



The Complete Lojban Language

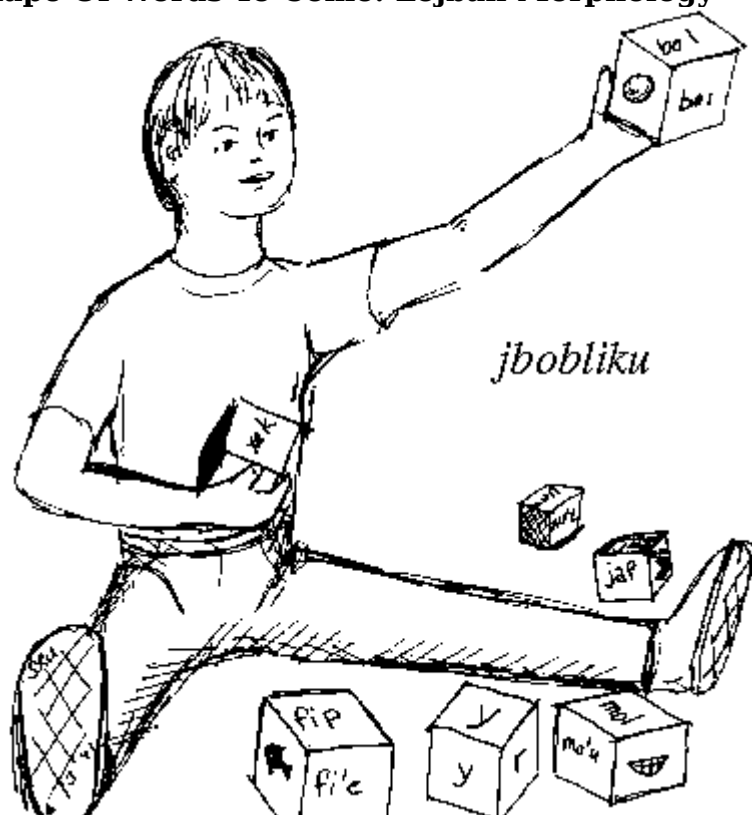
John Woldemar Cowan

A Logical Language Group Publication

Table of Contents

1. The Shape Of Words To Come: Lojban Morphology	4
1.1. Introductory	4
1.2. cmavo	6
1.3. brivla	8
1.4. gismu	8
1.5. lujvo	10
1.6. rafsi	12
1.7. fu'ivla and zi'evla	16
1.8. cmevla	19
1.9. Rules for inserting pauses	23
1.10. Considerations for making lujvo	23
1.11. The lujvo-making algorithm	24
1.12. The lujvo scoring algorithm	25
1.13. lujvo-making examples	26
1.14. The gismu creation algorithm	28
1.15. Cultural and other non-algorithmic gismu	30
1.16. rafsi fu'ivla: a proposal	33
Lojban Word Glossary	35
General Index	38
Lojban Words Index	43
Examples Index	45

Chapter 1 The Shape Of Words To Come: Lojban Morphology



1.1 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 it does 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. (A detailed algorithm for breaking sounds into words is part of the PEG grammar at the end of the book.) First, here are some conventions used to talk about groups of Lojban letters, including vowels and consonants.

1.1 Introductory

1. V represents any single Lojban vowel except *y*; that is, it represents *a*, *e*, *i*, *o*, or *u*.
2. VV represents a (falling) diphthong, one of the following:

<i>ai</i>	<i>ei</i>	<i>oi</i>	<i>au</i>
-----------	-----------	-----------	-----------
3. (“Vowel” as used in this chapter means any V or VV.)
4. V'V represents a two-syllable vowel pair with an apostrophe separating the vowels, one of the following:

<i>a'a</i>	<i>a'e</i>	<i>a'i</i>	<i>a'o</i>	<i>a'u</i>
<i>e'a</i>	<i>e'e</i>	<i>e'i</i>	<i>e'o</i>	<i>e'u</i>
<i>i'a</i>	<i>i'e</i>	<i>i'i</i>	<i>i'o</i>	<i>i'u</i>
<i>o'a</i>	<i>o'e</i>	<i>o'i</i>	<i>o'o</i>	<i>o'u</i>
<i>u'a</i>	<i>u'e</i>	<i>u'i</i>	<i>u'o</i>	<i>u'u</i>

5. 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* count as consonants for the purposes of this chapter.
6. G represents a single Lojban on-glide, one of *i*, or *u*.
7. CC represents two adjacent consonants of type C which constitute one of the 48 permissible initial consonant pairs:

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

8. An onset is a single C or G, a CC string as shown above, or a permissible initial consonant triple (CCC string), one of:

<i>cfr</i>	<i>cfl</i>	<i>sfr</i>	<i>sfl</i>	<i>jvr</i>	<i>jvl</i>	<i>zvr</i>	<i>zvl</i>
<i>cpr</i>	<i>cpl</i>	<i>spr</i>	<i>spl</i>	<i>jbr</i>	<i>jbl</i>	<i>zbr</i>	<i>zbl</i>
<i>ctr</i>	<i>str</i>	<i>jdr</i>	<i>zdr</i>				
<i>ckr</i>	<i>ckl</i>	<i>skr</i>	<i>skl</i>	<i>jgr</i>	<i>jgl</i>	<i>zgr</i>	<i>zgl</i>
<i>cmr</i>	<i>cml</i>	<i>smr</i>	<i>sml</i>	<i>jmr</i>	<i>jml</i>	<i>zmr</i>	<i>zml</i>

9. 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 Section 1.1 (p. 4). In brief, any consonant pair is permissible unless it: contains two identical letters, contains both a voiced (excluding *r*, *l*, *m*, *n*) and an unvoiced consonant, or is one of certain specified forbidden pairs.
10. 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 *cmevla*. 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

The Complete Lojban Language

content words, corresponding to English words like “come”, “red”, “doctor”, and “freely”; *cmevla* correspond to English proper names like “James”, “Afghanistan”, and “Pope John Paul II”.

1.2 *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 1 (p. 4) to Chapter 1 (p. 4) and summarized in Chapter 1 (p. 4).

A *cmavo* is composed of a vowel or *y*, with optionally a consonant or glide before it and any number of 'V, 'VV, or 'y syllables after it. Here are some *cmavo* of various shapes:

.a | *di* | *.oi* | *iu* | *.u'u* | *fau* | *mi'o* | *ja'ai* | *by*.

In addition, there is the *cmavo* *.ybu*, the only one of its shape.

There are some conventions by which a *cmavo*'s shape can hint at its grammar or meaning: for example, *cmavo* beginning with vowel letters are usually attitudinal. In general, though, the form of a *cmavo* tells you little or nothing about how it is used.

All *cmavo* longer than two syllables, and most two-syllable *cmavo* containing diphthongs, are not officially defined and are reserved for experimental use. “Experimental use” means that the language designers have not assigned any standard meaning or usage to these words. Experimental-use provide an escape hatch for adding grammatical mechanisms (as opposed to semantic concepts) the need for which was not foreseen.

cmavo of the form *Cy*, a consonant followed by the letter *y*, together with *bu* and *.ybu*, represent letters of the Lojban alphabet, and are discussed in detail in Chapter 1 (p. 4). Sequences of *cmavo* ending in *y* need to be separated from other words following them by a pause.

cmavo clusters are sequences of *cmavo* attached together to form a single written word. *cmavo* written together are always identical in meaning and in grammatical use to the corresponding separated sequence of simple *cmavo*. 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.

cmavo clusters, 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 1.1

.iseju
.i se ju

Example 1.2

punaijecanai
pu nai je ca nai

Example 1.3

ki'a.u'e
ki'a .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. On the other hand,

Example 1.4

ki'au'e and *ki'a'u'e*

are both single cmavo reserved for experimental purposes.

Example 1.5

cy.ibu.abu
cy .ibu .abu

Again the pauses are required (see Section 1.1 (p. 4)); the pauses after *cy.* and before *.ibu* merge together (pause length is not phonemic).

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 1.6

.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 word following it is a brivla, you must insert a pause or glottal stop between the two stressed syllables. Thus

Example 1.7

lo re nanmu

can be optionally pronounced

Example 1.8

lo 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 1.9

LO re NANmu

is also valid; this would apply secondary stress (used for purposes of emphasis, contrast or sentence rhythm) to *lo*, comparable in rhythmical effect to the English phrase "THE two men". In Example 1.8 (p. 7), 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 1.10

lo re NANmu

This would probably be the most common usage.

1.3 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 cmevla) 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: they

1. always end in a vowel;
2. contain a consonant cluster, or start with a CVCy string;
3. and are always 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. The final vowel distinguishes brivla from cmevla, which always end in a consonant. Thus *da'amei* must be a string of cmavo because it lacks a consonant pair; *.lojban.* must be a name because it lacks a final vowel.

The three subtypes of brivla are:

1. gismu, the Lojban primitive roots from which other brivla are built;
2. lujvo, the compounds of two or more gismu; and
3. zi'evla (literally “free-word”), the words that are not Lojban primitives or natural compounds, and are often borrowed from other languages.

1.4 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 1.1 (p. 4), all gismu are of the forms CVC/CV or CCVCV. They must meet the rules for all brivla given in Section 1.1 (p. 4); furthermore, they:

1. always have five letters;
2. always start with a consonant and end with a single vowel;

1.4 gismu

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).

In addition, no gismu contains ' or a diphthong.

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 1.1 (p. 4) for a full explanation of the algorithm.

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

Example 1.11

creka
shirt

Example 1.12

lijda
religion

Example 1.13

blanu
blue

Example 1.14

mamta
mother

Example 1.15

cukta
book

Example 1.16

patfu
father

Example 1.17

nanmu
man

Example 1.18

ninmu
woman

A small number of gismu were formed differently; see Section 1.1 (p. 4) for a list.

1.5 lujvo

When specifying a concept that is not found among existing brivla (or, more specifically, when the relevant word seems too general in meaning), a Lojbanist generally first attempts to express the concept as a tanru. 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 1.19

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 1 (p. 4).

To express a simple tanru, simply say the component gismu together. Thus “big boat” becomes the tanru

Example 1.20

barda bloti

representing roughly the same concept as the English word “ship”.

“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 1.21

patfu mamta

and

Example 1.22

mamta patfu

respectively.

The possibility of semantic ambiguity can easily be seen in the last case. To interpret Example 1.22 (p. 10), 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.

1.5 lujvo

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 1.19 (p. 10), we would probably choose “user of computers”, and form the new word

Example 1.23

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 1.20 (p. 10), would be:

Example 1.24

brablo

“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 1.22 (p. 10) is

Example 1.25

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 relationship between the terms. (See Chapter 1 (p. 4) 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*, has the same definition 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 1.1 (p. 4).

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 1.26

soirsai
from *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.

1.6 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 1.1 (p. 4)).

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 1.27

mamtypatfu
from *mamta patfu*
“mother father” or “maternal grandfather”

Example 1.28

lerfyliste
from *lerfu liste*
“letter list” or a “list of letters”
(letters of the alphabet)

Example 1.29

nancyprali
from *nanca prali*
“year profit” or “annual profit”

Example 1.30

prunyplipe
from *pruni plipe*
“elastic (springy) leap” or “spring” (the verb)

Example 1.31

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.

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:

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.)

The Complete Lojban Language

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 1.1 (p. 4)). 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.

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

Example 1.32

cumfri
from *cumki lifri*
“possible experience”

Example 1.33

klezba
from *klesi zbasu*
“category make”

Example 1.34

kixta'a
from *krixa tavla*
“cry-out talk”

Example 1.35

sniju'o
from *sinxa djuno*
“sign know”

In addition, the unreduced forms in Example 1.27 (p. 12) and Example 1.28 (p. 12) may be fully reduced to:

Example 1.36

mampa'u
from *mamta patfu*
“mother father” or “maternal grandfather”

Example 1.37

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 a vowel. As a brivla, a lujvo must also contain a consonant cluster within the first five letters – this ensures that they cannot be mistaken for 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 a *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.)

1.6 rafsi

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 1.27 (p. 12) to Example 1.31 (p. 13).)

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 1.38

bridi zei vals

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

Example 1.39

xy. zei kantu
X ray

Example 1.40

kulnrfarsi zei lolgai
“Farsi floor-cover”
Persian rug

Example 1.41

na'e zei .a zei na'e zei by. livgyterbilma
“non-A, non-B liver-disease”
non-A, non-B hepatitis

Example 1.42

.cerman. zei jamkarce
"Sherman war-car"
Sherman tank

Example 1.41 (p. 15) is particularly noteworthy because the phrase that would be produced by removing the *zeis* 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 1.43

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 1.43 (p. 16) 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; *rafsi* appear only as parts of *lujvo*.

1.7 fu'ivla and zi'evla

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 Section 1.1 (p. 4)):

Example 1.44

me la'o ly. spaghetti .ly.

is a predicate with the place structure "x1 is a quantity of spaghetti".

Stage 2 involves changing the foreign name to a Lojbanized name, as explained in Section 1.1 (p. 4):

Example 1.45

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 *brivla* form. A *rafsi* is often 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 word, since *brivla* are not permitted to have more than one definition.

1.7 fu'ivla and zi'evla

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 1.1 (p. 4) 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 vals*) 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 a vowel;
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”, not discussed further in this book);
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. Prefix the result of steps 1-4 with a 4-letter rafsi that categorizes the fu'ivla into a “topic area”. It is only safe to use a single 4-letter or CVC-form rafsi; other 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), or if the rafsi ends in “r” and the rest of the fu'ivla begins with “tc”, “ts”, “dj”, or “dz” (using “n” would result in a phonotactically impermissible cluster), use an l-hyphen. (This is the only use of l-hyphen in Lojban.)

The Complete Lojban Language

6. Remember that the stress necessarily appears on the penultimate (next-to-the-last) syllable.

In this section, the hyphen syllable is set off with commas in the examples, but these commas are not required in writing.

Here are a few examples:

Example 1.46

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 1.47

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 1.48

brie (from French)
bri (Lojbanize)
cir,lr,bri (prefix rafsi)

where *cirl-* represents *cirla* (“cheese”).

Example 1.49

cobra
kobra (Lojbanize)
sin,cr,kobra (prefix rafsi)

where *sinc-* represents *since* (“snake”).

Example 1.50

quark
ku'ark (Lojbanize)
ku'arka (add final vowel)
sas,kr,ku'arka (prefix rafsi)

where *sask-* represents *saske* (“science”). Note the extra vowel *a* added to the end of the word.

Example 1.51

자모 (from Korean)
djamo (Lojbanize)
ler,fr,djamo (prefix rafsi)
le,rl,djamo (prefix rafsi)

1.8 cmevla

where *ler-* represents *lerfu* (“letter”). Note the l-hyphen in “lerldjamo”, since “lerndjamo” contains the forbidden cluster “ndj”.

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 *ku'arka* are valid Stage 4 fu'ivla, but *xaceru* looks like a sequence of 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 *cmacrntintegrale*, 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 *dinjrntintegrale* or *tarmrntintegrale*, 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 1.52

ban,gr,blgara

Bulgarian (in language)

Example 1.53

kul,nr,blgara

Bulgarian (in culture)

Example 1.54

gug,dr,blgara

Bulgaria (the country)

Example 1.55

ban,gr,kore'a

Korean (the language)

Example 1.56

kul,nr,kore'a

Korean (the culture)

Note the apostrophes in Example 1.55 (p. 19) and Example 1.56 (p. 19), 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”.

1.8 cmevla

Lojbanized names, called *cmevla*, are mostly used to Lojbanize names, in other words, they are used to create 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

The Complete Lojban Language

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 1.57

.djim.
Jim

Example 1.58

.djein.
Jane

Example 1.59

.arnold.
Arnold

Example 1.60

.pit.
Pete

Example 1.61

.katrinas.
Katrina

Example 1.62

.katyrin.
Catherine

(Note that *y* is skipped in determining the stressed syllable, so Example 1.62 (p. 20) is stressed on the *ka*.)

Example 1.63

.katis.
Cathy

Example 1.64

.keit.
Kate

Names may have almost any form, but always end in a consonant, and are both preceded and followed by a pause. They are penultimately stressed, unless unusual stress is marked with capitalization. A name may have multiple parts; these can be expressed as multiple words each ending with a consonant and pause, or the parts may be combined into a single word. For example,

Example 1.65

.djan. braun.

and

Example 1.66

.djanbraun.

are both valid Lojbanizations of “John Brown”.

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

1.8 cmevla

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.

Lojban cmevla 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 cmevla, provided that each consonant pair (whether standing by itself, or as part of a larger cluster) is a permissible pair.
2. They are always preceded and followed in speech by pauses, written as ..
3. 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.

Cmevla meeting these criteria may be invented, Lojbanized from names in other languages, or formed by appending a consonant onto a cmavo or brivla. Some cmevla built from Lojban words are:

Example 1.67

.pav.

the One

from the cmavo *pa*, with rafsi *pav*, meaning “one”

Example 1.68

.sol.

the Sun

from the gismu *solri*, meaning “sun”

Example 1.69

.ral.

Chief (as a title)

from the gismu *ralju*, meaning “principal”.

Example 1.70

.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 capitalization in written Lojban when it is necessary to preserve non-standard stress. Do not capitalize names otherwise.
5. If the name contains an impermissible consonant pair, insert a vowel between the consonants: *y* is recommended.

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 Section 1.1 (p. 4). 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

The Complete Lojban Language

in converting Linnaean binomials 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:

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:

English	"Mary"	. <i>meris</i> . or . <i>meiris</i> .
English	"Smith"	. <i>smit</i> .
English	"Jones"	. <i>djonz</i> .
English	"John"	. <i>djan</i> . or . <i>jan</i> . (American) or . <i>djon</i> . or . <i>jon</i> . (British)
English	"Alice"	. <i>alis</i> .
English	"Elise"	. <i>eLIS</i> .
English	"Johnson"	. <i>djansn</i> .
English	"William"	. <i>uiliiam</i> . or . <i>uili'am</i> .
English	"Brown"	. <i>braun</i> .
English	"Charles"	. <i>tcarlz</i> .
French	"Charles"	. <i>carl</i> .
French	"De Gaulle"	. <i>dyGOL</i> .
German	"Heinrich"	. <i>xainrix</i> .
Spanish	"Joaquin"	. <i>xouaKIN</i> .
Russian	"Svetlana"	. <i>sfetlanys</i> .
Russian	"Khrushchev"	. <i>xrucTCOF</i> .
Hindi	"Krishna"	. <i>kricnas</i> .

1.9 Rules for inserting pauses

Polish	“Lech Walesa”	<i>.lex. vauensas.</i>
Spanish	“Don Quixote”	<i>.don. kicotes.</i> or modern Spanish: <i>.don. kixotes.</i> or Mexican dialect: <i>.don. ki'otes.</i>
Chinese	“Mao Zedong”	<i>.maudzydun.</i>
Japanese	“Fujiko”	<i>.fudjikos.</i> or <i>.fujikos.</i>

1.9 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 preceded and followed by pauses. Necessarily, all such words are *cmevla*.
3. Every word beginning with a vowel must be preceded by a pause. Such words are either *cmavo*, *fu'ivla*, or *cmevla*; all *gismu* and *lujvo* begin with consonants.
4. 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 *cmevla* with unusual stress (which already ends with a pause, of course).
5. A *cmavo* of the form “*Cy*” must be followed by a pause unless another “*Cy*”-form *cmavo* follows.
6. 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 Section 1.1 (p. 4).)

1.10 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 1.1 (p. 4).

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 1.1 (p. 4) 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 Section 1.1 (p. 4)) 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 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 1 (p. 4); abstraction is explained in Chapter 1 (p. 4).) Plausibility is the key to learning new ideas and to evaluating unfamiliar *lujvo*.

1.11 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*.

1.12 The lujvo scoring algorithm

3. Join the resulting string of rafsi, initially without hyphens.
4. Add hyphen letters where necessary. 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 up to the first y-hyphen, or up to the end of the word in case there are no y-hyphens. 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.

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

The Complete Lojban Language

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

The score is then:

$$(1000 * L) - (500 * A) + (100 * H) - (10 * R) - V$$

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 1.71

zbasai

zba + sai

$$(1000 * 6) - (500 * 0) + (100 * 0) - (10 * 15) - 3 = 5847$$

Example 1.72

nunynau

nun + y + nau

$$(1000 * 7) - (500 * 0) + (100 * 1) - (10 * 13) - 3 = 6967$$

Example 1.73

sairzbata'u

sai + r + zba + ta'u

$$(1000 * 11) - (500 * 1) + (100 * 1) - (10 * 21) - 5 = 10385$$

Example 1.74

zbazbasysarji

zba + zbas + y + sarji

$$(1000 * 13) - (500 * 0) + (100 * 1) - (10 * 12) - 4 = 12976$$

1.13 *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 1.1 (p. 4), using every possible rafsi.

The rafsi for *gerku* are:

-ger-, *-ge'u-*, *-gerk-*, *-gerku*

The rafsi for *zdani* are:

-zda-, *-zdan-*, *-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

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

1.13 lujvo-making examples

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:

<i>lotkle</i>	<i>blokle</i>	<i>lo'ikle</i>
<i>lotlei</i>	<i>blolei</i>	<i>lo'irlei</i>

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:

<i>lotkle</i>	5878
<i>blokle</i>	5858
<i>lo'ikle</i>	6367
<i>lotlei</i>	5867
<i>blolei</i>	5847
<i>lo'irlei</i>	7456

So the form *blolei* is preferred, but only by a tiny margin over *blokle*; "lotlei" and "lotkle" 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:

<i>loj -ban -gri</i>	<i>loj -bau -gri</i>	<i>loj -bang -gri</i>
<i>logj -ban -gri</i>	<i>logj -bau -gri</i>	<i>logj -bang -gri</i>
<i>loj -ban -girzu</i>	<i>loj -bau -girzu</i>	<i>loj -bang -girzu</i>
<i>logj -ban -girzu</i>	<i>logj -bau -girzu</i>	<i>logj -bang -girzu</i>

and the 12 name possibilities are:

<i>loj -ban -gir</i>	<i>loj -bau -gir</i>	<i>loj -bang -gir</i>
<i>logj -ban -gir</i>	<i>logj -bau -gir</i>	<i>logj -bang -gir</i>
<i>loj -ban -girz</i>	<i>loj -bau -girz</i>	<i>loj -bang -girz</i>
<i>logj -ban -girz</i>	<i>logj -bau -girz</i>	<i>logj -bang -girz</i>

The Complete Lojban Language

After hyphenation, we have:

<i>lojbangri</i>	<i>lojbaugri</i>	<i>lojbangygri</i>
<i>logjybangri</i>	<i>logjybaugri</i>	<i>logjybangygri</i>
<i>lojbangirzu</i>	<i>lojbaugirzu</i>	<i>lojbangygirzu</i>
<i>logjybangirzu</i>	<i>logjybaugirzu</i>	<i>logjybangygirzu</i>
<i>lojbangir</i>	<i>lojbaugir</i>	<i>lojbangygir</i>
<i>logjybangir</i>	<i>logjybaugir</i>	<i>logjybangygir</i>
<i>lojbangirz</i>	<i>lojbaugirz</i>	<i>lojbangygirz</i>
<i>logjybangirz</i>	<i>logjybaugirz</i>	<i>logjybangygirz</i>

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:

<i>nak -kem -cin -ctu</i>	<i>nakykemcinctu</i>
<i>nak -kem -cin -ctuca</i>	<i>nakykemcinctuca</i>
<i>nak -kem -cins -ctu</i>	<i>nakykemcinsyctu</i>
<i>nak -kem -cins -ctuca</i>	<i>nakykemcinsyctuca</i>
<i>nakn -kem -cin -ctu</i>	<i>naknykemcinctu</i>
<i>nakn -kem -cin -ctuca</i>	<i>naknykemcinctuca</i>
<i>nakn -kem -cins -ctu</i>	<i>naknykemcinsyctu</i>
<i>nakn -kem -cins -ctuca</i>	<i>naknykemcinsyctuca</i>

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.

1.14 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

1.14 The gismu creation algorithm

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.

proposed gismu	existing gismu
<i>b</i>	<i>p, v</i>
<i>c</i>	<i>j, s</i>
<i>d</i>	<i>t</i>
<i>f</i>	<i>p, v</i>
<i>g</i>	<i>k, x</i>
<i>j</i>	<i>c, z</i>
<i>k</i>	<i>g, x</i>
<i>l</i>	<i>r</i>
<i>m</i>	<i>n</i>
<i>n</i>	<i>m</i>
<i>p</i>	<i>b, f</i>
<i>r</i>	<i>l</i>
<i>s</i>	<i>c, z</i>
<i>t</i>	<i>d</i>
<i>v</i>	<i>b, f</i>
<i>x</i>	<i>g, k</i>
<i>z</i>	<i>j, s</i>

See Section 1.1 (p. 4) for an example.

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 were as follows:

Chinese	0.36
English	0.21
Hindi	0.16
Spanish	0.11
Russian	0.09
Arabic	0.07

reflecting 1985 number-of-speakers data. A few gismu were made much later using updated weights:

The Complete Lojban Language

Chinese	0.347
Hindi	0.196
English	0.160
Spanish	0.123
Russian	0.089
Arabic	0.085

(English and Hindi switched places due to demographic changes.)

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.

1.15 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 rafsai assigned to them for use in building more complex *lujvo*. Assigning gismu to these concepts helps to keep the resulting *lujvo* reasonably short.

<i>broda</i>	1st assignable predicate
<i>brode</i>	2nd assignable predicate
<i>brodi</i>	3rd assignable predicate
<i>brodo</i>	4th assignable predicate
<i>brodu</i>	5th assignable predicate
<i>cmavo</i>	structure word (from <i>cmalu vals</i>)
<i>lojbo</i>	Lojbanic (from <i>logji bangu</i>)
<i>lujvo</i>	compound word (from <i>pluja vals</i>)
<i>mekso</i>	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.

Small metric prefixes (values less than 1):

<i>decti</i>	.1	deci
<i>centi</i>	.01	centi
<i>milti</i>	.001	milli
<i>mikri</i>	10^{-6}	micro
<i>nanvi</i>	10^{-9}	nano
<i>picti</i>	10^{-12}	pico
<i>femti</i>	10^{-15}	femto
<i>xatsi</i>	10^{-18}	atto
<i>zepti</i>	10^{-21}	zepto

1.15 Cultural and other non-algorithmic gismu

gocti : 10⁻²⁴ : yocto

Large metric prefixes (values greater than 1):

<i>dekto</i>	10	deka
<i>xecto</i>	100	hecto
<i>kilto</i>	1000	kilo
<i>megdo</i>	10 ⁶	mega
<i>gigdo</i>	10 ⁹	giga
<i>terto</i>	10 ¹²	tera
<i>petso</i>	10 ¹⁵	peta
<i>xexso</i>	10 ¹⁸	exa
<i>zetro</i>	10 ²¹	zetta
<i>gotro</i>	10 ²⁴	yotta

Other scientific or mathematical terms:

<i>delno</i>	candela
<i>kelvo</i>	kelvin
<i>molro</i>	mole
<i>radno</i>	radian
<i>sinso</i>	sine
<i>stero</i>	steradian
<i>tanjo</i>	tangent
<i>xampo</i>	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.

The six Lojban source languages:

<i>jungo</i>	Chinese (from "Zhōngguó")
<i>glico</i>	English
<i>xindo</i>	Hindi
<i>spano</i>	Spanish
<i>rusko</i>	Russian
<i>xrabo</i>	Arabic

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

<i>bengo</i>	Bengali
<i>porto</i>	Portuguese
<i>baxso</i>	Bahasa Melayu/Bahasa Indonesia
<i>ponjo</i>	Japanese (from "Nippon")
<i>dotco</i>	German (from „Deutsch“)
<i>fraso</i>	French (from « Français »)
<i>xurdo</i>	Urdu

The Complete Lojban Language

(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):

English:

merko : American
brito : British
skoto : Scottish
sralo : Australian
kadno : Canadian

Spanish:

gento : Argentinian
mexno : Mexican

Russian:

softo : Soviet/USSR
vukro : Ukrainian

Arabic:

filso : Palestinian
jerxo : Algerian
jordo : Jordanian
libjo : Libyan
lubno : Lebanese
misro : Egyptian (from “Misr”)
morko : Moroccan
rakso : Iraqi
sadjjo : Saudi
sirxo : Syrian

Bahasa Melayu/Bahasa Indonesia:

bindo : Indonesian
meljo : Malaysian

Portuguese:

brazo : Brazilian

Urdu:

kisto : Pakistani

The continents (and oceanic regions) of the Earth:

bemro : North American (from *berti merko*)
dzipo : Antarctic (from *cadzu cipni*)
ketco : South American (from “Quechua”)
friko : African
polno : Polynesian/Oceanic
ropno : European
xazdo : Asiatic

A few smaller but historically important cultures:

latmo : Latin/Roman
srito : Sanskrit
xebro : Hebrew/Israeli/Jewish
xelso : Greek (from “Hellas”)

Major world religions:

1.16 rafsi fu'ivla: a proposal

<i>budjo</i>	Buddhist
<i>dadjo</i>	Taoist
<i>muslo</i>	Islamic/Moslem
<i>xriso</i>	Christian

A few terms that cover multiple groups of the above:

<i>jegvo</i>	Jehovist (Judeo-Christian-Moslem)
<i>semto</i>	Semitic
<i>slovo</i>	Slavic
<i>xispo</i>	Hispanic (New World Spanish)

1.16 rafsi fu'ivla: a proposal

The list of cultures represented by gismu, given in Section 1.1 (p. 4), 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 1.1 (p. 4). For example,

Example 1.75

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 1.76

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

could be represented by the lujvo

Example 1.77

tci'ilykemocantutra

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,

The Complete Lojban Language

it might become standard for all Type 4 fu'ivla, including those made for plants, animals, foodstuffs, and other things.

Lojban Word Glossary

All definitions in this glossary are brief and unofficial. Only the published dictionary is a truly official reference for word definitions. These definitions are here simply as a quick reference.

a
placeholder definition

bakri
placeholder definition

baxso
placeholder definition

bemro
placeholder definition

bengo
placeholder definition

bilma
placeholder definition

bindo
placeholder definition

blaci
placeholder definition

blolei
placeholder definition

bloti
placeholder definition

bradi
placeholder definition

brazo
placeholder definition

bređi
placeholder definition

bridi
placeholder definition

brito
placeholder definition

brivla
placeholder definition

broda
placeholder definition

brode
placeholder definition

brođi
placeholder definition

brodo
placeholder definition

brođu
placeholder definition

bu
placeholder definition

budjo
placeholder definition

by
placeholder definition

centi
placeholder definition

cidja
placeholder definition

cidjrspageti
placeholder definition

cirla
placeholder definition

cmaci
placeholder definition

cmavo
placeholder definition

cmevla
placeholder definition

cy
placeholder definition

dadjo
placeholder definition

decti
placeholder definition

dekto
placeholder definition

delno
placeholder definition

đi
placeholder definition

dinju
placeholder definition

dotco
placeholder definition

dzipo
placeholder definition

fau
placeholder definition

fenti
placeholder definition

filso
placeholder definition

fraso
placeholder definition

friko
placeholder definition

ge
placeholder definition

gento
placeholder definition

gerku
placeholder definition

gerzda
placeholder definition

gigdo
placeholder definition

girzu
placeholder definition

gismu
placeholder definition

glico
placeholder definition

gocti
placeholder definition

The Complete Lojban Language

gotro
placeholder definition

ibu
placeholder definition

iu
placeholder definition

ja'ai
placeholder definition

jegvo
placeholder definition

jerxo
placeholder definition

jordo
placeholder definition

jungo
placeholder definition

ka
placeholder definition

kadno
placeholder definition

ke
placeholder definition

kelvo
placeholder definition

ketco
placeholder definition

kilto
placeholder definition

kisto
placeholder definition

klesi
placeholder definition

ku'arka
placeholder definition

la'o
placeholder definition

latmo
placeholder definition

lerfu
placeholder definition

libjo
placeholder definition

lo
placeholder definition

lojban
placeholder definition

lojbangirz
placeholder definition

lojbaugri
placeholder definition

lojbo
placeholder definition

lubno
placeholder definition

lujvo
placeholder definition

mamta
placeholder definition

megdo
placeholder definition

mekso
placeholder definition

meljo
placeholder definition

merko
placeholder definition

mexno
placeholder definition

mi'o
placeholder definition

mikri
placeholder definition

milti
placeholder definition

misro
placeholder definition

molro
placeholder definition

morko
placeholder definition

muslo
placeholder definition

nakykemcinctu
placeholder definition

nanvi
placeholder definition

nobli
placeholder definition

oi
placeholder definition

pa
placeholder definition

patyta'a
placeholder definition

petso
placeholder definition

picti
placeholder definition

polno
placeholder definition

ponjo
placeholder definition

porto
placeholder definition

radno
placeholder definition

rafsi
placeholder definition

rakso
placeholder definition

ralju
placeholder definition

re
placeholder definition

ropno
placeholder definition

rusko
placeholder definition

sadjo
placeholder definition

Lojban Word Glossary

sakli
placeholder definition

salci
placeholder definition

saske
placeholder definition

selma'o
placeholder definition

semto
placeholder definition

since
placeholder definition

sinso
placeholder definition

sirxo
placeholder definition

skoto
placeholder definition

slovo
placeholder definition

softo
placeholder definition

solri
placeholder definition

spageti
placeholder definition

spano
placeholder definition

sralo
placeholder definition

srito
placeholder definition

stero
placeholder definition

tanjo
placeholder definition

tarmi
placeholder definition

terto
placeholder definition

tricu
placeholder definition

u'e
placeholder definition

u'u
placeholder definition

vukro
placeholder definition

xampo
placeholder definition

xatsi
placeholder definition

xazdo
placeholder definition

xebro
placeholder definition

xecto
placeholder definition

xelso
placeholder definition

xexso
placeholder definition

xindo
placeholder definition

xispo
placeholder definition

xrabo
placeholder definition

xriso
placeholder definition

xurdo
placeholder definition

ybu
placeholder definition

za'e
placeholder definition

zdani
placeholder definition

zei
placeholder definition

zepti
placeholder definition

zetro
placeholder definition

zmadu
placeholder definition

zuljma
placeholder definition

General Index

- ' symbol: 11
 - and consonant cluster determination in
lujvo: 11
- er: 13
 - use of zmadu in forming: 13
- 4-letter rafsi: 12
 - definition: 12
- 5-letter rafsi: 12
 - definition: 12
- adjectives: 8
 - brivla as Lojban equivalents: 8
- adverbs: 8
 - brivla as Lojban equivalents: 8
- allowable diphthongs: 18
 - in fu'ivla contrasted with in gismu and
lujvo: 18
 - in gismu and lujvo contrasted with in
fu'ivla: 18
- animals: 16
 - use of fu'ivla for specific: 16
- apostrophe: 5, 11
 - and consonant cluster determination in
lujvo: 11
 - as not a consonant for morphological
discussions: 5
- articles: 6
 - cmavo as Lojban equivalents: 6
- boat class: 27
 - example: 27
- borrowing: 16
 - four stages of: 16
- borrowings: 16-17
 - fu'ivla form with categorizing rafsi: 16
 - fu'ivla form without categorizing rafsi: 17
 - most common form for: 16
 - Stage 1: 16
 - Stage 2: 16
 - Stage 3: 16
 - Stage 3 contrasted with Stage 4 in ease of
construction: 17
 - Stage 4: 17
 - using foreign-language name: 16
 - using lojbanized name: 16
- brivla: 5, 8, 11
 - as one of the 3 basic word classes: 5
 - definition: 8
 - from tanru: 11
 - properties of: 8
 - recognition of: 8
 - subtypes of: 8
- brivla form: 8
 - contrasted with cmavo form: 8
 - contrasted with cmevla form: 8
- C string: 5
 - as a symbol for a single consonant: 5
- C/C string: 5
 - as a symbol for a permissible consonant
pair: 5
- C/CC string: 5
 - as a symbol for a consonant triple: 5
- capitalization: 21
 - for unusual stress in names: 21
 - use in names: 21
 - use of: 21
- CC string: 5
 - as a symbol for a permissible initial
consonant pair: 5
- CCC string: 5
 - as a symbol for a permissible initial
consonant triple: 5
- CCVVCV fu'ivla: 33
 - and rafsi fu'ivla proposal: 33
- Chilean desert: 33
 - example: 33
- cmavo: 5-7, 12, 16, 23
 - as one of the 3 basic word classes: 5
 - compound: 6
 - contrasted with rafsi in usage: 16
 - contrasted with same-form rafsi in
meaning: 12
 - definition: 6
 - experimental: 6
 - for experimental use: 6
 - lack of relation of form to grammatical use:
6
 - rules for pause after Cy-form: 23
 - stress on: 7
 - structure of: 6
- cmavo and gismu: 8
 - major: 8
- cmavo clusters: 6
 - compared with sequence of simple cmavo:
6
 - definition: 6
 - recognition of: 6
- cmavo form: 8
 - contrasted with brivla form: 8
- cmavo without rafsi: 15
 - method of including in lujvo: 15
- cmevla: 5, 15, 19-21, 23
 - algorithm for: 21
 - and analyzability of speech stream: 19
 - as one of the 3 basic word classes: 5
 - authority for: 20
 - avoiding impermissible consonant clusters
in: 21
 - consonant clusters permitted in: 21
 - definition: 19
 - examples of: 19
 - final letter in: 21
 - from Lojban words: 21
 - method of including in lujvo: 15

General Index

- purpose of: 19
- rationale for lojbanizing: 19
- requirement for pause before and after: 21
- rules for: 21
- rules for formation: 20
- rules for pause before: 23
- stress in: 20-21
- unusual stress in: 20
- cmevla form: 8
 - contrasted with brivla form: 8
- comparatives: 13
 - use of zmadu in forming: 13
- compound cmavo: 6
 - compared with sequence of simple cmavo: 6
 - definition: 6
 - recognition of: 6
- compound of gismu: 8
 - lujvo as: 8
- concrete terms: 16
 - use of fu'ivla for: 16
- conjunctions: 6
 - cmavo as Lojban equivalents: 6
- consonant-final words: 23
 - necessity for pause after: 23
- continents: 32
 - gismu for: 32
- cultural words: 32
 - rafsi fu'ivla proposal for: 32
- Cy-form cmavo: 23
 - rules for pause after: 23
- derivational morphology: 4
 - definition: 4
- diphthongs: 18
 - in fu'ivla: 18
- doghouse: 26
 - example: 26
- experimental cmavo: 6
 - definition: 6
 - forms for: 6
- final syllable stress: 23
 - rules for pause after: 23
- flexible vocabulary: 8
- food: 16
 - use of fu'ivla for specific: 16
- fu'ivla: 15-19, 33
 - algorithm for constructing: 17
 - as Stage 3 borrowings: 16
 - as Stage 4 borrowings: 17
 - categorized contrasted with uncategorized in ease of construction: 17
 - considerations for choosing basis word: 19
 - consonant clusters in: 17
 - construction of: 17
 - diphthongs in: 18
 - disambiguation of: 19
 - form for rafsi fu'ivla proposal: 33
 - form of: 17
 - initial consonant cluster in: 17
 - method of including in lujvo: 15
 - rules for formation of: 17
 - stress in: 17
 - uniqueness of meaning in: 16
 - use of: 16
 - with invalid diphthongs: 19
- fu'ivla categorizer: 16-17, 19
 - for distinguishing fu'ivla form: 19
 - for distinguishing specialized meanings: 19
 - selection consideration for: 17
- fully reduced lujvo: 14
 - definition: 14
- G string: 5
 - as a symbol for an on-glide: 5
- gismu: 8-9, 12, 28-32
 - algorithm for: 28
 - and cmavo: 8
 - major: 8
 - as a subtype of brivla: 8
 - as partitioning semantic space: 8
 - basic rafsi for: 12
 - coined: 30
 - conflicts between: 9
 - creation: 28-29
 - and transcription blunders: 29
 - considerations for selection after scoring: 29
 - proscribed gismu pairs: 29
 - scoring rules: 28
 - cultural: 31
 - definition: 8
 - ethnic: 32
 - examples of: 9
 - exceptions to gismu creation by algorithm: 30
 - for countries: 31
 - for languages: 31
 - for Lojban source languages: 31
 - geographical: 32
 - length of: 8
 - level of uniqueness of rafsi relating to: 12
 - Lojban-specific: 30
 - rationale for choice of: 8
 - religious: 32
 - rules for: 8
 - scientific-mathematical: 30
 - selection of: 8
 - source of: 9
 - source-language weights for: 29
 - special: 9
 - too-similar: 29
- glue in lujvo: 11
 - n-hyphen as: 11
 - r-hyphen as: 11
 - y-hyphen as: 11
- hierarchy of priorities for selecting lujvo form: 26
- hyphen letter: 14
 - definition: 14
- hyphens: 14
 - use of: 14

The Complete Lojban Language

- integral: 19
 - architectural concept: 19
 - example: 19
 - mathematical concept: 19
 - example: 19
- invalid diphthongs: 19
 - in fu'ivla: 19
- jargon: 16
 - use of fu'ivla for: 16
- l-hyphen: 17
 - use of: 17
- learning Lojban: 8
 - magnitude of task: 8
- linguistic drift in Lojban: 23
 - possible source of: 23
- Linnaean names: 21
 - rules for: 21
- Logical Language Group: 27
 - example: 27
- long rafsi: 12
 - definition: 12
- long rafsi form: 11
 - compared with short form in effect on lujvo meaning: 11
- lujvo: 8, 11-12, 14-16, 23-26
 - algorithm for: 24
 - and consonant pairs: 15
 - and plausibility: 24
 - and the listener: 24
 - as a subtype of brivla: 8
 - based on multiple tanru: 24
 - consideration in choosing meaning for: 23
 - considerations for retaining elements of: 24
 - construction of: 11
 - dropping elements of: 24
 - examples of making: 26
 - from cmavo with no rafsi: 15
 - from tanru: 11
 - fully reduced: 14
 - invention of: 12
 - meaning drift of: 23
 - meaning of: 11
 - multiple forms of: 11
 - recognizing: 14
 - rules for formation of: 11
 - scored examples of: 26
 - scoring of: 25
 - selection of best form of: 25
 - shorter for more general concepts: 24
 - summary of form characteristics: 14
 - ultimate guideline for choice of meaning/place-structure: 24
 - unambiguity of: 23
 - unambiguous decomposition of: 11
 - unreduced: 12
 - unsuitability of for concrete/specific terms and jargon: 16
 - with zei: 15
- lujvo form: 14-15, 26
 - consonant cluster requirement in: 14
 - final letter of: 14
 - hierarchy of priorities for selection of: 26
 - number of letters in: 14
 - requirements for hyphen insertion in: 14
 - requirements for n-hyphen insertion in: 15
 - requirements for r-hyphen insertion in: 15
 - requirements for y-hyphen insertion in: 15
- male sexual teacher: 28
 - example: 28
- modifying brivla (see also seltau): 10
- morphology: 4-5
 - conventions for: 4
 - definition: 4
 - derivational: 4
 - simplicity of: 4
 - symbolic conventions for discussing: 5
- n-hyphen: 11, 15
 - contrasted with r-hyphen in requirements for use: 15
 - use of: 11, 15
- names: 19-21
 - algorithm for: 21
 - authority for: 20
 - examples of: 19
 - from Lojban words: 21
 - purpose of: 19
 - rationale for lojbanizing: 19
 - requirement for pause before and after: 21
 - rules for: 21
 - rules for formation: 20
 - stress in: 20-21
 - unusual stress in: 20
- names in Lojban (see also cmevla): 19
- Navajo: 19
 - example: 19
- non-Lojban text: 23
 - rules for pause with: 23
- nouns: 8
 - brivla as Lojban equivalents: 8
- numbers: 6, 14
 - cmavo as Lojban equivalents: 6
 - rafsi for: 14
- onset: 5
 - of a syllable: 5
- parts of speech: 5
- paternal grandmother: 10
 - example: 10
- pause: 7, 23
 - and cmevla: 23
 - and consonant-final words: 23
 - and Cy-form cmavo: 23
 - and final-syllable stress: 23
 - and non-Lojban text: 23
 - and vowel-initial words: 23
 - between words: 23
 - proscribed within words: 23
 - requirement between stressed syllables: 7

General Index

- pauses: 7, 23
 - before vowels: 7
 - rules for: 23
- plants: 16
 - use of fu'ivla for specific: 16
- prepositions: 6
 - cmavo as Lojban equivalents: 6
- primitive roots: 8
 - gismu as: 8
- punctuation marks: 6
 - cmavo as Lojban equivalents: 6
- r-hyphen: 11, 15
 - contrasted with n-hyphen in requirements for use: 15
 - use of: 11, 15
- rafsi: 11-16, 23
 - as fu'ivla categorizer: 16
 - considerations restricting construction of: 14
 - contrasted with cmavo in usage: 16
 - contrasted with same-form cmavo in meaning: 12
 - contrasted with words: 16
 - definition: 11
 - forms of: 12
 - four-letter: 15
 - requirement for y-hyphen: 15
 - lack of: 15
 - effect on forming lujvo: 15
 - level of uniqueness of relation to gismu: 12
 - long: 12
 - multiple for each gismu: 23
 - multiplicity of for single gismu: 12
 - possible forms for construction of: 13
 - rationale for assignments of: 13
 - rules for combining to form lujvo: 11
 - selection considerations in making lujvo: 12
 - short: 13
 - uniqueness in gismu referent of: 12
 - use of: 12
- rafsi assignments: 13
 - non-reassignability of: 13
- rafsi for numbers: 14
- rafsi form: 11
 - effect of choice on meaning of lujvo: 11
- rafsi fu'ivla: 32
- rafsi space: 13
- scientific names: 21
 - rules for: 21
- selma'o: 6
 - definition: 6
- seltau: 10
 - compared with English adjective: 10
 - compared with English adverb: 10
- sexual teacher: 28
 - male: 28
 - example: 28
- short rafsi: 13
 - short rafsi form: 11
 - compared with long form in effect on lujvo meaning: 11
 - slinku'i test: 17
 - definition: 17
 - source languages: 28
 - use in creating gismu: 28
 - spaghetti: 16
 - specific terms: 16
 - use of fu'ivla for: 16
 - stress: 7, 23
 - final syllable: 23
 - rules for pause after: 23
 - on cmavo: 7
 - structure words: 6
 - subtypes of words: 8
 - syllabic l: 5
 - considered as a consonant for morphological discussions: 5
 - syllabic m: 5
 - as a consonant for morphological discussions: 5
 - syllabic n: 5
 - as a consonant for morphological discussions: 5
 - syllabic r: 5
 - as a consonant for morphological discussions: 5
- tanru: 10-11
 - ambiguity in: 10
 - ambiguity of: 10
 - and creativity: 10
 - combination of: 10
 - explanation of: 10
 - expression of: 10
 - to lujvo: 11
- tosmabru test: 25
- types and subtypes of words: 8
- unreduced lujvo: 12
 - definition: 12
- V: 5
 - as a symbol for a single vowel: 5
- V'V string: 5
 - as a symbol for a double vowel: 5
- verbs: 8
 - brivla as Lojban equivalents: 8
- vowel-initial words: 23
 - necessity for pause before: 23
- VV string: 5
 - as a symbol for a diphthong: 5
- word classes: 5
- word forms: 4
 - as related to grammatical uses: 4
 - in Lojban (see also morphology): 4
- y: 5, 17
 - considered not to be a vowel for morphological discussions: 5
 - letter: 17
 - prohibition from fu'ivla: 17

The Complete Lojban Language

y-hyphen: 11

and consonant cluster determination: 11

and stress determination: 11

use of: 11

za'e: 23

use to avoid lujvo misunderstandings: 23

zi'evla: 8

as a subtype of brivla: 8

Zipf's Law: 24

Lojban Words Index

.a: 6
 .ibu: 7
 .lojban.: 8
 .oi: 6
 .u'u: 6
 .ybu: 6
 bakri: 13
 baxso: 31
 bemro: 32
 bengo: 31
 bilma: 16
 bindo: 32
 blaci: 13
 blolei: 27
 bloti: 27
 bradi: 30
 brazo: 32
 bredi: 30
 bridi: 30
 brito: 32
 brivla: 8, 11, 15
 broda: 9, 30
 brode: 9, 30
 brodi: 9, 30
 brodo: 9, 30
 brodu: 9, 30
 bu: 6
 budjo: 33
 by.: 6, 16
 centi: 30
 cidja: 18
 cidjrspageti: 18
 cirla: 18
 cmaci: 19
 cmavo: 8, 30
 cmevla: 19
 cy.: 7
 dadjo: 33
 decti: 30
 dekto: 31
 delno: 31
 di: 6
 dinju: 19
 dotco: 31
 dzipo: 32
 fau: 6
 femti: 30
 filso: 32
 fraso: 31
 friko: 32
 ge: 26
 gento: 32
 gerku: 26
 gerzda: 27
 gigdo: 31
 girzu: 28
 gismu: 8-9, 29
 glico: 31
 gocti: 31
 gotro: 31
 iu: 6
 ja'ai: 6
 jegvo: 33
 jerxo: 32
 jordo: 32
 jungo: 31
 ka: 20
 kadno: 32
 ke: 28
 kelvo: 31
 ketco: 32
 kildo: 31
 kisto: 32
 klesi: 27
 ku'arka: 19
 la'o: 16, 21
 latmo: 32
 lerfu: 19
 libjo: 32
 lo: 7
 lojbangirz: 28
 lojbaugri: 28
 lojbo: 30
 lubno: 32
 lujvo: 11, 30
 mamta: 14
 megdo: 31
 mekso: 30
 meljo: 32
 merko: 32
 mexno: 32
 mi'o: 6
 mikri: 30
 milti: 30
 misro: 32
 molro: 31
 morko: 32
 muslo: 33
 nakykemcinctu: 28
 nanvi: 30
 nobli: 21
 pa: 21
 patyta'a: 15
 petso: 31
 picti: 30
 polno: 32
 ponjo: 31
 porto: 31
 radno: 31
 rafsi: 8, 11
 rakso: 32
 ralju: 21
 re: 7
 ropno: 32

The Complete Lojban Language

rusko: 31
sadjo: 32
sakli: 13
salci: 14
saske: 18
selma'o: 6
semta: 33
since: 18
sinso: 31
sirxo: 32
skoto: 32
slovo: 33
softo: 32
solri: 21
spageti: 19
spano: 31
sralo: 32
srito: 32
stero: 31
tanjo: 31
tarmi: 19
terto: 31

tricu: 18
u'e: 7
vukro: 32
xampo: 31
xatsi: 30
xazdo: 32
xebro: 32
xecto: 31
xelso: 32
xexo: 31
xindo: 31
xispo: 33
xrabo: 31
xriso: 33
xurdo: 31
za'e: 24
zdani: 26
zei: 15-16, 33
zepti: 30
zetro: 31
zmadu: 13
zuljma: 15

Examples Index

Examples Index

Acer: 18
Arnold: 20
big boat: 10
brie: 18
Brown: 20
 John: 20
Bulgarian: 19
Catherine: 20
Cathy: 20
Chief: 21
cobra: 18
father mother: 10
field rations: 12
hepatitis: 15-16
Jane: 20
Jim: 20
John Brown: 20
Kate: 20
Katrina: 20
Korean: 19
Lady: 21
Lord: 21
maple sugar: 18
maple trees: 18
mother father: 10
One: 21
 the: 21
Persian rug: 15
Pete: 20
quark: 18
rug: 15
 Persian: 15
Sherman tank: 16
spaghetti: 18
Sun: 21
 the: 21
supper: 13
syllabic pronunciations of consonants: 18-19
 in fu'ivla: 18-19
 in fu'ivla category attachment: 18-19
tank: 16
 Sherman: 16
X-ray: 15