The Complete Lojban Language

John Woldemar Cowan

Table of Contents

Chapter 1
The Hills Are Alive With The Sounds Of Lojban

.i .ai .i .ai .o
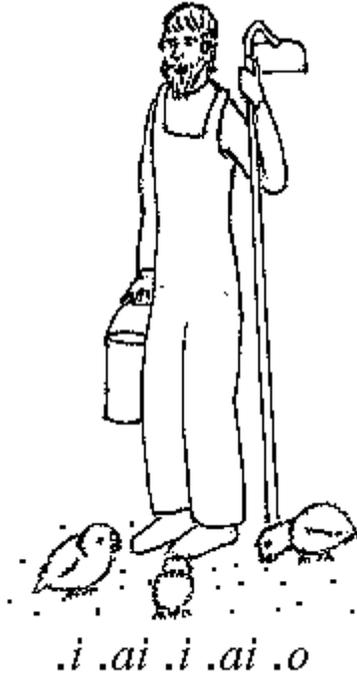
1.1 Orthography

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

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

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

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

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

4

Capital letters are used only to represent non-standard stress, which can appear only in the representation of cmevla. Thus the English name "Josephine", as normally pronounced, is Lojbanized as *.DJOsefin.*, pronounced [ʔˈdʒo.sɛ.finʔ]. (See Section 1.1 (p. 4) for an explanation of the symbols within square brackets.) Technically, it is sufficient to capitalize the vowel letter, in this case *O*, but it is easier on the reader to capitalize the whole syllable.

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

1.2 Basic Phonetics

Lojban pronunciations are defined using the International Phonetic Alphabet, or IPA, a standard method of transcribing pronunciations. By convention, IPA transcriptions are always within square brackets: for example, the word "cat" is pronounced (in General American pronunciation) [kæt]. Section 1.1 (p. 4) contains a brief explanation of the IPA characters used in this chapter, with their nearest analogues in English, and will be especially useful to those not familiar with the technical terms used in describing speech sounds.

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

Letter	IPA	X-SAMPA	Description
'	[h]	[h]	an unvoiced glottal spirant
,	[.]	[.]	the syllable separator
.	[ʔ]	[?]	a glottal stop or a pause
a	[a], [ɑ]	[a], [A]	an open vowel
b	[b]	[b]	a voiced bilabial stop
c	[ʃ], [ʂ]	[S], [s`]	an unvoiced coronal sibilant
d	[d]	[d]	a voiced dental/alveolar stop
e	[ɛ], [e]	[E], [e]	a front mid vowel
f	[f], [ɸ]	[f], [p\]	an unvoiced labial fricative
g	[g]	[g]	a voiced velar stop
i	[i]	[i]	a front close vowel
j	[ʒ], [ʐ]	[Z], [z`]	a voiced coronal sibilant
k	[k]	[k]	an unvoiced velar stop
l	[l], [l̩]	[l], [l=]	a voiced lateral approximant (may be syllabic)
m	[m], [m̩]	[m], [m=]	a voiced bilabial nasal (may be syllabic)
n	[n], [n̩], [ŋ], [ŋ̍]	[n], [n=], [N], [N=]	a voiced dental or velar nasal (may be syllabic)
o	[o], [ɔ]	[o], [O]	a back mid vowel
p	[p]	[p]	an unvoiced bilabial stop

Letter	IPA	X-SAMPA	Description
r	[r], [ɹ], [ɾ], [ʀ], [r̩], [ɹ̩], [ʀ̩]	[r], [r\], [4], [R\], [r=], [r\=], [4=], [R\=]	a rhotic sound
s	[s]	[s]	an unvoiced alveolar sibilant
t	[t]	[t]	an unvoiced dental/alveolar stop
u	[u]	[u]	a back close vowel
v	[v], [β]	[v], [B]	a voiced labial fricative
x	[x]	[x]	an unvoiced velar fricative
y	[ə]	[@]	a central mid vowel
z	[z]	[z]	a voiced alveolar sibilant

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

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

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

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

1.3 The Special Lojban Characters

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

The apostrophe represents a phoneme similar to a short, breathy English "h", (IPA [h]). The letter "h" is not used to represent this sound for two reasons: primarily in order to simplify explanations of the morphology, but also because the sound is very common, and the apostrophe is a visually lightweight representation of it. The apostrophe sound is a consonant in nature, but is not treated as either a consonant or a vowel for purposes of Lojban morphology (word-formation), which is explained in Chapter 1 (p. 4). In addition, the apostrophe visually parallels the comma and the period, which are also used (in different ways) to separate syllables.

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

The period represents a mandatory pause, with no specified length; a glottal stop (IPA [ʔ]) is considered a pause of shortest length. A pause (or glottal stop) may appear between any two words, and in certain cases – explained in detail in Section 1.1 (p. 4) – must occur. In particular, a word beginning with a vowel is always preceded by a pause, and a word ending in a consonant is always followed by a pause.

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

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

The comma is used to indicate a syllable break within a word, generally one that is not obvious to the reader. Such a comma is written to separate syllables, but indicates that there must be no pause between them, in contrast to the period. Removing a comma has no effect on how a text is pronounced or parsed.

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

Example 1.1
.i.ai.i.ai.o
[ʔi.ʔaj.ʔi.ʔaj.ʔo]
Ee! Eye! Ee! Eye! Oh!

However, this would sound clipped, staccato, and unmusical compared to the English. Furthermore, although Example 1.1 (p. 7) is a string of meaningful Lojban words, as a sentence it makes very little sense. (Note the use of periods embedded within the written word.)

If glides were used instead of glottal stops, we could represent the English string as a cmevla, ending in a consonant:

Example 1.2
.i,ia,ii,ia,ion.
[ʔi.ja.ji.ja.jon?]

If apostrophes were used instead of commas in Example 1.2 (p. 7), it would appear as:

Example 1.3
.i'ai'i'ai'on.
[ʔi.hai.hi.hai.hon?]

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

1.4 Diphthongs and Syllabic Consonants

There exist 16 diphthongs in the Lojban language. A diphthong is a vowel sound that consists of two elements, a short vowel sound and a glide, either a labial (IPA

[w]) or palatal (IPA [j]) glide, that either precedes (an on-glide) or follows (an off-glide) the main vowel. Diphthongs always constitute a single syllable.

For Lojban purposes, a vowel sound is a relatively long speech-sound that forms the nucleus of a syllable. Consonant sounds are relatively brief and normally require an accompanying vowel sound in order to be audible. Consonants may occur at the beginning or end of a syllable, around the vowel, and there may be several consonants in a cluster at the beginning. When multiple consonants appear between two vowels, as many of them as phonotactically allowed are assigned to the second syllable.

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

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

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

(Approximate English equivalents of most of these diphthongs exist: see Section 1.1 (p. 4) for examples.)

The first four diphthongs above (*ai*, *ei*, *oi*, and *au*, the ones with off-glides) are freely used in most types of Lojban words and behave similarly to pure vowels; the twelve following ones are absent from gismu-based lujvo and behave similarly to CV syllables.

The syllabic consonants of Lojban, [l̩], [m̩], [n̩], and [r̩], are variants of the non-syllabic [l], [m], [n], and [r] respectively. Although in principle any *l*, *m*, *n*, or *r* may be pronounced syllabically, only zi'evla can have syllabic consonants in their canonical pronunciation as defined by the grammar.

In terms of morphology, syllables with consonants as their nucleus (consonantal syllables) act similarly to syllables ending in -*y*, except that they do not signal the end of a rafsi. Cmevla, however, which are generally required to end in a consonant, are allowed to end with a syllabic consonant. An example is *.rl.*, which is an approximation of the English name "Earl".

Consonantal syllables are never stressed or counted when determining which syllables to stress (see Section 1.1 (p. 4)).

1.5 Vowel Pairs

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

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

The Lojban vowel pairs are:

a'a	a'e	a'i	a'o	a'u	a'y
e'a	e'e	e'i	e'o	e'u	e'y
i'a	i'e	i'i	i'o	i'u	i'y
o'a	o'e	o'i	o'o	o'u	o'y
u'a	u'e	u'i	u'o	u'u	u'y
y'a	y'e	y'i	y'o	y'u	y'y

When more than two vowels occur together in Lojban, they are grouped into syllables so that every syllable has an onset, as in the cmevla:

Example 1.4

.meiin.

.me,iin.

1.6 Consonant Clusters

A consonant sound is a relatively brief speech-sound that precedes or follows a vowel sound in a syllable; its presence either preceding or following does not add to the count of syllables. Each syllable that does not start a word begins with one or more consonants (or an on-glide) and may end with up to one consonant. Lojban has seventeen consonants: for the purposes of this section, the apostrophe is not counted as a consonant.

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

UNVOICED	VOICED
p	b
t	d
k	g
f	v
c	j
s	z
x	-

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

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

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

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

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

1. It is forbidden for both consonants to be the same, as this would violate the rule against double consonants.
2. It is forbidden for one consonant to be voiced and the other unvoiced. The consonants *l*, *m*, *n*, and *r* are exempt from this restriction. As a result, *bf* is forbidden, and so is *sd*, but both *fl* and *vl*, and both *ls* and *lz*, are permitted.
3. It is forbidden for both consonants to be drawn from the set *c*, *j*, *s*, *z*.
4. The specific pairs *cx*, *kx*, *xc*, *xk*, and *mz* are forbidden.

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

Example 1.5

.djeimyz.
[ʔdʒɛj.məzʔ]
James

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

1.7 Initial Consonant Pairs

The set of consonant pairs that may appear at the beginning of a syllable (excluding cmevla) is far more restricted than the fairly large group of permissible consonant pairs described in Section 1.1 (p. 4). Even so, it is more than English allows, although hopefully not more than English-speakers (and others) can learn to pronounce.

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

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

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

pl   pr   fl   fr
bl   br   vl   vr

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

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

Consonant triples can occur medially in Lojban words, straddling the boundaries of syllables. They are subject to the following rules:

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

A more limited set of consonant triples can appear at the beginning of syllables:

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

Clusters four consonants long appear where a syllable ending with a consonant comes before one beginning with an initial triple. In addidtion, cmevla can begin or end with any permissible consonant pair, not just the 48 initial consonant pairs listed above, and can have consonant triples in any location, as long as the pairs making up those triples are permissible. In addition, cmevla can contain consonant clusters with more than three consonants, again requiring that each pair within the cluster is valid.

1.8 Buffering Of Consonant Clusters

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

Example 1.6
    vrusi
    ['vru.si]
    or
    [vɪ.'ru.si]

Example 1.7

    .AMsterdam.
    [ˈʔam.ster.damʔ]
    or
    [ˈʔa.mɪ.sɪ.tɛ.rɪ.da.mɪʔ]

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

Here are more examples of unbuffered and buffered pronunciations:

Example 1.8

    klama
    [ˈkla.ma]
    [kɪ.ˈla.ma]

Example 1.9

    xapcke
    [ˈxap.ʃkɛ]
    [ˈxa.pɪ.ʃkɛ]
    [ˈxa.pɪ.ʃɪ.kɛ]

In Example 1.9 (p. 12), we see that buffering vowels can be used in just some, rather than all, of the possible places: the second pronunciation buffers the *pc* consonant pair but not the *ck*. The third pronunciation buffers both.

Example 1.10

    ponyni'u
    [po.nə.ˈni.hu]

Example 1.10 (p. 12) cannot contain any buffering vowel. It is important not to confuse the vowel *y*, which is pronounced [ə], with the buffer, which has a variety of possible pronunciations and is never written. Consider the contrast between

Example 1.11

    sobyladru
    [so.bə.ˈla.dru]

a Lojban compound word meaning "soy milk" and

Example 1.12

    so bladru
    [so.ˈbla.dru]

a sumti phrase meaning "nine blue roofs". If Example 1.12 (p. 12) were pronounced with buffering, as

Example 1.13

    [so.bɪ.ˈla.dru]

it would be very similar to Example 1.11 (p. 12). Only a clear distinction between *y* and any buffering vowel would keep the two utterances distinct.

Since buffering is done for the benefit of the speaker in order to aid pronounceability, there is no guarantee that the listener will not mistake a buffer vowel for one of the six regular Lojban vowels. The buffer vowel should be as laxly pronounced as possible, as central as possible, and as short as possible. Furthermore, it is worthwhile for speakers who use buffers to pronounce their regular vowels a bit longer than usual, to avoid confusion with buffer vowels.

The speakers of many languages will have trouble correctly hearing any of the suggested buffer vowels otherwise. By this guideline, Example 1.13 (p. 12) would be pronounced

Example 1.14

    [soː.bɪˈlaː.druː]

with lengthened vowels.

1.9 Syllabication and Stress

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

A single consonant always belongs to the following vowel. A consonant cluster is divided as far left as possible. Apostrophes also represent syllable breaks and belong to the following vowel. Syllabic consonants occur as the last consonant of syllables that are two consonants long.

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

Here are some examples of Lojban syllabication:

Example 1.15

    pujenaicajeba
    pu,je,nai,ca,je,ba

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

Example 1.16

    ninmu
    nin,mu

This word is split at a consonant pair.

Example 1.17

    gastro
    ga,stro

This word is canonically split before an initial consonant triple; another acceptable pronunciation is *gas,tro*.

Example 1.18

    fitpri
    fit,pri

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

Example 1.19

    sairgoi
    sair,goi

This word contains the consonant pair *rg*; the *r* may be pronounced syllabically or not. The non-syllabic version is canonical.

Example 1.20

    klezba
    kle,zba

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

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

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

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

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

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

The following are Lojban words with stress explicitly shown:

Example 1.21

    dikyjvo
    DI,ky,jvo

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

Example 1.22

    .armstrong.
    .A,rm,strong.

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

Example 1.23

    [ˈʔa.rɪ.mɪ.sɪ.tɪ.ro.nɪ.gɪʔ]

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

The English pronunciation of "Armstrong", as spelled in English, is not correct by Lojban standards; the letters "ng" in English represent a velar nasal (IPA [ŋ])

which is a single consonant. In Lojban, *ng* represents two separate consonants that must both be pronounced; you may not use [ŋ] to pronounce Lojban *ng*, although [ŋg] is acceptable. English speakers are likely to have to pronounce the ending with a buffer, as one of the following:

Example 1.24

    ['ʔa.rm.stron.gɪʔ]
    or
    ['ʔa.rm.stroŋ.gɪʔ]
    or even
    ['ʔa.rm.stro.nɪgʔ]

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

Example 1.25

    .armstron.

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

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

Example 1.26

    bisydja
    BI,sy,dja

This word is a compound word, or lujvo, built from the two affixes *bis* and *dja*. When they are joined, an impermissible consonant pair results: *sd*. In accordance with the algorithm for making lujvo, explained in Section 1.1 (p. 4), a *y* is inserted to separate the impermissible consonant pair; the *y* is not counted as a syllable for purposes of stress determination.

Example 1.27

    da'udja
    da'U,dja

da'U,dja and *da'UD,ja* sound the same to a Lojban listener – the association of unbuffered consonants in syllables is of no import in recognizing the word.

Example 1.28

    e'u bridi
    e'u BRIdi
    E'u BRIdi
    e'U.BRIdi

In Example 1.28 (p. 15), *e'u* is a cmavo and *bridi* is a brivla. Either of the first two pronunciations is permitted: no primary stress on either syllable of *e'u*, or primary stress on the first syllable. The third pronunciation, which places primary stress on the second syllable of the cmavo, requires that – since the following word is a brivla – the two words must be separated by a pause. Consider the following two cases:

Example 1.29

    le re nobli prenu
    le re NObli PREnu

Example 1.30

    le re no bliprenu
    le re no bliPREnu

If the cmavo *no* in Example 1.30 (p. 15) were to be stressed, the phrase would sound exactly like the given pronunciation of Example 1.29 (p. 15), which is unacceptable in Lojban: a single pronunciation cannot represent both.

1.10 IPA For English Speakers

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

[']	An IPA indicator of primary stress; the syllable which follows ['] receives primary stress.
[ʔ]	The usual realization of Lojban . in connected speech. This sound is not usually considered part of English. It is the catch in your throat that sometimes occurs prior to the beginning of a word (and sometimes a syllable) which starts with a vowel. In some dialects, like Cockney and some kinds of American English, it is used between vowels instead of "t": "bottle" [boʔl]. The English interjection "uh-oh!" almost always has it between the syllables.
[:]	A symbol indicating that the previous vowel is to be spoken for a longer time than usual. Lojban vowels can be pronounced long in order to make a greater contrast with buffer vowels.
[a]	The preferred pronunciation of Lojban *a*. This sound doesn't occur in GA, but sounds somewhat like the "ar" of "park", as spoken in RP or New England American. It is pronounced further forward in the mouth than [ɑ].
[ɑ]	An allowed variant of Lojban *a*. The "a" of GA "father". The sound [a] is preferred because GA speakers often relax an unstressed [ɑ] into a schwa [ə], as in the usual pronunciations of "about" and "sofa". Because schwa is a distinct vowel in Lojban, English speakers must either learn to avoid this shift or to use [a] instead: the Lojban word for "sofa" is *sfofa*, pronounced [sfofa] or [sfofɑ] but never [sfofə] which would be the non-word *sfofy*.
[æ]	Not a Lojban sound. The "a" of English "cat".
[b]	The preferred pronunciation of Lojban *b*. As in English "boy", "sober", or "job".
[β]	An allowed variant of Lojban *v*. Not an English sound; the Spanish "b" or "v" between vowels. This sound should not be used for Lojban *b*.
[d]	The preferred pronunciation of Lojban *d*. As in English "dog", "soda", or "mad".
[ɛ]	The preferred pronunciation of Lojban *e*. The "e" of English "met".
[e]	An allowed variant of Lojban *e*. This sound is not found in English, but is the Spanish "e", or the tense «e» of Italian. The vowel of English "say" is similar except for the off-glide: you can learn to make this sound by holding your tongue steady while saying the first part of the English vowel.
[ə]	The preferred pronunciation of Lojban *y*. As in the "a" of English "sofa" or "about". Schwa is generally unstressed in Lojban, as it is in English. It is a totally relaxed sound made with the tongue in the middle of the mouth.

[f]	The preferred pronunciation of Lojban *f*. As in "fee", "loafer", or "chef".
[ɸ]	An allowed variant of Lojban *f*. Not an English sound; the Japanese "f" sound.
[g]	The preferred pronunciation of Lojban *g*. As in English "go", "eagle", or "dog".
[h]	The preferred pronunciation of the Lojban apostrophe sound. As in English "aha" or the second "h" in "oh, hello".
[i]	The preferred pronunciation of Lojban *i*. Essentially like the English vowel of "pizza" or "machine", although the English vowel is sometimes pronounced with an off-glide, which should not be present in Lojban.
[ɪ]	A possible Lojban buffer vowel. The "i" of English "bit".
[ɨ]	A possible Lojban buffer vowel. The "u" of "just" in some varieties of GA, those which make the word sound more or less like "jist". Also Russian «ы» as in «byt'» (to be); like a schwa [ə], but higher in the mouth.
[j]	Used in Lojban diphthongs beginning or ending with *i*. Like the "y" in English "yard" or "say".
[k]	The preferred pronunciation of Lojban *k*. As in English "kill", "token", or "flak".
[l]	The preferred pronunciation of Lojban *l*. As in English "low", "nylon", or "excel".
[l̩]	The syllabic version of Lojban *l*, as in English "bottle" or "middle".
[m]	The preferred pronunciation of Lojban *m*. As in English "me", "humor", or "ham".
[m̩]	The syllabic version of Lojban *m*. As in English "catch 'em" or "bottom".
[n]	The preferred pronunciation of Lojban *n*. As in English *no*, "honor", or "son".
[n̩]	The syllabic version of Lojban *n*. As in English "button".
[ŋ]	An allowed variant of Lojban *n*, especially in cmevla and before *g* or *k*. As in English "sing" or "singer" (but not "finger" or "danger").
[ŋ̍]	An allowed variant of Lojban syllabic *n*, especially in cmevla.
[o]	The preferred pronunciation of Lojban *o*. As in the French « haute (cuisine) » or Spanish "como". There is no exact English equivalent of this sound. The nearest GA equivalent is the "o" of "goal" or "lore", but it is essential that the off-glide (a [w]-like sound) at the end of the vowel is not pronounced when speaking Lojban. The RP sound in these words is [əw] in IPA terms, and has no [o] in it at all; unless you can speak with a Scots, Irish, or American accent, you may have trouble with this sound.
[ɔ]	An allowed variant of Lojban *o*, especially before *r*. This sound is a shortened form of the "aw" in GA "dawn" (for those people who don't pronounce "dawn" and "Don" alike; if you do, you may have trouble with this sound). In RP, but not GA, it is the "o" of "hot".
[p]	The preferred pronunciation of Lojban *p*. As in English "pay", "super", or "up".
[r]	One version of Lojban *r*. Not an English sound. The Spanish "rr" and the Scots "r", a tongue-tip trill.
[ɹ]	One version of Lojban *r*. As in GA "right", "baron", or "car".
[ɾ]	One version of Lojban *r*. In GA, appears as a variant of "t" or "d" in the words "metal" and "medal" respectively. A tongue-tip tap.
[ʀ]	One version of Lojban *r*. Not an English sound. The French or German « r » in « reine » or „rot" respectively. A uvular trill.

[r̩], [ɹ̩], [r̩], [ʀ̩]	Syllabic versions of the above. [ɹ̩] appears in the GA (but not RP) pronunciation of "bird".
[s]	The preferred pronunciation of Lojban *s*. As in English "so", "basin", or "yes".
[ʃ]	The preferred pronunciation of Lojban *c*. The "sh" of English "ship", "ashen", or "dish".
[ʂ]	An allowed variant of Lojban *c*. Not an English sound. The Hindi retroflex "s" with dot below, or Klingon "S".
[t]	The preferred pronunciation of Lojban *t*. As in English "tea", "later", or "not". It is important to avoid the GA habit of pronouncing the "t" between vowels as [d] or [ɾ].
[θ]	Not normally a Lojban sound, but a possible variant of Lojban ' . The "th" of English "thin" (but not "then").
[v]	The preferred pronunciation of Lojban *v*. As in English "voice", "savor", or "live".
[w]	Used in Lojban diphthongs beginning or ending with *u*. Like the "w" in English "wet" [wɛt] or "cow" [kɑw].
[x]	The preferred pronunciation of Lojban *x*. Not normally an English sound, but used in some pronunciations of "loch" and "Bach"; "gh" in Scots "might" and "night". The German „Ach-Laut". To pronounce [x], force air through your throat without vibrating your vocal chords; there should be lots of scrape.
[ʏ]	A possible Lojban buffer vowel. Not an English sound: the „ü" of German „hübsch".
[z]	The preferred pronunciation of Lojban *z*. As in English "zoo", "hazard", or "fizz".
[ʒ]	The preferred pronunciation of Lojban *j*. The "si" of English "vision", or the consonant at the end of GA "garage".
[ʐ]	An allowed variant of Lojban *j*. Not an English sound. The voiced version of [ʂ].

1.11 English Analogues For Lojban Diphthongs

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

Lojban	English
ai	"pie"
ei	"pay"
oi	"boy"
au	"cow"
ia	"yard"
ie	"yes"
ii	"ye"
io	"yodel" (in GA only)
iu	"unicorn" or "few"
ua	"suave"
ue	"wet"
ui	"we"
uo	"woe" (in GA only)

Lojban	English
uu	"woo"
iy	"million" (the "io" part, that is)
uy	"was" (when unstressed)

1.12 Oddball Orthographies

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

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

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

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

Three points must be emphasized about this alternative orthography:

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

There also exists a Cyrillic orthography for Lojban which was designed when the introductory Lojban brochure was translated into Russian. It uses the "а", "б", "в", "г", "д", "е", "ж", "з", "и", "к", "л", "м", "н", "о", "п", "р", "с", "т", "у", "ф", "х", and "ш" in the obvious ways. The Latin letter "y" is mapped onto the hard sign "ъ", as in Bulgarian. The apostrophe, comma, and period are unchanged. Diphthongs are written as vowel pairs, as in the Roman representation.

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

tinco	calma	ando	anga
t	-	*d*	-

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

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

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

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.

bridi
    placeholder definition

brivla
    placeholder definition

cmavo
    placeholder definition

e'u
    placeholder definition

i
    placeholder definition

ky
    placeholder definition

no
    placeholder definition

se
    placeholder definition

sfofa
    placeholder definition

General Index

General Index

Lojban Words Index

Examples Index