Contents

1 Introduction 2

2 Phonology and Orthography 3
  2.1 Introduction .................................................. 3
  2.2 Alphabet and Capitalization ................................. 3
  2.3 A Note on Styles of Orthography ............................ 4
  2.4 Punctuation .................................................... 4
  2.5 Pronunciation .................................................. 5
    2.5.1 Regular Vowels ........................................... 5
    2.5.2 Grouping of Vowels ..................................... 6
    2.5.3 The Irregular Vowel ..................................... 7
    2.5.4 The Consonant Sounds ................................... 7
    2.5.5 Syllabic (“Vocalic”) Consonants ........................ 7
    2.5.6 Grouping of Consonants .................................. 7
    2.5.7 The Loglan Syllable ...................................... 8

3 Phonetic Classification of Words 9
  3.1 Structure Words ............................................... 9
  3.2 Names .......................................................... 10
  3.3 Predicates .................................................... 11
    3.3.1 Borrowings ............................................... 12
    3.3.2 Complexes ............................................... 12
    3.3.3 The zao construction .................................... 14

4 Word Forms 14
  4.1 Pauses ........................................................ 15
  4.2 Structure words .............................................. 15
    4.2.1 Logical connectives for sentence components ............ 15
    4.2.2 Sentence connectives and new utterance markers ......... 17
    4.2.3 Forethought logical and causal connectives ................ 17
    4.2.4 Numerals and quantifiers ................................ 17
    4.2.5 Letters, acronyms, and pronouns .......................... 20
    4.2.6 Tense/location/relation operators ....................... 22
1 Introduction

The purpose of this document is to give an independent description of the TLI Loglan language, in the provisional version embodied in my PEG parser, without PEG notations which are difficult for a nontechnical reader to follow. This does not mean that technicalities do not arise. One purpose of this is to give a clearly independent description of my intentions which can be used to double check the PEG parser. Another purpose is to give a venue for presenting material which is not in the purview of the parser, such as stating the semantics of grammatical words which are just items in lists for the parser.

In intention, this document is to be a complete description of the language. It does not represent a power grab on my part: this is to be modified in accordance with any decision of la Keugru. As always, I am well aware that my parser is not yet official. But I am not modest about the fact that I have definite ideas about how things will go, and I feel free to put them in here – but also obligated to point out proposals implicit in this text.

My intentions are conservative. I do not feel committed to not making changes as I work on this but I am committed to the language described being intelligible to a speaker of 1989 Loglan (if such a being existed).

The descriptions given in the phonology section are often far simpler than the PEG code makes them look. Some of the specifications are quite awkward to achieve with a PEG (and would be even more awkward with a BNF grammar). Of course at the moment it is unfinished – at this time I have covered the pronunciation and spelling, and the phonetic forms of words, and as of 12/15/2014 a first draft of the lexicography. The grammar proper calls! 12/18/2014 added option to close an untensed A connective with gu.

I note that a section on word building from djifoa is needed. The well-formedness of complex predicates is handled in the phonetics section, but the semantics and the pragmatics of building complexes need a venue.

3/28/2015 added an important statement about the concept of “word” to the foreword of the lexicography section and corrected some misleading statements in the phonetics section.
5/15/2015 global revisions. Added comments reflecting 5/15 revision of parser which forbids spaces internal to words.

5/16/2015 refinements of the 5/15 revision ruling out spaces inside of words. Removed capitalization restrictions in acronymic names, so la DaiNaizA is legal, for example.

5/20/2015 noticed that ii or uu as an explicit disyllable should be included in RepeatedVowel. This does rule out some forms for predicates with explicit syllable breaks in ii or uu diphthongs.

5/22/2015 an editing pass

2 Phonology and Orthography

2.1 Introduction

This section is about how Loglan is to be written (both letters and punctuation) and how it is to be spelled and pronounced.

2.2 Alphabet and Capitalization

The alphabet of Loglan has 23 letters abcddefghijklmnopqrstuvwxyz, the Latin alphabet without the letters qwx. Removing qwx is the content of a proposal before the Academy. aeiouy are the vowels; aeiou are the regular vowels; the non-vowels are consonants; the consonants mnir are continuants and may be used in a vocalic manner, in which situation they are referred to as syllabic consonants. The names of the consonants are formed by appending -al or -el to the letter (for uppercase or lowercase respectively). The names of the lowercase regular vowels are formed by prepending zi- to the vowel; the names of the upper case regular vowels are formed by appending -ma to the name of the corresponding lower case form. The renaming of the vowels is a current proposal of mine. The question of what the name of y is is open in my mind: ziy(ma) is accepted by my parser at the moment, but is irregular in form (liu ziy(ma) would not parse!) The language ought to have names for the foreign letters qwx but what these names are is open in my mind. (NOTE: names for qwx strike me as a definite (small) lack, writing 5/15/2015).

The capitalization convention of Loglan is that any uninterrupted sequences of letters (which may include syllable breaks and stress markers but not spaces or terminal punctuation) may have the initial letter either uppercase or lowercase, and must have all subsequent letters lowercase. The convention of allowing capitalization of the initial letter of the name of an upper case letter when it is embedded in an acronymic word without spacing is attested in existing Loglan text but not currently implemented in my parser. I am inclined to implement this eventually. (NOTE: I have this in mind for imminent handling, 5/15/2015)
2.3 A Note on Styles of Orthography

It was at one time a project to have a phonetic parser for Loglan. We would have thought of this as a separate gadget manipulating something like JCB’s phonetic notation in NB3 or Loglan 1. In the event, this is not what happened. In the course of writing my parser, it became clear that a careful analysis of syllables and stresses was implicit in correctly parsing the standard Loglan orthography from the level of letters upward, and adding explicit devices for indicating phonetic features such as pauses (already of course denoted by commas, but not all pauses can be represented by commas under LIP), stress (not explicitly notated at all in the standard orthography) and syllable breaks (represented explicitly by “close-commas” in some contexts in JCB’s notation; we have suppressed the close-comma and use hyphens for syllable breaks) seemed natural. We have produced a single parser which admits a continuum of styles ranging from the standard Loglan orthography to a style which I term “phonetic transcript” in which there are no spaces other than those which represent explicit pauses, and these are all marked with commas, and all stresses are shown explicitly (one could further show all syllable breaks explicitly, and for genuine phonetic transcript one should indicate explicitly how optional disyllables are being handled).

2.4 Punctuation

The comma , denotes a pause in speech. There are contexts in Loglan orthography where the presence of a mandatory pause can be deduced from the orthography though a comma is not present. It is a design goal in the parser, which I am not sure that I have achieved, to allow an explicit comma to be inserted in any place where a pause is allowed, and certainly wherever a pause is mandatory. A comma is always followed by a space. The use of a close comma to force a syllable break has been abandoned at least for now (if desired, it can be reintroduced as a variant of the syllable break - but I advise against it). Wherever a pause is intended, at least a space should be written. The 5/15 upgrade of the parser enforces the restriction that spaces cannot occur in the middle of words, so ideally wherever a space occurs it should be permissible to pause.

The period . denotes the termination of an utterance: other terminal punctuation marks ?!;; are exactly equivalent as far as the parser is concerned. Terminal punctuation ought to be followed by at least two spaces. The parser now enforces at least one space or end of text after a terminal punctuation mark. 5/15/2015: I am considering some other punctuation, such as ellipses, dashes, and (silent!) parentheses and brackets.

The hyphen - denotes a syllable break (it may not be pronounced as y as JCB proposed in Notebook 3). The apostrophe ’ marks a stressed syllable; it may be used instead of - (not in addition to the hyphen) to mark the syllable break after a stressed syllable, and it may be used after a final syllable before a non-letter or end of text to indicate that the final syllable is stressed.

The asterisk * may be used in two ways: in initial position, it marks utter-
stances which are deprecated or error-infested as Loglan utterances. It may also be used with the same grammar as the hyphen to denote emphatic stress.

Stress markers are always optional. There is no specific notation for a syllable which is not stressed, though it might be useful to add such a marker. (A note for those familiar with the language: of course, when we say that stress markers are optional, we should point out that indications of stress are not optional in some cases: the stress in a predicate word, if not signalled by an explicit stress marker, must be signalled by the end of the predicate word, indicated in this case by a space or terminal functional mark. In the presence of an explicit stress, the end of the predicate word can of course be determined without additional punctuation.)

I suggest using the stress markers to indicate rhetorical stress in a way the parser can handle. I have been doing this myself.

Spaces between words not occurring after a comma can in most cases be omitted; in certain cases they indicate mandatory pauses which can always be explicitly marked (insert missing comma), or in the case of spaces after predicates may serve to mark penultimate stresses in the preceding word, and can be eliminated if the stress on the penultimate syllable of the preceding word is explicitly marked. Places where spaces are written are often but not always places where pauses are possible in reading the given text, but one certainly should not presume that a space indicates a pause. The aim is that any space where the pause is optional can be omitted and any space where it is mandatory can have a comma added. The 5/15 parser upgrade forbids spaces in the middle of words, and makes it almost true that any place where a space is written is a place where one can pause (for those familiar with the language, there is an exception when an A word is followed by a PA word then a space: the final space cannot be converted to an explicit pause without breaking or changing the parse).

A specific style which should be possible to write for any Loglan utterance is the phonetic transcript, in which spaces appear only after commas and all stresses and syllable breaks are explicitly marked (or at least all stresses and syllable breaks of interest are marked; but if spaces are omitted stresses on predicates become mandatory).

2.5 Pronunciation

Input about actual pronunciation of Loglan sounds from other members of the community is actively solicited.

2.5.1 Regular Vowels

Vowels appearing singly (not adjacent to another vowel) are pronounced as follows:

- a is pronounced as in father
- e is pronounced as in bet
- i is pronounced as in machine
is pronounced as in *lost*

u is pronounced as *oo* in *poor*

All of these are pure sounds. They can generally pronounced as in most languages spoken in continental Europe (English is severely aberrant in its spelling).

### 2.5.2 Grouping of Vowels

Two-letter diphthongs pronounced monosyllabically are as follows:

- ai is English long *i* as in *pine*
- ei is English long *a* as in *pane*
- oi is as in English *boil*
- ao is as *ow* in English *cow* (this is an irregularity, but we are stuck with it).

These four are the mandatory monosyllables: where these letters are grouped together, they must be pronounced monosyllabically.

The pairs *ia, ie, ii, io, iu* are optional monosyllables. They may be pronounced as two syllables (smoothly moving from one vowel to the other without pause) or monosyllabically by pronouncing the initial *i* with the usual consonantal value of English *y*.

The pairs *ua, ue, ui, uo, uu* are optional monosyllables. They may be pronounced as two syllables (smoothly moving from one vowel to the other without pause) or monosyllabically by pronouncing the initial *u* with the usual consonantal value of English *w*.

These two classes are all the optional monosyllables. The disyllable pronunciation may be forced by an explicit syllable break (one of *‘*); some contexts without an explicit marker force the monosyllabic pronunciation, but I believe that no context forces the disyllable pronunciation in the absence of an explicit syllable break.

The other disyllables are obligatory disyllables: they should be pronounced with a smooth movement from one vowel to the other without pause. Pronunciation is assisted if one is stressed and one is not. In the repeated vowel disyllables *aa, ee, oo*, one of the syllables must be stressed and the other must be unstressed. An explicit stress marker is permitted to indicate which one is to be stressed, but is certainly not required. The same stress rule applies to *ii* and *uu* where these are pronounced disyllabically (remark added 5/20).

There are two different rules for grouping long strings of vowels. In a structure word (which will always be a compound attitudinal with an even number of vowels), the vowels are grouped in pairs and each pair is monosyllabic or disyllabic in a way compatible with the rules above. An odd length stream of vowels made of structure words will consist of a one-letter logical conjunction followed by a compound attitudinal.

In a name or predicate word, a long string of vowels without an explicitly given pause or syllable break is grouped using a priority scheme. If the first two letters of the stream make up a mandatory monosyllable, they are grouped together; otherwise, if the second and third letters make up a mandatory monosyllable, the first vowel is pronounced by itself and the second and third are
grouped together; otherwise, if the first two letters make up an optional disyllable they may be grouped together or pronounced separately (the parser, absent an explicit syllable break, prefers to group them together), and otherwise the first letter is pronounced by itself; in any of these cases, repeat the process with the remaining stream of vowels until it is exhausted. This is a new proposal of mine superseding rather different rules given in earlier texts.

2.5.3 The Irregular Vowel

The irregular vowel \( y \) has as its standard pronunciation the “schwa” sound found in English \( \texttt{sofa} \). John Cowan suggests that we might sometimes want to give it the value of \( \texttt{oo} \) in English \( \texttt{look} \), a suggestion which I rather like. It is very important to note that the English or Russian tendency to convert the vowel in an unstressed syllable to this sound must be resisted.

2.5.4 The Consonant Sounds

The consonants \( c \) and \( j \) have pronunciations which are unusual in English.

- \( c \) is pronounced as \( \texttt{sh} \) in \( \texttt{shoe} \).
- \( j \) is pronounced as \( \texttt{s} \) in \( \texttt{treasure} \).

The consonant \( g \) always has the hard pronunciation in \( \texttt{get} \).

The consonant \( h \) usually has its usual English pronunciation but may also be pronounced as \( \texttt{ch} \) in Scottish \( \texttt{loch} \) (the former pronunciation of \( \texttt{x} \)), as for example if it appears at the end of a syllable. (This alternative pronunciation is part of the proposal to eliminate the foreign consonants).

The consonant \( n \) has its usual pronunciation in English except before \( g \) or \( k \), where it is pronounced as \( \texttt{ng} \) in \( \texttt{song} \) (which is also quite usual in English!). The latter sound appears only as a pronunciation of \( n \) in such contexts. I note that this might also apply to \( \texttt{nh} \) if \( h \) has the alternative pronunciation.

The other consonants all have their principal pronunciations as in English.

It is worth noting that \( \texttt{tc} \) is English \( \texttt{ch} \) as in \( \texttt{chin} \) and \( \texttt{dj} \) is English \( \texttt{j} \) as in judge.

2.5.5 Syllabic (“Vocalic”) Consonants

No consonant ever appears doubled in Loglan, except the continuants \( \texttt{mnlr} \). Where the continuants appear doubled, they are used syllabically (in effect, as vowels). We call these doubled consonants \( \text{syllabic pairs} \). A syllabic pair is never adjacent to another occurrence of the same consonant. We require that a syllabic consonant (a consonant used as a vowel) must always be doubled: the main effect of this rule (which is suggested in Loglan 1) is that the spellings of some names must be changed.

2.5.6 Grouping of Consonants

Consonant clustering is governed by a number of rules.
There is a list of allowed initial pairs of consonants which may begin a syllable. An initial pair will not overlap with a syllabic consonant pair. A Loglan syllable will begin with a vowel or syllabic pair, or with a single consonant (not part of a syllabic pair), or with an initial pair (not overlapping a syllabic pair), or with an initial triple of consonants in which each of the two adjacent pairs of consonants is an initial pair.

The initial pairs are bl br ck cl cn cp cr ct dj dr dz fl fr gl gr jm kl kr mr pl pr sk sl sm sn sp sr st sv tc tr ts vl vr zb zl zv

There is a list of pairs of consonants which may not occur adjacent to one another, even across a syllable boundary. These are called impermissible medial pairs (and the other pairs are called permissible medial pairs).

The impermissible medial pairs consist of all doubled consonants, any pair beginning with h, any pair both of which are taken from cjsz, fv, kg, pb, td, any of fkpt followed by either of jz, bj, and sb.

There is a list of impermissible medial triples as well, consisting of cdz, cvl, ndj, ndz, dcm, dct, dts, pdz, gts, gzb, svl, jdj, jtc, jts, jvr, tvl, kz, vts, and mzb. All of these consist of a consonant followed by an initial pair, but they are not permitted to occur with the juncture between syllables in either of the two positions.

2.5.7 The Loglan Syllable

It is a curious feature of Loglan as presented heretofore that there is no precise definition of a permissible syllable, and in fact the places at which syllable junctures occur in JCB’s phonetic productions are sometimes quite odd. We are much more precise about this, though we believe that we have not thereby in principle much modified the set of allowed words (or in practice modified it at all).

A Loglan syllable consists of three components, only one of which is mandatory. It begins with an optional initial consonant group, continues with a mandatory vowel group, and terminates with a final consonant group.

The initial consonant group is either null, a single consonant, an initial pair or an initial triple in which both adjacent pairs are permitted initial pairs. This treatment of initial triples is stated in Notebook 3.

The vowel group is either a single vowel (y can occur), a mandatory or optional monosyllabic pair of regular vowels, or a syllabic pair of continuants. A syllabic pair of continuants cannot be adjacent to another occurrence of the same consonant.

The final consonant group is null or contains one or two consonants, neither of which can stand at the beginning of an impermissible medial pair or triple of consonants (possibly looking ahead into the next syllable). Further, if there are two consonants in the final group, they cannot consist of a non-continuant followed by a continuant (this is a new condition: such a combination would be forced to be pronounced as another syllable). The final consonant group cannot be or overlap a syllable pair.
A syllable break immediately preceding a vowel must also follow a vowel (any syllable following a consonant group picks up at least one of those consonants).

The position of the juncture in a group of three or more consonants at the end of a syllable may be optional; in no case is it phonemic (there cannot be two distinct words which differ only in the placement of such a juncture). Forcing a syllable break between vowels may change one word to another (this will only happen in proper names).

Of course, further conditions are imposed on syllables depending on the kind of word in which they appear.

It is interesting to note the maximum degree of consonant clustering: CC-CCC is possible in a name or borrowing.

3 Phonetic Classification of Words

There are three main classes of words in Loglan, structure words, names and predicates. This section is concerned with the phonetic shape of these words.

3.1 Structure Words

Structure words (in Loglan, cmapua), also sometimes called little words, are mostly grammatical particles. There are some words which are phonetically structure words and semantically predicates, and some which are semantically names (8/1 proposing to have acronymic names rather than predicates).

On the phonetic level, structure words are built out of phonetic units of one of the shapes V, VV, CV, CVV, Cvv-V, where V denotes a regular vowel, C denotes a consonant and vv denotes a monosyllable (optional or mandatory). The units of the shape Cvv-V are currently not used or very little used, but we leave the possibility open. The commonest cmapua consist of a single unit.

If any unit in a structure word is a VV, all units are VV’s. These words are the compound attitudinals already mentioned. The compound attitudinals are the only vowel-initial words which do not have to be preceded by a pause.

A unit of the shape V may only appear in initial position, and must be preceded by an explicit (that is, comma-marked) pause. A V by itself falls under this rule. These words are logical connectives of a certain class, and in fact all logical connectives of this class must be preceded by explicit pauses, though some of them have no phonetic reason to be so marked. This will be spelled out in more detail later.

Stress in a structure word is completely free – any, all or none of the syllables may be stressed. If the final syllable of a structure word is stressed and the following word is a predicate, there must be an explicit pause, comma-marked, between the two words. The phonetic cmapua which are semantically predicates are supposed to be penultimately stressed, like phonetic predicate words, but this is not currently enforced by my parser (it will be enforced when I get around to it!)
It is worth noting that the actual class of phonetic structure words has hardly any use in the grammar. Grammatically, this class is broken up into individual often very small classes each of which happens to satisfy its constraints. There is only one context in the PEG grammar, following the quotation article liu where the general phonetic class of structure words is used. I believe that in LIP it may never have been used at all, since after liu LIP accepts only actual structure words, not phonetically acceptable ones; the latter approach is the one we take.

It is also worth noting that the phonetic units of structure words are not individually necessarily single syllables: a VV or CVV might be a disyllable, and a Cvv-V unit is definitely not a single syllable.

3.2 Names

A name must be made up of valid Loglan syllables and must end with a consonant. The final consonant must be followed by either a comma-marked pause, end of text, a terminal punctuation mark, or a space followed by another name word or the name marker word ci (in which case there is a pause in the pronunciation). [It can also be followed by a stress marker followed by any of these things, if it is finally stressed]. I disagree with Steve Rice’s opinion in Loglan 3 that the comma after names can be omitted with experience; I believe it is important to reinforce it.

The requirement that names are formed of syllables is a new proposal. The original Loglan rule allowed any string of sounds ending in a consonant. In practice, nothing here has required changes in any Loglan name in use other than enforcing the rule that syllabic consonants must be doubled (which was actually suggested as an alternative by JCB in L1).

An acronymic name is not consonant final but must also have the following pause if not final, which must be explicitly comma-marked unless another name or ci follows.

We do not require that names be capitalized, but it is usual to capitalize them.

A name must be preceded either by a pause or by one of a class of name marker words. The name marker may optionally be followed by a pause. The class of name marker words is (currently) la hoi hue ci gao liu in orthography, the pauses mentioned here may or may not be comma-marked. Note that it is possible for a name to occur not preceded by a space, but only if it is preceded immediately by a name marker.

A vowel initial name must be preceded by a pause, which must be shown by at least a space and may or may not be comma-marked.

Names are the only consonant final words in Loglan. Thus the end of a name is always readily recognized.

A phonetic copy of a name marker word occurring in a name, such that the part of the name following the phonetic copy is a well-formed name, is called a false name marker. A false name marker presents a difficulty for the reader or auditor trying to determine the beginning of a name. The rule is that a name
begins as early as possible. To make a false name marker into a true one, follow it with a space or comma marked pause. A further important rule is that a name containing a false name marker cannot occur after another name without being marked (we will see that the marker used is ci).

An essay is needed about the pragmatics of avoiding false name marker problems. A possible additional rule is that any name with a false name marker in it must always be preceded by a pause, even when preceded by a name marker. Pragmatically, it is a good idea to pause after a vowel soon after the end of a name, if there is any danger of what follows being misconstrued as a second name. We believe that we have solved the problem of name resolution in our parser, but this did require attention and fine-tuning of a number of grammatical details.

3.3 Predicates

Predicate words fall into two classes, borrowings from other languages and complexes. We describe the class of borrowings first, but we note that when a predicate word is parsed, one first attempts to parse it as a complex, and only after that does one attempt to parse it as a borrowing.

A predicate must resolve into Loglan syllables.

All predicates have penultimate stress, meaning that they are always stressed on the last syllable but one, ignoring syllables not containing regular vowels (only one such syllable may intervene between the stressed syllable and the last syllable), and usually only on that syllable (note for those familiar with the language: it is permitted to stress the final syllable of a nonfinal borrowing djifoa before the y hyphen), so of course they have at least two syllables. This helps one to determine where a predicate word ends. All predicates are vowel-final (so they are not names). All predicates contain at least one occurrence of two adjacent consonants (so they are not structure words).

The rules governing the beginning of a predicate word are designed to prevent ambiguity between a predicate word and a structure word followed by a predicate word. If the word begins CC there is no difficulty. A predicate word cannot contain more than one consonant before the first CC junction, because then the first consonant plus the stream of following vowels could be peeled off as a structure word or words. So the general form of the beginning of a structure word is an optional single consonant, followed by an optional string of vowels, followed by CC, where the CC does not itself begin a valid predicate. The CC cannot be a permissible initial pair broken by a juncture.

A vowel initial predicate appearing in a noninitial position in a sentence must be preceded by a pause (which must be written at least as a space and can be comma marked).

A predicate cannot begin VCCV where the CC is a permissible initial pair.

No predicate can be of the form CCVV.

A predicate cannot begin with a syllable whose vowel segment is a syllabic pair. A syllabic pair cannot follow a vowel in a predicate.
The parser recognizes the end of a predicate either by noticing an explicit stress then counting syllables to the end of the word, or by seeing a space or punctuation ending the word and checking that the syllabification allows the syllable before the previous one (skipping a possible syllable without a regular vowel) to be stressed. This means that in the absence of an explicit stress, some spaces are mandatory which do not represent pauses in speech (though they always occur at points where it is permissible to pause), but rather signal the presence of a stress.

3.3.1 Borrowings

The additional features of a borrowing over and above the general features stated above are that it cannot contain any occurrence of y or of any of the disyllables aa, ee, oo which force stress on one of their components (nor of ii, uu as disyllables). Further, it cannot be a complex, but this is not enforced by the parser rule for borrowings directly, but by attempting to read any predicate as a complex first.

There are additional technical condition on explicit syllable breaks using -* in borrowings whose motivation is described below. The situation CV-CC is not permitted where the final CC is an initial pair, and a CC-C may not have the two consonants separated by the juncture forming a permissible initial pair.

3.3.2 Complexes

A complex is like a structure word in being composed of units which are not themselves syllables and whose interaction with syllabification can be tricky. These units are called combining forms officially; traditionally they have been called affixes, a deprecated usage. In Loglan they are called djifo.a.

It is required that a syllable does not overlap with more than one djifo.a. This is enforced by the restriction given above on junctures in borrowings: a string formed from a complex only by adding syllable breaks that violate djifo.a boundaries will exhibit one of the excluded behaviors, and so will not parse as a borrowing. The first of the two excluded behaviors for borrowings does occur in complexes.

The djifo.a are of the following basic forms (where C represents a consonant and V represents a regular vowel). Each of the djifo.a of one of the three letter forms is either an abbreviation for a five letter form (this information is in the dictionary) or is associated with a structure word (also in the dictionary):

**CVV:** Note that an initial CVV djifo.a cannot be followed by another CV- form due to the general rules of predicate formation. This is fixed by allowing an optional “hyphen” to be appended to the djifo.a. This hyphen may be r, or it may be a if followed immediately by r initial in the next affix, or it may be y. Note that CVV djifo.a where the VV is aa, ee, or oo can only occur in final or penultimate position among the djifo.a making up a complex, as one of the syllables of such a djifo.a must receive the main stress in the word (and if the VV is ii or uu it can only be pronounced disyllabically.
in final or penultimate position). The fact that a CVV appearing in final position where the VV is an optional monosyllable can be syllabified in two different ways may cause there to be two possible ways of stressing a complex.

CCV: A CCV never needs to be hyphenated (though my parser doesn’t know this).

CVC: Of course a CVC cannot appear in final position among the djifoa in a complex.

A CVC which is initial in a complex will be followed by a \textit{y} hyphen if the next affix begins with a consonant and the CC juncture created would be an initial pair, unless the word is CVCCVV or CVCCCV, which do not need this form of hyphenation (this avoids a CV form falling off the front of the word). The \textit{y} hyphen is a single syllable by itself; in terms of djifoa analysis, it is treated as part of the CVC djifoa. It should be noted that though this is related to the general rule that the part of a predicate beginning with the initial CC cannot be a predicate unless it is the whole predicate, it is not a mere case of this rule, as it is not necessarily the case that dropping the initial CV in such a predicate without the inserted \textit{y} would produce a legal predicate.

This rule is why TLI Loglan no longer has the \textit{slinkui} test.

A \textit{y} hyphen may also be appended to a CVC djifoa to prevent formation of an illegal medial pair or triple of consonants with the following affix.

CCVCV: In non-final position, the final V is replaced by \textit{y}. If a syllable break is expressed, it is CCV-CV.

CVCCV: In non-final position, the final V is replaced by \textit{y}. If a syllable break is expressed, it may be CV-CCV or CVC-CV – the former is of course allowed only if the CC is an initial pair.

\textbf{borrowing djifoa}: A borrowing djifoa is a complete borrowing plus hyphens; \textit{y} is added before it if it is not initial and after it if it is not final. A borrowing by itself is not a djifoa. These \textit{y} hyphens are regarded as part of the djifoa. A nonfinal borrowing djifoa is optionally stressed on its final syllable before the \textit{y} (which is not the same as the stress on the borrowing itself!); this stress must be expressed if the djifoa is followed by a final monosyllabic djifoa (because in this case the stress is the main penultimate stress on the whole predicate). It is permitted to pause after the subsequent \textit{y} hyphen if the stress is expressed (and to write a space or pause there). Recall that \textit{y} never appears in a borrowing; this makes it clear that borrowing djifoa can be resolved.

A complex is a word which satisfies the general conditions to be a predicate and resolves into djifoa (where any hyphens used are regarded as part of the affix). A single five letter djifoa is a complex (a primitive predicate).
3.3.3 The zao construction

John Cowan has proposed this as an alternative to the use of borrowing affixes, and I quite like it as an option. A sequence of predicate words separated by the word zao is grammatically a predicate word. There is no grouping in this construction any more than there is in the basic predicate construction. It is permissible for one or more of the initial items in a sequence of words linked with zao to form a predicate to be CV or CVV affixes.

I see one use of this construction as the ability to paraphrase a complex whose structure may be unclear to someone.

I have a preliminary suggestion that all CVh affixes can be regarded as associated with the CV structure word with which they begin.

4 Word Forms

This section deals with details of Loglan that are for the most part not manifest in the previous official formal grammar. The word class definitions are nowhere actually given formally; they are implicit in tables internal to the old interactive parser which are not human-readable and clearly have bugs.

Our program in designing the PEG parser was to parse Loglan from the level of letters upward, and as a result we have had to mandate exact formal definitions for these word classes, which in some cases are clearly not exactly the same as those implicit in LIP. Details will be seen below.

Quotation constructions and other constructions which import foreign text are handled in this section. My implementation of strong quotation is a completely new proposal.

3/28/2015 Cyril points out that it is very important to articulate the concept of “word” formally. As JCB says in NB3, the defining characteristic of a word is that one cannot pause in the middle of it. He says this in the abstract, but then does not give us any formal definition of cmapua words: the definitions of word classes in LIP are part of the internals not expressed in the formal grammar we inherited. We are told that in Lojban there are no cmapua words with more than one syllable in this sense. This is not the case in TLI Loglan. Certain cmapua classes are genuinely classes of words, in that one cannot explicitly pause in the middle of a production of this class. The 5/15 parser upgrade forbids words to be written with spaces in them (so one cannot write le na hasfa: lena is a word). Earlier versions of my parser and LIP permit le na hasfa but note that they did not permit le, na hasfa.

It may very well be that with further work we could achieve the situation reported in Lojban where a stream of one-syllable cmapua is understood without reference to any pauses that may occur between syllables, so that there are no multi-syllable cmapua which are words in this sense. We further note that we do regard it as unfortunate when the placement of a pause in a stream of cmapua syllables materially affects meaning, though we observe some situations where this seems difficult to avoid (the classic le, po problem has now been resolved,
though). We do not however regard the concept of multi-syllable cmapua word as alien to Loglan: JCB clearly envisaged there being such words.

4.1 Pauses

Some pauses in Loglan are syntactically significant. A pause cannot be significant if it is required for phonetic reasons. The pause after a name word is not significant. The pause before a vowel-initial word or an A series logical connective is not significant. A pause after a stressed syllable before a predicate is not significant. We will use the phrase “significant pause” to refer to pauses which can be syntactically significant. Further, a pause followed by y then a second pause or by cuu is not significant (these two constructions allow one to cancel a pause which might otherwise have grammatical effects – and of course these can be used to pause for effect!) The pause cancelling construction is a proposal.

There is a general rule that a pause, expressed or unexpressed, always terminates a word, and that spaces are not written in the middles of words.

4.2 Structure words

We begin by considering the many classes of structure words.

4.2.1 Logical connectives for sentence components

There are numerous parallel classes of logical connective words in Loglan. Here we are only talking about binary logical connectives like English “and”; the word no for the unary negation connective is the sole inhabitant of a separate word class of its own.

The basic series of connective roots is a, e, o, u, ha. These are words by themselves, but certain affixes can be attached to them to build a large class of words. One can add the prefix no and/or the suffix noi to an A root to obtain an A core.

We describe the class A of basic logical connectives. The prefix nu may appear initially to a logical connective word of the basic series; it may only appear if followed by u or no. The root taken from a, e, o, u, ha (possibly with prefixed no and/or affixed noi, i.e., an A core) follows this. Finally a complete PA word (a tense in the broadest sense) may follow as a suffix; if a PA suffix is present it must be followed by a significant pause or by gu: an optional significant pause or gu may appear in the absence of a PA word. An A word may not be followed by whitespace then a PA word then an explicit pause. If an A word is followed by a PA word then a space, this final space cannot represent a pause (this is the only exception to the general rule that where a space is written, a pause is permitted).

All A words are preceded by explicit comma-marked pauses (which cannot be significant pauses). The phonetic reason for this exists only when the words are vowel-initial, but the rule is enforced for all words of this class.
It should be noted that our treatment of APA words is a new proposal. These words present considerable difficulties in LIP, and have been abandoned entirely in Lojban. We have preserved them so far because they are common in the NB3 corpus, and because the related IPA words, which present much the same difficulties of termination, are clearly not dispensable without doing some violence to the corpus.

The ability to attach a PA word as a suffix is under review; it is clear that the following pause or gu is necessary to avoid ambiguity. The following pause or gu is a proposal; it seems that JCB was intermittently aware that such pauses were necessary to keep an IPA word from being confused with an A word followed by a PA clause. The pause or gu without accompanying PA word is an option which may be useful to disambiguate linked modifiers.

a means “or” (the inclusive and/or). e means “and”. o means “if and only if”. u means “whether or not”. nuu is the converse of u in the obvious sense. ha is the interrogative quantifier; an utterance with ha in it is a question which calls for an A word as an answer. Compounds built with ha are not excluded by the grammar but certainly would be odd.

Prefixing nu converts a logical connective to its converse. Prefixing no has the effect of negating the part of the logically connected utterance before the A word. Suffixing noi has the effect of negating the part of the logically connected utterance after the A word.

Suffixing a PA word has different semantics depending on whether or not the PA word is a KOU word. X, efa Y means X and then Y while X erau Y means X because Y, and careful analysis reveals that the first is fa X, Y while the second is X, rau Y. This is a slip, but we suggest following Lojban and keeping it this way. The alternative would be to have epa mean “and then”.

We now describe other series of connectives. The ACI and AGE connectives consist of an A connective, with any pause or gu after a PA word omitted, followed by ci, ge respectively. These connectives differ from A in precedence; their uses will be discussed in the grammar proper. They must be preceded by a pause, which is not a significant pause, just as in the case of A connectives.

The CA connectives are another related class. They are not preceded by pauses. The CA root forms are ca, ce, co, cu, ciha, ze. A CA root or a CA root with a prefix no and/or a suffix noi is a CA core. The semantics of ca, ce, co, cu, ciha are analogous to those of the A forms (and adding the no and/or noi has the same effect). ze builds composite objects or mixed predicates; its semantics are entirely different.

A CA connective word may take all the forms of an A connective with the A root component replaced by the corresponding CA component. A preceding pause is not required. The word ze has uses which a general CA word does not have (it can connect arguments). I am contemplating the formal possibility of zenoi and wondering if it might be useful.

The precise extent of the system of logical connective words here is not the same as that supported by LIP, but it is close. The scheme here allows more CA words; we will see if they are useful.
4.2.2 Sentence connectives and new utterance markers

The connectives given so far connect arguments and predicates. We now consider connectives which connect sentences.

The word i (always preceded by a pause) begins a new utterance, but can often be treated as if it were a high level logical connective meaning roughly e. Further words of the same class I can be constructed by appending a PA word as a suffix, followed by gu or a significant pause. All words of this class are preceded by a phonetically mandated non-significant pause.

A word of the class ICA consists of I followed by a CA connective word. This is a logical connective acting between sentences. Because it is vowel-initial, it must be preceded by a non-significant pause.

An I or ICA word cannot be followed by whitespace then a PA word (an explicit pause is needed to separate an utterance initial PA word from the I or ICA word).

There are further forms ICI and IGE constructed from words of class I or ICA by appending ci or ge (noting that a final PA word will not be closed with gu or a pause in this case).

The fact that a connective like irau or icafa which ends with a PA word must be followed by a significant pause or terminated with gu is a significant point of the grammar. The closure with gu is a new proposal here; the need to at least pause after these connectives seems to have been noticed by JCB in some contexts.

4.2.3 Forethought logical and causal connectives

The root forethought logical connective forms are ka, ke, ko, ku, nuku, kiha, each possibly followed by noi. The root KOU words are kou, moi, rau, soa (optionally prefixed with nu, no or nuno to give forms which we call KOU cores (roots are cores too)), of which we will have more to say later. The forethought logical connective words of class KA are either one of these root words, or a KOU core, followed by ki then possibly noi. These forms appear before the first of the two items connected, with ki or kinoi appearing between the two items. Forethought connectives can connect almost any grammatical structure that can be linked by logical connectives. Note that forethought analogues of APA words are not provided; they did exist in LIP and could easily be restored if wanted.

The force of the causal connectives such as kouki X ki Y is (for example) X and Y (because of X). nokouki X ki Y is (for example) X and Y (not because of (in spite of) X). Note that the initial no is not negating X or Y, they are both asserted!

How these words are used will be discussed below in the grammar.

4.2.4 Numerals and quantifiers

The numerals in Loglan are
ni: (0),
ne: (1),
to: (2),
te: (3),
fo: (4),
fe: (5),
so: (6),
se: (7),
vo: (8),
ve: (9).

Other words of the atomic quantifier word class NI0 are

kua: (division)
gie: (left bracket),
giu: (right bracket),
hie: (left parenthesis),
hiu: (right parenthesis),
kue: (inverse division),
nea: (unary minus sign),
nio: (subtraction),
pea: (unary plus sign),
pio: (addition),
suu: (root),
sua: (exponent),
tia: (times),
zoa: (double prime),
zoi: (prime),
pi: (decimal point),
re: (more than half of (quantifier)),

18
**ru:** (enough of (quantifier)),

**sa:** (about/approximately (prefix to a quantifier, by itself sara),

**si:** (at most, prefix to a quantifier, by itself sine),

**su:** (some/any/at least (quantifier prefix) by itself sune),

**hi:** (close comma),

**ho:** (interrogative quantifier)

The closely related RA class contains **ra** (all), **ri** (few), **ro** (many); these words are distinct because they have a different meaning when they appear as a suffix to a quantifier word (a quantifier word with a suffix with the phonetic shape of a RA word is a numerical predicate, for which see below).

We give semantics for these words briefly, but we do not envisage incorporating any official grammar of mathematical expressions into TLI Loglan; such a grammar might be desired by a group of users of the language, and they can develop their own for local use.

We handle the items **ma** and **moa** (00 and 000) differently than in earlier descriptions of the language. We define a class of numeral units consisting of a numeral (any word of class N10 but this really makes sense only for the digits) followed either by **ma** or by **moa** and an optional further digit. **D ma** means D followed by two zeroes; **D moa** means D followed by three zeroes. **D moa n** means D followed by 3n zeroes. Originally, **ma** and **mo** were words of class N10 meaning 00 and 000. **mo** is overused for other purposes, so we changed it to **moa**, and the use of an exponent seems better than repeating it.

A numeral block is a sequence of atomic quantifier words, numeral units and RA words in which a RA word if final is the only element.

A general quantifier word has a quite complex definition. It may optionally begin with **ie**. It is then followed by a numeral block, which may optionally be linked with more numeral blocks by CA cores (a CA root with optional negative prefix and/or suffix) [this structure of CA-linked numeral blocks is called a NI core]. This may optionally be followed by an acronym which must start with the marker **mue**; if this is present it is the last element in the word and is followed by end of text, terminal punctuation or an explicit pause. There is a final option of appending **cu**. A pause may be necessary to distinguish a NI word from a following NI word; to avoid this one may terminate a NI word with **pui** or a pause followed by **pui** (**pui** is always followed by another NI word.)

The word **ie** used as a prefix here is an interrogative word meaning which. The suffix **cu** (a late proposal of the last Keugru) generates indefinite mass or set descriptors from quantifiers (which are themselves grammatically a species of quantifier). I have to think carefully about whether this construction really describes a set as JCB says or a mass object; JCB, especially in later periods, tended to confuse the two.

The acronym suffixes create dimensioned numbers. The initial marker **mue** is a proposal of ours.
The device **pui** for making a break between adjacent NI words without a pause is a proposal.

Quantifiers have important grammatical uses in the language, to be revealed below. This is quite a separate issue from having a complex internal grammar of quantifiers/numerals, which we avoid.

### 4.2.5 Letters, acronyms, and pronouns

A Loglan upper case consonant letter is **Cai**. A Loglan lower case consonant letter is **Cei**. A third series **Ceo** is provided for lower case Greek letters.

A Loglan lower case vowel has the form **ziV**. The upper case form is **ziVma**. The old style forms **Vfi** and **Vma** are currently supported in the parser but deprecated, and it is proposed that they be eliminated.

The primary use of the letters in Loglan is **not** as names of phonemes but as **pronouns**. As a pronoun, a letter refers back to a recent argument with the same initial letter. There is a convention favoring using capital letters to refer back to proper names and lower case letters for general descriptions.

There is a further class of atomic pronoun words

- **tao**: (this [of situations],)
- **tio**: (that [of situations],)
- **tua**: (??tu ze da. this may be obsolete),
- **mio**: (we (first + third), independently),
- **miu**: (we (first + third) mass),
- **muo**: (we (first + second + third) independently),
- **muu**: (we (first + second + third) mass),
- **toa**: (this [of text]),
- **toi**: (that [of text]),
- **too**: (you, plural, independently),
- **tou**: (you, plural, jointly),
- **tuo**: (you and others independently (2+3)),
- **tuu**: (you and others (2+3) mass),
- **suo**: (self),
- **hu**: (interrogative pronoun),
- **(ba, be, bo, bu)**: series of indefinite [quantified] pronouns,
- **(da, de, di do du)**: the series of old-style definite pronouns,
The anaphora convention for the series \textit{da, de, di, do, du} can be read about in L1. The idea is that these words live on a stack in alphabetical order (those that are not already in use) and the nth description back in the text not already bound to a pronoun will be bound to the nth letter on this list when needed. It seems rather baroque but very simple cases can surely be used correctly.

The general class of pronoun words consists of letters or other pronouns, optionally suffixed with \textit{ci} followed by a quantifier word (usually a small numeral!!) It is very important to notice that for us a pronoun is a \textit{single letter}, possibly suffixed with a numeral. Multiletter variables lead to horrible ambiguities which do serious grammatical damage. Multiletter pronouns are in fact supported by LIP but there is language in NB3 which suggests that JCB did not intend to have them.

The reason that it is vitally important not to allow multiletter pronouns is that the use of a sequence of individual letters as a sequence of pronoun arguments without the inconvenience of having to pause is grammatically far more important than any use of sequences of letters as pronouns or acronyms.

Further letter words, which may be used as pronouns, but to which we may not attach numerical suffixes (I may want to allow this), are generated by \textit{gao} followed by a single well-formed word, either a name, a predicate, or a consonant initial unit cmapua (CVV or CV). This is a proposal of John Cowan, intended to provide names for letters in alien alphabets.

An acronym is a sequence of letter names (possibly abbreviated in the case of vowels to zV – not to just V – which eliminates distinctions of case of course), and number names (atomic quantifier words or numeral units), beginning either with the acronym marker \textit{mue} [a proposed feature] or a letter (possibly abbreviated) and having more than one component (the dummy \textit{mue} allows the formation of one letter acronyms and also of numeral initial acronyms without confusion with numerals or letterals). Acronyms are used to form dimensioned numbers (as noted above) and to form acronymic names (no longer acronymic predicates – a proposal of course). The initial marker \textit{mue} ensures that dimensioned number acronyms are not confused with sequences of pronouns, and the fact that acronymic names are \textit{names} ensures that they are head marked in a way which ensures that they cannot be confused with sequences of letter pronouns. A pause, terminal punctuation, or end of text is required after an acronym (so it can never attempt to consume a following letteral pronoun).
4.2.6 Tense/location/relation operators

The root words of this class (which we call PA words for short) are

- **gia**: (time free continuous tense, -ing),
- **gua**: (timeless habitual tense),
- **pia**: (past continuous tense, until [before terms]),
- **pua**: (was habitually -ing, continuous past tense),
- **nia**: (continuous present tense, during [before terms]),
- **nua**: (am now habitually -ing, continuous present tense),
- **biu**: (possibly, under conditions X [before terms]),
- **fea**: (???I cannot find any indication of what this means, though there is a reference to it in text of trial.85),
- **fia**: (will be -ing future continuous tense, since X [before terms]),
- **fua**: (will habitually be -ing, future continuous tense),
- **via**: (throughout a place of medium size),
- **vii**: (throughout a small place),
- **viu**: (throughout a large place),
- **ciu**: (X ga Y ciu Z means Z ga Y as much as X ga Y),
- **coi**: (according to rule X),
- **dau**: (probably, likely under conditions X),
- **dii**: (for, on behalf of X),
- **duo**: (by method X),
- **foi**: (X foi Y, X must Y, X ga Y foi Z, X must Y under conditions Z – Y a predicate),
- **fui**: (should, same structure as foi),
- **gau**: (can (same structure as foi?)),
- **hea**: (by, with the help of, X),
- **kau**: (can, is able to (structure of foi)),
- **kii**: (with/accompanied by X),
- **kui**: (???),

22
lia: (like, in the way that – I suggest that X ga Y lia Z means that X ga Z as Y ga Z, but X ga Y lia lepo Z ga W means X ga Y as Z ga W),
lui: (for, in order to please X),
mia: (subjective subjunctive, mia lepo X = were X the case),
mou: (more than, structure of ciu),
nui: (may/is permitted to, structure of foi),
peu: (as for/concerning X), roi (X roi Y = X intends to Y; X ga Y roi Z = X intends to Y under conditions Z),
rui: (???),
sea: (instead of X),
sio: (certainly, certain under conditions X [before terms]),
tie: (with/through/by means of instrument X),
va: (in the middle distance, near X),
vi: (here, at X),
vu: (far away, far from X),
na: (now, present tense, at the same time as X),
pa: (past tense, before X),
fa: (future tense, after X)

and the related small class of KOU roots kou (because (cause) of X),
moi (because/in order to (motive) of X), rau (because (reason) of X),
soa (because (logical premise) of X) which can be prefixed with nu, no, or nuno to give additional forms which we call KOU cores (a root is also a core).

It is important to notice that nokou lepo X does not deny X; in fact, it asserts X and says that the main event happened in spite of X. Forms like nukou are converses: they are versions of “therefore X”. Forms like nunokou are versions of “nevertheless X”; X happens, but not because of the main event, rather in spite of it.

A compound PA word begins with an optional numeral or quantifier, followed by a string of PA roots or KOU cores taken from the lists above (recalling that KOU cores may include certain prefixes), optionally linked to further strings of PA roots/KOU cores by CA cores, then may optionally be closed with one of the qualifiers za, zi, zu (to see the effects of these qualifiers on tense and location operators, see the dictionary) or with juu or a pause followed by juu (the juu forms must be followed by another PA word; this is a device to avoid necessarily having to pause when one complex PA word is followed by another).
The semantics of complex PA words will require a considerable essay, to be inserted here in due course. In particular, a summary of the location and tense words and their interaction with -zV suffixes is needed, since these have some ad hoc features. pazu a long time ago versus panazu in the past for a long time interval is an example I insert to remind myself.

These words can be used as prepositions (followed by an argument) or as tenses in the broadest sense (followed by a predicate). The word ga is a content free tense word not usable as a preposition. ga has other uses as well. Details of this will be seen in the grammar.

Where a PA word occurs as a suffix to another word form, it is generally illegal for it to be replaced by a space followed by a PA word: where a PA suffix is legal, it cannot be replaced by a following PA word (the rule for the optional PA suffix on an A word is different, and stated above; an A word without a PA suffix can be followed by a separate PA word as long as this is not in turn followed by GU or an explicit comma pause). It is illegal for a PA word without one of the optional closing components to be followed by a space then another PA word. These are all rules designed to defend the condition that spaces are forbidden to occur in the interiors of words.

The only spaces in the orthography which cannot be pronounced as explicit pauses are spaces following a PA word which follows an A word.

4.2.7 The system of tense and location words

Here we will lay out the system of compound tense and location words, indicating difficulties and possibly some suggestions for improvement.

The basic series of tense words is pa, na, fa, which mark present, past, future tense when they mark a predicate; pa X, na X, fa X mean before X, at the same time as X, after X, respectively.

A second series of tense words pia, nia, fia express continuous tenses. pia preda means “was preda-ing”. pia X means “until X”. fia preda means “will be preda-ing”. fia X means “since X”. pia preda means “was preda-ing”. pia X means “until X”. nia preda means “is preda-ing”. nia X means “during X (throughout)”. A third series of tense words pua, nua, fua express habitual tenses. Their meanings are similar to those of the previous series, but they refer to events which often or usually happen during an indicated period rather than events which happen continuously during an indicated period.

These words can be compounded. Here are the dictionary meanings of compound tenses.

papa: had (been)... ed, sign of the past perfect tense.
pana: was/were then... ing, sign of the past coincident tense.
pafa: was/were going to..., sign of past progressive tense, english inexact
napa: has/have (been).../a..., sign of the present perfect tense; already
**nana:** am/are/is now... ing, sign of the present coincident tense.

**nafa:** is/are going to..., sign of present progressive tense, English inexact.

**fapa:** will have... (been) ed, sign of the future perfect tense.

**fana:** shall/will be then... ing, sign of the future coincident tense.

**fafa:** will-be going to..., describes an action which takes place after the (future) time being recounted.

These words can be qualified with the suffixes $zV$. Here are the dictionary entries.

**pazi:** just... ed/was just (now a), a modified tense operator; just before..., before event terms.

**nazi:** at/coincident with..., an instant in time; at the time when, momentary event clauses.

**fazi:** will immediately (be a)..., modified tense operator; just after, before event terms.

**paza:** lately/newly/recently... ed, not too long ago, a modified tense operator; shortly before..., before event terms.

**naza:** during/in..., in some short interval, with terms.

**faza:** will soon (be)/be about to/just going to.; shortly after, with clauses.

**pazu:** long before, some event, before clauses.

**nazu:** during, in some long interval, with terms; while, during some long event.

**fazu:** will eventually (be a), a modified tense oper.; long after, some event, before terms.

The dictionary definitions are not fully systematic. Notice that **nia** and **nazu** express different meanings of “while, during”. I think in spite of some ambiguity about **nazV** forms, that the $zV$ operators do something uniform, qualifying the distance of the event from the argument (or the present in the case of tenses). **nazu** doesn’t say that the event actually is far from the present, but since it says the event is in a long interval around the present it permits a long distance from the present.

Continuous examples are also listed

**piazu:** for all that time until now, adverb and before preds; long-before then and until, with clauses.

**niaza:** while/throughout the short time, clauses.

**niazu:** while/throughout the long time, clauses.
**fiazu:** since, for a long time after, with clauses.

The basic series of location operators is **vi, va, vu**, at/near/far from.

The second series of location operators is **vii, via, viu**, throughout a small/medium/large sized place.

Here are the compounds listed in the dictionary.

**vivi:** around, in the place where, before terms.

**viva:** out of where, a short way, with clauses.

**vivu:** out of, for a long way, before terms.

**vavi:** into where, from nearby, before clauses.

**vava:** past where, nearby, before clauses.

**vavu:** away from, from near to far, before terms.

**vuvi:** into where, from far away, before clauses.

**vuva:** toward the place where, before clauses.

**vuvu:** past where, at a distance, before clauses.

Modifications with **zV** affixes:

**vizi:** right here/at this spot, before preds; at the spot where, with point like events.

**vazi:** near this spot/the spot where, of point like events, before predicates.

**vuzi:** far from this spot, before predicates; far from where, spatially limited events.

**viza:** in this place/small region, before preds; where, before spatially limited events.

**vaza:** near this place, before predicates; near the place where, of limited events.

**vuza:** far from this place, before predicates; far from where, of medium sized events.

**vizu:** in this place/big region, before preds; where, before spatially extensive events.

**vazu:** near this region, of extensive events, before predicates; near the place where, of extensive events.

**vuzu:** far from this region, before predicates; far from where, of extensive events.
The difficulty here is that there really isn’t a system as such – at least, if there is, it is only implicitly given. It is possible to extrapolate from this, and it is also possible to compare with the sister language Lojban, in which an effort has been made to systematize these issues.

Another point is the status of the qualifiers $zV$. These are affixes, and one of these terminates a PA word (this is true in my grammar, and experiment confirms that this happens with LIP as well). In a word such as $fanazu$, what does the $zu$ qualify? It seems most reasonable to suppose that in a word $pacenazu$, the $zu$ qualifies both conjuncts. The current grammar does not allow logical conjunction of PA cores with different $zV$ qualifiers to form words.

It is clear that a lot more words are formally possible, both for my grammar and for LIP.

4.2.8 Articles

The basic articles (constructors of definite arguments) are

- lea: article for sets: the set of all things with property ...
- leu: The particular set I have in mind of things with property...
- loe: The typical...
- lere: The one or more things I mean which actually are...
- laaa: The unique object which actually is... (the logical definite description).
- le: The default article. The objects(s) understood from context with property $X$...
- lo: The mass article (describes composite objects made of all the objects designated).
- la: The article for proper names.

These are the root words. Complex words of this class can be formed by following the root with an optional pronoun followed by an optional PA suffix. Spaces may not intervene between these components.

The name constructor $la$ and complex articles derived from it appear in the list above but appear in special constructions as well. The precise ways in which names are handled in this grammar involve new proposals.

There is a special class LEFORPO consisting of $le, lo$, and some quantifiers (the NI cores) which may appear followed by PO in the formation of abstract descriptions. Notice that no new words are involved. It is worth noting that $lepo$ and related forms are not single words, though they are often written without a space, and so can be written $le po$ or even $le, po$.

Details of the use of these classes belong in the grammar below.

$la$, $lua$ and $lou, lou$ are paired forms beginning and ending unordered and ordered lists, respectively. These are forms to which I admit having paid very little attention.
4.2.9 Constructions involving alien text and related articles

In this subsection we introduce the articles which handle quotations and imported foreign text, and we also give the full constructions of arguments (and predicates) of this kind. The strong quotation construction that we give is a completely new proposal.

Any well-formed Loglan utterance X can be quoted li X lu. X may be preceded and followed by explicit pauses (commas) if desired (this is not required).

A single Loglan word X may be quoted liu X. This is the only context in the grammar where the phonetic class of structure words plays any role. In LIP it plays no role even here, as LIP apparently only allows liu for actual cmapua of the various classes in this section. Lojban I believe only allows unit cmapua to be quoted; we admit that there are compound words, so we allow them to be quoted. A pause may sometimes be required to terminate a quoted word where you want it terminated. niu may be used instead of liu to explicitly signal that a quoted word, though phonetically acceptable, is not a Loglan word. I have just installed the correction allowing liu ziy and liu ziyma to parse.

One may refer to a letter (rather than use it as a pronoun) using the form lii X.

The further forms discussed here operate on alien text. Alien text will be a block of text beginning with whitespace or an explicit pause and ending with whitespace, an explicit pause (comma), or before terminal punctuation or end of text, and containing no commas or terminal punctuation otherwise. It may contain other symbols or non-Loglan letters. Initial and final whitespace must be expressed phonetically as a pause.

The article lao followed by one or more blocks of alien text, with blocks being separated by y set off with spaces (which must be pronounced as explicit pauses) if there is more than one block, forms a foreign name. Wherever names are to be written by “look” rather than as they are to be read phonetically in Loglan, lao should probably be used. This construction was originally presented as a construction for the Linnaean names of biology; it is a valuable observation due to Steve Rice that it has a far more general usefulness. We abandon all other aspects of JCB’s discussion of Linnaean names as such: the details of scientific terminology are not part of the purview of the Loglan grammarian.

sao followed by alien text forms a predicate. This is a way to import a foreign word directly. sue followed by foreign text intended to transcribe or suggest a sound forms a predicate meaning “makes that sound”. sue miao is to meow.

Now we present our strong quotation proposal. The basic idea is that a series of blocks of alien text separated by whitespace is quoted by placing lie before the first block and ci before each subsequent block. The level of quotation (quotation, quotation of a quotation, etc) can be modified by suffixing a numeral directly to lie or ci (and it can change from block to block; where a change occurs the numeral must be explicitly given (the only case where ne is required)). This device was added because this technique of quotation would become really burdensome if nested quotations were made explicitly. This is an entirely new
proposal. The original strong quotation method is not PEG parsable (it is not even BNF parsable) and I think has other weaknesses. The bit in Alice with the multifariously nested quotation marks must be translated into Loglan using this quotation style!

We support in the grammar without necessarily approving (also without necessarily disapproving; I know some Keugru members do not like them) the qualifiers za (text) and zi (sound) for quoted forms. The qualifier is to be affixed to the initial quotation article (after any suffixed numeral in the case of lie).

4.2.10 Assorted grammatical particles, somewhat classified

Here is a list of terminators and boundary markers: ci, cui, ga, ge, geu (cue), gi, go, gu, gui, guo, guu, gue

The particles je and jue mark tightly bound arguments (or modifiers, according to a proposal).

The words

jie: (restrictive set membership),
jae: (nonrestrictive set membership),
pe: (general possessive),
ji: (which/that (is) (identifying),
ja: (which/that (is) nonidentifying

construct subordinate clauses from arguments, modifiers or predicates.

The words jio, jao construct subordinate clauses from sentences.

The case tags, including the positional ones are listed:

beu: (patients/parts),
cau: (quantities/amounts/values),
dio: (destinations/receivers),
foa: (wholes/sets/collectives),
kao: (actors/agents/doers),
jiu: (lessers),
neu: (conditions/circumstances/fields),
pou: (products/purposes),
goa: (greaters),
sau: (sources/reasons/causes),
veu: (effects/states/effects/deeds/means/routes),
zua: (first argument),
zue: (second argument),
zui: (third argument),
zuo: (fourth argument),
zuu: (fifth argument),
lae: (lae X = what is referred to by X),
lue: (lue X = something which refers to X)

The operators of indirect reference lae and lue are a different sort of creature, but they have the same grammar as case tags. The latter two operators can be iterated (and so can case tags, probably indicating that more than one applies to the same argument).

My opinion of the optional case tag system is that I would never have installed it myself, and it represents an extra layer of work for dictionary maintenance, but it is potentially usable and represents a large amount of work by our predecessors, so my intention is to leave it in place (and try to be good about assigning tags when I define predicates) and maybe maybe some day actually learn the case tags! The whole scheme is quite optional for speakers, though pressure to learn them would be imposed on a hypothetical Loglan community if many speakers actually used them.

The particle me constructs predicates from arguments. I believe the addition of mea was a mistake, as me properly understood, already served its exact function. I’ll write an essay on this eventually.

The particles nu, fu, ju interchange the 2nd, 3rd, 4th argument of a predicate respectively with the first. These are called conversion operators.

The particles nuo, fuo, juo eliminate the 2nd, 3rd, 4th argument place of a predicate respectively, stipulating that it is occupied by the same object that occupies the first argument place (these are reflexives).

More conversion and reflexive words are formed by suffixing a quantifier. The only meaningful ones as far as I can see would be numerals larger than 4 and ra, which would choose the last argument place.

Yet more words of this class can be formed by concatenating conversion operators and reflexives; they simply compose, allowing complex reordering and identification of arguments. Words which form abstraction predicates are the short-scope poi, puu, zoo and the long-scope po, pu, zo. In each set, the words form predicates for events, properties, and quantities respectively.

The uses of all these words will be revealed by the grammar.

4.2.11 Words which form free modifiers

The register markers indicate attitude toward the person addressed:

die: (dear),
fie: (comrade/brother/sister),
kae: (gentle as in gentle reader to an equal at a certain distance),
nue: (Mr Ms Mrs neutral and at a distance),
rie: (Sir, Madam, Sire, Honorable – to a superior)

They can be negated.

The vocative marker is hoi. The inverse vocative marker (indicating the speaker or author) is hue.

The “right scare quote” is jo, which may be prefixed with a numeral. It indicates that previous text is not to be taken quite literally; the numeral would indicate how many words are in the scope of the jo. I notice that if a scare quote were to be applied to a quantity, it would have to be nejo.

The paired words kie and kiu serve as spoken parentheses: include a well-formed Loglan utterance between them to form a free modifier.

Smilies can be spoken in Loglan: soi X, where X is a predicate, forms a free modifier inviting the auditor to imagine the speaker doing X. soi crano is literally :-)?

The freestanding attitudinal words of the original VV flavor, generally expressing emotions or attitudes, are

ua: (there! thats it! done! satisfaction),
ue: (indeed! oh! surprise),
ui: (fine! good! (pleasure)),
uo: (come now! look here! (annoyance)),
uuu: (Alas! Sorry! sadness/sympathy/regret/not apology, that is sie),
oa: (moral obligation – it must be),
oe: (preferably),
oi: (permissibly, you may),
oo: (disapproving hmmm) [to be added!],
ou: (no matter (ethical indifference)),
ia: (yes), agreement),
ii: 9maybe (tentative belief),
io: (I expect that, apparently, moderate belief),
iu: (I have no idea!, ignorance, lack of belief or knowledge),
ea: (let’s, I suggest...),

31
ee: (caution! careful! take care! [to be added]),
ei: (is it true that? forms yes/no questions),
eo: (please? will you? asks permission),
eu: (let us suppose that...(subjunctive)),
aa: (I see (what you mean)),
ae: (yes, I wish to (hope or weak intention)),
ai: (I intend to...Definitely...(strong intention)),
ao: (Yes, I want to, Ill try to...(moderate intention)),
au: (I dont care...indifference, absence of intention)

ie is not really an attitudinal, but an interrogative meaning “which”. (the words
aa, ee, oo are not in the trial.85 list of UI words, though likely the preparser
handles them fine in LIP; I have added them).

Additional words with the same grammar are

bea: (for example),
buo: (however, on the contrary, but),
cea: (in other words, namely),
cia: (similarly), coa (in short, briefly),
dou: (given, by hypothesis),
fae: (and vice versa),
fao: (finally, in conclusion),
feu: (in fact, actually),
gea: (again, I repeat),
kuo: (usually, customarily),
kuu: (generally),
rea: (clearly, obviously, of course),
nao: (now, next, new paragraph),
nie: (in detail, looking closely),
pae: (etc., and so forth),
piu: (in particular),
saa: (roughly, simplifying),
sui: (also, as well, furthermore),
taa: (in turn, sequence),
toe: (respectively),
voi: (skipping details),
zou: (by the way, incidentally),
ceu: (anyhow),
sii: (evidently)

These words are discourse operators, comments on the way we are speaking.

The word cao emphasizes the next word. The grammar will not show this, as it associates attitudinals with the previous word or construction! Notice that one can use the phonetic stress markers to indicate stress in writing.

Finally, we have words of social lubrication, loi (hello), loa (goodbye), sia (thank you), siu (you’re welcome, don’t mention it), sie (sorry (apology)) The word sie (to be distinguished from uu, sorry in the sense of regret but not apology) is new. Cyril and I believe it reasonably that siu be a polite answer to sie as well as sia. [I have it in mind to make these words be additional name markers, so that one can say Loa Djan as well as Loa, hoi Djan; this would also require new freemod forms]

The attitudinal, discourse and social words (class UI) can be negated by preceding them with no or following them with noi.

In addition, there are discursive operators firstly, secondly, lastly formed by suffixing quantity words with fi.

4.2.12 Negation

The word no is the logical negation operator. Initial no in attitudinal forms, KOU words, and subordinate clauses (as well as occurrences internal to some compound structure words) must be excluded from this grammatical class.

4.3 The Large Word Classes

There remain the large classes of predicate and name words.

4.3.1 Predicate words

The words bia (is part of), bie (is a member of (a set)), cie (is less than (math)), cio (is greater than (math)), bi (is defined as) are all predicates semantically, though they are structure words phonetically. They form a grammatical class BI of identity predicates.

The words
**he:** (interrogative predicate; a sentence with a he in it is a question with a predicate answer),

**dua:** (first free predicate variable),

**dui:** (second free predicate variable),

**bua:** (first bound predicate variable),

**bui:** (second bound predicate variable)

are grammatically ordinary predicates, though phonetically structure words. None of them are really very ordinary predicates! (some essay will be needed here).

The class PREDA of predicate words includes the last list.

The class PREDA includes quantity words suffixed with ra, ri, ro to form numerical predicates (cardinal, ordinal, quality ordinal, respectively). A predicate *tora* is a two place predicate, X is a two element subset of Y; *tora* is a two place predicate, X is the second term in series Y. I do not know what the quality ordinal predicates are supposed to be like (homework for me to do). The numerical predicates should be penultimately stressed just as ordinary predicates are [the grammar currently does not enforce this].

And of course the class PREDA includes the predicate words in the phonetic sense of the first section.

### 4.3.2 Borrowing predicates

A section which is owed.

### 4.3.3 Making complex predicates

A section which is owed.

### 4.3.4 Name words

The name words consist of the name words in the phonetic sense of the first section and the acronyms. One is required to pause after an acronym used as a name, and one is permitted to omit the explicit comma in writing under exactly the same conditions as after an ordinary name word. It is worth noting that a pause is also required after an acronym when it is used as a dimension in a quantity.

Contrary to statement in L1, we maintain that a Loglan name word should always be written as it is to be pronounced. Names written to look visually like their forms in other languages should be treated as alien text and turned into grammatical proper names with lao. Thus, la Ainctain is the native version of Einstein’s name, but we can of course also write lao Einstein. The first must usually be followed by an explicit pause, while the latter may be followed by an innocent space — which will also be a pause, as stated in the rules for alien text. la Einstein is a legal Loglan name, but would be pronounced quite oddly.