanai
pu nai je ca nai

**Example 4.2.3**
ki'e.u'e
ki'e .u'e

The cmavo *.u'e* begins with a vowel, and like all words beginning with a vowel, requires a pause (represented by .) before it. This pause cannot be omitted simply because the cmavo is incorporated into a compound cmavo. On the other hand,

**Example 4.2.4**
ki'e'u'e

is a single cmavo reserved for experimental purposes: it has four vowels.

**Example 4.2.5**
cy.ibu.abu
cy. .ibu .abu

Again the pauses are required (see Section 4.9); the pause after *cy.* merges with the pause before *.ibu*.

There is no particular stress required in cmavo or their compounds. Some conventions do exist that are not mandatory. For two-syllable cmavo, for example, stress is typically placed on the first vowel; an example is

**Example 4.2.6**
.e'o ko ko kurji
.E'o ko ko KURji

This convention results in a consistent rhythm to the language, since brivla are required to have penultimate stress; some find this esthetically pleasing.

If the final syllable of one word is stressed, and the first syllable of the next word is stressed, you must insert a pause or glottal stop between the two stressed syllables. Thus

**Example 4.2.7**
le re nanmu

can be optionally pronounced

**Example 4.2.8**
le RE. NANmu

since there are no rules forcing stress on either of the first two words; the stress on *re*, though, demands that a pause separate *re* from the following syllable *nan* to ensure that the stress on *nan* is properly heard as a stressed syllable. The alternative pronunciation

**Example 4.2.9**
LE re NANmu

is also valid; this would apply secondary stress (used for purposes of emphasis, contrast or sentence rhythm) to *le*, comparable in rhythmical effect to the English phrase "THE two men". In Example 4.2.8, the secondary stress on *re* would be similar to that in the English phrase "the TWO men".

Both cmavo may also be left unstressed, thus:

**Example 4.2.10**
le re NANmu

This would probably be the most common usage.

## 4.3 brivla

Predicate words, called *brivla*, are at the core of Lojban. They carry most of the semantic information in the language. They serve as the equivalent of English nouns, verbs, adjectives, and adverbs, all in a single part of speech.

Every brivla belongs to one of three major subtypes. These subtypes are defined by the form, or morphology, of the word — all words of a particular structure can be assigned by sight or sound to a particular type (cmavo, brivla, or cmene) and subtype. Knowing the type and subtype then gives you, the reader or listener, significant clues to the meaning and the origin of the word, even if you have never heard the word before.

The same principle allows you, when speaking or writing, to invent new brivla for new concepts *on the fly*; yet it offers people that you are trying to communicate with a good chance to figure out your meaning. In this way, Lojban has a flexible vocabulary which can be expanded indefinitely.

All brivla have the following properties:

1. always end in a vowel;
2. always contain a consonant pair in the first five letters, where *y* and apostrophe are not counted as letters for this purpose. (See Section 4.6.)
3. always are stressed on the next-to-the-last (penultimate) syllable; this implies that they have two or more syllables.

The presence of a consonant pair distinguishes brivla from cmavo and their compounds. The final vowel distinguishes brivla from cmene, which always end in a consonant. Thus *da'amei* must be a compound cmavo because it lacks a consonant pair; *lojban.* must be a name because it lacks a final vowel.

Thus, *bisycla* has the consonant pair *sc* in the first five non-*y* letters even though the *sc* actually appears in the form of *syc*. Similarly, the word *ro'inre'o* contains *nr* in the first five letters because the apostrophes are not counted for this purpose.

The three subtypes of brivla are:

1. gismu, the Lojban primitive roots from which all other brivla are built;
2. lujvo, the compounds of two or more gismu; and
3. fu'ivla (literally *copy-word*), the specialized words that are not Lojban primitives or natural compounds, and are therefore borrowed from other languages.

## 4.4 gismu

The gismu, or Lojban root words, are those brivla representing concepts most basic to the language. The gismu were chosen for various reasons: some represent concepts that are very familiar and basic; some represent concepts that are frequently used in other languages; some were added because they would be helpful in constructing more complex words; some because they represent fundamental Lojban concepts (like *cmavo* and *gismu* themselves).

The gismu do not represent any sort of systematic partitioning of semantic space. Some gismu may be superfluous, or appear for historical reasons: the gismu list was being collected for almost 35 years and was only weeded out once. Instead, the intention is that the gismu blanket semantic space: they make it possible to talk about the entire range of human concerns.

There are about 1350 gismu. In learning Lojban, you need only to learn most of these gismu and their combining forms (known as *rafsi*) as well as perhaps 200 major cmavo, and you will be able to communicate effectively in the language. This may sound like a lot, but it is a small number compared to the vocabulary needed for similar communications in other languages.

All gismu have very strong form restrictions. Using the conventions defined in Section 4.1, all gismu are of the forms CVC/CV or CCVCV. They must meet the rules for all brivla given in Section 4.3; furthermore, they:

1. always have five letters
2. always start with a consonant and end with a single vowel
3. always contain exactly one consonant pair, which is a permissible initial pair (CC) if it's at the beginning of the gismu, but otherwise only has to be a permissible pair (C/C);
4. are always stressed on the first syllable (since that is penultimate).

The five letter length distinguishes gismu from lujvo and fu'ivla. (It is possible to have fu'ivla like *spa'i* that are five letters long, but they must have '; no gismu contains '.)

With the exception of five special brivla variables, *broda*, *brode*, *brodi*, *brodo*, and *brodu*, no two gismu differ only in the final vowel. Furthermore, the set of gismu was specifically designed to reduce the likelihood that two similar sounding gismu could be confused. For example, because *gismu* is in the set of gismu, *kismu*, *xismu*, *gicmu*, *gizmu*, and *gisnu* cannot be.

Almost all Lojban gismu are constructed from pieces of words drawn from other languages, specifically Chinese, English, Hindi, Spanish, Russian, and Arabic, the six most widely spoken natural languages. For a given concept, words in the six languages that represent that concept were written in Lojban phonetics. Then a gismu was selected to maximize the recognizability of the Lojban word for speakers of the six languages by weighting the inclusion of the sounds drawn from each language by the number of speakers of that language. See Section 4.14 for a full explanation of the algorithm.

Here are a few examples of gismu, with rough English equivalents (not definitions):

**Example 4.4.1**
creka
shirt

**Example 4.4.2**
lijda
religion

**Example 4.4.3**
blanu
blue

**Example 4.4.4**
mamta
mother

**Example 4.4.5**
cukta
book

**Example 4.4.6**
patfu
father

**Example 4.4.7**
nanmu
man

**Example 4.4.8**
ninmu
woman

A small number of gismu were formed differently; see Section 4.15 for a list.

## 4.5 lujvo

When specifying a concept that is not found among the gismu (or, more specifically, when the relevant gismu seems too general in meaning), a Lojbanist generally attempts to express the concept as a tanru. Lojban tanru are an elaboration of the concept of *metaphor* used in English. In Lojban, any brivla can be used to modify another brivla. The first of the pair modifies the second. This modification is usually restrictive — the modifying brivla reduces the broader sense of the modified brivla to form a more narrow, concrete, or specific concept. Modifying brivla may thus be seen as acting like English adverbs or adjectives. For example,

**Example 4.5.1**
skami pilno

is the tanru which expresses the concept of *computer user*.

The simplest Lojban tanru are pairings of two concepts or ideas. Such tanru take two simpler ideas that can be represented by gismu and combine them into a single

more complex idea. Two-part tanru may then be recombined in pairs with other tanru, or with individual gismu, to form more complex or more specific ideas, and so on.

The meaning of a tanru is usually at least partly ambiguous: *skami pilno* could refer to a computer that is a user, or to a user of computers. There are a variety of ways that the modifier component can be related to the modified component. It is also possible to use cmavo within tanru to provide variations (or to prevent ambiguities) of meaning.

Making tanru is essentially a poetic or creative act, not a science. While the syntax expressing the grouping relationships within tanru is unambiguous, tanru are still semantically ambiguous, since the rules defining the relationships between the gismu are flexible. The process of devising a new tanru is dealt with in detail in Chapter 5.

To express a simple tanru, simply say the component gismu together. Thus the binary metaphor *big boat* becomes the tanru

**Example 4.5.2**
barda bloti

representing roughly the same concept as the English word *ship*.

The binary metaphor *father mother* can refer to a paternal grandmother ("a father-ly type of mother"), while *mother father* can refer to a maternal grandfather ("a mother-ly type of father"). In Lojban, these become the tanru

**Example 4.5.3**
patfu mamta

and

**Example 4.5.4**
mamta patfu

respectively.

The possibility of semantic ambiguity can easily be seen in the last case. To interpret Example 4.5.4, the listener must determine what type of motherliness pertains to the father being referred to. In an appropriate context, *mamta patfu* could mean not *grandfather* but simply *father with some motherly attributes*, depending on the culture. If absolute clarity is required, there are ways to expand upon and explain the exact interrelationship between the components; but such detail is usually not needed.

When a concept expressed in a tanru proves useful, or is frequently expressed, it is desirable to choose one of the possible meanings of the tanru and assign it to a new brivla. For Example 4.5.1, we would probably choose *user of computers*, and form the new word

**Example 4.5.5**
sampli

Such a brivla, built from the rafsi which represent its component words, is called a *lujvo*. Another example, corresponding to the tanru of Example 4.5.2, would be:

**Example 4.5.6**
bralo'i
big-boat
ship

The lujvo representing a given tanru is built from units representing the component gismu. These units are called *rafsi* in Lojban. Each rafsi represents only one gismu. The rafsi are attached together in the order of the words in the tanru, occasionally inserting so-called *hyphen* letters to ensure that the pieces stick together as a single word and cannot accidentally be broken apart into cmavo, gismu, or other word forms. As a result, each lujvo can be readily and accurately recognized, allowing a listener to pick out the word from a string of spoken Lojban, and if necessary, unambiguously decompose the word to a unique source tanru, thus providing a strong clue to its meaning.

The lujvo that can be built from the tanru *mamta patfu* in Example 4.5.4 is

**Example 4.5.7**
mampa'u

which refers specifically to the concept "maternal grandfather". The two gismu that constitute the tanru are represented in *mampa'u* by the rafsi *mam-* and *-pa'u*, respectively; these two rafsi are then concatenated together to form *mampa'u*.

Like gismu, lujvo have only one meaning. When a lujvo is formally entered into a dictionary of the language, a specific definition will be assigned based on one particular interrelationship between the terms. (See Chapter 12 for how this has been done.) Unlike gismu, lujvo may have more than one form. This is because there is no difference in meaning between the various rafsi for a gismu when they are used to build a lujvo. A long rafsi may be used, especially in noisy environments, in place of a short rafsi; the result is considered the same lujvo, even though the word is spelled and pronounced differently. Thus the word *brivla*, built from the tanru *bridi valsi*, is the same lujvo as *brivalsi*, *bridyvla*, and *bridyvalsi*, each of which uses a different combination of rafsi.

When assembling rafsi together into lujvo, the rules for valid brivla must be followed: a consonant cluster must occur in the first five letters (excluding *y* and '), and the lujvo must end in a vowel.

A *y* (which is ignored in determining stress or consonant clusters) is inserted in the middle of the consonant cluster to glue the word together when the resulting cluster is either not permissible or the word is likely to break up. There are specific rules describing these conditions, detailed in Section 4.6.

An *r* (in some cases, an *n*) is inserted when a CVV-form rafsi attaches to the beginning of a lujvo in such a way that there is no consonant cluster. For example, in the lujvo

**Example 4.5.8**
soirsai
sonci sanmi
soldier meal
field rations

the rafsi *soi-* and *-sai* are joined, with the additional *r* making up the *rs* consonant pair needed to make the word a brivla. Without the *r*, the word would break up into *soi sai*, two cmavo. The pair of cmavo have no relation to their rafsi lookalikes; they will either be ungrammatical (as in this case), or will express a different meaning from what was intended.

Learning rafsi and the rules for assembling them into lujvo is clearly seen to be necessary for fully using the potential Lojban vocabulary.

Most important, it is possible to invent new lujvo while you speak or write in order to represent a new or unfamiliar concept, one for which you do not know any existing Lojban word. As long as you follow the rules for building these compounds, there is a good chance that you will be understood without explanation.

## 4.6 rafsi

Every gismu has from two to five rafsi, each of a different form, but each such rafsi represents only one gismu. It is valid to use any of the rafsi forms in building lujvo — whichever the reader or listener will most easily understand, or whichever is most pleasing — subject to the rules of lujvo making. There is a scoring algorithm which is intended to determine which of the possible and legal lujvo forms will be the standard dictionary form (see Section 4.12).

Each gismu always has at least two rafsi forms; one is the gismu itself (used only at the end of a lujvo), and one is the gismu without its final vowel (used only at the beginning or middle of a lujvo). These forms are represented as -CVC/CV or -CCVCV (called *the 5-letter rafsi*), and -CVC/C- or -CCVC- (called *the 4-letter rafsi*) respectively. The dashes in these rafsi form representations show where other rafsi may be attached to form a valid lujvo. When lujvo are formed only from 4-letter and 5-letter rafsi, known collectively as "long rafsi", they are called *unreduced lujvo*.

Some examples of unreduced lujvo forms are:

**Example 4.6.1**
mamtypatfu
from *mamta patfu*
*mother father* or *maternal grandfather*

**Example 4.6.2**
lerfyliste
from *lerfu liste*
*letter list* or a *list of letters*
(letters of the alphabet)

**Example 4.6.3**
nancyprali
from *nanca prali*
*year profit* or *annual profit*

**Example 4.6.4**
prunyplipe
from *pruni plipe*
*elastic (springy) leap* or *spring* (the verb)

**Example 4.6.5**
vancysanmi
from *vanci sanmi*
*evening meal* or *supper*

In addition to these two forms, each gismu may have up to three additional short rafsi, three letters long. All short rafsi have one of the forms -CVC-, -CCV-, or -CVV-.

The total number of rafsi forms that are assigned to a gismu depends on how useful the gismu is, or is presumed to be, in making lujvo, when compared to other gismu that could be assigned the rafsi.

For example, *zmadu* ("more than") has the two short rafsi *-zma-* and *-mau-* (in addition to its unreduced rafsi *-zmad-* and *-zmadu*), because a vast number of lujvo have been created based on *zmadu*, corresponding in general to English comparative adjectives ending in *-er* such as *whiter* (Lojban *labmau*). On the other hand, *bakri* ("chalk") has no short rafsi and few lujvo.

There are at most one CVC-form, one CCV-form, and one CVV-form rafsi per gismu. In fact, only a tiny handful of gismu have both a CCV-form and a CVV-form rafsi assigned, and still fewer have all three forms of short rafsi. However, gismu with both a CVC-form and another short rafsi are fairly common, partly because more possible CVC-form rafsi exist. Yet CVC-form rafsi, even though they are fairly easy to remember, cannot be used at the end of a lujvo (because lujvo must end in vowels), so justifying the assignment of an additional short rafsi to many gismu.

The intention was to use the available *rafsi space* — the set of all possible short rafsi forms — in the most efficient way possible; the goal is to make the most-used lujvo as short as possible (thus maximizing the use of short rafsi), while keeping the rafsi very recognizable to anyone who knows the source gismu. For this reason, the letters in a rafsi have always been chosen from among the five letters of the corresponding gismu. As a result, there are a limited set of short rafsi available for assignment to each gismu. At most seven possible short rafsi are available for consideration (of which at most three can be used, as explained above).

Here are the only short rafsi forms that can possibly exist for gismu of the form CVC/CV, like *sakli*. The digits in the second column represent the gismu letters used to form the rafsi.

| Form | Letters Used | Rafsi |
|---|---|---|
| CVC | 123 | -sak- |
| CVC | 124 | -sal- |
| CVV | 12'5 | -sa'i- |
| CVV | 125 | -sai- |
| CCV | 345 | -kli- |
| CCV | 132 | -ska- |

(The only actual short rafsi for *sakli* is *-sal-*.)

For gismu of the form CCVCV, like *blaci*, the only short rafsi forms that can exist are:

| Form | Letters Used | Rafsi |
|---|---|---|
| CVC | 134 | -bac- |
| CVC | 234 | -lac |
| CVV | 13'5 | -ba'i- |
| CVV | 135 | -bai- |
| CVV | 23'5 | -la'i- |
| CVV | 235 | -lai- |
| CCV | 123 | -bla- |

(In fact, *blaci* has none of these short rafsi; they are all assigned to other gismu. Lojban speakers are not free to reassign any of the rafsi; the tables shown here are to help understand how the rafsi were chosen in the first place.)

There are a few restrictions: a CVV-form rafsi without an apostrophe cannot exist unless the vowels make up one of the four diphthongs *ai*, *ei*, *oi*, or *au*; and a CCV-form rafsi is possible only if the two consonants form a permissible initial consonant pair (see Section 4.1). Thus *mamta*, which has the same form as *salci*, can only have *mam*, *mat*, and *ma'a* as possible rafsi: in fact, only *mam* is assigned to it.

Some cmavo also have associated rafsi, usually CVC-form. For example, the ten common numerical digits, which are all CV form cmavo, each have a CVC-form rafsi formed by adding a consonant to the cmavo. Most cmavo that have rafsi are ones used in composing tanru (for a complete list, see Chapter 12).

The term for a lujvo made up solely of short rafsi is *fully reduced lujvo*. Here are some examples of fully reduced lujvo:

**Example 4.6.6**
cumfri
from *cumki lifri*
*possible experience*

**Example 4.6.7**
klezba
from *klesi zbasu*
*category make*

**Example 4.6.8**
kixta'a
from *krixa tavla*
*cry-out talk*

**Example 4.6.9**
sniju'o
from *sinxa djuno*
*sign know*

In addition, some of the unreduced forms in the previous example may be fully reduced to:

**Example 4.6.10**
mampa'u
from *mamta patfu*
*mother father* or *maternal grandfather*

**Example 4.6.11**
lerste
from *lerfu liste*
*letter list* or a *list of letters*

As noted above, CVC-form rafsi cannot appear as the final rafsi in a lujvo, because all lujvo must end with one or two vowels. As a brivla, a lujvo must also contain a consonant cluster within the first five letters — this ensures that they cannot be mistaken for compound cmavo. Of course, all lujvo have at least six letters since they have two or more rafsi, each at least three letters long; hence they cannot be confused with gismu.

When attaching two rafsi together, it may be necessary to insert a hyphen letter. In Lojban, the term *hyphen* always refers to a letter, either the vowel *y* or one of the consonants *r* and *n*. (The letter *l* can also be a hyphen, but is not used as one in lujvo.)

The *y*-hyphen is used after a CVC-form rafsi when joining it with the following rafsi could result in an impermissible consonant pair, or when the resulting lujvo could fall apart into two or more words (either cmavo or gismu).

Thus, the tanru *pante tavla* ("protest talk") cannot produce the lujvo *patta'a*, because *tt* is not a permissible consonant pair; the lujvo must be *patyta'a*. Similarly, the tanru *mudri siclu* ("wooden whistle") cannot form the lujvo *mudsiclu*; instead, *mudysiclu* must be used. (Remember that *y* is not counted in determining whether the first five letters of a brivla contain a consonant cluster: this is why.)

The *y*-hyphen is also used to attach a 4-letter rafsi, formed by dropping the final vowel of a gismu, to the following rafsi. (This procedure was shown, but not explained, in Example 4.6.1 to Example 4.6.5.) The lujvo forms *zunlyjamfu*, *zunlyjma*, *zuljamfu*, and *zuljma* are all legitimate and equivalent forms made from the tanru *zunle jamfu* ("left foot"). Of these, *zuljma* is the preferred one since it is the shortest; it thus is likely to be the form listed in a Lojban dictionary.

The *r*-hyphen and its close relative, the *n*-hyphen, are used in lujvo only after CVV-form rafsi. A hyphen is always required in a two-part lujvo of the form CVV-CVV, since otherwise there would be no consonant cluster.

An *r*-hyphen or *n*-hyphen is also required after the CVV-form rafsi of any lujvo of the form CVV-CVC/CV or CVV-CCVCV since it would otherwise fall apart into a CVV-form cmavo and a gismu. In any lujvo with more than two parts, a CVV-form rafsi in the initial position must always be followed by a hyphen. If the hyphen were to be omitted, the supposed lujvo could be broken into smaller words without the hyphen: because the CVV-form rafsi would be interpreted as a cmavo, and the remainder of the word as a valid lujvo that is one rafsi shorter.

An *n*-hyphen is only used in place of an *r*-hyphen when the following rafsi begins with *r*. For example, the tanru *rokci renro* ("rock throw") cannot be expressed as *ro'ire'o* (which breaks up into two cmavo), nor can it be *ro'irre'o* (which has an impermissible double consonant); the *n*-hyphen is required, and the correct form of the hyphenated lujvo is *ro'inre'o*. The same lujvo could also be expressed without hyphenation as *rokre'o*.

There is also a different way of building lujvo, or rather phrases which are grammatically and semantically equivalent to lujvo. You can make a phrase containing any desired words, joining each pair of them with the special cmavo *zei*. Thus,

**Example 4.6.12**
bridi zei valsi

is the exact equivalent of *brivla* (but not necessarily the same as the underlying tanru *bridi valsi*, which could have other meanings.) Using *zei* is the only way to get a cmavo lacking a rafsi, a cmene, or a fu'ivla into a lujvo:

**Example 4.6.13**
xy. zei kantu
X ray

**Example 4.6.14**
kulnr,farsi zei lolgai
Farsi floor-cover
Persian rug

**Example 4.6.15**
na'e zei .a zei na'e zei by. livgyterbilma
non-A, non-B liver-disease
non-A, non-B hepatitis

**Example 4.6.16**
.cerman. zei xarnykarce
Sherman war-car
Sherman tank

Example 4.6.15 is particularly noteworthy because the phrase that would be produced by removing the *zei*s from it doesn't end with a brivla, and in fact is not even grammatical. As written, the example is a tanru with two components, but by adding a *zei* between *by.* and *livgyterbilma* to produce

**Example 4.6.17**
na'e zei .a zei na'e zei by. zei livgyterbilma
non-A-non-B-hepatitis

the whole phrase would become a single lujvo. The longer lujvo of Example 4.6.17 may be preferable, because its place structure can be built from that of *bilma*, whereas the place structure of a lujvo without a brivla must be constructed ad hoc.

Note that rafsi may not be used in *zei* phrases, because they are not words. CVV rafsi look like words (specifically cmavo) but there can be no confusion between the two uses of the same letters, because cmavo appear only as separate words or in compound cmavo (which are really just a notation for writing separate but closely related words as if they were one); rafsi appear only as parts of lujvo.

## 4.7 fu'ivla

The use of tanru or lujvo is not always appropriate for very concrete or specific terms (e.g. *brie* or *cobra*), or for jargon words specialized to a narrow field (e.g. *quark*, *integral*, or *iambic pentameter*). These words are in effect names for concepts, and the names were invented by speakers of another language. The vast majority of words referring to plants, animals, foods, and scientific terminology cannot be easily expressed as tanru. They thus must be borrowed (actually *copied*) into Lojban from the original language.

There are four stages of borrowing in Lojban, as words become more and more modified (but shorter and easier to use). Stage 1 is the use of a foreign name quoted with the cmavo *la'o* (explained in full in Chapter 19):

**Example 4.7.1**
me la'o ly. spaghetti .ly.

is a predicate with the place structure "$x_1$ is a quantity of spaghetti".

Stage 2 involves changing the foreign name to a Lojbanized name, as explained in Section 4.8:

**Example 4.7.2**
me la spagetis.

One of these expedients is often quite sufficient when you need a word quickly in conversation. (This can make it easier to get by when you do not yet have full command of the Lojban vocabulary, provided you are talking to someone who will recognize the borrowing.)

Where a little more universality is desired, the word to be borrowed must be Lojbanized into one of several permitted forms. A rafsi is then usually attached to the beginning of the Lojbanized form, using a hyphen to ensure that the resulting word doesn't fall apart.

The rafsi categorizes or limits the meaning of the fu'ivla; otherwise a word having several different jargon meanings in other languages would require the word-inventor to choose which meaning should be assigned to the fu'ivla, since fu'ivla (like other brivla) are not permitted to have more than one definition. Such a Stage 3 borrowing is the most common kind of fu'ivla.

Finally, Stage 4 fu'ivla do not have any rafsi classifier, and are used where a fu'ivla has become so common or so important that it must be made as short as possible. (See Section 4.16 for a proposal concerning Stage 4 fu'ivla.)

The form of a fu'ivla reliably distinguishes it from both the gismu and the cmavo. Like cultural gismu, fu'ivla are generally based on a word from a single non-Lojban language. The word is *borrowed* (actually *copied*, hence the Lojban tanru *fukpi valsi*) from the other language and Lojbanized — the phonemes are converted to their closest Lojban equivalent and modifications are made as necessary to make the word a legitimate Lojban fu'ivla-form word. All fu'ivla:

1. must contain a consonant cluster in the first five letters of the word; if this consonant cluster is at the beginning, it must either be a permissible initial consonant pair, or a longer cluster such that each pair of adjacent consonants in the cluster is a permissible initial consonant pair: *spraile* is acceptable, but not *ktraile* or *trkaile*;

2. must end in one or more vowels;

3. must not be gismu or lujvo, or any combination of cmavo, gismu, and lujvo; furthermore, a fu'ivla with a CV cmavo joined to the front of it must not have the form of a lujvo (the so-called *slinku'i test*);

4. cannot contain *y*, although they may contain syllabic pronunciations of Lojban consonants;

5. like other brivla, are stressed on the penultimate syllable.

Note that consonant triples or larger clusters that are not at the beginning of a fu'ivla can be quite flexible, as long as all consonant pairs are permissible. There is no need to restrict fu'ivla clusters to permissible initial pairs except at the beginning.

This is a fairly liberal definition and allows quite a lot of possibilities within *fu'ivla space*. Stage 3 fu'ivla can be made easily on the fly, as lujvo can, because the procedure for forming them always guarantees a word that cannot violate any of the rules. Stage 4 fu'ivla require running tests that are not simple to characterize or perform, and should be made only after deliberation and by someone knowledgeable about all the considerations that apply.

Here is a simple and reliable procedure for making a non-Lojban word into a valid Stage 3 fu'ivla:

1. Eliminate all double consonants and silent letters.
2. Convert all sounds to their closest Lojban equivalents. Lojban *y*, however, may not be used in any fu'ivla.
3. If the last letter is not a vowel, modify the ending so that the word ends in a vowel, either by removing a final consonant or by adding a suggestively chosen final vowel.
4. If the first letter is not a consonant, modify the beginning so that the word begins with a consonant, either by removing an initial vowel or adding a suggestively chosen initial consonant.
5. (a) Prefix the result of steps 1-5 with a 4-letter rafsi that categorizes the fu'ivla into a *topic area*. It is only safe to use a 4-letter rafsi; short rafsi sometimes produce invalid fu'ivla. Hyphenate the rafsi to the rest of the fu'ivla with an *r*-hyphen; if that would produce a double *r*, use an *n*-hyphen instead; if the rafsi ends in *r* and the rest of the fu'ivla begins with *n* (or vice versa) use an *l*-hyphen. (This is the only use of *l*-hyphen in Lojban.)

   (b) Alternatively, if a CVC-form short rafsi is available it can be used instead of the long rafsi.
6. Remember that the stress necessarily appears on the penultimate (next-to-the-last) syllable.

In this section, the hyphen is set off with commas in the examples, but these commas are not required in writing, and the hyphen need not be pronounced as a separate syllable.

Here are a few examples:

**Example 4.7.3**
spaghetti (from English or Italian)
spageti (Lojbanize)
cidj,r,spageti (prefix long rafsi)
dja,r,spageti (prefix short rafsi)

where *cidj-* is the 4-letter rafsi for *cidja*, the Lojban gismu for *food*, thus categorizing *cidjrspageti* as a kind of food. The form with the short rafsi happens to work, but such good fortune cannot be relied on: in any event, it means the same thing.

**Example 4.7.4**
Acer (the scientific name of maple trees)
acer (Lojbanize)
xaceru (add initial consonant and final vowel)
tric,r,xaceru (prefix rafsi)
ric,r,xaceru (prefix short rafsi)

where *tric-* and *ric-* are rafsi for *tricu*, the gismu for "tree". Note that by the same principles, *maple sugar* could get the fu'ivla *saktrxaceru*, or could be represented by the tanru *tricrxaceru sakta*. Technically, *ricrxaceru* and *tricrxaceru* are distinct fu'ivla, but they would surely be given the same meanings if both happened to be in use.

**Example 4.7.5**
brie (from French)
bri (Lojbanize)
cirl,r,bri (prefix rafsi)

where *cirl-* represents *cirla* ("cheese").

**Example 4.7.6**
cobra
kobra (Lojbanize)
sinc,r,kobra (prefix rafsi)

where *sinc-* represents *since* ("snake").

**Example 4.7.7**
quark
kuark (Lojbanize)
kuarka (add final vowel)
sask,r,kuarka (prefix rafsi)

where *sask-* represents *saske* ("science"). Note the extra vowel *a* added to the end of the word, and the diphthong *ua*, which never appears in gismu or lujvo, but may appear in fu'ivla.

The use of the prefix helps distinguish among the many possible meanings of the borrowed word, depending on the field. As it happens, *spageti* and *kuarka* are valid Stage 4 fu'ivla, but *xaceru* looks like a compound cmavo, and *kobra* like a gismu.

For another example, *integral* has a specific meaning to a mathematician. But the Lojban fu'ivla *integrale*, which is a valid Stage 4 fu'ivla, does not convey that mathematical sense to a non-mathematical listener, even one with an English-speaking background; its source — the English word *integral* — has various other specialized meanings in other fields.

Left uncontrolled, *integrale* almost certainly would eventually come to mean the same collection of loosely related concepts that English associates with *integral*, with only the context to indicate (possibly) that the mathematical term is meant.

The prefix method would render the mathematical concept as *cmacrntegrale*, if the *i* of *integrale* is removed, or something like *cmacrnintegrale*, if a new consonant is added to the beginning; *cmac-* is the rafsi for *cmaci* ("mathematics"). The architectural sense of *integral* might be conveyed with *djinrnintegrale* or *tarmrnintegrale*, where *dinju* and *tarmi* mean "building" and "form" respectively.

Here are some fu'ivla representing cultures and related things, shown with more than one rafsi prefix:

**Example 4.7.8**
bang,r,blgaria
Bulgarian (in language)

**Example 4.7.9**
kuln,r,blgaria
Bulgarian (in culture)

**Example 4.7.10**
gugd,r,blgaria
Bulgaria (the country)

**Example 4.7.11**
bang,r,kore,a
Korean (the language)

**Example 4.7.12**
kuln,r,kore,a
Korean (the culture)

Note the commas in Example 4.7.11 and Example 4.7.12, used because *ea* is not a valid diphthong in Lojban. Arguably, some form of the native name *Chosen* should have been used instead of the internationally known *Korea*; this is a recurring problem in all borrowings. In general, it is better to use the native name unless using it will severely impede understanding: *Navajo* is far more widely known than *Dine'e*.

## 4.8 cmene

Lojbanized names, called *cmene*, are very much like their counterparts in other languages. They are labels applied to things (or people) to stand for them in descriptions or in direct address. They may convey meaning in themselves, but do not necessarily do so.

Because names are often highly personal and individual, Lojban attempts to allow native language names to be used with a minimum of modification. The requirement that the Lojban speech stream be unambiguously analyzable, however, means that most names must be modified somewhat when they are Lojbanized. Here are a few examples of English names and possible Lojban equivalents:

**Example 4.8.1**
djim.
Jim

**Example 4.8.2**
djein.
Jane

**Example 4.8.3**
.arnold.
Arnold

**Example 4.8.4**
pit.
Pete

**Example 4.8.5**
katrinas.
Katrina

**Example 4.8.6**
kat,r,in.
Catherine

(Note that syllabic *r* is skipped in determining the stressed syllable, so Example 4.8.6 is stressed on the *ka*.)

**Example 4.8.7**
katis.
Cathy

**Example 4.8.8**
keit.
Kate

Names may have almost any form, but always end in a consonant, and are followed by a pause. They are penultimately stressed, unless unusual stress is marked with capitalization. A name may have multiple parts, each ending with a consonant and pause, or the parts may be combined into a single word with no pause. For example,

**Example 4.8.9**
djan. djonz.

and

**Example 4.8.10**
djandjonz.

are both valid Lojbanizations of *John Jones*.

The final arbiter of the correct form of a name is the person doing the naming, although most cultures grant people the right to determine how they want their own name to be spelled and pronounced. The English name *Mary* can thus be Lojbanized as *meris.*, *maris.*, *meiris.*, *merix.*, or even *marys.*. The last alternative is not pronounced much like its English equivalent, but may be desirable to someone who values spelling over pronunciation. The final consonant need not be an *s*; there must, however, be some Lojban consonant at the end.

Names are not permitted to have the sequences *la*, *lai*, or *doi* embedded in them, unless the sequence is immediately preceded by a consonant. These minor restrictions are due to the fact that all Lojban cmene embedded in a speech stream will be preceded

by one of these words or by a pause. With one of these words embedded, the cmene might break up into valid Lojban words followed by a shorter cmene. However, break-up cannot happen after a consonant, because that would imply that the word before the *la*, or whatever, ended in a consonant without pause, which is impossible.

For example, the invalid name *laplas.* would look like the Lojban words *la plas.*, and *ilanas.* would be misunderstood as *.i la nas.*. However, *nederlants.* cannot be misheard as *neder lants.*, because *neder* with no following pause is not a possible Lojban word.

There are close alternatives to these forbidden sequences that can be used in Lojbanizing names, such as *ly*, *lei*, and *dai* or *do'i*, that do not cause these problems.

Lojban cmene are identifiable as word forms by the following characteristics:

1. They must end in one or more consonants. There are no rules about how many consonants may appear in a cluster in cmene, provided that each consonant pair (whether standing by itself, or as part of a larger cluster) is a permissible pair.

2. They may contain the letter y as a normal, non-hyphenating vowel. They are the only kind of Lojban word that may contain the two diphthongs *iy* and *uy*.

3. They are always followed in speech by a pause after the final consonant, written as ..

4. They may be stressed on any syllable; if this syllable is not the penultimate one, it must be capitalized when writing. Neither names nor words that begin sentences are capitalized in Lojban, so this is the only use of capital letters.

Names meeting these criteria may be invented, Lojbanized from names in other languages, or formed by appending a consonant onto a cmavo, a gismu, a fu'ivla or a lujvo. Some cmene built from Lojban words are:

**Example 4.8.11**
pav.
the One
from the cmavo *pa*, with rafsi *pav*, meaning *one*

**Example 4.8.12**
sol.
the Sun
from the gismu *solri*, meaning *solar*, or actually
*pertaining to the Sun*

**Example 4.8.13**
ralj.
Chief (as a title)
from the gismu *ralju*, meaning *principal*.

**Example 4.8.14**
nol.
Lord/Lady
from the gismu *nobli*, with rafsi *nol*, meaning *noble*.

To Lojbanize a name from the various natural languages, apply the following rules:

1. Eliminate double consonants and silent letters.
2. Add a final *s* or *n* (or some other consonant that sounds good) if the name ends in a vowel.
3. Convert all sounds to their closest Lojban equivalents.
4. If possible and acceptable, shift the stress to the penultimate (next-to-the-last) syllable. Use commas and capitalization in written Lojban when it is necessary to preserve non-standard syllabication or stress. Do not capitalize names otherwise.
5. If the name contains an impermissible consonant pair, insert a vowel between the consonants: *y* is recommended.
6. No cmene may have the syllables *la*, *lai*, or *doi* in them, unless immediately preceded by a consonant. If these combinations are present, they must be converted to something else. Possible substitutions include *ly*, *ly'i*, and *dai* or *do'i*, respectively.

There are some additional rules for Lojbanizing the scientific names (technically known as *Linnaean binomials* after their inventor) which are internationally applied to each species of animal or plant. Where precision is essential, these names need not be Lojbanized, but can be directly inserted into Lojban text using the cmavo *la'o*, explained in Chapter 19. Using this cmavo makes the already lengthy Latinized names at least four syllables longer, however, and leaves the pronunciation in doubt. The following suggestions, though incomplete, will assist in converting Linnaean binomals to valid Lojban names. They can also help to create fu'ivla based on Linnaean binomials or other words of the international scientific vocabulary. The term "back vowel" in the following list refers to any of the letters *a*, *o*, or *u*; the term *front vowel* correspondingly refers to any of the letters *e*, *i*, or *y*.

1. Change double consonants other than *cc* to single consonants.
2. Change *cc* before a front vowel to *kc*, but otherwise to *k*.
3. Change *c* before a back vowel and final *c* to *k*.
4. Change *ng* before a consonant (other than *h*) and final *ng* to *n*.
5. Change *x* to *z* initially, but otherwise to *ks*.
6. Change *pn* to *n* initially.
7. Change final *ie* and *ii* to *i*.
8. Make the following idiosyncratic substitutions:

| **Latin** | | **Lojban** |
|---|---|---|
| aa | ⇒ | a |
| ae | ⇒ | e |
| ch | ⇒ | k |
| ee | ⇒ | i |
| eigh | ⇒ | ei |

| | | |
|---|---|---|
| ew | ⇒ | u |
| igh | ⇒ | ai |
| oo | ⇒ | u |
| ou | ⇒ | u |
| ow | ⇒ | au |
| ph | ⇒ | f |
| q | ⇒ | k |
| sc | ⇒ | sk |
| w | ⇒ | u |
| y | ⇒ | i |

However, the diphthong substitutions should not be done if the two vowels are in two different syllables.

9. Change *h* between two vowels to ', but otherwise remove it completely. If preservation of the *h* seems essential, change it to *x* instead.

10. Place ' between any remaining vowel pairs that do not form Lojban diphthongs.

Some further examples of Lojbanized names are:

| **Langauge** | **Original** | **Lojbanized** |
|---|---|---|
| English | *Mary* | meris. or meiris. |
| English | *Smith* | smit. |
| English | *Jones* | djonz. |
| English | *John* | djan. or jan. (American)<br>or djon. or jon. (British) |
| English | *Alice* | .alis. |
| English | *Elise* | .eLIS. |
| English | *Johnson* | djansn. |
| English | *William* | .uiliam. or .uil,iam. |
| English | *Brown* | braun. |
| English | *Charles* | tcarlz. |
| French | *Charles* | carl. |
| French | *De Gaulle* | dyGOL. |
| German | *Heinrich* | xainrix. |
| Spanish | *Joaquin* | xuaKIN. |
| Russian | *Svetlana* | sfietlanys. |
| Russian | *Khrushchev* | xrucTCOF. |
| Hindi | *Krishna* | kricnas. |
| Polish | *Lech Walesa* | lex. va,uensas. |
| Spanish | *Don Quixote* | don. kicotes.<br>or don. kixotes. (Modern Spanish)<br>or don. ki'otes. (Mexican dialect) |
| Chinese | *Mao Zedong* | maudzydyn. |
| Japanese | *Fujiko* | fudjikos. or fujikos. |

## 4.9 Rules for inserting pauses

Summarized in one place, here are the rules for inserting pauses between Lojban words:

1. Any two words may have a pause between them; it is always illegal to pause in the middle of a word, because that breaks up the word into two words.

2. Every word ending in a consonant must be followed by a pause. Necessarily, all such words are cmene.

3. Every word beginning with a vowel must be preceded by a pause. Such words are either cmavo, fu'ivla, or cmene; all gismu and lujvo begin with consonants.

4. Every cmene must be preceded by a pause, unless the immediately preceding word is one of the cmavo *la*, *lai*, *la'i*, or *doi* (which is why those strings are forbidden in cmene). However, the situation triggering this rule rarely occurs.

5. If the last syllable of a word bears the stress, and a brivla follows, the two must be separated by a pause, to prevent confusion with the primary stress of the brivla. In this case, the first word must be either a cmavo or a cmene with unusual stress (which already ends with a pause, of course).

6. A cmavo of the form Cy must be followed by a pause unless another Cy-form cmavo follows.

7. When non-Lojban text is embedded in Lojban, it must be preceded and followed by pauses. (How to embed non-Lojban text is explained in Chapter 19.)

## 4.10 Considerations for making lujvo

Given a tanru which expresses an idea to be used frequently, it can be turned into a lujvo by following the lujvo-making algorithm which is given in Section 4.11.

In building a lujvo, the first step is to replace each gismu with a rafsi that uniquely represents that gismu. These rafsi are then attached together by fixed rules that allow the resulting compound to be recognized as a single word and to be analyzed in only one way.

There are three other complications; only one is serious.

The first is that there is usually more than one rafsi that can be used for each gismu. The one to be used is simply whichever one sounds or looks best to the speaker or writer. There are usually many valid combinations of possible rafsi. They all are equally valid, and all of them mean exactly the same thing. (The scoring algorithm given in Section 4.12 is used to choose the standard form of the lujvo — the version which would be entered into a dictionary.)

The second complication is the serious one. Remember that a tanru is ambiguous — it has several possible meanings. A lujvo, or at least one that would be put into the dictionary, has just a single meaning. Like a gismu, a lujvo is a predicate which encompasses one area of the semantic universe, with one set of places. Hopefully the meaning chosen is the most useful of the possible semantic spaces. A possible source of linguistic drift in Lojban is that as Lojbanic society evolves, the concept that seems the most useful one may change.

You must also be aware of the possibility of some prior meaning of a new lujvo, especially if you are writing for posterity. If a lujvo is invented which involves the same tanru as one that is in the dictionary, and is assigned a different meaning (or even just a different place structure), linguistic drift results. This isn't necessarily bad. Every natural language does it. But in communication, when you use a meaning different from the dictionary definition, someone else may use the dictionary and therefore misunderstand you. You can use the cmavo *za'e* (explained in Chapter 19) before a newly coined lujvo to indicate that it may have a non-dictionary meaning.

The essential nature of human communication is that if the listener understands, then all is well. Let this be the ultimate guideline for choosing meanings and place structures for invented lujvo.

The third complication is also simple, but tends to scare new Lojbanists with its implications. It is based on Zipf's Law, which says that the length of words is inversely proportional to their usage. The shortest words are those which are used more; the longest ones are used less. Conversely, commonly used concepts will be tend to be abbreviated. In English, we have abbreviations and acronyms and jargon, all of which represent complex ideas that are used often by small groups of people, so they shortened them to convey more information more rapidly.

Therefore, given a complicated tanru with grouping markers, abstraction markers, and other cmavo in it to make it syntactically unambiguous, the psychological basis of Zipf's Law may compel the lujvo-maker to drop some of the cmavo to make a shorter (technically incorrect) tanru, and then use that tanru to make the lujvo.

This doesn't lead to ambiguity, as it might seem to. A given lujvo still has exactly one meaning and place structure. It is just that more than one tanru is competing for the same lujvo. But more than one meaning for the tanru was already competing for the *right* to define the meaning of the lujvo. Someone has to use judgment in deciding which one meaning is to be chosen over the others.

If the lujvo made by a shorter form of tanru is in use, or is likely to be useful for another meaning, the decider then retains one or more of the cmavo, preferably ones that set this meaning apart from the shorter form meaning that is used or anticipated. As a rule, therefore, the shorter lujvo will be used for a more general concept, possibly even instead of a more frequent word. If both words are needed, the simpler one should be shorter. It is easier to add a cmavo to clarify the meaning of the more complex term than it is to find a good alternate tanru for the simpler term.

And of course, we have to consider the listener. On hearing an unknown word, the listener will decompose it and get a tanru that makes no sense or the wrong sense for the context. If the listener realizes that the grouping operators may have been dropped out, he or she may try alternate groupings, or try inserting an abstraction operator if that seems plausible. (The grouping of tanru is explained in Chapter 5; abstraction is explained in Chapter 11.) Plausibility is the key to learning new ideas and to evaluating unfamiliar lujvo.

## 4.11 The lujvo-making algorithm

The following is the current algorithm for generating Lojban lujvo given a known tanru and a complete list of gismu and their assigned rafsi. The algorithm was designed by Bob LeChevalier and Dr. James Cooke Brown for computer program implementation.

It was modified in 1989 with the assistance of Nora LeChevalier, who detected a flaw in the original *tosmabru test*.

Given a tanru that is to be made into a lujvo:

1. Choose a 3-letter or 4-letter rafsi for each of the gismu and cmavo in the tanru except the last.

2. Choose a 3-letter (CVV-form or CCV-form) or 5-letter rafsi for the final gismu in the tanru.

3. Join the resulting string of rafsi, initially without hyphens.

4. Add hyphen letters where necessary. It is illegal to add a hyphen at a place that is not required by this algorithm. Right-to-left tests are recommended, for reasons discussed below.

   (a) If there are more than two words in the tanru, put an *r*-hyphen (or an *n*-hyphen) after the first rafsi if it is CVV-form. If there are exactly two words, then put an *r*-hyphen (or an *n*-hyphen) between the two rafsi if the first rafsi is CVV-form, unless the second rafsi is CCV-form (for example, *saicli* requires no hyphen). Use an *r*-hyphen unless the letter after the hyphen is *r*, in which case use an *n*-hyphen. Never use an *n*-hyphen unless it is required.

   (b) Put a *y*-hyphen between the consonants of any impermissible consonant pair. This will always appear between rafsi.

   (c) Put a *y*-hyphen after any 4-letter rafsi form.

5. Test all forms with one or more initial CVC-form rafsi — with the pattern CVC ... CVC + X — for *tosmabru failure*. X must either be a CVCCV long rafsi that happens to have a permissible initial pair as the consonant cluster, or is something which has caused a *y*-hyphen to be installed between the previous CVC and itself by one of the above rules. The test is as follows:

   (a) Examine all the C/C consonant pairs that join the CVC rafsi, and also the pair between the last CVC and the X portion, ignoring any *y*-hyphen before the X. These consonant pairs are called *joints*.

   (b) If all of those joints are permissible initials, then the trial word will break up into a cmavo and a shorter brivla. If not, the word will not break up, and no further hyphens are needed.

   (c) Install a *y*-hyphen at the first such joint.

Note that the *tosmabru test* implies that the algorithm will be more efficient if rafsi junctures are tested for required hyphens from right to left, instead of from left to right; when the test is required, it cannot be completed until hyphenation to the right has been determined.

## 4.12 The lujvo scoring algorithm

This algorithm was devised by Bob and Nora LeChevalier in 1989. It is not the only possible algorithm, but it usually gives a choice that people find preferable. The algorithm may be changed in the future. The lowest-scoring variant will usually be the dictionary form of the lujvo. (In previous versions, it was the highest-scoring variant.)

1. Count the total number of letters, including hyphens and apostrophes; call it $L$.
2. Count the number of apostrophes; call it $A$.
3. Count the number of *y*-, *r*-, and *n*-hyphens; call it $H$.
4. For each rafsi, find the value in the following table. Sum this value over all rafsi; call it $R$:

| | | |
|---|---|---|
| CVC/CV | (final) (-sarji) | 1 |
| CVC/C | (-sarj-) | 2 |
| CCVCV | (final) (-zbasu) | 3 |
| CCVC | (-zbas-) | 4 |
| CVC | (-nun-) | 5 |
| CVV | with an apostrophe (-ta'u-) | 6 |
| CCV | (-zba-) | 7 |
| CVV | with no apostrophe (-sai-) | 8 |

5. Count the number of vowels, not including *y*; call it $V$.

The score is then: $(1000 \times L) - (500 \times A) + (100 \times H) - (10 \times R) - V$

In case of ties, there is no preference. This should be rare. Note that the algorithm essentially encodes a hierarchy of priorities: short words are preferred (counting apostrophes as half a letter), then words with fewer hyphens, words with more pleasing rafsi (this judgment is subjective), and finally words with more vowels are chosen. Each decision principle is applied in turn if the ones before it have failed to choose; it is possible that a lower-ranked principle might dominate a higher-ranked one if it is ten times better than the alternative.

Here are some lujvo with their scores (not necessarily the lowest scoring forms for these lujvo, nor even necessarily sensible lujvo):

**Example 4.12.1**
zbasai
zba + sai
$(1000 \times 6) - (500 \times 0) + (100 \times 0) - (10 \times 15) - 3 = 5847$

**Example 4.12.2**
nunynau
nun + y + nau
$32500 - (1000 \times 7) + (500 \times 0) - (100 \times 1) + (10 \times 13) + 3 = 6967$

**Example 4.12.3**
sairzbata'u
sai + r + zba + ta'u

$$32500 - (1000 \times 11) + (500 \times 1) - (100 \times 1) + (10 \times 21) + 5 = 10385$$

**Example 4.12.4**
zbazbasysarji
zba + zbas + y + sarji
$32500 - (1000 \times 13) + (500 \times 0) - (100 \times 1) + (10 \times 12) + 4 = 12976$

## 4.13 lujvo-making examples

This section contains examples of making and scoring lujvo. First, we will start with the tanru *gerku zdani* ("dog house") and construct a lujvo meaning *doghouse*, that is, a house where a dog lives. We will use a brute-force application of the algorithm in Section 4.12, using every possible rafsi.

The rafsi for *gerku* are: *-ger-*, *-ge'u-*, *-gerk-*, and *-gerku*; the rafsi for *zdani* are: *-zda-*, *-zdan-*, and *-zdani*.

Step 1 of the algorithm directs us to use *-ger-*, *-ge'u-*, and *-gerk-* as possible rafsi for *gerku*; Step 2 directs us to use *-zda-* and *-zdani* as possible rafsi for *zdani*. The six possible forms of the lujvo are then: *ger-zda*, *ger-zdani*, *ge'u-zda*, *ge'u-zdani*, *gerk-zda*, and *gerk-zdani*.

We must then insert appropriate hyphens in each case. The first two forms need no hyphenation: *ge* cannot fall off the front, because the following word would begin with *rz*, which is not a permissible initial consonant pair. So the lujvo forms are *gerzda* and *gerzdani*.

The third form, *ge'u-zda*, needs no hyphen, because even though the first rafsi is CVV, the second one is CCV, so there is a consonant cluster in the first five letters. So *ge'uzda* is this form of the lujvo.

The fourth form, *ge'u-zdani*, however, requires an *r*-hyphen; otherwise, the *ge'u-* part would fall off as a cmavo. So this form of the lujvo is *ge'urzdani*.

The last two forms require *y*-hyphens, as all 4-letter rafsi do, and so are *gerkyzda* and *gerkyzdani* respectively.

The scoring algorithm is heavily weighted in favor of short lujvo, so we might expect that *gerzda* would win. Its *L* score is 6, its *A* score is 0, its *H* score is 0, its *R* score is 12, and its *V* score is 3, for a final score of 5878. The other forms have scores of 7917, 6367, 9506, 8008, and 10047 respectively. Consequently, this lujvo would probably appear in the dictionary in the form *gerzda*.

For the next example, we will use the tanru *bloti klesi* ("boat class") presumably referring to the category (rowboat, motorboat, cruise liner) into which a boat falls. We will omit the long rafsi from the process, since lujvo containing long rafsi are almost never preferred by the scoring algorithm when there are short rafsi available.

The rafsi for *bloti* are *-lot-*, *-blo-*, and *-lo'i-*; for *klesi* they are *-kle-* and *-lei-*. Both these gismu are among the handful which have both CVV-form and CCV-form rafsi, so there is an unusual number of possibilities available for a two-part tanru:

| | | |
|---|---|---|
| lotkle | blokle | lo'ikle |
| lotlei | blolei | lo'irlei |

Only *lo'irlei* requires hyphenation (to avoid confusion with the cmavo sequence *lo'i lei*). All six forms are valid versions of the lujvo, as are the six further forms using

long rafsi; however, the scoring algorithm produces the following results:

| **lujvo** | **Score** |
|---|---|
| lotkle | 5878 |
| blokle | 5858 |
| lo'ikle | 6367 |
| lotlei | 5867 |
| blolei | 5847 |
| lo'irlei | 7456 |

So the form *blolei* is preferred, but only by a tiny margin over *blokle*; the next two forms are only slightly worse; *lo'ikle* suffers because of its apostrophe, and *lo'irlei* because of having both apostrophe and hyphen.

Our third example will result in forming both a lujvo and a name from the tanru *logji bangu girzu*, or *logical-language group* in English. ("The Logical Language Group" is the name of the publisher of this book and the organization for the promotion of Lojban.) The available rafsi are *-loj-* and *-logj-*; *-ban-*, *-bau-*, and *-bang-*; and *-gri-* and *-girzu*, and (for name purposes only) *-gir-* and *-girz-*. The resulting 12 lujvo possibilities are:

| | | |
|---|---|---|
| loj-ban-gri | loj-bau-gri | loj-bang-gri |
| logj-ban-gri | logj-bau-gri | logj-bang-gri |
| loj-ban-girzu | loj-bau-girzu | loj-bang-girzu |
| logj-ban-girzu | logj-bau-girzu | logj-bang-girzu |

and the 12 name possibilities are:

| | | |
|---|---|---|
| loj-ban-gir. | loj-bau-gir. | loj-bang-gir. |
| logj-ban-gir. | logj-bau-gir. | logj-bang-gir. |
| loj-ban-girz. | loj-bau-girz. | loj-bang-girz. |
| logj-ban-girz. | logj-bau-girz. | logj-bang-girz. |

After hyphenation, we have:

| | | |
|---|---|---|
| lojbangri | lojbaugri | lojbangygri |
| logjybangri | logjybaugri | logjybangygri |
| lojbangirzu | lojbaugirzu | lojbangygirzu |
| logjybangirzu | logjybaugirzu | logjybangygirzu |

| | | |
|---|---|---|
| lojbangir. | lojbaugir. | lojbangygir. |
| logjybangir. | logjybaugir. | logjybangygir. |
| lojbangirz. | lojbaugirz. | lojbangygirz. |
| logjybangirz. | logjybaugirz. | logjyban gygirz. |

The only fully reduced lujvo forms are *lojbangri* and *lojbaugri*, of which the latter has a slightly lower score: 8827 versus 8796, respectively. However, for the name of the organization, we chose to make sure the name of the language was embedded in it, and to use the clearer long-form rafsi for *girzu*, producing *lojbangirz.*

Finally, here is a four-part lujvo with a cmavo in it, based on the tanru *nakni ke cinse ctuca* or "male (sexual teacher)". The *ke* cmavo ensures the interpretation "teacher of sexuality who is male", rather than "teacher of male sexuality". Here are the possible forms of the lujvo, both before and after hyphenation:

| | |
|---|---|
| nak-kem-cin-ctu | nakykemcinctu |
| nak-kem-cin-ctuca | nakykemcinctuca |
| nak-kem-cins-ctu | nakykemcinsyctu |
| nak-kem-cins-ctuca | nakykemcinsyctuca |
| nakn-kem-cin-ctu | naknykemcinctu |
| nakn-kem-cin-ctuca | naknykemcinctuca |
| nakn-kem-cins-ctu | naknykemcinsyctu |
| nakn-kem-cins-ctuca | naknykemcinsyctuca |

Of these forms, *nakykemcinctu* is the shortest and is preferred by the scoring algorithm. On the whole, however, it might be better to just make a lujvo for *cinse ctuca* (which would be *cinctu*) since the sex of the teacher is rarely important. If there was a reason to specify *male*, then the simpler tanru *nakni cinctu* ("male sexual-teacher") would be appropriate. This tanru is actually shorter than the four-part lujvo, since the *ke* required for grouping need not be expressed.

## 4.14 The gismu creation algorithm

The gismu were created through the following process:

1. At least one word was found in each of the six source languages (Chinese, English, Hindi, Spanish, Russian, Arabic) corresponding to the proposed gismu. This word was rendered into Lojban phonetics rather liberally: consonant clusters consisting of a stop and the corresponding fricative were simplified to just the fricative (*tc* became *c*, *dj* became *j*) and non-Lojban vowels were mapped onto Lojban ones. Furthermore, morphological endings were dropped. The same mapping rules were applied to all six languages for the sake of consistency.

2. All possible gismu forms were matched against the six source-language forms. The matches were scored as follows:

   (a) If three or more letters were the same in the proposed gismu and the source-language word, and appeared in the same order, the score was equal to the number of letters that were the same. Intervening letters, if any, did not matter.

   (b) If exactly two letters were the same in the proposed gismu and the source-language word, and either the two letters were consecutive in both words, or were separated by a single letter in both words, the score was 2. Letters in reversed order got no score.

   (c) Otherwise, the score was 0.

3. The scores were divided by the length of the source-language word in its Lojbanized form, and then multiplied by a weighting value specific to each language, reflecting the proportional number of first-language and second-language speakers

of the language. (Second-language speakers were reckoned at half their actual numbers.) The weights were chosen to sum to 1.00. The sum of the weighted scores was the total score for the proposed gismu form.

4. Any gismu forms that conflicted with existing gismu were removed. Obviously, being identical with an existing gismu constitutes a conflict. In addition, a proposed gismu that was identical to an existing gismu except for the final vowel was considered a conflict, since two such gismu would have identical 4-letter rafsi.

   More subtly: If the proposed gismu was identical to an existing gismu except for a single consonant, and the consonant was "too similar" based on the following table, then the proposed gismu was rejected. (See Section 4.4 for an example.)

| **Proposed gismu** | **Existing gismu** |
|---|---|
| b | p |
| v | c j |
| s | d t f p |
| v | g k |
| x | j c |
| z | k g |
| x | l r m n n m p b |
| f | r l s c |
| z | t d v b |
| f | x g |
| k | z j |
| s | |

5. The gismu form with the highest score usually became the actual gismu. Sometimes a lower-scoring form was used to provide a better rafsi. A few gismu were changed in error as a result of transcription blunders (for example, the gismu *gismu* should have been *gicmu*, but it's too late to fix it now).

The language weights used to make most of the gismu reflected 1985 number-of-speakers data. A few gismu were made much later using updated weights.

| **Language** | **Weight** | **Language** | **Weight** |
|---|---|---|---|
| 1985 | | Later | |
| Chinese | 0.36 | Chinese | 0.347 |
| English | 0.21 | Hindi | 0.196 |
| Hindi | 0.16 | English | 0.160 |
| Spanish | 0.11 | Spanish | 0.123 |
| Russian | 0.09 | Russian | 0.089 |
| Arabic | 0.07 | Arabic | 0.085 |

Note that the stressed vowel of the gismu was considered sufficiently distinctive that two or more gismu may differ only in this vowel; as an extreme example, *bradi*,

*bredi*, *bridi*, and *brodi* (but fortunately not *brudi*) are all existing gismu.

## 4.15 Cultural and other non-algorithmic gismu

The following gismu were not made by the gismu creation algorithm. They are, in effect, coined words similar to fu'ivla. They are exceptions to the otherwise mandatory gismu creation algorithm where there was sufficient justification for such exceptions. Except for the small metric prefixes and the assignable predicates beginning with *brod-*, they all end in the letter *o*, which is otherwise a rare letter in Lojban gismu.

The following gismu represent concepts that are sufficiently unique to Lojban that they were either coined from combining forms of other gismu, or else made up out of whole cloth. These gismu are thus conceptually similar to lujvo even though they are only five letters long; however, unlike lujvo, they have rafsi assigned to them for use in building more complex lujvo. Assigning gismu to these concepts helps to keep the resulting lujvo reasonably short.

**Definition 4.1**

**Non-algorithmic gismu**:

| | |
|---|---|
| **broda** | 1st assignable predicate |
| **brode** | 2nd assignable predicate |
| **brodi** | 3rd assignable predicate |
| **brodo** | 4th assignable predicate |
| **brodu** | 5th assignable predicate |
| **cmavo** | structure word (from *cmalu valsi*) |
| **lojbo** | Lojbanic (from *logji bangu*) |
| **lujvo** | compound word (from *pluja valsi*) |
| **mekso** | Mathematical EXpression |

It is important to understand that even though *cmavo*, *lojbo*, and *lujvo* were made up from parts of other gismu, they are now full-fledged gismu used in exactly the same way as all other gismu, both in grammar and in word formation.

The following three groups of gismu represent concepts drawn from the international language of science and mathematics. They are used for concepts that are represented in most languages by a root which is recognized internationally.

| gismu | Meaning | gismu | Meaning |
|---|---|---|---|
| decti | 0.1 (*deci*) | dekto | 10 (*deka*) |
| centi | 0.01 (*centi*) | xecto | 100 (*hecto*) |
| milti | 0.00 (*milli*) | kilto | 1000 (*kilo*) |
| mikri | $1 \times 10^{-6}$ (*micro*) | megdo | $1 \times 10^{6}$ (*mega*) |
| nanvi | $1 \times 10^{-9}$ (*nano*) | gigdo | $1 \times 10^{9}$ (*giga*) |
| picti | $1 \times 10^{-12}$ (*pico*) | terto | $1 \times 10^{12}$ (*tera*) |
| femti | $1 \times 10^{-15}$ (*femto*) | petso | $1 \times 10^{15}$ (*peta*) |
| xatsi | $1 \times 10^{-18}$ (*atto*) | xexso | $1 \times 10^{18}$ (*exa*) |
| zepti | $1 \times 10^{-21}$ (*zepto*) | zetro | $1 \times 10^{21}$ (*zetta*) |
| gocti | $1 \times 10^{-24}$ (*yocto*) | gotro | $1 \times 10^{24}$ (*yotta*) |

**Definition 4.2**
**Other scientific or mathematical terms**:

| | |
|---|---|
| **delno** | candela |
| **kelvo** | kelvin |
| **molro** | mole |
| **radno** | radian |
| **sinso** | sine |
| **stero** | steradian |
| **tanjo** | tangent |
| **xampo** | ampere |

The gismu *sinso* and *tanjo* were only made non-algorithmically because they were identical (having been borrowed from a common source) in all the dictionaries that had translations. The other terms in this group are units in the international metric system; some metric units, however, were made by the ordinary process (usually because they are different in Chinese).

Finally, there are the cultural gismu, which are also borrowed, but by modifying a word from one particular language, instead of using the multi-lingual gismu creation algorithm. Cultural gismu are used for words that have local importance to a particular culture; other cultures or languages may have no word for the concept at all, or may borrow the word from its home culture, just as Lojban does. In such a case, the gismu algorithm, which uses weighted averages, doesn't accurately represent the frequency of usage of the individual concept. Cultural gismu are not even required to be based on the six major languages.

**Definition 4.3**
**The six Lojban source languages**:

| | |
|---|---|
| **jungo** | Chinese (from *Zhong*$^1$ *guo*$^2$) |
| **glico** | English |
| **xindo** | Hindi |
| **spano** | Spanish |
| **rusko** | Russian |
| **xrabo** | Arabic |

Seven other widely spoken languages that were on the list of candidates for gismu-making, but weren't used:

**Definition 4.4**

| | |
|---|---|
| **bengo** | Bengali |
| **porto** | Portuguese |
| **baxso** | Bahasa Melayu/Bahasa Indonesia |
| **ponjo** | Japanese (from *Nippon*) |
| **dotco** | German (from *Deutsch*) |
| **fraso** | French (from *Françis*) |
| **xurdo** | Urdu |

(Urdu and Hindi began as the same language with different writing systems, but have now become somewhat different principally in borrowed vocabulary. Urdu-speakers

were counted along with Hindi-speakers when weights were assigned for gismu-making purposes.)

Countries with a large number of speakers of any of the above languages (where the meaning of *large* is dependent on the specific language):

**Definition 4.5**
**English**:

| | |
|---|---|
| **merko** | American |
| **brito** | British |
| **skoto** | Scottish |
| **sralo** | Australian |
| **kadno** | Canadian |

**Definition 4.6**
**Spanish**:

| | |
|---|---|
| **gento** | Argentinian |
| **mexno** | Mexican |

**Definition 4.7**
**Russian**:

| | |
|---|---|
| **softo** | Soviet/USSR |
| **vukro** | Ukrainian |

**Definition 4.8**
**Arabic**:

| | |
|---|---|
| **filso** | Palestinian |
| **jerxo** | Algerian |
| **jordo** | Jordanian |
| **libjo** | Libyan |
| **lubno** | Lebanese |
| **misro** | Egyptian (from *Mizraim* |
| **morko** | Moroccan |
| **rakso** | Iraqi |
| **sadjo** | Saudi |
| **sirxo** | Syrian |

**Definition 4.9**
**Bahasa Melayu/Bahasa Indonesia**:

| | |
|---|---|
| **bindo** | Indonesian |
| **meljo** | Malaysian |

**Definition 4.10**
**Portuguese**:

| | |
|---|---|
| **brazo** | Brazilian |

**Definition 4.11**
**Urdu**:

| | |
|---|---|
| **kisto** | Pakistani |

**Definition 4.12**
**The continents (and oceanic regions) of the Earth**:

| | |
|---|---|
| **bemro** | North American (from *berti merko*) |
| **dzipo** | Antarctican (from *cadzu cipni*) |
| **ketco** | South American (from *Quechua*) |
| **friko** | African |
| **polno** | Polynesian/Oceanic |
| **ropno** | European |
| **xazdo** | Asiatic |

**Definition 4.13**
**A few smaller but historically important cultures**:

| | |
|---|---|
| **latmo** | Latin/Roman |
| **srito** | Sanskrit |
| **xebro** | Hebrew/Israeli |
| **xelso** | Greek (from *Hellas*) |

**Definition 4.14**
**Major world religions**:

| | |
|---|---|
| **budjo** | Buddhist |
| **dadjo** | Taoist |
| **muslo** | Islamic/Moslem |
| **xriso** | Christian |

**Definition 4.15**
**A few terms that cover multiple groups of the above**:

| | |
|---|---|
| **jegvo** | Jehovist (Judeo-Christian-Moslem) |
| **semto** | Semitic |
| **slovo** | Slavic |
| **xispo** | Hispanic (New World Spanish) |

## 4.16 rafsi fu'ivla: a proposal

The list of cultures represented by gismu, given in Section 4.15, is unavoidably controversial. Much time has been spent debating whether this or that culture "deserves a gismu" or "must languish in fu'ivla space". To help defuse this argument, a last-minute proposal was made when this book was already substantially complete. I have added it here with experimental status: it is not yet a standard part of Lojban, since all its implications have not been tested in open debate, and it affects a part of the language (lujvo-making) that has long been stable, but is known to be fragile in the face of small changes. (Many attempts were made to add general mechanisms for making lujvo that contained fu'ivla, but all failed on obvious or obscure counterexamples; finally the general *zei* mechanism was devised instead.)

The first part of the proposal is uncontroversial and involves no change to the language mechanisms. All valid Type 4 fu'ivla of the form CCVVCV would be reserved for cultural brivla analogous to those described in Section 4.15. For example,

**Example 4.16.1**
tci'ile
Chilean

is of the appropriate form, and passes all tests required of a Stage 4 fu'ivla. No two fu'ivla of this form would be allowed to coexist if they differed only in the final vowel; this rule was applied to gismu, but does not apply to other fu'ivla or to lujvo.

The second, and fully experimental, part of the proposal is to allow rafsi to be formed from these cultural fu'ivla by removing the final vowel and treating the result as a 4-letter rafsi (although it would contain five letters, not four). These rafsi could then be used on a par with all other rafsi in forming lujvo. The tanru

**Example 4.16.2**
tci'ile ke canre tutra
Chilean type-of (sand territory)
Chilean desert

could be represented by the lujvo

**Example 4.16.3**
tci'ilykemcantutra

which is an illegal word in standard Lojban, but a valid lujvo under this proposal. There would be no short rafsi or 5-letter rafsi assigned to any fu'ivla, so no fu'ivla could appear as the last element of a lujvo.

The cultural fu'ivla introduced under this proposal are called *rafsi fu'ivla*, since they are distinguished from other Type 4 fu'ivla by the property of having rafsi. If this proposal is workable and introduces no problems into Lojban morphology, it might become standard for all Type 4 fu'ivla, including those made for plants, animals, food-stuffs, and other things.

# Chapter 5

# *Pretty Little Girls' School*: The Structure of Lojban selbri

## 5.1 Lojban content words: brivla

At the center, logically and often physically, of every Lojban bridi is one or more words which constitute the selbri. A bridi expresses a relationship between things: the selbri specifies which relationship is referred to. The difference between:

**Example 5.1.1**
do mamta mi
You are-a-mother-of me.
You are my mother.

and

**Example 5.1.2**
do patfu mi
You are-a-father-of me.
You are my father.

lies in the different selbri.

The simplest kind of selbri is a single Lojban content word: a brivla. There are three different varieties of brivla: those which are built into the language (the gismu), those which are derived from combinations of the gismu (the lujvo), and those which are taken (usually in a modified form) from other languages (the fu'ivla). In addition, there are a few cmavo that can act like brivla; these are mentioned in Section 5.9, and discussed in full in Chapter 7.

For the purposes of this chapter, however, all brivla are alike. For example,

**Example 5.1.3**
ta bloti
That is-a-boat.
That is a boat.

**Example 5.1.4**
ta brablo
that is-a-large-boat.
That is a ship.

**Example 5.1.5**
ta blotrskunri
That is-a-(boat)-schooner.
That is a schooner.

illustrate the three types of brivla (gismu, lujvo, and fu'ivla respectively), but in each case the selbri is composed of a single word whose meaning can be learned independent of its origins.

The remainder of this chapter will mostly use gismu as example brivla, because they are short. However, it is important to keep in mind that wherever a gismu appears, it could be replaced by any other kind of brivla.

## 5.2 Simple tanru

Beyond the single brivla, a selbri may consist of two brivla placed together. When a selbri is built in this way from more than one brivla, it is called a tanru, a word with no single English equivalent. The nearest analogue to tanru in English are combinations of two nouns such as *lemon tree*. There is no way to tell just by looking at the phrase *lemon tree* exactly what it refers to, even if you know the meanings of *lemon* and *tree* by themselves. As English-speakers, we must simply know that it refers to "a tree which bears lemons as fruits". A person who didn't know English very well might think of it as analogous to *brown tree* and wonder, "What kind of tree is lemon-colored?"

In Lojban, tanru are also used for the same purposes as English adjective- noun combinations like *big boy* and adverb-verb combinations like *quickly run*. This is a consequence of Lojban not having any such categories as *noun*, *verb*, *adjective*, or *adverb*. English words belonging to any of these categories are translated by simple brivla in Lojban. Here are some examples of tanru:

**Example 5.2.1**
tu pelnimre tricu
that-yonder is-a-(lemon tree).
That is a lemon tree.

**Example 5.2.2**
la djan. barda nanla
John is-a-big boy.
John is a big boy.

**Example 5.2.3**
mi sutra bajra
I quick run.
I quickly run/I run quickly.

Note that *pelnimre* is a lujvo for *lemon*; it is derived from the gismu *pelxu*, yellow, and *nimre*, citrus. Note also that *sutra* can mean *fast/quick* or *quickly* depending on its use:

**Example 5.2.4**
mi sutra
I am-fast/quick.

shows *sutra* used to translate an adjective, whereas in Example 5.2.3 it is translating an adverb. (Another correct translation of Example 5.2.3, however, would be *I am a quick runner*.)

There are special Lojban terms for the two components of a tanru, derived from the place structure of the word *tanru*. The first component is called the *seltau*, and the second component is called the *tertau*.

The most important rule for use in interpreting tanru is that the tertau carries the primary meaning. A "pelnimre tricu" is primarily a tree, and only secondarily is it connected with lemons in some way. For this reason, an alternative translation of Example 5.2.1 would be:

**Example 5.2.5**
That is a lemon type of tree.

This *type of* relationship between the components of a tanru is fundamental to the tanru concept.

We may also say that the seltau modifies the meaning of the tertau:

**Example 5.2.6**
That is a tree which is lemon-ish
(in the way appropriate to trees)

would be another possible translation of Example 5.2.1. In the same way, a more explicit translation of Example 5.2.2 might be:

**Example 5.2.7**
John is a boy who is big in the way that boys are big.

This *way that boys are big* would be quite different from the way in which elephants are big; big-for-a-boy is small-for-an-elephant.

All tanru are ambiguous semantically. Possible translations of:

**Example 5.2.8**
ta klama jubme
That is-a-goer type-of-table.

include:

**Definition 5.1**
That is a table which goes (a wheeled table, perhaps). That is a table owned by one who goes. That is a table used by those who go (a sports doctor's table?). That is a table when it goes (otherwise it is a chair?).

In each case the object referred to is a "goer type of table", but the ambiguous *type of* relationship can mean one of many things. A speaker who uses tanru (and pragmatically all speakers must) takes the risk of being misunderstood. Using tanru

is convenient because they are short and expressive; the circumlocution required to squeeze out all ambiguity can require too much effort.

No general theory covering the meaning of all possible tanru exists; probably no such theory can exist. However, some regularities obviously do exist:

**Example 5.2.9**
do barda prenu
You are-a-large person.

**Example 5.2.10**
do cmalu prenu
You are-a-small person.

are parallel tanru, in the sense that the relationship between *barda* and *prenu* is the same as that between *cmalu* and *prenu*. Section 5.14 and Section 5.15 contain a partial listing of some types of tanru, with examples.

## 5.3 Three-part tanru grouping with *bo*

The following cmavo is discussed in this section:

`bo  BO  closest scope grouping`

Consider the English sentence:

**Example 5.3.1**
That's a little girls' school.

What does it mean? Two possible readings are:

**Example 5.3.2**
That's a little school for girls.

**Example 5.3.3**
That's a school for little girls.

This ambiguity is quite different from the simple tanru ambiguity described in Section 5.2. We understand that *girls' school* means "a school where girls are the students", and not "a school where girls are the teachers" or "a school which is a girl" (!). Likewise, we understand that *little girl* means "girl who is small". This is an ambiguity of grouping. Is *girls' school* to be taken as a unit, with *little* specifying the type of girls' school? Or is *little girl* to be taken as a unit, specifying the type of school? In English speech, different tones of voice, or exaggerated speech rhythm showing the grouping, are used to make the distinction; English writing usually leaves it unrepresented.

Lojban makes no use of tones of voice for any purpose; explicit words are used to do the work. The cmavo *bo* (which belongs to selma'o BO) may be placed between the two brivla which are most closely associated. Therefore, a Lojban translation of Example 5.3.2 would be:

**Example 5.3.4**
ta cmalu nixli bo ckule
That is-a-small girl – school.

Example 5.3.3 might be translated:

**Example 5.3.5**
ta cmalu bo nixli ckule
That is-a-small – girl school.

The *bo* is represented in the literal translation by a hyphen because in written English a hyphen is sometimes used for the same purpose: *a big dog-catcher* would be quite different from a *big-dog catcher* (presumably someone who catches only big dogs).

Analysis of Example 5.3.4 and Example 5.3.5 reveals a tanru nested within a tanru. In Example 5.3.4, the main tanru has a seltau of *cmalu* and a tertau of *nixli bo ckule*; the tertau is itself a tanru with *nixli* as the seltau and *ckule* as the tertau. In Example 5.3.5, on the other hand, the seltau is *cmalu bo nixli* (itself a tanru), whereas the tertau is *ckule*. This structure of tanru nested within tanru forms the basis for all the more complex types of selbri that will be explained below.

What about Example 5.3.6? What does it mean?

**Example 5.3.6**
ta cmalu nixli ckule
That is-a-small girl school.

The rules of Lojban do not leave this sentence ambiguous, as the rules of English do with Example 5.3.1. The choice made by the language designers is to say that Example 5.3.6 means the same as Example 5.3.5. This is true no matter what three brivla are used: the leftmost two are always grouped together. This rule is called the *left-grouping rule*. Left-grouping in seemingly ambiguous structures is quite common — though not universal — in other contexts in Lojban.

Another way to express the English meaning of Example 5.3.4 and Example 5.3.5, using parentheses to mark grouping, is:

**Example 5.3.7**
ta cmalu nixli bo ckule
That is-a-small type-of (girl type-of school).

**Example 5.3.8**
ta cmalu bo nixli ckule
That is-a-(small type-of girl) type-of school.

Because *type-of* is implicit in the Lojban tanru form, it has no Lojban equivalent.

Note: It is perfectly legal, though pointless, to insert *bo* into a simple tanru:

**Example 5.3.9**
ta klama bo jubme
That is-a goer – table

is a legal Lojban bridi that means exactly the same thing as Example 5.2.8, and is ambiguous in exactly the same ways. The cmavo *bo* serves only to resolve grouping ambiguity: it says nothing about the more basic ambiguity present in all tanru.

## 5.4 Complex tanru grouping

If one element of a tanru can be another tanru, why not both elements?

**Example 5.4.1**
do mutce bo barda gerku bo kavbu
You are-a-(very type-of large) (dog type-of capturer).
You are a very large dog-catcher.

In Example 5.4.1, the selbri is a tanru with seltau *mutce bo barda* and tertau *gerku bo kavbu*. It is worth emphasizing once again that this tanru has the same fundamental ambiguity as all other Lojban tanru: the sense in which the *dog type-of capturer* is said to be "very type-of large" is not precisely specified. Presumably it is his body which is large, but theoretically it could be one of his other properties.

We will now justify the title of this chapter by exploring the ramifications of the phrase "pretty little girls' school", an expansion of the tanru used in Section 5.3 to four brivla. (Although this example has been used in the Loglan Project almost since the beginning — it first appeared in Quine's book *Word and Object* (1960) — it is actually a mediocre example because of the ambiguity of English *pretty*; it can mean *beautiful*, the sense intended here, or it can mean *very*. Lojban *melbi* is not subject to this ambiguity: it means only *beautiful*.) Here are four ways to group this phrase:

**Example 5.4.2**
ta melbi cmalu
    nixli ckule
That is-a-((pretty type-of little)
    type-of girl) type-of school.
That is a school for girls who are beautifully small.

**Example 5.4.3**
ta melbi cmalu nixli bo ckule
That is-a-(pretty type-of little) (girl type-of school).
That is a girls' school which is beautifully small.

**Example 5.4.4**
ta melbi cmalu bo nixli
    ckule
That is-a-(pretty type-of (little type-of girl))
    type-of school.
That is a school for small girls who are beautiful.

**Example 5.4.5**
ta melbi cmalu bo
    nixli bo ckule
That is-a-pretty type-of (little type-of
    (girl type-of school)).
That is a small school for girls which is beautiful.

Example 5.4.5 uses a construction which has not been seen before: *cmalu bo nixli bo ckule*, with two consecutive uses of *bo* between brivla. The rule for multiple *bo*

constructions is the opposite of the rule when no *bo* is present at all: the last two are grouped together. Not surprisingly, this is called the *right-grouping rule*, and it is associated with every use of *bo* in the language. Therefore,

**Example 5.4.6**
ta cmalu bo nixli bo ckule
That is-a-little type-of (girl type-of school).

means the same as Example 5.3.4, not Example 5.3.5. This rule may seem peculiar at first, but one of its consequences is that *bo* is never necessary between the first two elements of any of the complex tanru presented so far: all of Examples 4.2 through 4.5 could have *bo* inserted between *melbi* and *cmalu* with no change in meaning.

## 5.5 Complex tanru with *ke* and *ke'e*

The following cmavo are discussed in this section:

```
ke  KE  start grouping
 ke'e    KEhE    end grouping
```

There is, in fact, a fifth grouping of "pretty little girls' school" that cannot be expressed with the resources explained so far. To handle it, we must introduce the grouping parentheses cmavo, *ke* and *ke'e* (belonging to selma'o KE and KEhE respectively). Any portion of a selbri sandwiched between these two cmavo is taken to be a single tanru component, independently of what is adjacent to it. Thus, Example 5.4.2 can be rewritten in any of the following ways:

**Example 5.5.1**
ta ke melbi cmalu ke'e nixli ckule
That is-a-( pretty little ) girl school.

**Example 5.5.2**
ta ke ke melbi cmalu ke'e nixli ke'e ckule
That is-a-( ( pretty little ) girl ) school.

**Example 5.5.3**
ta ke ke ke melbi cmalu ke'e nixli ke'e ckule ke'e
That is-a-( ( ( pretty little ) girl ) school ).

Even more versions could be created simply by placing any number of *ke* cmavo at the beginning of the selbri, and a like number of *ke'e* cmavo at its end. Obviously, all of these are a waste of breath once the left-grouping rule has been grasped. However, the following is equivalent to Example 5.4.4 and may be easier to understand:

**Example 5.5.4**
ta melbi ke cmalu nixli ke'e
ckule
That is-a-(pretty type-of ( little type-of girl ))
type-of school.

Likewise, a *ke* and *ke'e* version of Example 5.4.3 would be:

**Example 5.5.5**
ta melbi cmalu
ke nixli ckule ⟨*ke'e*⟩
That is-a-(pretty type-of little)
( girl type-of school ).

The final *ke'e* is given in square brackets here to indicate that it can be elided. It is always possible to elide *ke'e* at the end of the selbri, making Example 5.5.5 as terse as Example 5.4.3.

Now how about that fifth grouping? It is

**Example 5.5.6**
ta melbi
ke cmalu nixli ckule ⟨*ke'e*⟩
That is-a-pretty type-of
( (little type-of girl) type-of school )
That is a beautiful school for small girls.

Example 5.5.6 is distinctly different in meaning from any of Examples 4.2 through 4.5. Note that within the *ke* ... *ke'e* parentheses, the left-grouping rule is applied to *cmalu nixli ckule*.

It is perfectly all right to mix *bo* and *ke* ... *ke'e* in a single selbri. For instance, Example 5.4.5, which in pure *ke* ... *ke'e* form is

**Example 5.5.7**
ta melbi
ke cmalu ke nixli ckule ⟨*ke'e*⟩ [ke'e]
That is-a-pretty type-of
( little type-of ( girl type-of school ) ).

can equivalently be expressed as:

**Example 5.5.8**
ta melbi
ke cmalu nixli bo ckule ⟨*ke'e*⟩
That is-a-pretty
type-of ( little type-of (girl type-of school) ).

and in many other different forms as well.

## 5.6 Logical connection within tanru

The following cmavo are discussed in this section:

```
je  JA  tanru logical \q{and}
 ja  JA  tanru logical \q{or}
 joi JOI mixed mass \q{and}
 gu'e    GUhA    tanru forethought logical \q{and}
 gi  GI  forethought connection separator
```

Consider the English phrase *big red dog*. How shall this be rendered as a Lojban tanru? The naive attempt:

**Example 5.6.1**
barda xunre gerku
(big type-of red) type-of dog

will not do, as it means a dog whose redness is big, in whatever way redness might be described as *big*. Nor is

**Example 5.6.2**
barda xunre bo gerku
big type-of (red type-of dog)

much better. After all, the straightforward understanding of the English phrase is that the dog is big as compared with other dogs, not merely as compared with other red dogs. In fact, the bigness and redness are independent properties of the dog, and only obscure rules of English adjective ordering prevent us from saying *red big dog*.

The Lojban approach to this problem is to introduce the cmavo *je*, which is one of the many equivalents of English *and*. A big red dog is one that is both big and red, and we can say:

**Example 5.6.3**
barda je xunre gerku
(big and red) type-of dog

Of course,

**Example 5.6.4**
xunre je barda gerku
(red and big) type-of dog

is equally satisfactory and means the same thing. As these examples indicate, joining two brivla with *je* makes them a unit for tanru purposes. However, explicit grouping with *bo* or *ke ... ke'e* associates brivla more closely than *je* does:

**Example 5.6.5**
barda je pelxu bo xunre gerku
barda je ke pelxu xunre ke'e gerku
(big and (yellow type-of red)) dog
big yellowish-red dog

With no grouping indicators, we get:

**Example 5.6.6**
barda je pelxu xunre gerku
((big and yellow) type-of red) type-of dog
biggish- and yellowish-red dog

which again raises the question of Example 5.6.1: what is does *biggish-red* mean?

Unlike *bo* and *ke ... ke'e*, *je* is useful as well as merely legal within simple tanru. It may be used to partly resolve the ambiguity of simple tanru:

**Example 5.6.7**
ta blanu je zdani
that is-blue and is-a-house

definitely refers to something which is both blue and is a house, and not to any of the other possible interpretations of simple *blanu zdani*. Furthermore, *blanu zdani* refers to something which is blue in the way that houses are blue; *blanu je zdani* has no such implication — the blueness of a *blanu je zdani* is independent of its houseness.

With the addition of *je*, many more versions of "pretty little girls' school" are made possible: see Section 5.16 for a complete list.

A subtle point in the semantics of tanru like Example 5.6.3 needs special elucidation. There are at least two possible interpretations of:

**Example 5.6.8**
ta melbi je nixli ckule
That is-a-(beautiful and girl) type-of school.

It can be understood as:

**Example 5.6.9**
That is a girls' school and a beautiful school.

or as:

**Example 5.6.10**
That is a school for things
which are both girls and beautiful.

The interpretation specified by Example 5.6.9 treats the tanru as a sort of abbreviation for:

**Example 5.6.11**
ta ke melbi ckule ke'e
je ke nixli ckule ⟨*ke'e*⟩
That is-a-( beautiful type-of school )
and ( girl type-of school )

whereas the interpretation specified by Example 5.6.10 does not. This is a kind of semantic ambiguity for which Lojban does not compel a firm resolution. The way in which the school is said to be of type *beautiful and girl* may entail that it is separately a beautiful school and a girls' school; but the alternative interpretation, that the members of the school are beautiful and girls, is also possible. Still another interpretation is:

**Example 5.6.12**
That is a school for beautiful things
and also for girls.

so while the logical connectives help to resolve the meaning of tanru, they by no means compel a single meaning in and of themselves.

In general, logical connectives within tanru cannot undergo the formal manipulations that are possible with the related logical connectives that exist outside tanru; see Chapter 14 for further details.

The logical connective *je* is only one of the fourteen logical connectives that Lojban provides. Here are a few examples of some of the others:

**Example 5.6.13**
le bajra cu jinga ja te jinga
the runner(s) is/are winner(s) or loser(s).

**Example 5.6.14**
blanu naja lenku skapi
(blue only-if cold) skin
skin which is blue only if it is cold

**Example 5.6.15**
xamgu jo cortu nuntavla
(good if-and-only-if short) speech
speech which is good if (and only if) it is short

**Example 5.6.16**
vajni ju pluka nuntavla
(important whether-or-not pleasing) event-of-talking
speech which is important, whether or not it is pleasing

In Example 5.6.13, *ja* is grammatically equivalent to *je* but means *or* (more precisely, *and/or*). Likewise, *naja* means *only if* inExample 5.6.14, *jo* means "if and only if" in Example 5.6.15, and *ju* means *whether or not* in Example 5.6.16.

Now consider the following example:

**Example 5.6.17**
ricfu je blanu jabo crino
rich and (blue or green)

which illustrates a new grammatical feature: the use of both *ja* and *bo* between tanru components. The two cmavo combine to form a compound whose meaning is that of *ja* but which groups more closely; *jabo* is to *ja* as plain *bo* is to no cmavo at all. However, both *ja* and *jabo* group less closely than *bo* does:

**Example 5.6.18**
ricfu je blanu jabo crino bo blanu
rich and (blue or green – blue)
rich and (blue or greenish-blue)

An alternative form of Example 5.6.17 is:

**Example 5.6.19**
ricfu je ke blanu ja crino ⟨*ke'e*⟩
rich and ( blue or green )

In addition to the logical connectives, there are also a variety of non-logical connectives, grammatically equivalent to the logical ones. The only one with a well-understood meaning in tanru contexts is *joi*, which is the kind of *and* that denotes a mixture:

**Example 5.6.20**
ti blanu joi xunre bolci
This is-a-(blue and red) ball.

The ball described is neither solely red nor solely blue, but probably striped or in some other way exhibiting a combination of the two colors. Example 5.6.20 is distinct from:

**Example 5.6.21**
ti blanu xunre bolci
This is a bluish-red ball

which would be a ball whose color is some sort of purple tending toward red, since *xunre* is the more important of the two components. On the other hand,

**Example 5.6.22**
ti blanu je xunre bolci
This is a (blue and red) ball

is probably self-contradictory, seeming to claim that the ball is independently both blu and red at the same time, although some sensible interpretation may exist.

Finally, just as English *and* has the variant form "both ... and", so *je* between tanru components has the variant form *gu'e ... gi*, where *gu'e* is placed before the components and *gi* between them:

**Example 5.6.23**
gu'e barda gi xunre gerku
(both big and red) type-of dog

is equivalent in meaning to Example 5.6.3. For each logical connective related to *je*, there is a corresponding connective related to *gu'e ... gi* in a systematic way.

The portion of a *gu'e ... gi* construction before the *gi* is a full selbri, and may use any of the selbri resources including *je* logical connections. After the *gi*, logical connections are taken to be wider in scope than the *gu'e ... gi*, which has in effect the same scope as *bo*:

**Example 5.6.24**
gu'e barda je xunre gi gerku ja mlatu
(both (big and red) and dog) or cat
something which is either big, red, and a dog,
or else a cat

leaves *mlatu* outside the *gu'e–gi* construction. The scope of the *gi* arm extends only to a single brivla or to two or more brivla connected with *bo* or *ke–ke'e*.

## 5.7 Linked sumti: *be–bei–be'o*

The following cmavo are discussed in this section:

```
be  BE  linked sumti marker
 bei BEI linked sumti separator
 be'o    BEh0    linked sumti terminator
```

The question of the place structures of selbri has been glossed over so far. This chapter does not attempt to treat place structure issues in detail; they are discussed in Chapter 9. One grammatical structure related to places belongs here, however. In simple sentences such as Example 5.1.1, the place structure of the selbri is simply the defined place structure of the gismu *mamta*. What about more complex selbri?

For tanru, the place structure rule is simple: the place structure of a tanru is always the place structure of its tertau. Thus, the place structure of *blanu zdani* is that of *zdani*: the $x_1$ place is a house or nest, and the $x_2$ place is its occupants.

What about the places of *blanu*? Is there any way to get them into the act? In fact, *blanu* has only one place, and this is merged, as it were, with the $x_1$ place of *zdani*. It is whatever is in the $x_1$ place that is being characterized as blue-for-a-house. But if we replace *blanu* with *xamgu*, we get:

**Example 5.7.1**
ti xamgu zdani
this is-a-good house.
This is a good (for someone, by some standard) house.

Since *xamgu* has three places ($x_1$, the good thing; $x_2$, the person for whom it is good; and $x_3$, the standard of goodness), Example 5.7.1 necessarily omits information about the last two: there is no room for them. Room can be made, however!

**Example 5.7.2**
ti xamgu be do bei mi ⟨*be'o*⟩ zdani
this is-a-good ( for you by-standard me ) house.
This is a house that is good for you by my standards.

Here, the gismu *xamgu* has been followed by the cmavo *be* (of selma'o BE), which signals that one or more sumti follows. These sumti are not part of the overall bridi place structure, but fill the places of the brivla they are attached to, starting with $x_2$. If there is more than one sumti, they are separated by the cmavo *bei* (of selma'o BEI), and the list of sumti is terminated by the elidable terminator *be'o* (of selma'o BEhO).

Grammatically, a brivla with sumti linked to it in this fashion plays the same role in tanru as a simple brivla. To illustrate, here is a fully fleshed-out version of Example 5.3.4, with all places filled in:

**Example 5.7.3**
ti cmalu be le ka canlu
bei lo'e ckule be'o
nixli be li mu
bei lo merko be'o bo
ckule la bryklyn. loi pemci
le mela nu,IORK. prenu
le jecta
This is a small (in-dimension the property-of volume
by-standard the-typical school)
(girl (of-years the-number five
by-standard some American-thing)
school) in-Brooklyn with-subject poems
for-audience New-York persons

with-operator the state.
This is a school, small in volume compared to the
typical school, pertaining to five-year-old
girls (by American standards), in Brooklyn,
teaching poetry to the New York community
and operated by the state.

Here the three places of *cmalu*, the three of *nixli*, and the four of *ckule* are fully specified. Since the places of *ckule* are the places of the bridi as a whole, it was not necessary to link the sumti which follow *ckule*. It would have been legal to do so, however:

**Example 5.7.4**
mi klama be le zarci bei le zdani ⟨*be'o*⟩
I go (to-the market from-the house).

means the same as

**Example 5.7.5**
mi klama le zarci le zdani
I go to-the market from-the house.

No matter how complex a tanru gets, the last brivla always dictates the place structure: the place structure of

**Example 5.7.6**
melbi je cmalu nixli bo ckule
a (pretty and little) (girl school)
a school for girls which is both beautiful and small

is simply that of *ckule*. (The sole exception to this rule is discussed in Section 5.8.)

It is possible to precede linked sumti by the place structure ordering tags *fe*, *fi*, *fo*, and *fu* (of selma'o FA, discussed further in Chapter 9), which serve to explicitly specify the $x_2$, $x_3$, $x_4$, and $x_5$ places respectively. Normally, the place following the *be* is the $x_2$ place and the other places follow in order. If it seems convenient to change the order, however, it can be accomplished as follows:

**Example 5.7.7**
ti xamgu be fi mi bei fe do ⟨*be'o*⟩ zdani
this is-a-good ( by-standard me for you ) house

which is equivalent in meaning to Example 5.7.2. Note that the order of *be*, *bei*, and *be'o* does not change; only the inserted *fi* tells us that *mi* is the $x_3$ place (and correspondingly, the inserted *fe* tells us that *do* is the $x_2$ place). Changing the order of sumti is often done to match the order of another language, or for emphasis or rhythm.

Of course, using FA cmavo makes it easy to specify one place while omitting a previous place:

**Example 5.7.8**
ti xamgu be fi mi ⟨*be'o*⟩ zdani
this is-a-good (by-standard me) house
This is a good house by my standards.

Similarly, sumti labeled by modal or tense tags can be inserted into strings of linked sumti just as they can into bridi:

**Example 5.7.9**
ta blanu be ga'a mi ⟨*be'o*⟩ zdani
That is-a-blue ( to-observer me ) house.
That is a blue, as I see it, house.

The meaning of Example 5.7.9 is slightly different from:

**Example 5.7.10**
ta blanu zdani ga'a mi
That is-a-blue house to-observer me.
That is a blue house, as I see it.

See discussions in Chapter 9 of modals and in Chapter 10 of tenses for more explanations.

The terminator *be'o* is almost always elidable: however, if the selbri belongs to a description, then a relative clause following it will attach to the last linked sumti unless *be'o* is used, in which case it will attach to the outer description:

**Example 5.7.11**
le xamgu be do noi barda cu zdani
The good-thing for you (who are-large) is-a-house.

**Example 5.7.12**
le xamgu be do be'o noi barda cu zdani
The (good-thing for you) (which is-large) is-a-house

(Relative clauses are explained in Chapter 8.)

In other cases, however, *be'o* cannot be elided if *ku* has also been elided:

**Example 5.7.13**
le xamgu be le ctuca ⟨*ku*⟩ be'o zdani
the good (for the teacher ) house

requires either *ku* or *be'o*, and since there is only one occurrence of *be*, the *be'o* must match it, whereas it may be confusing which occurrence of *le* the *ku* terminates (in fact the second one is correct).

## 5.8 Inversion of tanru: *co*

The following cmavo is discussed in this section:

```
co  CO  tanru inversion marker
```

The standard order of Lojban tanru, whereby the modifier precedes what it modifies, is very natural to English-speakers: we talk of *blue houses*, not of *houses blue*. In other languages, however, such matters are differently arranged, and Lojban supports this reverse order (tertau before seltau) by inserting the particle *co*. Example 5.8.1 and Example 5.8.2 mean exactly the same thing:

**Example 5.8.1**
ta blanu zdani
That is-a-blue type-of-house.
That is a blue house.

**Example 5.8.2**
ta zdani co blanu
That is-a-house of-type blue.
That is a blue house.

This change is called *tanru inversion*. In tanru inversion, the element before *co* (*zdani* in Example 5.8.2) is the tertau, and the element following *co* (*blanu*) in Example 5.8.2) is the seltau.

The meaning, and more specifically, the place structure, of a tanru is not affected by inversion: the place structure of *zdani co blanu* is still that of *zdani*. However, the existence of inversion in a selbri has a very special effect on any sumti which follow that selbri. Instead of being interpreted as filling places of the selbri, they actually fill the places (starting with $x_2$) of the seltau. In Section 5.7, we saw how to fill interior places with *be ... bei ... be'o*, and in fact Example 5.8.3 and Example 5.8.4 have the same meaning:

**Example 5.8.3**
mi klama be le zarci bei le zdani be'o
troci
I am-a-(goer to the market from the house)
type-of trier.
I try to go to the market from the house.

**Example 5.8.4**
mi troci co klama le zarci le zdani
I am-a-trier
of-type (goer to-the market from-the house).
I try to go to the market from the house.

Example 5.8.4 is a less deeply nested construction, requiring fewer cmavo. As a result it is probably easier to understand.

Note that in Lojban *trying to go* is expressed using *troci* as the tertau. The reason is that *trying to go* is a *going type of trying*, not a *trying type of going*. The trying is more fundamental than the going — if the attempt fails, we may not have a going at all.

Any sumti which precede a selbri with an inverted tanru fill the places of the selbri (i.e., the places of the tertau) in the ordinary way. In Example 5.8.4, *mi* fills the $x_1$ place of *troci co klama*, which is the $x_1$ place of *troci*. The other places of the selbri remain unfilled. The trailing sumti *le zarci* and *le zdani* do not occupy selbri places, despite appearances.

As a result, the regular mechanisms (involving selma'o VOhA and GOhI, explained in Chapter 7) for referring to individual sumti of a bridi cannot refer to any of the trailing places of Example 5.8.4, because they are not really *sumti of the bridi* at all.

When inverting a more complex tanru, it is possible to invert it only at the most general modifier-modified pair. The only possible inversion of Example 5.3.4, for instance, is:

**Example 5.8.5**
ta nixli ⟨*bo*⟩ ckule co cmalu
that (is-a-girl type-of school) of-type little.
That's a girls' school which is small.

Note that the *bo* of Example 5.3.4 is optional in Example 5.8.5, because *co* groups more loosely than any other cmavo used in tanru, including none at all. Not even *ke* ... *ke'e* parentheses can encompass a *co*:

**Example 5.8.6**
ta cmalu ke nixli ckule ⟨*ke'e*⟩
co melbi
that is-a-(little type-of (girl type-of school))
of-type pretty.
That's a small school for girls which is beautiful.

In Example 5.8.6, the *ke'e* is automatically inserted before the *co* rather than at its usual place at the end of the selbri. As a result, there is a simple and mechanical rule for removing *co* from any selbri: change *A co B* to *ke B ke'e A*. (At the same time, any sumti following the selbri must be transformed into "be ... bei ... be'o" form and attached following B.) Therefore,

**Example 5.8.7**
ckule co melbi nixli
school of-type pretty girl
school for beautiful girls

means the same as:

**Example 5.8.8**
ke melbi nixli ke'e ckule
( pretty girl ) school

Multiple *co* cmavo can appear within a selbri, indicating multiple inversions: a right-grouping rule is employed, as for *bo*. The above rule can be applied to interpret such selbri, but all *co* cmavo must be removed simultaneously:

**Example 5.8.9**
ckule co nixli co cmalu
school of-type (girl of-type little)

becomes formally

**Example 5.8.10**
ke ke cmalu ke'e nixli ke'e ckule
( ( little ) girl ) school

which by the left-grouping rule is simply

**Example 5.8.11**
cmalu nixli ckule
little girl school
school for little girls

As stated above, the selbri places, other than the first, of

**Example 5.8.12**
mi klama co sutra
I am-a-goer of-type quick
I go quickly

cannot be filled by placing sumti after the selbri, because any sumti in that position fill the places of *sutra*, the seltau. However, the tertau places (which means in effect the selbri places) can be filled with *be*:

**Example 5.8.13**
mi klama be le zarci co sutra
I am-a-goer (to the store) of-type quick.
I go to the store quickly.

## 5.9 Other kinds of simple selbri

The following cmavo are discussed in this section:

```
go'i    GOhA    repeats the previous bridi
 du  GOhA    equality
 nu'a    NUhA    math operator to selbri
 moi MOI changes number to ordinal selbri
 mei MOI changes number to cardinal selbri
 nu  NU  event abstraction
 kei KEI terminator for \q{nu}
```

So far we have only discussed brivla and tanru built up from brivla as possible selbri. In fact, there are a few other constructions in Lojban which are grammatically equivalent to brivla: they can be used either directly as selbri, or as components in tanru. Some of these types of simple selbri are discussed at length in Chapter 7, Chapter 11, and Chapter 18; but for completeness these types are mentioned here with a brief explanation and an example of their use in selbri.

The cmavo of selma'o GOhA (with one exception) serve as pro-bridi, providing a reference to the content of other bridi; none of them has a fixed meaning. The most commonly used member of GOhA is probably *go'i*, which amounts to a repetition of the previous bridi, or part of it. If I say:

**Example 5.9.1**
la djan. klama le zarci
John goes-to the market.

you may retort:

**Example 5.9.2**
la djan. go'i troci
John ⟨*repeat last*⟩ are-a-tryer
John tries to.

Example 5.9.2 is short for:

**Example 5.9.3**
la djan. klama be le zarci be'o troci
John is-a-goer (to the market) type-of trier.

because the whole bridi of Example 5.9.1 has been packaged up into the single word *go'i* and inserted into Example 5.9.2.

The exceptional member of GOhA is *du*, which represents the relation of identity. Its place structure is:

**Definition 5.2**
$x_1$ is identical with $x_2$, $x_3$, ...

for as many places as are given. More information on selma'o GOhA is available in Chapter 7.

Lojban mathematical expressions (mekso) can be incorporated into selbri in two different ways. Mathematical operators such as *su'i*, meaning *plus*, can be transformed into selbri by prefixing them with *nu'a* (of selma'o NUhA). The resulting place structure is:

**Definition 5.3**
$x_1$ is the result of applying (the operator) to arguments $x_2$, $x_3$, etc.

for as many arguments as are required. (The result goes in the $x_1$ place because the number of following places may be indefinite.) For example:

**Example 5.9.4**
li vo nu'a su'i li re li re
The-number 4 is-the-sum-of the-number 2 and-the-number 2.

A possible tanru example might be:

**Example 5.9.5**
mi jimpe tu'a nu'a su'i nabmi
I understand something-about the-mass-of is-the-sum-of problems.
I understand addition problems.

More usefully, it is possible to combine a mathematical expression with a cmavo of selma'o MOI to create one of various numerical selbri. Details are available in Chapter 18. Here are a few tanru:

**Example 5.9.6**
la prim. palvr. pamoi cusku
Preem Palver is-the-1-th speaker.
Preem Palver is the first speaker.

**Example 5.9.7**
la an,iis. joi la .asun. bruna remei
Anyi massed-with Asun are-a-brother type-of-twosome.
Anyi and Asun are two brothers.

Finally, an important type of simple selbri which is not a brivla is the abstraction. Grammatically, abstractions are simple: a cmavo of selma'o NU, followed by a bridi, followed by the elidable terminator *kei* of selma'o KEI. Semantically, abstractions are an extremely subtle and powerful feature of Lojban whose full ramifications are documented in Chapter 11. A few examples:

**Example 5.9.8**
ti nu zdile kei kumfa
This is-an-event-of amusement room.
This is an amusement room.

Example 5.9.8 is quite distinct in meaning from:

**Example 5.9.9**
ti zdile kumfa
This is-an-amuser room.

which suggests the meaning *a room that amuses someone*.

## 5.10 selbri based on sumti: *me*

The following cmavo are discussed in this section:

```
me  ME  changes sumti to simple selbri
 me'u    MEhU    terminator for \q{me}
```

A sumti can be made into a simple selbri by preceding it with *me* (of selma'o ME) and following it with the elidable terminator *me'u* (of selma'o MEhU). This makes a selbri with the place structure

**Definition 5.4**
$x_1$ is one of the referents of *[the sumti]*

which is true of the thing, or things, that are the referents of the sumti, and not of anything else. For example, consider the sumti

**Example 5.10.1**
le ci nolraitru
the three noblest-governors
the three kings

If these are understood to be the Three Kings of Christian tradition, who arrive every year on January 6, then we may say:

**Example 5.10.2**
la BALtazar. cu me le ci nolraitru
Balthazar is one-of-the-referents-of *the three kings*.
Balthazar is one of the three kings.

and likewise

**Example 5.10.3**
la kaspar. cu me le ci nolraitru
Caspar is one of the three kings.

and

**Example 5.10.4**
la melxi,or. cu me le ci nolraitru
Melchior is one of the three kings.

If the sumti refers to a single object, then the effect of *me* is much like that of *du*:

**Example 5.10.5**
do du la djan.
You are-identical-with the-one-called *John*.
You are John.

means the same as

**Example 5.10.6**
do me la djan.
You are-the-referent-of *the-one-called 'John*'.
You are John.

It is common to use *me* selbri, especially those based on name sumti using *la*, as seltau. For example:

**Example 5.10.7**
ta me lai kraislr. ⟨*me'u*⟩ karce
That (is-a-referent of *the-mass-called 'Chrysler*') car.
That is a Chrysler car.

The elidable terminator *me'u* can usually be omitted. It is absolutely required only if the *me* selbri is being used in an indefinite description (a type of sumti explained in Chapter 6), and if the indefinite description is followed by a relative clause (explained in Chapter 8) or a sumti logical connective (explained in Chapter 14). Without a *me'u*, the relative clause or logical connective would appear to belong to the sumti embedded in the *me* expression. Here is a contrasting pair of sentences:

**Example 5.10.8**
re me le ci nolraitru .e la djan. ⟨*me'u*⟩ cu blabi
Two of the group *the three kings and John* are white.

**Example 5.10.9**
re me le ci nolraitru me'u .e la djan. cu blabi
Two of the three kings, and John, are white.

In Example 5.10.8 the *me* selbri covers the three kings plus John, and the indefinite description picks out two of them that are said to be white: we cannot say which two. In Example 5.10.9, though, the *me* selbri covers only the three kings: two of them are said to be white, and so is John.

Finally, here is another example requiring *me'u*:

**Example 5.10.10**
ta me la'e le se cusku be do me'u cukta
That is-a-(what-you-said) type of book.
That is the kind of book you were talking about.

There are other sentences where either *me'u* or some other elidable terminator must be expressed:

**Example 5.10.11**
le me le ci nolraitru ⟨*ku*⟩ me'u nunsalci
the (the three kings) type-of-event-of-celebrating
the Three Kings celebration

requires either *ku* or *me'u* to be explicit, and (as with *be'o* in Section 5.7) the *me'u* leaves no doubt which cmavo it is paired with.

## 5.11 Conversion of simple selbri

Conversion is the process of changing a selbri so that its places appear in a different order. This is not the same as labeling the sumti with the cmavo of FA, as mentioned in Section 5.7, and then rearranging the order in which the sumti are spoken or written. Conversion transforms the selbri into a distinct, though closely related, selbri with renumbered places.

In Lojban, conversion is accomplished by placing a cmavo of selma'o SE before the selbri:

**Example 5.11.1**
mi prami do
I love you.

is equivalent in meaning to:

**Example 5.11.2**
do se prami mi
You ⟨*swap* $x_1$ *and* $x_2$⟩ love me.
You are loved by me.

Conversion is fully explained in Chapter 9. For the purposes of this chapter, the important point about conversion is that it applies only to the following simple selbri. When trying to convert a tanru, therefore, it is necessary to be careful! Consider Example 5.11.3:

**Example 5.11.3**
la .alis. cu cadzu klama le zarci
Alice is-a-walker type-of goer to-the market.
Alice walkingly goes to the market.
Alice walks to the market.

To convert this sentence so that *le zarci* is in the $x_1$ place, one correct way is:

**Example 5.11.4**
le zarci cu se ke cadzu klama ⟨*ke'e*⟩ la .alis.
The market is-a-[swap $x_1$/$x_2$] ( walker type-of goer) Alice.
The market is-walkingly gone-to by-Alice.

The *ke ... ke'e* brackets cause the entire tanru to be converted by the *se*, which would otherwise convert only *cadzu*, leading to:

**Example 5.11.5**
le zarci cu se cadzu
klama la .alis.
The market (is-a-⟨*swap* $x_1$/$x_2$⟩ walker)
type-of goer to Alice.
The market is-a-walking-surface type-of goer to Alice.

whatever that might mean. An alternative approach, since the place structure of *cadzu klama* is that of *klama* alone, is to convert only the latter:

**Example 5.11.6**
le zarci cu cadzu se klama la .alis.
The market walkingly is-gone-to by-Alice.

But the tanru in Example 5.11.6 may or may not have the same meaning as that in Example 5.11.3; in particular, because *cadzu* is not converted, there is a suggestion that although Alice is the goer, the market is the walker. With a different sumti as $x_1$, this seemingly odd interpretation might make considerable sense:

**Example 5.11.7**
la djan. cu cadzu se klama la .alis
John walkingly is-gone-to by Alice

suggests that Alice is going to John, who is a moving target.

There is an alternative type of conversion, using the cmavo *jai* of selma'o JAI optionally followed by a modal or tense construction. Grammatically, such a combination behaves exactly like conversion using SE. More details can be found in Chapter 9.

## 5.12 Scalar negation of selbri

Negation is too large and complex a topic to explain fully in this chapter; see Chapter 15. In brief, there are two main types of negation in Lojban. This section is concerned with so-called *scalar negation*, which is used to state that a true relation between the sumti is something other than what the selbri specifies. Scalar negation is expressed by cmavo of selma'o NAhE:

**Example 5.12.1**
la .alis. cu na'e ke cadzu klama ⟨*ke'e*⟩ le zarci
Alice non- (walkingly goes) to-the market.
Alice other-than (walkingly goes) to-the market.
Alice doesn't walk to the market.

meaning that Alice's relationship to the market is something other than that of walking there. But if the *ke* were omitted, the result would be:

**Example 5.12.2**
la .alis. cu na'e cadzu klama le zarci
Alice non- walkingly goes to-the market.
Alice doesn't walk to the market.

meaning that Alice does go there in some way (*klama* is not negated), but by a means other than that of walking. Example 5.12.1 negates both *cadzu* and *klama*, suggesting that Alice's relation to the market is something different from walkingly-going; it might be walking without going, or going without walking, or neither.

Of course, any of the simple selbri types explained in Section 5.9 may be used in place of brivla in any of these examples:

**Example 5.12.3**
la djonz. cu na'e pamoi cusku
Jones is non-1st speaker
Jones is not the first speaker.

Since only *pamoi* is negated, an appropriate inference is that he is some other kind of speaker.

Here is an assortment of more complex examples showing the interaction of scalar negation with *bo* grouping, *ke* and *ke'e* grouping, logical connection, and sumti linked with *be* and *bei*:

**Example 5.12.4**
mi na'e sutra cadzu be fi le birka be'o
klama le zarci
I ((non-quickly) ( walking using the arms))
go-to the market.
I go to the market, walking using my arms
other than quickly.

In Example 5.12.4, *na'e* negates only *sutra*. Contrast Example 5.12.5:

**Example 5.12.5**
mi na'e ke sutra cadzu be fi le birka ⟨*be'o*⟩ ke'e
klama le zarci
I non- ( quickly (walking using the arms) )
go-to the market.
I go to the market, other than by walking
quickly on my arms.

Now consider Example 5.12.6 and Example 5.12.7, which are equivalent in meaning, but use *ke* grouping and *bo* grouping respectively:

**Example 5.12.6**
mi sutra cadzu be fi le birka be'o je masno
klama le zarci
I (quickly – (walking using the arms) and slowly)
go-to the market.
I go to the market, both quickly walking
using my arms and slowly.

**Example 5.12.7**
mi ke sutra cadzu be fi le birka ⟨*be'o*⟩ ke'e
je masno klama le zarci
I ((quickly (walking using the arms))
and slowly) go-to the market.
I go to the market, both quickly walking
using my arms and slowly.

However, if we place a *na'e* at the beginning of the selbri in both Example 5.12.6 and Example 5.12.7, we get different results:

**Example 5.12.8**
mi na'e sutra cadzu be fi le birka be'o
je masno klama le zarci
I ((non- quickly) – (walking using the arms)
and slowly) go-to the market.
I go to the market, both walking using my arms
other than quickly, and also slowly.

**Example 5.12.9**
mi na'e ke sutra cadzu be fi le birka ⟨*be'o*⟩ ke'e
je masno klama le zarci
I (non-(quickly (walking using the arms))
and slowly) go-to the market.
I go to the market, both other than quickly
walking using my arms, and also slowly.

The difference arises because the *na'e* in Example 5.12.9 negates the whole construction from *ke* to *ke'e*, whereas in Example 5.12.8 it negates *sutra* alone.

Beware of omitting terminators in these complex examples! If the explicit *ke'e* is left out in Example 5.12.9, it is transformed into:

**Example 5.12.10**
mi na'e ke sutra cadzu be fi le birka be'o
je masno klama ⟨*ke'e*⟩ le zarci
I non-(quickly ((walking using the arms))
and slowly) go-to) the market.
I do something other than quickly both
going to the market walking using my arms
and slowly going to the market.

And if both *ke'e* and *be'o* are omitted, the results are even sillier:

**Example 5.12.11**
mi na'e ke sutra cadzu be fi le birka
je masno klama ⟨*be'o*⟩ [ke'e] le zarci
I non-(quickly walk on my (arm-type
and slow) goers) on the market.
I do something other than quickly walking using the
goers, both arm-type and slow, relative-to the market.

In Example 5.12.11, everything after *be* is a linked sumti, so the place structure is that of *cadzu*, whose $x_2$ place is the surface walked upon. It is less than clear what an *arm-type goer* might be. Furthermore, since the $x_3$ place has been occupied by the linked sumti, the *le zarci* following the selbri falls into the nonexistent $x_4$ place of *cadzu*. As a result, the whole example, though grammatical, is complete nonsense. (The bracketed Lojban words appear where a fluent Lojbanist would understand them to be implied.)

Finally, it is also possible to place *na'e* before a *gu'e ... gi* logically connected tanru construction. The meaning of this usage has not yet been firmly established.

## 5.13 Tenses and bridi negation

A bridi can have cmavo associated with it which specify the time, place, or mode of action. For example, in

**Example 5.13.1**
mi pu klama le zarci
I ⟨*past*⟩ go to-the market.
I went to the market.

the cmavo *pu* specifies that the action of the speaker going to the market takes place in the past. Tenses are explained in full detail in Chapter 10. Tense is semantically a property of the entire bridi; however, the usual syntax for tenses attaches them at the front of the selbri, as in Example 5.13.1. There are alternative ways of expressing tense information as well. Modals, which are explained in Chapter 9, behave in the same way as tenses.

Similarly, a bridi may have the particle *na* (of selma'o NA) attached to the beginning of the selbri to negate the bridi. A negated bridi expresses what is false without saying anything about what is true. Do not confuse this usage with the scalar negation of Section 5.12. For example:

**Example 5.13.2**
la djonz. na pamoi cusku
Jones (Not!) is-the-first speaker
It is not true that Jones is the first speaker.
Jones isn't the first speaker.

Jones may be the second speaker, or not a speaker at all; Example 5.13.2 doesn't say. There are other ways of expressing bridi negation as well; the topic is explained fully in Chapter 15.

Various combinations of tense and bridi negation cmavo are permitted. If both are expressed, either order is permissible with no change in meaning:

**Example 5.13.3**
mi na pu klama le zarci
mi pu na klama le zarci
It is false that I went to the market.
I didn't go to the market.

It is also possible to have more than one *na*, in which case pairs of *na* cmavo cancel out:

**Example 5.13.4**
mi na na klama le zarci
It is false that it is false that I go to the market.
I go to the market.

It is even possible, though somewhat pointless, to have multiple *na* cmavo and tense cmavo mixed together, subject to the limitation that two adjacent tense cmavo will be understood as a compound tense, and must fit the grammar of tenses as explained in Chapter 10.

**Example 5.13.5**
mi na pu na ca klama le zarci
I ⟨*not*⟩ [past]
[not] [present] go to-the market
It is not the case that in the past it was not
the case that in the present I went
to the market.
I didn't not go to the market.
I went to the market.

Tense, modal, and negation cmavo can appear only at the beginning of the selbri. They cannot be embedded within it.

## 5.14 Some types of asymmetrical tanru

This section and Section 5.15 contain some example tanru classified into groups based on the type of relationship between the modifying seltau and the modified tertau. All the examples are paralleled by compounds actually observed in various natural languages. In the tables which follow, each group is preceded by a brief explanation of the relationship. The tables themselves contain a tanru, a literal gloss, an indication of the languages which exhibit a compound analogous to this tanru, and (for those tanru with no English parallel) a translation.

Here are the 3-letter abbreviations used for the various languages (it is presumed to be obvious whether a compound is found in English or not, so English is not explicitly noted):

**Definition 5.5**
Aba = Abazin Chi = Chinese Eng = English Ewe = Ewe Fin = Finnish Geo = Georgian Gua = Guarani Hop = Hopi Hun = Hungarian Imb = Imbabura Quechua Kar = Karaitic Kaz = Kazakh Kor = Korean Mon = Mongolian Qab = Qabardian Que = Quechua Rus = Russian Skt = Sanskrit Swe = Swedish Tur = Turkish Udm = Udmurt

Any lujvo or fu'ivla used in a group are glossed at the end of that group.

The tanru discussed in this section are asymmetrical tanru; that is, ones in which the order of the terms is fundamental to the meaning of the tanru. For example, *junla dadylsi*, or *clock pendulum*, is the kind of pendulum used in a clock, whereas *dadysli junla*, or *pendulum clock*, is the kind of clock that employs a pendulum. Most tanru are asymmetrical in this sense. Symmetrical tanru are discussed in Section 5.15.

The tertau represents an action, and the seltau then represents the object of that action:

**Definition 5.6**

**pinsi kilbra** pencil sharpener (Hun) zgike nunctu music instruction (Hun) mirli nunkalte deer hunting (Hun) finpe nunkalte fish hunting (Tur,Kor,Udm,Aba = fishing) smacu terkavbu mousetrap (Tur,Kor,Hun,Udm,Aba) zdani turni house ruler (Kar = host) zerle'a nunte'a thief fear (Skt = fear of thieves) cevni zekri god crime (Skt = offense against the gods) kilbra = sharp-apparatus nunctu = event-of-teaching nunkalte = event-of-hunting terkavbu = trap zerle'a = crime-taker nunte'a = event-of-fearing

The tertau represents a set, and the seltau the type of the elements contained in that set:

**Definition 5.7**

**zdani lijgri** house row selci lamgri cell block karda mulgri card pack (Swe) rokci derxi stone heap (Swe) tadni girzu student group (Hun) remna girzu human-being group (Qab = group of people) cpumi'i lijgri tractor column (Qab) cevni jenmi god army (Skt) cevni prenu god folk (Skt) lijgri = line-group lamgri = adjacent-group mulgri = complete-group cpami'i = pull-machine

Conversely: the tertau is an element, and the seltau represents a set in which that element is contained. Implicitly, the meaning of the tertau is restricted from its usual general meaning to the specific meaning appropriate for elements in the given set. Note the opposition between *zdani linji* in the previous group, and *linji zdani* in this one, which shows why this kind of tanru is called *asymmetrical*.

**Definition 5.8**

**carvi dirgo** raindrop (Tur,Kor,Hun,Udm,Aba) linji zdani row house

The seltau specifies an object and the tertau a component or detail of that object; the tanru as a whole refers to the detail, specifying that it is a detail of that whole and not some other.

**Definition 5.9**

**junla dadysli** clock pendulum (Hun) purdi vorme garden door (Qab) purdi bitmu garden wall (Que) moklu skapi mouth skin (Imb = lips) nazbi kevna nose hole (Imb = nostril) karce xislu automobile wheel (Chi) jipci pimlu chicken feather (Chi) vinji rebla airplane tail (Chi) dadysli = hang-oscillator

Conversely: the seltau specifies a characteristic or important detail of the object described by the tertau; objects described by the tanru as a whole are differentiated from other similar objects by this detail.

**Definition 5.10**

**pixra cukta** picture book kerfa silka hair silk (Kar = velvet) plise tapla apple cake (Tur) dadysli junla pendulum clock (Hun) dadysli = hang-oscillator

The tertau specifies a general class of object (a genus), and the seltau specifies a sub-class of that class (a species):

**Definition 5.11**

**ckunu tricu** pine tree (Hun,Tur,Hop)

The tertau specifies an object of possession, and the seltau may specify the possessor (the possession may be intrinsic or otherwise). In English, these compounds have an explicit possessive element in them: *lion's mane*, *child's foot*, *noble's cow*.

**Definition 5.12**

**cinfo kerfa** lion mane (Kor,Tur,Hun,Udm,Qab) verba jamfu child foot (Swe) nixli tuple girl leg (Swe) cinfo jamfu lion foot (Que) danlu skapi animal skin (Ewe) ralju zdani chief house (Ewe) jmive munje living world (Skt) nobli bakni noble cow (Skt) nolraitru ralju king chief (Skt = emperor) nolraitru = nobly-superlative-ruler

The tertau specifies a habitat, and the seltau specifies the inhabitant:

**Definition 5.13**

**lanzu tumla** family land

The tertau specifies a causative agent, and the seltau specifies the effect of that cause:

**Definition 5.14**

**kalselvi'i gapci** tear gas (Hun) terbi'a jurme disease germ (Tur) fenki litki crazy liquid (Hop = whisky) pinca litki urine liquid (Hop = beer) kalselvi'i = eye-excreted-thing terbi'a = disease

Conversely: the tertau specifies an effect, and the seltau specifies its cause.

**Definition 5.15**

**djacu barna** water mark (Chi)

The tertau specifies an instrument, and the seltau specifies the purpose of that instrument:

**Definition 5.16**

**taxfu dadgreku** garment rack (Chi) tergu'i ti'otci lamp shade (Chi) xirma zdani horse house (Chi = stall) nuzba tanbo news board (Chi = bulletin board) dadgreku = hang-frame tergu'i = source of illumination ti'otci = shadow-tool

More vaguely: the tertau specifies an instrument, and the seltau specifies the object of the purpose for which that instrument is used:

**Definition 5.17**

**cpina rokci** pepper stone (Que = stone for grinding pepper) jamfu djacu foot water (Skt = water for washing the feet) grana mudri post wood (Skt = wood for making a post) moklu djacu mouth water (Hun = water for washing the mouth) lanme gerku sheep dog (dog for working sheep)

The tertau specifies a product from some source, and the seltau specifies the source of the product:

**Definition 5.18**

**moklu djacu** mouth water (Aba,Qab = saliva) ractu mapku rabbit hat (Rus) jipci sovda chicken egg (Chi) sikcurnu silka silkworm silk (Chi) mlatu kalci cat feces (Chi) bifce lakse bee wax (Chi = beeswax) cribe rectu bear meat (Tur,Kor,Hun,Udm,Aba) solxrula grasu sunflower oil (Tur,Kor,Hun,Udm,Aba) bifce jisra bee juice (Hop = honey) tatru litki breast liquid (Hop = milk) kanla djacu eye water (Kor = tear) sikcurnu = silk-worm solxrula = solar-flower

Conversely: the tertau specifies the source of a product, and the seltau specifies the product:

**Definition 5.19**

**silna jinto** salt well (Chi) kolme terkakpa coal mine (Chi) ctile jinto oil well (Chi) terkakpa = source of digging

The tertau specifies an object, and the seltau specifies the material from which the object is made. This case is especially interesting, because the referent of the tertau may normally be made from just one kind of material, which is then overridden in the tanru.

**Definition 5.20**

**rokci cinfo** stone lion snime nanmu snow man (Hun) kliti cipni clay bird blaci kanla glass eye (Hun) blaci kanla glass eye (Que = spectacles) solji sicni gold coin (Tur) solji junla gold watch (Tur,Kor,Hun) solji djine gold ring (Udm,Aba,Que) rokci zdani stone house (Imb) mudri zdani wood house (Ewe = wooden house) rokci bitmu stone wall (Ewe) solji carce gold chariot (Skt) mudri xarci wood weapon (Skt = wooden weapon) cmaro'i dargu pebble road (Chi) sudysrasu cutci straw shoe (Chi) cmaro'i = small-rock sudysrasu = dry-grass

Note: the two senses of *blaci kanla* can be discriminated as:

**Definition 5.21**

**blaci kanla bo tarmi** glass (eye shape) = glass eye blaci kanla bo sidju glass (eye helper) = spectacles

The tertau specifies a typical object used to measure a quantity and the seltau specifies something measured. The tanru as a whole refers to a given quantity of the thing being measured. English does not have compounds of this form, as a rule.

**Definition 5.22**

**tumla spisa** land piece (Tur = piece of land) tcati kabri tea cup (Kor,Aba = cup of tea) nanba spisa bread piece (Kor = piece of bread) bukpu spisa cloth piece (Udm,Aba = piece of cloth) djacu calkyguzme water calabash (Ewe = calabash of water) calkyguzme = shell-fruit, calabash

The tertau specifies an object with certain implicit properties, and the seltau overrides one of those implicit properties:

**Definition 5.23**

**kensa bloti** spaceship bakni verba cattle child (Ewe = calf)

The seltau specifies a whole, and the tertau specifies a part which normally is associated with a different whole. The tanru then refers to a part of the seltau which stands in the same relationship to the whole seltau as the tertau stands to its typical whole.

**Definition 5.24**

**kosta degji** coat finger (Hun = coat sleeve) denci genja tooth root (Imb) tricu stedu tree head (Imb = treetop)

The tertau specifies the producer of a certain product, and the seltau specifies the product. In this way, the tanru as a whole distinguishes its referents from other referents of the tertau which do not produce the product.

**Definition 5.25**

**silka curnu** silkworm (Tur,Hun,Aba)

The tertau specifies an object, and the seltau specifies another object which has a characteristic property. The tanru as a whole refers to those referents of the tertau which possess the property.

**Definition 5.26**

**sonci manti** soldier ant ninmu bakni woman cattle (Imb = cow) mamta degji mother finger (Imb = thumb) cifnu degji baby finger (Imb = pinky) pacraistu zdani hell house (Skt) fagri dapma fire curse (Skt = curse destructive as fire) pacraistu = evil-superlative-site

As a particular case (when the property is that of resemblance): the seltau specifies an object which the referent of the tanru resembles.

**Definition 5.27**

**grutrceraso jbama** cherry bomb solji kerfa gold hair (Hun = golden hair) kanla djacu eye water (Kar = spring) bakni rokci bull stone (Mon = boulder) grutrceraso = fu'ivla for *cherry* based on Linnean name sorprema'e = many-person-vehicle

The seltau specifies a place, and the tertau an object characteristically located in or at that place.

**Definition 5.28**

**ckana boxfo** bed sheet (Chi) mrostu mojysu'a tomb monument (Chi = tombstone) jubme tergusni table lamp (Chi) foldi smacu field mouse (Chi) briju ci'ajbu office desk (Chi) rirxe xirma river horse (Chi = hippopotamus) xamsi gerku sea dog (Chi = seal) cagyce'u zdani village house (Skt) mrostu = dead-site mojysu'a = remember-structure ci'ajbu = write-table cagyce'u = farm-community

Specifically: the tertau is a place where the seltau is sold or made available to the public.

**Definition 5.29**

**cidja barja** food bar (Chi = restaurant) cukta barja book bar (Chi = library)

The seltau specifies the locus of application of the tertau.

**Definition 5.30**

**kanla velmikce** eye medicine (Chi) jgalu grasu nail oil (Chi = nail polish) denci pesxu tooth paste (Chi) velmikce = treatment used by doctor

The tertau specifies an implement used in the activity denoted by the seltau.

**Definition 5.31**

**me la pinpan. bolci** Ping-Pong ball (Chi)

The tertau specifies a protective device against the undesirable features of the referent of the seltau.

**Definition 5.32**

**carvi mapku** rain cap (Chi) carvi taxfu rain garment (Chi = raincoat) vindu firgai poison mask (Chi = gas mask) firgai = face-cover

The tertau specifies a container characteristically used to hold the referent of the seltau.

**Definition 5.33**

**cukta vasru** book vessel (Chi = satchel) vanju kabri wine cup (Chi) spatrkoka lanka coca basket (Que) djacu calkyzme water calabash (Ewe) rismi dakli rice bag (Ewe,Chi) tcati kabri tea cup (Chi) ladru botpi milk bottle (Chi) rismi patxu rice pot (Chi) festi lante trash can (Chi) bifce zdani bee house (Kor = beehive) cladakyxa'i zdani sword house (Kor = sheath) manti zdani ant nest (Gua = anthill) spatrkoka = fu'ivla for *coca* calkyzme = shell-fruit, calabash cladakyxa'i = (long-knife)-weapon

The seltau specifies the characteristic time of the event specified by the tertau.

**Definition 5.34**

**vensa djedi** spring day (Chi) crisa citsi summer season (Chi) cerni bumru morning fog (Chi) critu lunra autumn moon (Chi) dunra nicte winter night (Chi) nicte ckule night school (Chi)

The seltau specifies a source of energy for the referent of the tertau.

**Definition 5.35**

**dikca tergusni** electric lamp (Chi) ratni nejni atom energy (Chi) brife molki windmill (Tur,Kor,Hun,Udm,Aba) tergusni = illumination-source

Finally, some tanru which don't fall into any of the above categories.

**Definition 5.36**

**ladru denci** milk tooth (Tur,Hun,Udm,Qab) kanla denci eye tooth

It is clear that *tooth* is being specified, and that *milk* and *eye* act as modifiers. However, the relationship between *ladru* and *denci* is something like "tooth which one has when one is drinking milk from one's mother", a relationship certainly present nowhere except in this particular concept. As for *kanla denci*, the relationship is not only not present on the surface, it is hardly possible to formulate it at all.

## 5.15 Some types of symmetrical tanru

This section deals with symmetrical tanru, where order is not important. Many of these tanru can be expressed with a logical or non-logical connective between the components.

The tanru may refer to things which are correctly specified by both tanru components. Some of these instances may also be seen as asymmetrical tanru where the seltau specifies a material. The connective *je* is appropriate:

**Definition 5.37**

**cipnrstrigi pacru'i** owl demon (Skt) nolraitru prije royal sage (Skt) remna nakni human-being male (Qab = man) remna fetsi human-being female (Qab = woman) sonci tolvri soldier coward (Que) panzi nanmu offspring man (Ewe = son) panzi ninmu offspring woman (Ewe = daughter) solji sicni gold coin (Tur) solji junla gold watch (Tur,Kor,Hun) solji djine gold ring (Udm,Aba,Que) rokci zdani stone house (Imb) mudri zdani wooden house (Ewe) rokci bitmu stone wall (Ewe) solji carce gold chariot (Skt) mudri xarci wooden weapon (Skt) zdani tcadu home town (Chi) cipnrstrigi = fu'ivla for *owl* based on Linnean name pacru'i = evil-spirit tolvri = opposite-of-brave

The tanru may refer to all things which are specified by either of the tanru components. The connective *ja* is appropriate:

**Definition 5.38**

**nunji'a nunterji'a** victory defeat (Skt = victory or defeat) donri nicte day night (Skt = day and night) lunra tarci moon stars (Skt = moon and stars) patfu mamta father mother (Imb,Kaz,Chi = parents) tuple birka leg arm (Kaz = extremity) nuncti nunpinxe eating drinking (Udm = cuisine) bersa tixnu son daughter (Chi = children) nunji'a = event-of-winning nunterji'a = event-of-losing nuncti = event-of-eating nunpinxe = event-of-drinking

Alternatively, the tanru may refer to things which are specified by either of the tanru components or by some more inclusive class of things which the components typify:

**Definition 5.39**

**curnu jalra** worm beetle (Mon = insect) jalra curnu beetle worm (Mon = insect) kabri palta cup plate (Kaz = crockery) jipci gunse hen goose (Qab = housefowl) xrula tricu flower tree (Chi = vegetation)

The tanru components specify crucial or typical parts of the referent of the tanru as a whole:

**Definition 5.40**

**tumla vacri** land air (Fin = world) moklu stedu mouth head (Aba = face) sudysrasu cunmi hay millet (Qab = agriculture) gugde ciste state system (Mon = politics) prenu so'imei people multitude (Mon = masses) djacu dertu water earth (Chi = climate) sudysrasu = dry-grass so'imei = manysome

<h3> 16. *Pretty little girls' school*: forty ways to say it</h3>

The following examples show every possible grouping arrangement of *melbi cmalu nixli ckule* using *bo* or "ke ... ke'e" for grouping and *je* or *jebo* for logical connection. Most of these are definitely not plausible interpretations of the English phrase "pretty little girls' school", especially those which describe something which is both a girl and a school.

Examples 4.2, 4.3, 4.4, 4.5, and 5.6 are repeated here as Examples 16.1, 16.9, 16.17, 16.25, and 16.33 respectively. The seven examples following each of these share the same grouping pattern, but differ in the presence or absence of *je* at each possible site. Some of the examples have more than one Lojban version. In that case, they differ only in grouping mechanism, and are always equivalent in meaning.

The logical connective *je* is associative: that is, "A and (B and C)" is the same as *(A and B) and C*. Therefore, some of the examples have the same meaning as others. In particular, 16.8, 16.16, 16.24, 16.32, and 16.40 all have the same meaning because all four brivla are logically connected and the grouping is simply irrelevant. Other equivalent forms are noted in the examples themselves. However, if *je* were replaced by *naja* or *jo* or most of the other logical connectives, the meanings would become distinct.

It must be emphasized that, because of the ambiguity of all tanru, the English translations are by no means definitive — they represent only one possible interpretation of the corresponding Lojban sentence.

**Example 5.15.1**
melbi cmalu nixli ckule
((pretty type-of little) type-of girl) type-of school
school for girls who are beautifully small

**Example 5.15.2**
melbi je cmalu nixli ckule
((pretty and little) type-of girl) type-of school
school for girls who are beautiful and small

**Example 5.15.3**
melbi bo cmalu je nixli ckule
((pretty type-of little) and girl) type-of school
school for girls and for beautifully small things

**Example 5.15.4**
ke melbi cmalu nixli ke'e je ckule
((pretty type-of little) type-of girl) and school
thing which is a school and a beautifully small girl

**Example 5.15.5**
melbi je cmalu je nixli ckule
((pretty and little) and girl) type-of school
school for things which are beautiful, small, and girls
Note: same as 16.21

**Example 5.15.6**
melbi bo cmalu je nixli je ckule
((pretty type-of little) and girl) and school
thing which is beautifully small, a school, and a girl
Note: same as 16.14

**Example 5.15.7**
ke melbi je cmalu nixli ke'e je ckule
((pretty and little) type-of girl) and school
thing which is a school and a girl who is both
beautiful and small

**Example 5.15.8**
melbi je cmalu je nixli je ckule
((pretty and little) and girl) and school
thing which is beautiful, small, a girl, and a school

**Example 5.15.9**
melbi cmalu nixli bo ckule
(pretty type-of little) type-of (girl type-of school)
girls' school which is beautifully small

**Example 5.15.10**
melbi je cmalu nixli bo ckule
(pretty and little) type-of (girl type-of school)
girls' school which is beautiful and small

**Example 5.15.11**
melbi cmalu nixli je ckule
(pretty type-of little) type-of (girl and school)
something which is a girl and a school
which is beautifully small

**Example 5.15.12**
melbi bo cmalu je nixli bo ckule
(pretty type-of little) and (girl type-of school)
something which is beautifully small and a girls' school

**Example 5.15.13**
melbi je cmalu nixli je ckule
(pretty and little) type-of (girl and school)
a pretty and little type of thing which is
both a girl and a school

**Example 5.15.14**
melbi bo cmalu je nixli jebo ckule
(pretty type-of little) and (girl and school)
thing which is beautifully small, a school, and a girl
Note: same as 16.6

**Example 5.15.15**
melbi jebo cmalu je nixli bo ckule
(pretty and little) and (girl type-of school)
thing which is beautiful and small and a girl's school
Note: same as 16.30

**Example 5.15.16**
melbi jebo cmalu je nixli jebo ckule
(pretty and little) and (girl and school)
thing which is beautiful, small, a girl, and a school

**Example 5.15.17**
melbi cmalu bo nixli ckule
(pretty type-of (little type-of girl)) type-of school
school for beautiful girls who are small

**Example 5.15.18**
melbi cmalu je nixli ckule
(pretty type-of (little and girl)) type-of school
school for beautiful things which are small and are girls

**Example 5.15.19**
melbi je cmalu bo nixli ckule
(pretty and (little type-of girl)) type-of school
school for things which are beautiful and are small girls

**Example 5.15.20**
ke melbi cmalu bo nixli ke'e je ckule
melbi bo cmalu bo nixli je ckule
(pretty type-of (little type-of girl)) and school
thing which is a school and a small girl who is beautiful

**Example 5.15.21**
melbi je cmalu jebo nixli ckule
(pretty and (little and girl)) type-of school
school for things which are beautiful, small, and girls
Note: same as 16.5

**Example 5.15.22**
melbi je cmalu bo nixli je ckule
(pretty and (little type-of girl)) and school
thing which is beautiful, a small girl, and a school
Note: same as 16.38

**Example 5.15.23**
ke melbi cmalu je nixli ke'e je ckule
(pretty type-of (little and girl)) and school
thing which is beautifully small, a beautiful girl,
and a school

**Example 5.15.24**
melbi je cmalu jebo nixli je ckule
(pretty and (little and girl)) and school
thing which is beautiful, small, a girl, and a school

**Example 5.15.25**
melbi cmalu bo nixli bo ckule
melbi ke cmalu ke nixli ckule ⟨*ke'e*⟩ [ke'e]
pretty type-of (little type-of (girl type-of school))
small school for girls which is beautiful

**Example 5.15.26**
melbi ke cmalu nixli je ckule ⟨*ke'e*⟩
pretty type-of (little type-of (girl and school))
small thing, both a girl and a school, which is beautiful

**Example 5.15.27**
melbi cmalu je nixli bo ckule
pretty type-of (little and (girl type-of school))
thing which is beautifully small
and a girls' school that is beautiful

**Example 5.15.28**
melbi je cmalu bo nixli bo ckule
melbi je ke cmalu nixli bo ckule ⟨*ke'e*⟩
melbi je ke cmalu ke nixli ckule [ke'e] [ke'e]
pretty and (little type-of (girl type-of school))
thing which is beautiful and a small type of
girls' school

**Example 5.15.29**
melbi cmalu je nixli jebo ckule
melbi cmalu je ke nixli je ckule ⟨*ke'e*⟩
pretty type-of (little and (girl and school))
thing which is beautifully small, a beautiful girl,
and a beautiful school
Note: same as 16.37

**Example 5.15.30**
melbi je cmalu jebo nixli bo ckule
melbi je ke cmalu je nixli bo ckule ⟨*ke'e*⟩
pretty and (little and (girl type-of school))
thing which is beautiful, small and a girls' school
Note: same as 16.15

**Example 5.15.31**
melbi je ke cmalu nixli je ckule ⟨*ke'e*⟩
pretty and (little type-of (girl and school))
beautiful thing which is a small girl and a small school

**Example 5.15.32**
melbi jebo cmalu jebo nixli jebo ckule
pretty and (little and (girl and school))
thing which is beautiful, small, a girl, and a school

**Example 5.15.33**
melbi ke cmalu nixli ckule ⟨*ke'e*⟩
pretty type-of ((little type-of girl) type-of school)
beautiful school for small girls

**Example 5.15.34**
melbi ke cmalu je nixli ckule ⟨*ke'e*⟩
pretty type-of ((little and girl) type-of school
beautiful school for things which are small
and are girls

**Example 5.15.35**
melbi ke cmalu bo nixli je ckule ⟨*ke'e*⟩
pretty type-of ((little type-of girl) and school)
beautiful thing which is a small girl and a school

**Example 5.15.36**
melbi je ke cmalu nixli ckule ⟨*ke'e*⟩
pretty and ((little type-of girl) type-of school)
thing which is beautiful and a school for small girls

**Example 5.15.37**
melbi cmalu je nixli je ckule
pretty type-of ((little and girl) and school)
thing which is beautifully small, a beautiful girl,
and a beautiful school
Note: same as 16.29

**Example 5.15.38**
melbi je ke cmalu bo nixli je ckule ⟨*ke'e*⟩
pretty and ((little type-of girl) and school)
thing which is beautiful, a small girl and a school
Note: same as 16.22

**Example 5.15.39**
melbi je ke cmalu je nixli ckule ⟨*ke'e*⟩
pretty and ((little and girl) type-of school)
thing which is beautiful and is a small school
and a girls' school

**Example 5.15.40**
melbi je ke cmalu je nixli je ckule ⟨*ke'e*⟩
pretty and ((little and girl) and school)
thing which is beautiful, small, a girl, and a school