

Notebook 3

THE PRESENT STATE  
OF  
THE LOGLAN™ LANGUAGE

by

James Cooke Brown

Copyright© 1987 by  
The Loglan™ Institute, Inc.  
A Non-Profit Research Corporation  
1701 Northeast 75th Street  
Gainesville FL 32601  
U. S. A.

Note added 7/28/2014: this document remains intellectual property of the Loglan Institute. It is licensed freely for private non-commercial use by people interested in Loglan or Lojban; please contact us about any other use.

## PREFACE

The present work is a revision and extension of The Institute's two previous notebooks, both published in 1982, and the two special issues of **The Loglanist**, **TL6/1** (1983) and **TL7/1** (1984), and it incorporates some material from the latter two works. The scope and organization of this present work is quite new, however; in particular it is the first complete description of the language to be published since 1975.

This account of New Loglan is long overdue and I apologize. Both financial reasons and reasons of personal health have slowed The Institute's work down since the early 1980's, when so much seemed to be being accomplished. One of the reasons is that that certain parlor game which had supported the Loglan Project for so much of its life was withdrawn from the market in 1983, and Loglan has had to go it alone ever since. Fortunately, we are now on the verge of Going Public Again; so the long, dry period of Loglan's being everybody's "poor relation" may soon be over. If, as everybody seems to think, Loglan is about to become at least a modest commercial success, the project may at last become financially self-supporting.

I wish to take this occasion to acknowledge the stalwart few who have contributed to the work of The Institute over these last few, difficult years. Faith Rich has made a large contribution to the next Loglan dictionary by completing the Eaton Interface. She was ably helped in doing so by Jeffrey Taylor, Kieran Carroll and Robert McIvor. Unfortunately, their work is not quite ready to be published. But it will, as I say, form the bulk of the next dictionary of the Loglan language whenever that is published.

My daughter Jennifer Fuller Brown managed to bring the Optional Case Tags Project to a happy conclusion this Spring; and the fruits of her work are in this notebook. Paloma Ibanez ably assisted me in bringing the Scientific Borrowings Project very nearly to an algorithmic conclusion; but the fruits of that project are, as explained elsewhere, not quite ready to be published. Glen Haydon has helped me put together the two MacTeach programs that are now available. Bill Greenhood has counselled me from time to time on the proprieties of scientific word-making. And Scott Layson has made yet another extraordinary gift to the project by updating all the Lyces software—which is the tool with which I do my grammatical work—for the more capacious environment of our new Zenith 100 computer.

Users of this notebook are invited to send in (1) notices of whatever errata they may find, and (2) proposals for improving the language by adding to, changing, or deleting any of the provisions described in this notebook. Please keep these two kinds of contributions separate, however. Formal proposals will go to the Loglan Academy for assessment when they meet in the early Spring of 1988; and the format for making proposals formally to the Academy has been described in a recent Lognet. Notices of errata should also be kept separate from the covering letter. Preferably they should be on sheets or cards that may be filed separately from correspondence.

We at The Institute look forward to a vigorous testing of the language described in these pages, and to GPA-ing with it in the very near future.

JCB  
Gainesville  
July 1987

## TABLE OF CONTENTS

Foreward	7
CHAPTER 1. PHONOLOGY (SOUNDS & SOUND-VARIANTS)	9
1.1 Definitions and Conventions	9
1.2 Two Types of Phonemes	10
1.3 Regular Phonemes	10
1.4 The 6 Regular Vowels	10
1.5 The Advantages of Romance (aa)	11
1.6 The TWO Spellings of ‘e’ of ‘met’	12
1.7 The Odd Spellings of /i/ and /y/	12
1.8 /y/ as a Hyphen	12
1.9 /y/ as a Buffer	13
1.10 /iy/ as a Hyphen in Buffered Dialects	13
1.11 The Effect of Hyphenating and Buffering on Stress	13
1.12 /y/ in Names	14
Table 1.1 Permissible Pronunciations of the Twenty-Five Loglan Vowel-Pairs	14
1.13 The 25 Vowel-Pairs	14
1.14 The 10 Optionally Disyllabic Vowel-Pairs	15
1.15 The Pair-from-the-Left-Rule	15
1.16 Indications of Syllabicity	16
1.17 The 17 Regular Consonants	17
1.18 The Odd Sounds of [C c] and [J j]	18
1.19 The Four Vocalic Consonants	18
1.20 The Unfamiliar Consonant Pairs	18
1.21 The Three Irregular Phonemes /q w x/	19
1.22 The Use of Irregular Phonemes in Names	19
1.23 Three Stress Phonemes	20
1.24 One Pause Phoneme	20
1.25 Intonation	21
1.26 Buffered Dialects	21
CHAPTER 2. MORPHOLOGY (WORDS & WORD-FORMS)	23
2.1 Design Objectives	23
2.2 Definitions and Conventions	23
Table 2.1 The Two Partitions and Three Classes of Loglan Words	25
2.3 Two Major Partitions and Three Word-Classes	25
A. THE MORPHOLOGY OF NAMES	27
2.4 The Forms of Names	27
2.5 The Pause Before Vowel-Initial Names	27

## Notebook 3 - Contents

2.6	The Name-Marker Restriction on Names	27
2.7	Working Around the Name-Marker Restriction	28
2.8	Derivations of Names	28
2.9	Internal Names	28
2.10	External Names	29
2.11	Auditorily-Modeled External Names	29
2.12	Visually-Modeled External Names	29
2.13	The Linnaean Polynomials	30
2.14	Pronunciation of the Linnaean Polynomials	30
2.15	Writing Linnaean Names	31
2.16	The Post-Nominal Pause	32
2.17	Resolving Names	32
<b>B.</b>	<b>THE MORPHOLOGY OF STRUCTURE WORDS</b>	<b>34</b>
2.18	The Functions of Structure Words	34
2.19	The Four Little-Word Forms	34
2.20	Compound Little Words	35
2.21	Letter-Words	35
2.22	Suffixes for the 52 Latin Letter-Words	35
2.23	Suffixes for the 48 Greek Letter-Words	36
2.24	Uses of Letter-Words	37
2.25	Spelling Aloud	37
2.26	Little Word Predicates	37
2.27	Mathematical Predicates	37
2.28	The "No Pausing Inside Words" Rule	37\
2.29	Acronymic Predicates	38
2.30	Pause and Stress Around Acronymic Words and Letter-Words	39
2.31	Pause and Stress Around Dimensioned Numbers	40
2.32	Acronym Recovery Rules	40
2.33	Resolving Structure Words	41
<b>C.</b>	<b>THE RESOLUTION OF PREDICATES</b>	<b>43</b>
2.34	The Functions of Predicates	43
2.35	A Temporary Stress-Marking Convention	44
2.36	The Post-Emphatic and Intervocalic Pauses	44
2.37	Stress in Predicates	44
2.38	The Forms of Predicates	45
2.39	Three Kinds of Predicates	45
2.40	Primitives	46
2.41	Complexes	46
2.42	Borrowings	46
2.43	Consonant-Pairs	47
2.44	Permissible Medials	47
2.45	Intelligibility at the C/CC-Joint	48
2.46	Hyphenation	48
2.47	Permissible Initials	48
2.48	The Decipherability of Complexes	49
2.49	Affix-Length and Frequency of Use	49
2.50	Term-Reduction	49
2.51	Long Affixes	49
2.52	Short Affixes & Their Derivations	50

## Notebook 3 - Contents

2.53	Affix-Assignment & Coverage	50
2.54	Pre-empted CVr Affixes	50
2.55	Making Complexes	51
2.56	The 'Tosmabru Test'	51
2.57	Allowable Borrowings	52
2.58	The 'Slinkui Test'	52
2.59	The Resolution & Partial Classification of Predicates	53
2.60	The Predicate Resolution	53
2.61	Term-Resolution	54
2.62	The Recognition of Borrowings	54
2.63	Making Borrowed Predicates	55
 CHAPTER 3. LEXICON (WORDS & SPEECH PARTS)		 57
3.1.	Definitions and Conventions	57
 (The *-ed lexemes are machine oriented)		
Lexeme A:	Afterthought Connectives (Eks)	58
Lexeme ACI:	Hyphenating Eks	58
Lexeme AGE:	Right-Grouping Eks	59
Lexeme BI:	Identity Operators	59
*Lexeme BAD	59	
Lexeme CA:	Predicate Word Connectives (Sheks)	60
Lexeme CI:	The Interverbal Hyphen	60
Lexeme CUI:	The Shek Left-Parenthesis	60
Lexeme DA:	Variables	60
Lexeme DIO:	Argument Tags	61
Lexeme DJAN:	Name Words	62
*Lexeme END	62	
*Lexeme FI:	The Utterance Ordinal Suffix	62
Lexeme GE:	The Grouping Operator	62
Lexeme GI:	The Fronting Operator	62
Lexeme GO:	The Inversion Operator	63
Lexeme GU:	The Optional Right Boundary Marker ("Comma")	63\
Lexeme GUE:	The GE-Group Optional Terminator	63
Lexeme HOI:	The Vocative Marker	64
Lexeme HU:	The Interrogative Argument	64
Lexeme I:	Sentence Connectives (Eesheks)	64
Lexeme ICI:	Hyphenating Eesheks	65
Lexeme IE:	The Identity Interrogative	65
Lexeme IGE:	Right-Grouping Eesheks	66
Lexeme JE:	The First Linking Operator	66
Lexeme JI:	Argument Modification Links	66
Lexeme JIO:	Subordinate Clause Links	66
Lexeme JO:	Metaphorizers	66
Lexeme JUE:	The Second Linking Operator	67
Lexeme KA:	Prefix Members of Forethought Connectives (Keks)	67
Lexeme KI:	Infixes for Forethought Connectives (Keks)	69
Lexeme KIE:	The Left-Parenthesis	69
Lexeme KIU:	The Right-Parenthesis	69
Lexeme LAE:	Indirect Designation Operators	69
Lexeme LE:	Descriptors	70

## Notebook 3 - Contents

Lexeme LEPO: Event Operators	70
Lexeme LI: The Left Quotation Operator	71
Lexeme LI: The Left Quotation Operator	71
Lexeme LIO: The Number Designator	71
Lexeme LIU: The Single-Word Quotation Operator	71
Lexeme LU: The Right Quotation Operator	72
Lexemes M1* through M11*	72
Lexeme ME: The Predifying Operator	72
Lexeme NI: Quantifiers	72
Lexeme NO: The Negation Operator	73
*Lexeme NOI: The Negation Suffix	74
Lexeme NU: Conversion Operators	74
Lexeme PA: Inflectors/Adverbs/Prepositions	75
Tense Operators	75
Location Operators	76
Modal Operators	76
Causal Operators	76
The Ga Operator	77
Lexeme PAUSE: The Pause-Comma	77
Lexeme PO: Abstraction Operators	78
Lexeme PREDA: Predicate Words	79
*Lexeme RA: Numerical Predicate Suffixes	80
Lexeme TAI: Letter Variables	80
Lexeme UI: Free Modifiers	81
Attitudinals	82
Discursives	83
Relative Interrogatives	83
Utterance Ordinals	84
Salutations	84
A Note on Other Free Modifiers	84
Lexeme ZE: The Joining Operator	84
*Lexeme ZI: Magnitude Suffixes	84
*Lexeme ZO: The Quantity Abstractor	85
CHAPTER 4. GRAMMAR (UTTERANCE FORMS)	86
4.1 Design Objectives	86
4.2 Definitions and Conventions	86
4.3 The Structure of Loglan Grammars	88
Group A. The Optional Punctuators, Rules 1-7	88
Group B. Linked Arguments, Rules 8-19	89
Group C. Predicate Units, Rules 20-33	91
Group D. Descriptive Predicates, Rules 34-48	92
Group E. Sentence Predicates, Rules 49-58	94
Group F. Modifiers, Rules 59-67	95
Group G. Arguments, Rules 68-116	96
Group H. Terms & Term Sets, Rules 117-127	100

## Notebook 3 - Contents

Group I.	Predicates, Rules 128-154	102
Group J.	Sentences, Rules 155-164	104
Group K.	Utterances, Rules 165-194	105
LIST 5 THE TEACHING CORPUS		109
A.	Imperatives & Responses ( <b>eo ao ai ae ti tu mi no</b> )	109
B.	Address & Response; Offers & Replies ( <b>loi loa sia siu ea oi mu</b> )	109
C.	Addressing vs. Naming ( <b>la ta e hoi</b> )	110
D.	Descriptions ( <b>le ne su gu</b> )	111
E.	Questions with he; Demonstratives & Plurals; Replacement with da ( <b>he da na ri ro levi leva</b> )	112
F.	Identity Questions & Sentences; Replacement with de & dui ( <b>ie bi hu i de dui</b> )	113
G.	Yes/No Questions & Answers; Utterance Demonstratives ( <b>ei ia toi toa</b> )	114
H.	Tenses; Time Questions & Answers; Local Modification; Punctuation ( <b>pa fa ji ipou nahu</b> )	115
I.	Time Phrases ( <b>pahu fahu fazi pazu tiu</b> )	116
J.	Space Questions & Answers; Space Phrases ( <b>vi va vu vihu vahu vuhu</b> )	117
K.	Existentials & Universals; Completion ( <b>ba be bo bu raba rabe rabo nibe ifeu inusoa vina uu</b> )	117
L.	Predicate Strings; Grouping, Hyphenation, Connection & Inversion ( <b>ge go ci ce ke ki</b> )	119
M.	More Connections & Groupings in Predicate Strings ( <b>gue cui canoi ka kanoi</b> )	120
N.	Event/State Predicates; Other Abstractions ( <b>po pu zo di</b> )	122
O.	Mass & Event Descriptions; Mass Event Descriptions ( <b>lo lovi lepo lopo</b> )	122
P.	Specified & Nested Event Descriptions ( <b>No new LWs.</b> )	124
Q.	Attitude Indication; Conversion, Negation & Superlatives ( <b>uo ue ua uu ui nu fu</b> )	125
R.	Counting, Quantifying & Numerical Questions ( <b>to te fo fe so se vo ve iesu iene ho hoba toba teba foba soba</b> )	127
S.	Quantified Descriptions & Questions ( <b>iete iefo ieho</b> )	128
T.	Measurement, Dimensioned Numbers & Numerical Description ( <b>lio lepa -ma -mei -dai</b> )	129
U.	Linked Description; Identity Clauses; Replacement with Letter-Words; Mixed Predicates and Arguments ( <b>je jue ze sui -mo -ai -ei (dai/dei, etc.)</b> )	129

## Notebook 3 - Contents

V.	Identifying vs. Claiming Subordinate Clauses ( <b>jio jia</b> )	131
W.	Sentence, Predicate & Argument Negation ( <b>ni</b> )	132
X.	Quotation of Loglan; Fronted Arguments ( <b>li lu liu gi</b> )	133
Y.	Predicates from Arguments and Prenex Quantifiers ( <b>me me- goi</b> )	133
Z.	Prenex Quantifiers ( <b>goi</b> )	134
AA.	Connected Arguments & Predicates; Joint Argument Sets ( <b>a anoi onoi noa efa epa gugu do</b> )	134
BB.	Causal Inflectors, Modifiers & Phrases ( <b>moi soa kou moipa numoi kouhu moihu nukouhu</b> )	136
CC.	Compound Term Connectives ( <b>enumoi enukou efa eva epa</b> )	138
DD.	Connective Questions ( <b>ha enoi noenoi</b> )	138
EE.	Internal Arguments ( <b>No new LWs.</b> )	139
FF.	Argument Ordinals (HB-tags) ( <b>pua pue pui puo puu</b> )	140
GG.	Compound & Connected Tenses ( <b>-fa- -pa- -na- ra- ne- ni- -noi-</b> )	141
HH.	Logically Connected Clauses ( <b>inoca icanoi ica Ice</b> )	142
II.	Causally-Connected Clauses ( <b>i- ki-</b> )	143
JJ.	Indirect Designation; Foreign Quotation ( <b>lae sae lie</b> )	143
KK.	Metaphor-Marking or "Figurative Quotation" ( <b>ja</b> )	144\
LL.	Letter-Variables and Acronyms ( <b>-z-</b> )	144
MM.	Predicates as Names & Vocatives ( <b>No new LWs.</b> )	145
NN.	Grouped & Ungrouped Afterthought Connections ( <b>i- -ge -ci</b> )	146
OO.	Spelling ( <b>No new LWs</b> )	147
PP.	Sentences in VOS Order ( <b>goa</b> )	147



## FOREWARD

The objectives of the present notebook are three. The first is to provide users of the 1975 language with a description of the present language which will allow them to become competent in Loglan once again. Once that is accomplished, I would hope these rearmed loglanists would then use the enlarged domain of modern Loglan in creative and testing ways, and communicate to The Institute their results. But there is a second objective. I have also tried to create a document that will serve as a teaching text—buttressed, as it now can be, by the two "MacTeach" (computerized flashcard) programs that have recently become available for learning primitives and affixes—but intended mainly for those who wish to learn the current language more or less from scratch. The third objective is to provide a technically complete description of the language that will serve as the easily updated reference manual we will soon need to back up the less formal and more popular publications which The Institute plans to offer to the general public when we go public again... a development of which, we trust, this notebook will be the final forerunner.

Current Loglan has emerged over the last four or five years from the word-making, grammar-expanding, and translating activities of a very few people. Their work has enlarged the language considerably, both in vocabulary, in grammatical domain, and in usage, and is now ready to be reported out. The translating and word-making activities were outgrowths of—actually, they were deliberately undertaken engineering tests of—our more publically-conducted 1976-1982 design studies of usage, grammar, and morphology... the last two having been called affectionately the MacGram ("machine grammar") and the GMR ("Great Morphological Revolution") projects, respectively, while they were still underway.

But well before these various engineering projects had boosted Loglan into a new and higher state, an active corps of competent users, albeit a small one, had developed by 1978 or '79 out of our 1975 publications. To be sure, their competence was in a language—or rather, in what were sometimes highly personal extrapolations from a plan for a language—which was substantially but incompletely described in the two 1975 books, **Loglan 1** and **Loglan 4 & 5**, and in the first four volumes of **The Loglanist**, 1976 through 1980, most notably in the **Supplement to Loglan 1**, a special issue of **TL** published in November 1980 which was the capstone of the first four years of public discussion. Incomplete as those earlier documentations of the language were even then, however, they are now, in addition, very badly out of date. And while there have been two subsequent special issues of **The Loglanist**—**TL6/1** in 1983 and **TL7/1** in 1984, issues designed to help people catch up with the then-current states of the language—even these two documents largely antedate the recent word-making, translating and grammar-expanding activities and so no longer tell the whole story.

Thus the first goal of the present notebook is simply to update the documentation of the language and make it whole. If that could be done well, I reckoned, then this third notebook would provide a tool with which once-competent loglanists would be able rapidly to restore their competence should they wish to do so. To serve their more sophisticated and often technical purposes, therefore, I have striven mainly to produce a description of the present language as I know it that would be as complete, as technically exact, and as conveniently cross-referenced as I have been able to make it.

It was during the early days of writing for experts in the Spring of 1986 that the notebook acquired its second purpose. A large number of the current partisans of Loglan, I had been learning, happen to have joined the project well after the creative ferment of the late '70's, and so did not participate in it. Moreover, there are many current loglanists who, although "old hands" in the historical sense, had never actually mastered the old language before it disappeared again into the engineering laboratory. Both kinds of potential users of the notebook began to write me. They, too, hoped to get some mileage out of the new notebook, especially now that developmental research on the language appeared to be slowing down. For these relative newcomers, then, but also for those old hands who have been until now only onlookers, I have tried to erect a second kind of document on the substructure provided by the first. In addition to a technical description of the current language, I have tried to produce a didactically useful, amply-illustrated account of the language from the point of view of the second-language learner. I have tried, in short, to provide these two kinds of sometime students of the language with the means by which, with some personal effort, they may at last become its masters.

These two objectives have not always been easy bedfellows. As the second one began to press itself upon me last Spring I had to admit that a book that promised also to be a reference manual for one-time

## Notebook 3 – Forward

experts is not an ideal place in which to teach a second language to completely innocent newcomers. Even so, examples are necessary even for experts. And such newcomers as choose to consort with experts can endure an algorithm or two. So I have attempted to select the examples and illustrations in this book in such a way that they will, of themselves, constitute a gradual climb through the structure of the language, starting at ground level with the utter simplicity of its phonology, rising through morphological and lexical materials of middle difficulty, and ending with what may, I fear, be found the stratospheric intricacies of the machine grammar. It is I trust a compact account, but it does move through these several levels of intellectual difficulty. (The language itself, of course, remains refreshingly simple...as I trust the reader will soon rediscover. It is just these increasingly exact scientific descriptions of it—which have been made possible and in some sense necessary by our increasingly exact understanding of it—that sometimes border on the intricate.)

There is a third objective of which I have only recently become aware; and that is the possibility that a second edition of **Notebook 3** may even now be looming. Suitably retitled, the next update of this notebook may very well be the one that accompanies the fourth edition of **Loglan 1** to the marketplace. This will probably be in the Spring or Summer of 1988; for it is then that The Institute presently plans to "go public again" with the language. If these plans do indeed develop in this way, then **Notebook 3** may be the first in a long series of continuously updated technical manuals, the purpose of which will be to describe in a single place the current state of the whole language. None of the specific requirements of that looming reference manual have, however, shaped the writing of the notebook...except of course for that ubiquitous canon of completeness, which has been dictated by the first objective as well.

A final note, and an apology. Earlier accounts of the contents planned for **Notebook 3** announced that it would include a small but exemplary vocabulary of scientific borrowings, as well as the algorithm that made them. I meant also to include the translation forays I had made into the international literature of science; for these had provided the test words in the first place and were meant, in the end, to contain them. These translation materials are not included. The latest reasonable date for the publication of this notebook—already twice delayed—was Mid-Summer 1987. I could not make the algorithm for the construction of "best scientific words"—a process that involved, as usual, a statistical analysis of the many judges' opinions I have collected—in time to include it, and the vocabulary it was intended to make exemplary, in the notebook. And to publish my translations with a non-exemplary vocabulary seemed counter-productive. I am sorry to disappoint those readers who expected to find this textual material in this notebook. Perhaps another notebook will be in order after this one has had its day. On the other hand, it seems increasingly likely that the next large task for The Institute, after the loglanists have made whatever use they wish to make of this one, will be GPA (The Institute's acronym for Going Public Again).

But even without the scientific word-list, and the translations that evoked them, the contents of **Notebook 3** will, I trust, be found substantial. It not only contains the most complete description of this developing language that has ever yet been published, it is the first publication since 1975 that even purports to describe the language as a whole. I trust, therefore, that everyone who endures the long march through its lists and pages will have a reasonably good chance of learning to use, for da's own purposes, the extraordinarily rich creative instrument that Loglan has lately become.

JCB  
Gainesville  
July 1987

## CHAPTER 1

### PHONOLOGY (SOUNDS & SOUND-VARIANTS)

**1.1 Definitions and Conventions:** We require a small technical vocabulary. A **phoneme** is a class of one or more speech-sounds all of which are regarded as “instances of the same sound” by the speakers of some language. The individual sounds so classified are called **phones**. All the phones which are members of a given phoneme are called its **allophones**. I shall use strokes [/] to mark off phonemes, sets of phonemes, and phonemic transcriptions. Thus /a/ is a phoneme, /a e i o u/ is a set of phonemes, and /eAmuGODzi/ is a **phonemic transcription**. I shall use parentheses to mark off phones, sets of phones, and phonetic transcriptions. Thus (ah) is a phone, (ah aa) are the two allophones of phoneme /a/, and (ey-AH-moo-GOHD-zee) is the **phonetic transcription** of the utterance phonemically transcribed as /eAmuGODzi/. The utterance itself is **Ea mu godzi**; it means ‘Let’s go!’ Phonetic transcriptions of utterances will sometimes be called **guides**; they are phonetic guides to at least one way of pronouncing those utterances. Phonemic transcriptions are often referred to simply as **productions**. They exhibit one of the ways an utterance might be produced in speech.

**Stressed syllables** are shown in uppercase letters in both the guides and productions. Hyphens are used to show **syllabification** in guides. Stress rise and fall, shown by the shift from lower to upper case characters, are normally sufficient to show syllabification in productions. But when there are strings of unstressed vowels it is sometimes necessary to show syllable breaks in a production by using **close-commas** (commas without an adjacent space), e.g., /i,aimiGODzi/. This production shows us that the utterance involved was **I ai mi godzi** = ‘And yes I intend to go’. The close-comma between them puts /i/ and /ai/ in separate syllables. Without the close-comma, the **default syllabification rule** would pair the vowels from the left. Thus /i,aimiGODzi/ syllabifies as (yah-ee-mee-GOHD-zee), and this gives quite a different resolution, namely **\*Ia i mi godzi** = ‘Certainly; and I go’.

The **asterisk** [\*] on **\*Ia i mi godzi** shows that this utterance is wrongly pronounced. I shall use leading asterisks [\*] to mark all expressions which are illegitimate at some level of correctness, whether that level is phonological, morphological or grammatical. Even bad usages are so marked. **\*Ia i mi godzi** is morphologically incorrect because it is missing an obligatory pause before the little word **I**; such pauses are sometimes necessary for resolution. I shall use leading **question-marks** [?] to show that an expression is of unknown legitimacy—unknown, that is, to the reader. Thus **?Ia i mi godzi** is how we would mark this utterance before the reader is expected to know that it is malformed.

Pauses are shown in written specimens either by a **comma-space**, e.g., **Ia, ice mi godzi** (‘Certainly; and I go’), or by **period-spaces** which require two following spaces, e.g., **Ia. I mi godzi** (‘Certainly. I go’). The second marking is equivalent to the full-stop in English. Both pauses are shown in productions by **close-periods**, e.g., /ia.imiGODzi/, and in the guides by **open-periods** (space-period-space). e.g., (yah . ee-mee-GOHD-zee). So the correct pronunciation of these last two specimens is given by the pair of guides (yah . ee-mee-GOHD-zee) for the one that means ‘Certainly. I go’ and (yah . EE-sheh-mee-GOHD-zee) for the one that means ‘Certainly; I go’. A **silence** is the long pause before an utterance begins or after one ends; it is not usually useful to show silences in productions, but when it is useful to do so they are shown by pound-signs [#]. A **breathgroup** is a pauseless string of phonemes lying between pauses, silences, or pauses and silences in either order.

I shall use **square brackets** to enclose textual **characters**, or sets or strings of characters. Thus [a] and [l] are visual characters; [, ] is the comma-space and [ . ] is the period-space; [ . ] is the open-period used in guides; while [, ] [.] and [#] are the close-comma, close-period and silence, respectively, used in phonemic transcriptions. Although we will rarely have occasion to use bracketed utterances, we can also say that [Ea mu godzi] and [Let us go.] are visually perceived utterances, that is, pieces of text.

Sometimes I shall wish to refer to characters that are not on the fonts available to me for printing this document. I shall say, for example, that in lieu of the proper characters the printed expression [alpha] stands for the Greek lower-case character called alpha while [Alpha] refers to the upper case version of that character.

I shall, as is customary in Institute publications, use **bolding** to mark Loglan specimens whenever their status as text or speech is immaterial. Thus **Ea mu godzi** is the utterance variously represented above. I shall also use bolding to indicate **emphatic stress** in both phonemic transcriptions and guides: /eAmu**GODzi**/

## Notebook 3 – Phonology

= /ey-AH-moo-**GOHD**-zee/ = ‘Let’s go!’ I shall use **single quotes** to mark non-Loglan (usually English) specimens when the mode of delivery is immaterial, and I shall use **underlining** for emphasis in both such non-Loglan specimens and in ordinary text. Thus ‘Let’s go!’ is a translation of **Ea mu godzi** pronounced as /eAmuGODzi/ with no special stress on any word, while ‘Let us go!’ is a translation of the same utterance when pronounced /eAmu**GODzi**/ with **godzi** emphasized. (Bolding is also used for titles, technical terms about to be defined, and for chapter, page and section headings in this document...I trust without confusion.) I shall use **double-quotes** for English expressions in the text which are either not being used literally or which have not been defined yet. Thus the allophones of a phoneme may be said to be “instances of the same sound” but are not literally.

The Loglan words which I customarily use as English words when I am addressing loglanists remain unmarked. For example, ‘sutori’, which means ‘at least second’ or ‘second and subsequent’, as in ‘the sutori places of the predicate’, has become a quasi-English word in my idiolect and so is undistinguished in my English prose...just as the French, German or Latin words and phrases which have been taken into English are now usually unmarked in written English (‘That’s a priori reasoning’ ‘He’s a gestalt psychologist’). I expect loglandic readers to know the meanings of these borrowings from Loglan, as from German or Latin, or to catch up by looking them up. Besides, they impart a usefully loglandic flavor to one’s English thought and speech. The borrowings from Loglan which I notice I use most frequently are ‘da’ and its kin. I treat these words as genderless, numberless, and caseless 3rd person English pronouns.

I occasionally use the sign ‘=>’ to mean ‘is the source of’, ‘<=’ to mean ‘is derived from’, and ‘=’ between a specimen and its translation to mean ‘has approximately the same meaning as’. I use parentheses for other uses than Loglan phonetics, of course, as in the next paragraph. I trust this will cause no confusion.

I have used **least equivocal spellings** in ordinary English letters (a procedure made familiar by the Berlitz people) to represent all the Loglan phones we’ll need. Thus, I’ve used (sh) to represent the sound of Loglan /c/, (ee) for the principal sound of Loglan /i/, (oo) for the main value of its /u/, (igh) for the otherwise hopelessly equivocally-spelled ‘eye/aye/I’-sound of English, which in Loglan is the sound of the diphthong /ai/, and I’ve used (eigh) for the sound it invariably spells in such words as ‘eight’ and ‘freight’ for one of the two allophones of Loglan /e/; and I have chosen to do this rather than use the special symbols of the International Phonetic Alphabet. The IPA symbols are not widely understood except among linguists; changing fonts to print them would slow down the production of this notebook by about a factor of ten; and our discussion of the sounds of Loglan loses little in technical precision by our adoption of this humbling convention. Phonetic issues in Loglan are not so wide-ranging that they require the full armamentarium of modern phonetic scripts.

**1.2 Two Types of Phonemes:** There are two types of phonemes in Loglan: the **regular** phonemes that are used throughout the language, and the **irregular** ones that may be used only in (proper) names. The term ‘name’ will always mean a proper name in our discussion of Loglan.

**1.3 Regular Phonemes:** There are 23 regular phonemes in Loglan: /a b c d e f g h i j k l m n o p r s t u v y z/; so to represent them, Loglan uses the entire English alphabet less the letters [q w x]. Two of these phonemes, /h y/, are new since 1975. /h/ was added in 1981 because it was present in all the target languages and adding it made some much needed word-space available for new little words. Unexpectedly, incorporating /h/ also increased the average recognition scores of Loglan composite primitives (C-Prims) due to the large improvement in the recognizability of the 100 or so primitives that were remade with /h/. Schwa was added in 1981 as the buffer in buffered dialects and was assigned the unused English letter [y]. In 1986 the sound /y/ was given the even more important function of serving as the intraverbal hyphen in regular words; see Sec. 1.8.

**1.4 The 6 Regular Vowels:** The 6 regular vowels are pronounced as in the following table. The primary allophone is given first, the second, if one is recognized, appears on following lines:

/a/	As in ‘father’ in the Germanic dialects of Loglan, but as in S. ‘casa’, F. ‘la’ and E. ‘palm’ in the Romance dialects. In the “standard” (Gainesville) dialect, the Romance pronunciation is preferred.	
	Romance /a/, as in S. ‘casa’, F. ‘la’ 6c E. ‘palm’	(aa)

## Notebook 3 – Phonology

	Germanic /a/, as in G. 'vater' 6c E. 'father'	(ah)
	<b>Problems:</b> In unstressed syllables, /a/ is occasionally but incorrectly pronounced as in 'about' or 'sofa', which permits it to be confused with the (uh) sound of /y/; see below.	
/e/	As in 'met'	(e) or (eh)
	Before vowels, the sound of 'eigh' in 'freight' or 'a' in 'ate'; it is also the first part of the diphthong 'ey' in 'grey'	(eigh)
	<b>Problems:</b> There is a strong tendency for English-speaking persons to use either the monophthong (eigh) as in 'eight' or the diphthong (ay) as in 'day' for /e/ in V-final monosyllables, i.e., to say 'day' for <b>de</b> and 'say' for <b>se</b> . This must be resisted, if only because Loglan words like <b>dei</b> and <b>sei</b> also exist in which the diphthong (ay) is genuinely present. This tendency can be resisted by making sure that <b>de se</b> are pronounced like the first parts of 'debt' 'set'. These then contrast sharply with <b>dei sei</b> as properly pronounced (day say).	
/i/	As in 'machine'	(ee)
	Before vowels, 'y' as in 'yet'	(y)
/o/	As in 'note' but shorter; resist the tendency to say 'oh-oo'	(oh)
	Before /r i/, as in 'more' or 'noise'	(aw)
/u/	As in 'lute' but shorter; resist the tendency to say 'oo-oo'	(oo)
	Before vowels, 'w' as in 'woo'	(w)
/y/	'e' as in unstressed 'the', 'a' in 'sofa' and 'above', 'u' in 'up' and 'under'	(uh)

**1.5 The Advantages of Romance (aa):** Loglan /a/ has two permissible sounds, the Germanic (ah) of 'father' and the Romance (aa) of 'palm' and Spanish 'casa'. I shall usually use the Romance (aa) of 'casa' in the guides simply because I wish to remind the reader that, in the regular words of Loglan (although not necessarily in its names), the Romance pronunciation is preferred over Germanic (ah). Thus, I will usually guide the reader toward the preferred pronunciation (MAAT-maa) of **matma** = 'mother', despite the fact that the Germanic pronunciations (MAHT-mah) and \*(MAHT-muh), the second one erroneous, are much more common among current loglanists, who are largely anglophones. (English, along with Dutch, German and the several Scandinavian languages, is a Germanic language.) But remember that you may always substitute the more familiar Germanic sound of (ah) for any of my (aa)s. \*(MAHT-muh), however, is not an acceptable pronunciation of **matma**. If you're going to use Germanic (ah), you must do so in both syllables, saying (MAHT-mah) with the (ah) of 'father' in both places.

The chief difficulty with Germanic (ah) is that it is a slack sound that easily degenerates into (uh) in unstressed syllables...especially in the mouths of native speakers of Germanic or Slavic languages. Speakers of such languages regularly and unconsciously allow this slackening to happen to many of their unstressed vowels. So, if you do elect to try to pronounce **matma** consistently as (MAHT-mah), the second (ah) is very likely to slip into a grunted (uh) when you're not listening. Unfortunately, hard as you try, you'll probably end up saying \*(MAHT-muh)...which may cause trouble for your human auditors and eventually for your computer. If, on the other hand, you at least try for the tense Spanish (aa) at the outset, trying to say (MAAT-maa) with the stretched lips, wide-open mouth and higher pitch that the production of this Romance (aa) requires—think of a Spaniard saying 'casa'—then your Loglan /a/ is much less likely to slump off into a grunted (uh) even when it is unstressed.

## Notebook 3 – Phonology

That's the main reason, then, why you should consider trying to speak the Romance dialect of Loglan: your unstressed /a/s are much more likely to remain intelligible. If you know any Romance languages, you can easily find your models for this (aa). Spanish especially is noteworthy for them. But if you are a monolingual speaker of American English, then the Boston Irish version of 'faather'—if you can conjure it up from a recent movie—is probably your best model...unless, of course, you have The Institute cassettes, which abound in Romance (aa)s.

**1.6 The TWO Spellings of 'e' of 'met':** Notice that the principal allophone of /e/, which is the 'e' of 'met', has two spellings in the pronunciation guides, namely (e) and (eh). These expressions spell exactly the same sound, namely that same 'e' of 'met'. Two spellings are necessary because there is a pair of phonetic contexts in which each spelling would be misleading if it were used in both contexts, yet each is a good guide in one of them. So both spellings are required, each in its own context. Thus (e) is used to guide pronunciation when /e/ just precedes a consonant, as it does in 'met' itself, unless that consonant is /r/. Thus the 'e' of 'met' in **metro** 'meter' is spelled (e) in the guide (MET-roh). But the same 'e' of 'met' when found with following /r/ is spelled (eh), as it is in (KEHR-tee), for example, which is a guide to the pronunciation of **kerti** = 'air'. The same sound is also spelled (eh) when /e/ is final in a word or syllable. This doesn't happen in English, so we can give no English examples. But the 'e' of 'met' is final in many Loglan words. Thus, **ne se pe** all end with the 'e' of 'met'. So the guides for these three words are (neh seh peh), which happen to be the sounds of English 'net set pet' with the 't's omitted.

To take another example, the pronunciation of the Loglan word **ie**—which means 'Which?'—has the value of the first two sounds in English 'ye(t)'. As /e/ is final in **ie** it will be written with (eh) in the pronunciation guide, which is therefore (yeh). Notice that using (e) alone in the guide to **ie**—that is to say, giving its pronunciation as \*(ye)—would create quite a different, and for some speakers an erroneous, impression of the sound of this Loglan word. Once again, **ie** is pronounced (yeh), and this is identical to the first two-thirds of the word 'ye(t)'. This is a very un-English way of pronouncing a final /e/, but with a little effort you can learn to do it.

Loglan /e/ has a second allophone, of course. This allophone is the sound of (eigh) in 'eight' which you may have been tempted to give /e/ in words like **ne se pe**, thus pronouncing them like 'nay' 'say' 'pay' incorrectly. But in Loglan, the (eigh) allophone of /e/ is reserved for positions before vowels, and it never occurs anywhere else. Thus (eigh) occurs in words like English 'Mayo' (MEIGH-oh) and Loglan **eo** (EIGH-oh) ('Please') because this /e/ precedes a vowel; but (eigh) is never correct for /e/ when /e/ precedes a consonant or is final in a word or syllable. Native Romance speakers are inclined to misuse Loglan (eigh) in just this way.

**1.7 The Odd Spellings of /i/ and /y/:** Two of the 6 vowel phonemes will seem oddly-spelled to the English-reading eye, namely /i/ and /y/. These phonemes have been given letters which commonly have quite different phonetic values in English. But the sounds of these Loglan phonemes—which are (ee y) in the case of /i/ and the "grunted" (uh) in the case of /y/—are certainly not odd to the English ear. The sounds (ee y) are in fact very common in English. Indeed, the letter [i] has just these values in many European languages, including but not only the Romance ones. On the other hand, the short, "grunted" vowel (uh) that is spelled with the letter [y] in Loglan is not consistently represented in any European spelling, the (uh)-sound being usually regarded as a "slighted", or even a "degenerate", version of some other vowel. Thus, the German letter [e] when final is always sounded as (uh), even though [e] has other values in other contexts. In English, (uh) is sometimes spelled with [e] as in 'the' (thuh), sometimes with an [a] as in 'sofa' (SOH-fuh), sometimes with a [u] as in 'upon' (uh-PAHN), and sometimes with an [o] as in 'phonetic' (fuh-NEH-tik); so it does not appear to "have a letter of its own" in English. On the other hand, the sound (uh) is extremely common in English. So having such a familiar sound uniquely represented by the same single letter, even an odd one, may actually be something of a relief for English-reading eyes. Note that the work the letter [y] most commonly does in English is performed in Loglan by another letter, namely [i]. This frees the letter [y] to do its present work in Loglan as (uh).

**1.8 /y/ as a Hyphen:** As we have seen, the sound of Loglan /y/, which is often called schwa (shvuh) by linguists, is a very short, usually unstressed, "grunted" vowel that occurs in all Germanic and Slavic languages, and in many other consonant-rich languages as well, /y/ has a very odd distribution in Loglan in

## Notebook 3 – Phonology

that in the regular words of **unbuffered dialects**—see the next section for its role in the buffered ones—it occurs only as an **intraverbal hyphen**. In these dialects—which include the standard one—/y/ is the pronunciation of the hyphen-like letter [y] that is used to "glue" the parts of complex words together. An example is **mekykiu** (MEK-uh-kyoo), which means 'ophthalmologist' or 'eye-doctor'. In this word the hyphenated elements /mek/ and /kiu/ are not words, but **affixes**. An affix is a usually shortened representative of one of the words in the **defining metaphor** used to build a complex word; see Sec. 2.55 on Making Complexes. In this case, the defining metaphor is **menki kicmu** (MENG-kee-KEESH-moo), and **mek** and **kiu** are derived from these words by processes that we will consider later; Sec. 2.52. The segment /y/ between them turns them into a single word. The phrase **menki kicmu** also means 'eye doctor' but this time without the English hyphen. When /y/ is used as a hyphen in regular words, it is always a never-stressed sound between affixes. Because the sound (uh) is a hyphen it does not even count as a syllable in locating stress; see Sec. 1.11.

In the regular words of unbuffered dialects the phoneme /y/ has no other use than as a hyphen. That is to say, /y/ is always the sign of the intraverbal hyphen in these dialects whenever it appears in a non-name; see Sec. 1.12 for the uses of schwa in names and below for its uses in buffered dialects.

**1.9 /y/ as a Buffer:** In the **buffered dialects** of Loglan, /y/ may be used as a buffer between any pair of "difficult" consonants, even when that pair is wholly within an affix or a simple word. For example, in the Japanese dialect of Loglan, the word for 'mother' will probably be buffered. It will become **matyma**, a 3-syllable word pronounced (MAAT-uh-maa). Here the phoneme /y/ is not used as a hyphen. Indeed, **matma** is a simple word without any constituent affixes, so it could hardly contain a hyphen. Instead the /y/ in **matyma** is a buffer between the two "difficult" consonants /t/ and /m/, that is, it makes them pronounceable. (Consonant pairs are **difficult** when the speakers of a buffered dialect choose to buffer them by introducing /y/ between the members of the difficult pair.) If this happens, **matyma** will become a **dialect word**, that is, a harmless variant of some standard word. It will have exactly the same meaning as the standard word **matma**. Moreover, we expect (MAAT-uh-maa) to be as easily recognized by a speaker of the standard dialect as (MAAT-maa) is likely to be recognized by speakers of the Japanese dialect as a "variant" of their own **matyma**.

**1.10 /iy/ as a Hyphen in Buffered Dialects:** Notice that the speakers of a buffered dialect may not use /y/ as their intraverbal hyphen. /y/ has already been preempted as their buffer, and to use it for the hyphen as well would lead to confusion for the resolver as to where the "joints" in a complex word actually were. This is contrary to the spirit of Loglan, in which not only are all words uniquely resolvable, but also all terms within words. Using /y/ for both hyphen and buffer in a buffered dialect would make the term-resolver's task impossible in that dialect.

So the hyphen adopted for use in these dialects is /iy/ (pronounced (yuh)). Thus the Japanese speaker of Loglan is obliged to use **mekiykiu** pronounced (MEK-yuh-kyoo) for **mekykiu** (MEK-uh-kyoo) 'ophthalmologist' if da buffers any consonants at all. Of course, if dab doesn't buffer da's consonants, da is in that respect speaking the standard dialect. In that case da may use the standard hyphen /y/ without fear of spoiling the resolvability of his speech; see Sec. 2.1 on resolution.

**1.11 The Effect of Hyphenating and Buffering on Stress:** In the regular words of any dialect the buffer /y/ and the hyphens /y/ and /iy/ are always unstressed. Hyphens and buffers are "so unstressed", in fact, that they do not even count as syllables in locating stress. This rule permits the same syllable to be stressed in both the buffered and the standard versions of a given word, e.g., (MAAT-uh-maa) and (MAAT-maa). Indeed, because of this rule both versions of the word **matma** may be said to be "penultimately stressed"; the penultimate stress rule is discussed in Sec. 2.37. Although **mekykiu** and **mekiykiu** are both stressed on the syllable (MEK) in (MEK-uh-kyoo) and (MEK-yuh-kyoo), they may both be said to be penultimately stressed; for neither hyphen counts as a syllable.

## Notebook 3 – Phonology

**Table 1.1 Permissible Pronunciations of the Twenty-Five Loglan Vowel-Pairs**

**1.12 /y/ in Names:** The phonology of /y/ in names is, as usual, a different matter. In names /y/ may occur in either stressed or unstressed positions in all dialects. For example, /y/ appears in a stressed syllable in 'Hunter' => **Hyntr** (HUHN-trr) and in an unstressed one in 'Washington' => **Ua'cintyn** (WAH-sheen-tuhn); see Secs. 2.4-12 on making names.

But what are we going to do with the Japanese version of the name for Amsterdam, which my English-Japanese dictionary lists as 'Amusuterudamu'? Like all loglanists, Japanese loglanists will no doubt be making the effort to pronounce the name of this city in the way its inhabitants do. In standard Loglan, that effort produces **A'msterdam** (AHM-stehr-dahm), for we anglophones who speak standard Loglan are as tolerant of consonant-clustering as the Dutch are. But not so the Japanese. The habit of buffering, which is encouraged elsewhere in their dialect of Loglan, may well cause them to buffer their Loglan names too. In that case, the word for **A'msterdam** in the Japanese dialect of Loglan is likely to come out **A'mysyterydam** (AH-muh-suh-tehr-uh-dahm) or even **A'mysyterydamy** (AH-muh-suh-tehr-uh-dah-muh). Unfortunately this seems to generate two, or even three, different Loglan names for the same place. But the "misfortune" is only apparent. Having one, two or a dozen variant names for Amsterdam, along with the Loglan standard one with which all the others may be equated by an easy algorithm, is as unlikely to be troublesome for the mutual intelligibility of the two dialects as the fact that buffering creates two words for mothers. Each phonological variant is a synonym of the standard, and once the pattern of buffering in any given dialect has stabilized, each member of such synonym pairs will easily be recaptured from the other. So each dialect word will be as easily understood by the speakers of the standard dialect as the dialect speakers will understand the standard word. Mutual intelligibility of dialects is the aim. We think that consonant-buffering will advance this aim rather than retard it.

	-a	-e	-i	-o	-u
a-	—	—	(igh)	(ow)	—
	(aa-aa) or (ah-ah)	(aa-eh) or (ah-eh)	—	—	(aa-oo) or (ah-oo)
e-	—	—	(ey)	—	—
	(eigh-aa) or (eigh-ah)	(eigh-eh)	—	(eigh-oh)	(eigh-oo)
i-	(yaa)	(yeh)	(yee)	(yoh)	(you)
	(ee-aa) or (ee-ah)	(ee-eh)	(ee-ee)	(ee-oh)	(ee-oo)
o-	—	—	(oy)	—	—
	(oh-aa) or (oh-ah)	(oh-eh)	—	(oh-oh)	(oh-oo)
u-	(waa)	(weh)	(wee)	(woe)	(woo)
	(oo-aa) or (oo-ah)	(oo-eh)	(oo-ee)	(oo-oh)	(oo-oo)

**1.13 The 25 Vowel-Pairs:** Loglan abounds in vowel-pairs. Not being allowed to enter the syllables of regular words, schwa or /y/ does not figure in any of them. But all possible pairings of the other five vowels /a e i o u/ occur in some words of the language. This creates the 25 vowel-pairs shown in Table 1, which provides a “phonological map” of their pronunciations. The upper entries in each cell of the table are the monosyllabic pronunciations of that particular pair if such a pronunciation is permitted. A ‘—’ is entered on the first line if it is not. The lower entries in each cell are that pair’s pronunciation as a disyllable, again if that is permitted.



## Notebook 3 – Phonology

Notice that some pairs in the table may be pronounced both monosyllabically and disyllabically. Thus /ii/ may be (yee) or (ee-ee). Others, like /aa/, may only be pronounced disyllabically. Still others, like /ai/, apparently occur only monosyllabically. That is, /ai/ may be pronounced only as the (igh) of **hai** and **sai** and never as \*(ah-ee). Incidentally, these two Loglan words are identical in sound to English 'high' and 'sigh'.

In using this table, try to pronounce the pairs with hyphenated guides as two distinct syllables with a "glide" between. It is easier to do this if you stress one of the two syllables relative to the other. While both stress patterns are possible, stressing the first syllable in a vowel-pair is a bit more common in Loglan. As an anglophone, you are likely to find stressing the first syllable a bit more "natural" as well.

We observe that the **always-monosyllabic** pairs are the four "natural" diphthongs /ai ao ei oi/. These can easily be pronounced monosyllabically as (igh ow ay oy) in any context; and so they are.

The next group are the eleven **always-disyllabic** pairs. These are /aa ae au ea ee eo eu oa oe oo ou/. Except for /au/ and /ou/ these pairs cannot be pronounced as diphthongs; and so they never are. /au/ and /ou/ are included among the pairs that are always disyllables to prevent their monosyllabic pronunciations from being confused with one another and with the natural diphthong /ao/ = (ow). So /au/ and /ou/ are always pronounced disyllabically in Loglan, that is as (aa-oo) and (oh-oo) (or (ah-oo) and (oh-oo) if the Germanic /a/ of 'father' is preferred); and /ao/ is always pronounced (ow). This keeps them acoustically quite separate from one another.

The remaining ten pairs are the optional disyllables and are discussed separately in the next section.

**1.14 The 10 Optionally Disyllabic Vowel-Pairs:** Notice that the 10 vowel-pairs which have both monosyllabic and disyllabic pronunciations in the Table 1 are all those that begin with either /i-/ (**ia ie ii io iu**) or /u-/ (**ua ue ui uo uu**). (In English, we would think of them as the 'y'- and 'w'-words, respectively.) They are therefore quite easy to remember. In certain contexts these optional disyllables are all easily pronounced as monosyllables. For example when they stand alone they are always given the values (yah yeh yee yoh yoo) and (wah weh wee woh woo). But there exist contexts in which these vowel-pairs are difficult to pronounce monosyllabically. Following /r l m n/ seems to be four of these contexts. So in these four contexts, and perhaps some others not yet identified, the loglanist is free to pronounce the **i-** and **u-**words either as monosyllables or as disyllables, that is, as (ee-ah) and (oo-ah), etc. Thus, the Loglan word **lui** ('to please...') may be pronounced either as the English name 'Louie' (LOO-ee) or as the French name 'Louis' (l'wee); and which is chosen is largely a matter of personal preference. The word will resolve as **lui** either way. In the Gainesville dialect, the swift French monosyllables seem currently to be preferred.

Uncertainty about the pronunciation of the optional disyllables reflects the fact that The Institute does not know yet how loglanists of a sufficiently wide variety of native-language backgrounds are going to treat them. So we will wait to describe a general pronunciation pattern here, preferring to observe the speech of active loglanists a little longer. In the meantime, no morphological confusion is generated by keeping these options open.

**1.15 The Pair-from-the-Left-Rule:** The default rule for grouping a written string of vowels into pairs is to start pairing them from the left and to restart the pairing process at any marked pause or syllable-break encountered. Once this is done we have to examine the pairs so made to see whether they are obligatory monosyllables, obligatory disyllables, or optional disyllables. For example, suppose we encounter the written word [aiuiaoea], which is an implausibly long "compound attitudinal"; see Lexeme UI. How do we pronounce it? The first part of the answer is given by the pairing rule. Starting from the left we can try first to syllabify the string as ?/ai,ui,ao,ea/. (We don't know yet whether this is a correct transcription of this new word, so we mark it with leading [?].) We now see that /ui/ is an optional disyllable, and so may be pronounced either (wee) or (oo-ee), and that /ea/ is an obligate disyllable that must be pronounced either (eigh-aa) or (eigh-ah) depending on our personal choice between Romance or Germanic /a/. The other two pairs are the obligate monosyllables (igh) and (ow). We choose to pronounce /ui/ as (wee). This gives us our second transcription, namely ?/ai,ui,ao,e,a/. But what about stress? We decide to use penultimate stress for this new word (which we probably haven't heard yet), for this is another good default rule; see Sec. 2.37. So we produce ?/ai,ui,aoEa/ as our next transcription, this time exploiting the economizing convention that wherever there is a stress-rise or -fall in a transcription, a syllable break is inferrable. Is this the most economical way we can transcribe the pronunciation (igh-wee-ow-EIGH-aa) phonemically? No; it is not.

## Notebook 3 – Phonology

The very pair-from-the-left rule we have been studying allows us to remove the two remaining close-commas. Thus /aiuiaoea/ transcribes the pronunciation (igh-wee-ow-EIGH-aa) accurately and compactly. It does so because we can infer that there are syllable breaks between the vowel-pairs as counted from the left.

What if someone elects to exercise da's privilege of pronouncing the optional disyllable [u] in the word [aiuiaoea] disyllabically? This rather odd pronunciation—odd because this [ui] is not in the difficult context of leading [l r m n] in which we might expect some readers to prefer the disyllable—can nevertheless be shown by the insertion of two close-commas, one on either side of the second [i]: /aiu,i,aoEa/. This does indeed give (igh-oo-ee-ow-EIGH-aa). One close-comma would not be sufficient here; the pairing rule starts over again with any pause or inferrable syllable break. As a consequence, the transcription /aiu,iaoEa/ would have to be pronounced (igh-oo-yah-oh-EIGH-aa). The interested reader is invited to work out other details of this transcription system.

**1.16 Indications of Syllabicity:** The novice is of course not expected to remember which pairs of vowel letters in a given string are to be pronounced monosyllabically, or to carry out all the above inferences. So at first we need to tell da. In the pronunciation guides the "syllabicity" of a vowel-pair—i.e., whether it is to be pronounced as a monosyllable or as a disyllable—is always plain. All syllable-joints within a breathgroup are indicated by hyphens [-]. So if a pair of vowels is shown with a hyphenated joint, it's a disyllable; if it is not, it's a monosyllable. Thus, (LOO-ee) and (lwee) are plainly the disyllabic and monosyllabic productions, respectively, of **lui**. (They are also the pronunciations of English 'Louie' and French 'Louis'.)

But in phonemic transcriptions, where syllable structure has been played-down so as to lift other features of Loglan speech into prominence, we have seen that certain inferences may be drawn from the transcriptions themselves that will satisfactorily distinguish disyllabic from monosyllabic productions of all vowel-pairs. By way of summary, these inferences proceed as follows.

First, if the two vowels of some pair are shown as having different levels of stress—as they are shown to have in both /eAmuSUCmi/ **Ea mu sucmi** = 'Let's swim' and /muSUCmiEa/ **Mu sucmi ea** = 'We swim, I suggest'—then they are obviously disyllabic in both productions. Indeed **ea** is one of those eleven vowel-pairs, see Sec. 1.13, that are always disyllabic. So what we are doing in these two transcriptions is showing which of the two vowels in **ea** is being stressed: (eigh-AA-moo-SOOSH-mee) vs. (moo-SOOSH-mee-EIGH-aa).

Second, if both elements of an identifiable vowel-pair are shown as having stress—both emphatic or both normal stress—then that pair of vowels is to be pronounced as a monosyllable. Thus in /AImuSUCmi/ **Ai mu sucmi** = 'Yes, we'll swim' the /AI/ is obviously being pronounced monosyllabically, that is, as the (IGH) of 'high'. Loglan **ai** is the 'Aye' of intention, compliance or consent...as in the nautical 'Aye, aye, sir!'. We can of course write the pronunciation guide for this utterance as (IGH-moo-SOOSH-mee). (But is it possible that you are beginning not to need the guides?)

This leaves pairs of unstressed vowels as the only unexamined Case. For example, in /eituMREnu/ = (ay-too-MREH-noo) for **Ei tu mre nu** = 'Are you a man?'—literally, 'Eh, you man?'—both phonemes in the interrogative **ei** (which is pronounced, remember, as the English diphthong (ay) in 'say') are shown to be without stress. The interrogative **ei** is often unstressed when it is in utterance-initial position. In utterance-final position, however, **ei** is usually stressed. We can show this by writing /tuMREnuEI/ (too-MRE-noo-AY). We know from Sec. 1.13 that /ei/ is one of the four vowel-pairs that are always pronounced monosyllabically, namely /ai ao ei oi/. But even if we didn't know this, or had forgotten, we could infer the monosyllabicity of **ei** from the transcription /eituMREnu/ itself. On the other hand if we wish to show that two adjacent unstressed vowels are being pronounced disyllabically, as for example if someone spoke the invitation **Ea mu sucmi** without stressing either syllable of the disyllabic /ea/, then we could show that production with a close-comma in the transcription: /e,amuSUCmi/. So we can report the stress possibilities for **ea** in this sentence as the set of three productions: /eAmuSUCmi/ /EamuSUCmi/ and /e,amuSUCmi/. Make sure you can pronounce them all. The fourth logical possibility, /E,AmuSUCmi/, is not really a lively phonological one.

We can now write the third inference as follows:-

Third, if two adjacent unstressed vowels are separated by no mark, and they are indeed members of the same vowel-pair, then they are to be pronounced monosyllabically. Thus, /ei/ in /eituMREnu/ can be inferred to be a monosyllable because no mark separates the unstressed phonemes /e/ and /i/ from each other. (Note that from this point of view, a **stress-rise**, such as occurs in /eA/ (eigh-AA), and a **stress-fall**,

## Notebook 3 – Phonology

such as occurs in /Ea/ (EIGH-aa), are both "marks".) If and only if the two unstressed adjacent vowels are separated by a close comma /,/—as in /e,amuSUCmi/—are they intended to be pronounced disyllabically. Such a close-comma is not a pause; but it does indicate elongation of the time interval occupied by these two unstressed phonemes just as pausing between them would do.

It is important that we transcribe these ambiguous unstressed cases in this explicit way and not rely on the reader's memory of what kind of vowel-pairs they are. Besides, we have those ten optional i- and u-initial pairs to worry about, those for which pronunciation has still not settled down. So we must be especially careful to convey in our guides and transcriptions just which of the several options we are asking the reader to consider.

We have now finished with the Loglan vowel sounds and can address the consonants.

**1.17 The 17 Regular Consonants:** There are 17 consonants among the **regular phonemes**, which means the sounds used in non-names. Their pronunciations are given here by English examples. Fortunately, all regular Loglan sounds occur in English, so good examples can always be found. The secondary and tertiary allophones of some phonemes are given on the sutori lines of their entries. The parenthetic expressions in the right hand column of the table are phones. The spellings given in this column are the least equivocal English spellings used in the pronunciation guides:

/b/	as in 'boy'	(b)
/c/	as 'sh' in 'sheep'	(sh)
/d/	as in 'dog'	(d)
/f/	as in 'fat'	(f)
/g/	as in 'goat'	(g)
/h/	as in 'hat'	(h)
/j/	as 'z' in 'azure', '-ge' in 'garage'	(zh)
/k/	as in 'king'	(k)
/l/	as in 'lake'	(l)
	'-le' in 'kettle' when vocalic	(ll)
/m/	as in 'make'	(m)
	as in 'chasm' when vocalic	(mm)
/n/	as in 'net'	(n)
	before /k g/, as in 'sank'	(ng)
	'-ain' in 'certain' when vocalic	(nn)
/p/	as in 'pot'	(p)
/r/	as in 'rot'	(r)
	'-er' in 'father' when vocalic	(rr)
/s/	as in 'sat'	(s)
/t/	as in 'top'	(t)
/v/	as in 'vet'	(v)
/z/	as in 'zoo'	(z)

All phones are single sounds even if the guides to them are written between parentheses with two letters, e.g., (ng). Also, note that the sounds represented by (ll mm nn rr) are what are produced when the normally consonantal continuants (l m n r) are sustained, e.g., as in 'brrr' where (r) is sustained to become the vowel-like (rr). Observe that the sound after (b) in (brr) is long but that it is a single sound. That is to say, no articulatory change occurs within it. When used in this way a consonant is said to be **vocalic**. We shall have more to say on the **vocalic consonants** in Sec. 1.19.

Notice that, except for /c/ and /j/, the letters associated with these 17 Loglan consonants are pronounced exactly as they most commonly are in English. Thus [b d f g h k l m n p r s t v z], while sometimes ambiguous in English, have a most common sound. That most common sound is always at least one of the sounds that letter will have in Loglan. This is true of allophones as well. Thus Loglan /n/ has the three allophones (n ng nn), all of which occur in similar contexts in English. English speakers who repeat the phrase 'sing out for sin gout in the fountain' will hear all three sounds of /n/. The difference between Loglan

## Notebook 3 – Phonology

and English /n/ is that (ng) is one of its allophones in Loglan, whereas /ng/ is phonemic in English. Thus the 'sin/sing' distinction exists in English but not in Loglan. But in both languages the sound that is written as [n] is pronounced (ng) whenever it immediately precedes a /g/ or a /k/. Observe that the English word 'sink' has the sounds of 'sing' in it. The same is true of Loglan. Thus **manko** = 'mouth' is standardly pronounced (MAHNG-koh) in Loglan because the (ng) allophone of /n/ is required before /k/.

**1.18 The Odd Sounds of [C c] and [J j]:** Only the letters used to represent /c/ = (sh) and /j/ = (zh) seem to be oddly paired with their sounds in Loglan. The letters [c j] have been assigned to these two phonemes, first, because we need these sounds in this international language—(sh zh) are especially common in Chinese, for example—and, second, because no "digraphs" are permitted in Loglan. (A **digraph** is a sequence of two letters used to represent a single sound; thus [ck kn ng ph pn sh th] are all English digraphs.) But the Loglan rule is that each simple sound must be represented in written Loglan by a single letter. The letter [c] happened to be unemployed, [k] and [s] were doing all its (English) work in Loglan; and [c] is at least weakly associated with the (sh) sound in other languages. For example, French 'chat' is (shaa) and Italian 'ciao' is (chow). It turns out that the (ch)-sound in this Italian word may be further analyzed as (t) + (sh). So (ch) is in fact written in Loglan phonetics as (tsh), and phonemically this is of course /tc/. Thus, with [c] given the value (sh), the Italian word 'ciao' can now be written in Loglan phonemics as /tcao/, and English 'chew' as /tcu/.

Loglan /j/ is parallel to Loglan /c/. Thus (zh) is the **voiced**, that is, the unwhispered, version of **unvoiced** or whispered (sh), and it occurs commonly in French as the French /j/ of 'Jean' (zhaa<sup>n</sup>). (We shall show the nasalization of a preceding vowel by superscript <sup>n</sup>.) (zh) also occurs in certain words of French origin in English, such as 'azure' and 'garage'. But (zh) is also one component of a very common consonant of English, one that is also represented by the letter [j]. For just as the (ch) sound may be analyzed as (t) + (sh) in Loglan, so the sound of English [j] may be analyzed as (d) + (zh). Thus English 'joke' may be written (dzhohk) in Loglan phonetics, and /djok/ in Loglan phonemics. By a precise parallel, 'choke' may be written (tshohk) phonetically and /tcok/ phonemically. This explains why the Loglan word for 'chain' is written **tcena** and that for 'judge' is written **djudi** (Try to pronounce these as (CHEH-naa) and (JOO-dee), not as (CHEIGH-naa) and (JUH-dee).)

**1.19 The Four Vocalic Consonants:** The vocalic values of the four consonants /l m n r/, namely (ll mm nn rr), are virtually confined to borrowed names in Loglan. Examples are English 'Earl' which is respelled in Loglan as **R1** but is still pronounced (RR-ll); 'Burton' => **Brtn** pronounced (BRR-tnn) or sometimes (BRR-ty); and 'Herbert' => **Hrbt** pronounced (HRR-brt). In such English names, as in similar names in other consonant-rich languages, these four **vocalic consonants** play the roles of vowels. So they have vocalic values in the Loglan loan-words **R1**, **Brtn** and **Hrbt** which are derived from these natural names, and which thus require no other vowels.

These same four consonants, however, are seldom pronounced as vowels when they occur in the regular words of Loglan, i.e., in its non-names. Thus the instances of /r/ in the predicates **rodlu** (ROHD-loo) 'road', **farfu** (FAAR-foo) 'father', and **brudi** (BROO-dee) all have consonantal values. In contexts like **mrenu** 'man', however, the /r/ is sometimes vocalic, especially in the speech of newcomers to the language. That is, **mrenu** is sometimes pronounced as the trisyllable (mrr-EN-oo) rather than as the disyllable (MREH-noo). But the word is still spelled **mrenu** and transcribed as /MREnu/ in Loglan phonemics; for in regular words (rr) is simply an allophone of /r/. That is to say, (rr) is a legitimate, contextual variation of /r/ but does not have separate phonemic status in the language.

What this means is that hearing (mrr-EN-oo) rather than (MRE-noo) does not cause a loglanist to think da's heard a different word; but only that da is listening to someone who is not comfortable with this Loglan consonant combination yet.

**1.20 The Unfamiliar Consonant Pairs:** /mr/ is just one of a handful of Loglan consonant combinations which will be unfamiliar to monolingual anglophones. Some others are /dz/ as in **dzoru** (DZAW-roo) = 'walk'; /ts/ as in **tsero** (TSEH-roh) = 'error'; /cl/ as in **cluva** (SHLOO-vaa) = 'love'; /ct/ as in **ctifu** (SHTEE-

## Notebook 3 – Phonology

foo) = 'stuff/matter/material'; /ck/ as in **cko**zu (SHKOH-zoo) = 'cause'; /cm/ as in **cmeni** (SHME-nee) = 'money'. Despite their curious appearance, none of these consonant-pairs will be difficult for an anglophone to pronounce.

The two pairs /dj/ and /tc/ look difficult but are not. The sounds they represent are actually very common in English, being usually spelled [j] and [ch] as discussed above.

**1.21 The Three Irregular Phonemes /q w x/:** These three sounds—one of them, /w/, a vowel, and the other two, /q x/, being consonants—occur only in names and letter-words, and then only rarely. All three are, by world standards, "difficult" phonemes. That is to say, each is found in only a small subset of the world's languages and is therefore unfamiliar to most human ears and tongues. But the letters assigned to these three sounds in Loglan, namely [Q q W w X x], occur as **letter-words** (see Sec. 2.21) or as parts of **acronyms** (2.29) in many scientific contexts—for example, the chemical symbol for tungsten is 'W', and the 'X' in 'X-ray' must somehow be accommodated—and they are also employed, although not frequently, in writing one of the most extensive vocabularies of science, namely the Linean binomials of biology. These two-term species names, like '**Homo sapiens**' and '**Escherichia coli**', are by international convention always either italicized or bolded in text and spelled identically in all languages...even in Chinese, Japanese, Hindi, and Russian, languages which do not use the Latin alphabet in any other context. It is therefore advisable that Loglan, too, provide the means, at least, for the transcription of these binomials, and in ways that preserve as much as possible of their standard visual character. To this end, once the 23 regular Loglan sounds had been given letters, the 3 letters left over from the standard Western European alphabet were given to these three fairly commonly used European sounds. Their phonetic values are as follows:

/q/	'th' as in E. 'thin' (unvoiced)	(th)
	'th' as in E. 'then' (voiced)	(dh)
/w/	'eu' as in F. 'bleu', umlaut 'u' in G. 'München'	(eu)
/x/	'ch' as in G. 'Bach', 'k' in R. 'Kruschev'	(kh)

As an anglophone, you will know the sounds (th) and (dh), of course; both are spelled in English with [th]. But you may balk for awhile at spelling them with [Q q]. One useful mnemonic is that upper case [Q] looks quite a lot like upper case Greek theta [θ] but with its bar slipped down and tipped a little, whence [Q]. Theta is, of course, the letter used in Greek for its (th) sound. The other two irregular sounds are non-English, but probably familiar enough to those linguaphiles who make an effort to pronounce foreign words correctly.

**1.22 The Use of Irregular Phonemes in Names:** As suggested by the examples given above, the three irregular sounds of Loglan may also be used, at the name-maker's discretion, to produce better imitations of natural place and person names than would otherwise be possible. Thus German 'Bach' may now be imported into Loglan without change as **Bax** (bahkh). Similarly, 'München' (Munich) may be exactly reproduced as **Mwnxen** (MEUN-khen). The sounds of English 'Theodore' may be exactly reproduced as **Qi'ydor** (THEE-uh-dohr), and 'Kruschev' may be well-approximated as **Xrucyf** (KHROO- shuhf). On the other hand, if the word-maker is concerned about minimizing pronunciation difficulties for the large number of persons, on a world-wide basis, to whom one or more of these three irregular sounds is completely alien, and if the respelling of the natural name in regular Loglan sounds is both sufficiently distinct from other names and a recognizable transformation of its source word, the builder may opt for less than perfect imitations of the natural language sources and use regular sounds entirely. This would give **Bak** (bahk), **Muncen** (MOON-

## Notebook 3 – Phonology

shen), **Tiador** (TYAH-dohr), and **Krucyf** (KROO-shuhf) for the four natural names in question; and at present this is perfectly acceptable.

At the moment, a mixed strategy seems to be most attractive. Thus with the arrival of /x/, 'Bach' is now being enthusiastically rendered by **Bax** and 'Krushev' by **Xrucyf**. On the other hand, **Tiador** and **Muncen** still seem to be acceptable variants of 'Theodore' and 'München' which avoid the formidable oddities of **Qi'ydor** and **Mwnxen**. It may be that these oddities will turn out to be mainly visual. In that case, they are likely to become less objectionable with the passage of time. So the policy of whether or not to use the irregular sounds of Loglan in contriving the Loglan forms of imported natural names is still an open one. It is likely to remain so for some time.

**1.23 Three Stress Phonemes:** We have already indicated how stress is written in the guides and productions, namely **no stress** by lower-case letters, **normal syllable stress** by upper-case letters, and **emphatic stress** by both bolding and upper-casing the normally stressed syllable of the emphatically stressed word. It remains to mention that these three values are sufficient to describe all meaningful differences in Loglan stress contours.

For example, the pronunciation guide for the Loglan equivalent of 'You too came!' is

(too-**SWEE** . paa-KAAM-laa)

The Loglan sentence whose production is here being guided is of course **Tu sui pa kamla**. Word-for-word this means 'You also before come'. As we will see in the chapter on grammar, the Loglan tense operator is a separate word; it precedes the expression to be tensed.

Acoustically, stress is an increase in both the amplitude (loudness) and the duration of the sounds that make up a stressed syllable or syllables. Both are measured relative to the other sounds of its utterance. Thus (**SWEE**) is not only the loudest syllable in the above utterance, it is also the longest.

**1.24 One Pause Phoneme:** The pauses that occur in Loglan, as in any language, vary markedly in both length and significance: from long hesitations to brief phrasing pauses. Nevertheless, for morphological and grammatical purposes, it is sufficient to classify all Loglan pauses as members of a single phoneme. This makes the Loglan pause phoneme a single one with many allophones. For example, the glottal stop in Loglan is an allophone of pause. So those instances of exceedingly brief glottal stops that occur, for instance, between the vowels of a disyllabically pronounced vowel-pair are not counted as pauses at all. In short, as in most languages, Loglan pauses are those silences that occur in the midst of speech that are perceptible as such. In acoustic practice this will mean that there is a number of milliseconds above which a pause, say a glottal stop, is perceptible by most auditors as such and below which it is usually not.

A pause is always marked by a period, or full stop, in the guides and productions. In the guides, open-periods [ . ] are used; in the productions close-periods [.] are used. The silences that precede and follow utterances are not pauses; they are either not marked at all or marked by the pound-sign [#] in the productions. An utterance so marked is an **isolated utterance**. Thus

/#tu**SUI**.pa**KAM**la#/#

is the full phonemic transcription of one way in which the utterance **Tu sui pa kamla** may be produced, namely as an isolated utterance with emphatic **sui** and an optional pause after it. Contrast the phonemic transcription of this production with the phonetic guide to the pronunciation of the same utterance in the previous section. Notice that the phonetic detail has disappeared, leaving the stress/pause "contour" of the production plainly revealed. Because stress and pause together describe a pattern that overlies the whole utterance, the stress and pause phonemes of a language are often called its **suprasegmental** phonemes. Its vowels and consonants are called its **segmental** phonemes.)

## Notebook 3 – Phonology

**1.25 Intonation:** Intonation is not a phonemic quality of Loglan speech, and so is not indicated in our phonemic transcriptions. Intonation in Loglan is in fact free to vary in any way a speaker or a dialect group finds natural. It is true that in the Loglan speech we have observed so far questions are generally accompanied by rising tone, and the approaching end of a declaration by falling tone. But this may only be because the Loglan speakers we have observed so far are native speakers of English; and in English this is what happens. Still, in deference to the tonal habits of this potentially "ancestral" group of speakers, a questioning or exclamatory intonation is sometimes shown in the pronunciation guides by the use of [?] or [!].

It will be interesting to see how the tonal dimension of Loglan develops as speakers become more fluent. Will it only reflect the dialectical origins of different groups of speakers in different natural languages? Or will other, language-wide tonal features eventually develop? If so, will these be the result of "universal mechanisms", that is to say, of mechanisms which are inferrably fixed in the human gene pool and so operate in human languages generally? Or will they be only "founder effects", that is, the accidental consequences of the fact that the language started in a certain place—Gainesville, Florida, U.S.A., to be exact—and was first taught by a certain retired college professor who happened to be a native speaker of English, but who also happened to be foreign-born, to have lived many years abroad, and to have a small acquaintance with certain other European languages? Complete tonelessness, such as often characterizes the speech of badly programmed computers, is probably not even a possibility for human speakers.

Incidentally, toneless computer speech—of the sort that can now be produced by Vocoders and other devices—is surprisingly intelligible when the language is Loglan. This is probably because all the syntactical work normally done by tones in human languages is done by particles in Loglan. Thus, we use the question particle *ei* (ay) (as in 'lay') to transform the declarative sentence **Tu mre<sup>n</sup>u** /tuMREnu/ ('You are a man') into the question **Tu mre<sup>n</sup>u ei** /tuMREnuEI/ ('Are you a man?'). (**Ei tu mre<sup>n</sup>u** /eituMREnu/ works just as well of course, and is slightly to be preferred grammatically.) Notice that the phonetic difference between the two English utterances is far greater than that between the two Loglan ones, and that the English difference is largely in the word-order and intonation. In Loglan, our recognition of the fact that the second utterance is a question could be made to depend entirely on our hearing and understanding the extra word **ei**. So both questions /tuMREnuEI/ and /eituMREnu/ may be spoken quite tonelessly and still be recognized as questions by a loglanist...or by a suitably programmed loglaphone computer.

**1.26 Buffered Dialects:** Anticipating that many who will eventually speak Loglan will find some of the consonant clusters in its standard, anglophone dialect awkward to produce, we expect the sixth vowel, schwa or what is now Loglan /y/, to be used for "buffering" any or all of those clusters: that is, to separate difficult consonant pairs by interposing this neutral vowel between them. Dialects which used /y/ in this way would then be known as the "buffered" dialects of Loglan.

It is likely that the Japanese dialect of Loglan will be buffered. While Japanese has some consonant pairs, they together comprise only a small fraction of those that occur in standard Loglan. Here's a partial list of permissible consonant-pairs in Japanese: **dj tc ts; n** followed by **c d g k r s t z** and by **dj tc ts; mb mp**. Nearly any other consonant-pair is going to be awkward for a Japanese-speaking loglanist to produce.

As an example of what the Japanese already do about such matters, here is their rendering of Dutch 'Amsterdam': 'Amusuterudamu'. [u] is obviously the letter by which they represent their buffering vowel in Latin transcription. But phonetically the Japanese buffering vowel has a value much closer to schwa /y/ than to our Loglan /u/, being approximately the sound of English 'u' in 'full'. Thus a Japanese dialect of Loglan could well employ /y/, or some suitably neutral local allophone of it, to stand between any pair of consonants that was not on some official list of "pronounceable pairs", a completion of the list begun above. The result would be easy for a Japanese to produce and yet be fully understandable to speakers of the standard, European dialect of Loglan...as 'Amusuterudamu' is understandable to us now. Thus with a little practice in listening to them, **ma'ty<sup>m</sup>a**, **fa'ry<sup>f</sup>u**, **so'ry<sup>m</sup>e**, and **byru'di** would soon be intelligible Loglan words to English-speaking loglanists, whether in speech or in print. Many predicate words would be the same in both dialects, e.g., **ti'tci**, **da'm<sup>p</sup>a** and **tse'ro**; and of course none of the Loglan little word sequences—which offer few if any pronunciation difficulties to the Japanese—would require buffering; and so this extensive region of the language would also remain substantially invariant across dialects.

### Notebook 3 – Phonology

The question arises: If we are going to use /y/ for buffering consonants in the buffered dialects, what are we going to use as the intraverbal hyphen in those dialects? The two functions must be kept separate; to let one operator serve them both would be to give wrong information about term-joints to the resolver, /iy/ pronounced (yuh) has suggested itself as the hyphenating syllable to be used in buffered dialects; see Sec. 1.10. /iy/ is easy to pronounce, phonemically similar to /y/, and can easily—one might say harmlessly—be used in all positions in which /y/ is used in the standard dialect. As an example of a word that would then contain both /y/ and /iy/, take **sanpyse'nsi**, the new Loglan word for 'semiotics' in the standard dialect. The consonant pair /ns/ is permissible in Japanese and needs no buffering; but /np/ is not, and so requires buffering. Replacing /y/ with /iy/ and buffering /np/ with /y/ produces **sanypiyse'nsi** (saan-uhp-yuh-SEN-see), an eminently pronounceable word for 'semiotics' in Japanese Loglan...and one which is recognizably the "same" as (saan-puh-SEN-see) in standard Loglan.



## CHAPTER 2

### MORPHOLOGY (WORDS & WORD-FORMS)

**2.1 Design Objectives:** The principal objective to be met in designing Loglan morphology was to ensure the **audio-visual isomorphism** of the language. By this is meant that the written and spoken forms of every well-formed Loglan utterance, as those forms would be seen and heard by their readers and listeners, respectively, would in all non-trivial respects be reciprocally inferable from one another. In short, that the spoken string /la.ARqr.paceNOInaBRAgai/ be convertible always and everywhere into the written string **La Arqr, pacenoia bragai** ('Arthur is no longer king'), and vice versa. It was hoped that this feature would contribute not only to Loglan's utility as a man-machine interface language, where the contribution of isomorphism is obvious, but also to its learnability as a second language by adults, in which the transition from writing to speaking is often a difficult one, and finally also to its acquisition as a second tongue by children, in which the reverse transition is often the difficult one.

Perhaps the most important task to be performed in insuring audio-visual isomorphism is that of **word resolution**, a task of the listener. For the listener to perform this task as swiftly as speech is uttered, and also impeccably, there must be a reasonably simple resolving algorithm—learnable by humans and programmable for computers—that is capable of transforming any heard utterance which has been grammatically composed of well-formed but largely pauselessly-spoken Loglan words, into just one string of properly spaced, capitalized and punctuated written words. Loglan morphology makes such an algorithm both easy to write and to execute in real time. Even a small computer can execute it routinely.

We must note one exception to Loglan's audio-visual isomorphism. If an utterance being read aloud from text involves a Linnaean binomial, that is, one of the standardly spelled pairs of genus and species names which are universally used as naming terminology in biological science (e.g., **Australopithecus afarensis** and **Escherichia coli**), then that utterance will not always be simultaneously transcribable in a form that matches the original text in all particulars. The Linnaean binomials are regularly pronounceable in Loglan phonology; see Sec. 2.13. As such they may be regularly resolved and transcribed; but those transcriptions will not be infallible reproductions of the original Linnaean words. To recapture those originals exactly, thus completing the isomorphism, will require in most cases a tabular lookup that will probably not be executable in real time.

**2.2 Definitions and Conventions:** First, we need a small technical vocabulary and a number of expository conventions. A **morph** is a simple word or a simple component of a larger word that has a distinct meaning or makes a distinct contribution to the meaning of the embedding word. A morph, in short, is the smallest meaning-bearing component of speech. For example, **pa** is both a simple word meaning (roughly) 'previously' and a component of compounds like **pacenoia**, where it means 'before'. A **morpheme** is a set of one or more morphs with the same or related meanings. If a morpheme has two or more morphs as members, they are said to be its **allomorphs**. For example, both **no** and **-noi(-)** convey the meaning of negation in Loglan ('no' 'not' 'non-' etc.), and the second morph is plainly a derivative of the first; so each is an allomorph of the morpheme **no**. A morpheme is named for its **free allomorph**, i.e., the one that is derivationally most basic and also capable of functioning as a simple word. In Loglan, these are always the same morph; and in this case it is **no**. Morphs that are not free are **bound**. Thus **-noi(-)** is the **bound allomorph** of **no** because it occurs only in compounds. The leading '-' in the expression '-noi(-)' means that **-noi(-)** is never initial in its embedding word; the trailing '-' means that **-noi(-)** is sometimes medial and sometimes final. In an alternative terminology, the markings on **-noi(-)** show that it is either an infix or a suffix but never a **prefix**. In this terminology a bound morph is called an **affix**.

A **word** is a string of one or morphs that is treated as a single element by the grammar, just as 'nevertheless' is treated as a single element by English grammar. Words are **simple** if they contain exactly one morph, and **compound** otherwise. Boundaries between words are called **junctures**. It is the business of the **resolver**—a computer program or some subsystem of the human central nervous system—to discover the junctures in a string of pauselessly pronounced words. When it is important to show junctures, we will represent them by the boldface vertical bar [ | ]. The joints between morphs in a compound or complex word will be shown by the boldface slash [ / ].

A set of morphs or morphemes whose phonological constituents may be specified in a relatively compact formula is said to be a **form-class** or simply a **form**. We shall use the following conventions to

### Notebook 3 – Morphology

represent the constituents of such forms: 'X' will represent any (segmental) phoneme, that is, any vowel or consonant including the irregular ones /q w x/ and /y/. 'V' will represent any of the **primary vowels** /a e i o u/ or any vowel except /y w/. 'C' will represent any **regular consonant**, that is, any consonant except /q x/. If it is unimportant whether a pair of primaries is pronounced monosyllabically or not, we shall write its formula as 'VV'; if an obligatory monosyllable is called for, we shall write it 'vv'; if we require an obligate disyllable, we shall write 'V,V'.

A **permissible initial** consonant pair is a pair of regular consonants that may be initial in a morph, i.e., in an affix or a word; a **permissible medial** is a pair of consonants that may be adjacent in a morph. Clearly, the permissible initials are a proper subset of the permissible medials. When it is unimportant whether a pair of adjacent consonants is a permissible initial, we shall write it as 'CC'. When a consonant pair in a given formula must be a permissible initial, we will write it 'cc'.

As before, we shall mark proscribed words or forms with a leading asterisk; for example, **\*spai** is not a permissible word because \*ccvv is not a permissible form. In a demonstration sequence, if the reader is not expected to know yet whether a given form is acceptable, for example, ?ccV,V, it will be marked by a leading question-mark. **?Spea** pronounced (SPEIGH-aa) is one of the instances of this questioned form. Actually, this form too is unacceptable...for reasons that we will find out later.

Formulas and their ingredients are usually distinctive enough to require no special marking. Thus VV is the formula for the 25 possible pairs of the 5 primary vowels. **Silences, pauses, syllable breaks** and **stresses** on preceding vowels or vowel pairs are shown in the formulas by their phonological marks [# . , ']. If a specific phoneme is required in some formula, it will be shown in the formula by the same lower-case letter that represents it phonemically. Thus the formula dV is the set of 5 morphs generated by putting each of the 5 primaries after phoneme /d/, in particular, it is the set **da de di do du**. The ambiguity this convention seems to create between 'vv', which could mean a pair of /v/s but doesn't, and 'vv', a monosyllabic vowel-pair, is only apparent. If it were a pair of v's, they would be bolded, stroked or bracketed. Besides, "double consonants", or adjacent instances of the same consonant, do not exist in Loglan. This renders 'cc' unambiguous as well.

To specify a set of equally permissible alternatives in a formula we shall separate them by strokes and enclose the set in square brackets. Thus, [CCV'/CV'C] says that either of these two kinds of stressed syllables is acceptable here. The order in which the alternatives are listed is unimportant. A plus-sign is used as a sign of concatenation between the parts of a word. Thus [CCV'/CV'C]+CV is the complete formula for the form-class to which both **brudi** and **matma** belong. (They are pronounced (BROO-dee) and (MAAT-maa) and mean 'brother' and 'mother'.) Parts of a form which may be iterated one or more times—that is to say, may be present in a larger form as a string of one or more instances of that part—are shown in parentheses in morphological formulas. Thus, (X)+[C/q/x]+. is a formula for the form composed of a string of one or more consonants and vowels of any sort and in any order followed by either a regular consonant or one of the two irregular consonants /q x/, the whole followed by a pause.

## Notebook 3 – Morphology

**Table 2.1 The Two Partitions and Three Classes of Loglan Words**

Feature	Irregular Words		Regular Words	
	Names	Structure Words	Predicates	
Ending	Always C + ./		Always V	
/q w x/	May		Never	
Stress	Any	Penultimate or none	Always penultimate	
/y/	May	Never	As hyphen only	
CC	May	No	At least one	

**2.3 Two Major Partitions and Three Word-Classes:** As Table 2.1 shows, the two major partitions of Loglan morphology are (1) between **regular** and **irregular** words, and, among regular words, (2) between **structure words** and **predicates**. There is only one kind of irregular words, namely **names**. Thus the two morphological partitions generate three major word classes: (i) names, (ii) structure words, and (iii) predicates.

**Names** are the relatively permanent designations of places, persons and things, which, in any given context, are always meant to be unique designations. In Loglan most but not all names are borrowed from the natural languages. The **structure words** of any language establish the grammatical structure of its utterances and are generally quite short. They tend to be "little words", like 'the' and 'now' in English and **le** and **na** in Loglan, but structure words may also be compounds of short words, like 'nevertheless' in English and **Lena** = 'the-present...' in Loglan. **Compound structure words** may sometimes be quite long, for example, **pacenoia** = 'before-and-not-now', or 'no longer'. Occasionally we will speak of simple structure words as **little words**. The words in the third class, **predicates**, are the words of reference in any language. For example, predicates are the nouns, verbs and adjectives of English. Unlike English predicates Loglan predicates are grammatically undifferentiated and have a simple, readily identifiable form.

Table 2.1 gives the features that form the two partitions and identify and characterize the three classes of words. The first two features differentiate irregular from regular words. All irregular words are names, and these and only these may contain the irregular sounds /q w x/. Furthermore, only names end with consonants followed by an obligatory pause; all other Loglan words end in vowels. The next feature, stress, also helps to differentiate names from non-names. Stress may occur anywhere in a name, and if stress in a name is "irregular", that is, if it falls on any syllable except the penultimate one, the stressed vowel is marked with an apostrophe in the written form, e.g., **Pari's**. The next feature also helps to differentiate names from non-names. The grunted vowel schwa or /y/ may occur anywhere in a name and even be stressed, e.g., **Hyntr** (HUHN-trr); but it does not occur in structure words at all and occurs only as a hyphen in predicates, e.g., **mekykiu** (MEK-uh-kyoo). The final feature, the presence or absence of consonant-pairs, differentiates the two kinds of regular words from one another. Structure words never contain adjacent consonants; predicates always contain at least one pair. Names, in keeping with their general flexibility, may or may not contain adjacent consonants.

Because of these deep morphological "furrows" in the vocabulary of Loglan, the learner can tell the broad morphological class of each new Loglan word da encounters at a glance. Take the sentence **La Arqr, pacenoia bragai** (laa . AAR-thrr . paa-ceh-NOY-naa-BRAA-gigh). **La** must be a little word; it is V-final

### Notebook 3 – Morphology

and contains no CC. **Arqr** must be a name; that it is C-final is sufficient to tell us this, but there is another clue. **Arqr** also contains the irregular phoneme /q/ (th). **Pacenoina**, although clearly too long to be a "little" structure word, must nevertheless be a compound one, i.e., one that is composed of the little words **pa** + **ce** + **noi** + ..., etc. For it, too, is V-final and free of CC's. Only **bragai** (BRAA-gigh) is a predicate; while it is also V-final, it contains the consonant-pair /br/. Eventually you will be able to translate this sentence as 'Arthur is no longer king'. But right now we are concerned only with its morphology.

The forms of names are discussed in the next section. The morphology of structure words begins with Sec. 2.18, and of predicates, with Sec. 2.34.

## A. THE MORPHOLOGY OF NAMES

**2.4 The Forms of Names:** A name is any sequence of two or more sounds, including irregular ones, in any order except that the last must be a consonant and the whole followed by a pause. The formula for names is therefore (X)+[C/q/x]+. There is one restriction and one exception to this formula. The restriction will be discussed in the next paragraph. The exception is that there is one very special class of scientific names, the so-called "Linnaean binomials" of biology, whose spellings are fixed by international convention and may not be altered. Occasionally these genus and/or species names end in vowels, e.g., **Escherichia coli**. Occasionally they end in consonants, E.g., **Australopithecus afarensis**. Thus the formula for Linnaean names is (X)X.. As these exceptional names are irregular beyond the hope of any ordinary morphological redemption, certain extraordinary arrangements must be made to place them in special contexts from which their resolution will be possible in Loglan. These arrangements are described in Sec. 2.13-15. In all intervening sections the word 'name' will refer to non-Linnaean names.

**2.5 The Pause Before Vowel-Initial Names:** If a name commences with a vowel, then it must also be preceded by a pause. Thus the English names 'Ellen' 'Eileen' and 'Iona' go into Loglan as **Elyn, Aili'n** and **Ai,onas**. So 'Hello, Ellen', 'Come in, Eileen' and 'Give it to Iona' all require pauses in Loglan: /LOI.ELyn/ **Loi Elyn**, /nenKAa.aiLIN/ **Nenkaa, Aili'n** and /DONsula.aiONas.da/ **Donsu la Ai,onas, da**. (Note the final /s/ added to /aiONa/ to make it C-final; there is more on this move in Sec. 2.11.) If the preceding word is a "name-marker", like **loi** or **la** (see next section), or another name, then the pre-nominal pause is not represented by a comma in text. If the preceding word is anything other than a name-marker or a name, as in **Nenkaa, Aili'n**, then the pre-nominal pause is represented by a comma in text.

The pre-nominal pause is almost always "intervocalic"; see Sec. 2.36 for a discussion of intervocalic pauses.

**2.6 The Name-Marker Restriction on Names:** The restriction on the (X)+[C/q/x]+. name-making formula arises out of the fact that Loglan names are morphologically so irregular that in order for them to be uniquely resolvable in the Loglan speechstream they must not only be followed by pauses but surrounded by them—as in **Takna, Djan, mi** = 'Talk, John, to me!'—whether they are V-initial or not, unless they are initial in an utterance (**Djan, takna mi** = 'John, talk to me!') or preceded by a "name-marker" (**Takna mi Hoi Djan** = 'Talk to me, O John!'). **Name-markers** like **Hoi** are thus a morphologically privileged class of words. They serve to link C-initial names pauselessly to the rest of the utterance. For example, in the production /TAKnalaDJAN/ (TAAK-naa-laa-JAAN) the name **Djan** is not preceded by a pause. The reason it isn't is because it is preceded by the name operator **la**, which is another member of the class of name-markers that make pausing before C-initial names unnecessary. This last production resolves as **Takna la Djan** and means 'Talk to John!'. It is clear that the name-marker **la** is crucial to its resolution; see Sec. 2.17.

There are just seven simple little words that have this pause-blanketing privilege, namely **I /i/ Hoi la loa loi sia siu**, plus any of the compounds formed with initial /i-/, e.g., **Ice** = 'And'. Notice that **I**-words including **I** itself are themselves preceded by a pause and so must be initial in their breathgroups. But the other six name-markers do not require pausing either before or after they are spoken and so may occur anywhere in a breathgroup. Only **la**, by the way, must be followed by a name; the other name-markers may or may not be. This set of seven simple words and the **I**- compounds are both all the words and the only words which may precede C-initial names pauselessly.

Restriction is the other side of privilege. Obviously no copy of a name-marker may occur either initially or with a resolvable "prequel" within the first N-2 phonemes of a name. (The last two phonemes of a C-final breathgroup are bound to be all or part of the resident-name, and so do not have to be included in the search for name-markers.) If **\*Taknaladja'n**, for example, were offered to us as a name, we would protest that its first part is not a name, but the incomplete phrase **Takna la...**, because that is how we have already resolved it. What we would have to do if it were new to us is first spot the name-marker **la** or a copy of it, and then discover that its prequel, **takna**, resolved. The **prequel** of a copy of a name-marker is that portion of the

## Notebook 3 – Morphology

embedding breathgroup that lies to its left. A prequel **resolves** if the resolver can resolve it completely into regular words, that is, into structure words and/or predicates. Obviously it can do this with /TAKna/.

So it is clear why this restriction is necessary. Indeed if we wanted to make a name out of the utterance **Takna la Djan**—as we are free to try to do in Loglan—all we have to do is remove that offending sequence /la/. Is **?Taknadja'n** then a name? Yes it is. It meets all the requirements for names including the name-marker restriction. For no copy of any of the sequences /i... la hoi loa loi sia siu/ occurs within its production as /taknaDJAN/. We might translate such a name into English as 'Talkerjohn'...which opens up some interesting possibilities.

**2.7 Working Around the Name-Marker Restriction:** The name-marker restriction sounds as if it might have eliminated a large proportion of the objects otherwise usable as names in Loglan. This is apparently not the case. In the many years I have been living with this restriction, I have found only two proscribed sequences which occur in natural names with sufficient frequency to justify a routine response. These are /i/ initially and /la/ either initially or in the context /VlaCX/; and both problems are easily dealt with. All other proscribed sequences are exceedingly rare. In fact I remember only two occasions on which a transcription of a natural name had to be rejected because a copy of some name-marker other than /i la/ was resolvably found in it; but the transcriptions were easily modified and I have now forgotten what the problem words were.

The routine solutions I use for modifying /i/- and /la/-containing transcriptions of natural names are as follows. In names like 'Ibanez' 'Isabel' and 'Ibiza'—Spanish is especially rich in /i/-initial names—I replace the /i/ (ee) in the transcription with /ii/ (yee). Thus /iBANieq/ /isaBEL/ and /iBIQa/ are the transcriptions and **iibanieq** (yee-BAAN-yeth), **iisabe'l** (yee-saa-BEL) and **iibiqas** (yee-BEETH-aas) are the Loglan names. /la/-containing names are perhaps a little more common and many come from French. I used to replace the /a/ in /la/ with /e/; but these days I find the /la/ to /ly/ ((laa) to (luh)) as somehow closer to the French. So I now import French names like 'La Fontaine' (laa-fo<sup>n</sup>-TEN) and 'de la Roche' (duh-laa-ROSH) as Loglan **Lyfonte'n** (luh-fohn-TEN) and **Dylyro'c** (duh-luh-ROHSH).

No doubt other languages and problems will inspire other solutions.

**2.8 Derivations of Names:** All Loglan names that have been coined so far have been derived from some pre-existing linguistic source. Some sources have been **internal**, from within the Loglan language; others **external**, from outside it. Internal names are discussed in the next section; external ones in Sec. 2.10.

**2.9 Internal Names:** These are the names that have sources within the language. For example, our family has long had a dog named **Cimr** (SHEEM-rr); her name is derived from the Loglan word for 'summer', which is **cimra**, and certainly fits her disposition. (**Cimra** means 'is the summer of year...'.) So a predicate was converted into a name of a puppy by dropping its final vowel, following it with a pause in speech or a comma in text, and capitalizing its initial letter in text. The result is **Cimr** as in **Cimr, kamla** (SHEEM-rr . KAAM-laa), or, for that matter, in 'Cimr, come!' Another example. We once had a cat with the uninspired name **Gro'katm**, which means Big Cat (from **groda katma**). We pronounced it (GROH-kaat-mm); but we probably shouldn't have. We were deriving the name from the complex predicate **groatma** (groh-KAAT-maa) in which (KAAT) is the stressed syllable; so the name could have been regularly stressed as **Groatm** (groh-KAAT-mm). Still, the apostrophe in the written form **Gro'katm** would have made it clear in text—had it ever gotten into text—that we were stressing it irregularly...as one is free to do with names.

Any predicate may be used to yield a name in this way. Thus, if one wants to imitate the Caribbean custom of saying 'Man!' vocatively, Loglan **Mren**, used in such expressions as **Mren, ea mu safgoi** /MREN.eAmuSAFgoi/ = 'Man, let's sail!', does the job nicely. (**Safgoi** is derived from **salfa godzi**, which means 'sail-go'.) **Fum** is an equally useful vocative, coming as it does from **fumna** = 'is a woman'. Some other internally derived names come from structure words. These usually require that a consonant be added to the V-final word to make the name, and by convention the consonant added to make internal names is /n/. Thus **Tun** comes from **tu**, which, when used as a vocative, has the sense of 'You!' or 'Hey, you!'.

## Notebook 3 – Morphology

**2.10 External Names:** Externally derived names are the closest possible imitations in Loglan of either (a) the **sounds** or (b) the **appearances** of certain natural language names which are to be imported into Loglan. It is generally not possible to do both.

**2.11 Auditorily-Modeled External Names:** When the things named are persons or places, it is conventional in Loglan to imitate the sounds of the natural name as closely as this may be done in Loglan phonetics. The preferred auditory model is the way the name in question is, or would have been, pronounced by the person named, or by the people who live in the place named. If that natural pronunciation does not end in a consonant, then it is conventional to add the phoneme /s/ to external names. Thus if Peter is an Englishman, his own pronunciation of 'Peter' will sound something very like (PEET-uh) in Loglan phonetics; whence **Pityts** is his Loglan name. But if Peter is an American, we will hear (PEED-rr) when he tells us who he is, so no /s/ will be required to complete his Loglan name. **Pidr** and **Pityts** are two very different-sounding names; and in Loglan they are spelled differently as well.

The procedure is similar for names of places and famous personages. 'Bach' is **Bax** and 'Berlin' is **Berli'n** in Loglan because (bahkh) is a good copy of the German name of that German composer and (behr-LEEN) is an almost perfect reproduction of how Berliners speak the name of their city. 'London', on the other hand, is **Lyndn** because (LUHN-dnn) is another very good copy of the sounds one hears on the lips of Londoners despite the unsettling departure from English spelling that is caused by writing it phonetically. (English, as everyone knows, is an unphonetically spelled language.) 'Paris' suffers hardly any visual change because in being phonetically rewritten as **?Pari'** (paa-REE) it becomes clear to the word-maker that it requires the conventional addition of final /s/; thus **Pari's** is Loglan for 'Paris', an adventitious similitude. 'Roma' is **Romas** because again the local pronunciation (ROH-maa) is V-final and again /s/ is conventionally added to the original, this time without an adventitious matching of the local spelling.

In general, when borrowing place and person names, loglanists try to make the best imitations possible with their limited phoneme kit unless the result is thought to be too difficult for loglanists to use on a world-wide basis. Thus, should German 'München' (Munich) be exactly reproduced as **Mwnxen** (MEUN-khen) (as it can be by using the irregular phonemes /w/ = (eu) and /x/ = (kh) which happen to be German)? Or should we adopt the more forgiving **Muncen** (MOON-shen), which although quite close to the original is still pronounceable by that large population of potential loglanists whose tongues could not reproduce the German sounds exactly? This is not a morphological question but a broader linguistic one on which The Institute has no policy at present. Until it is settled by an international congress of loglanists, current loglanists are free to adopt whatever degree of imitation they can manage. Thus either **Mwnxen** or **Muncen** but certainly not **\*Miunek**.

**2.12 Visually-Modeled External Names:** When the name to be imported is a scientific or scholarly one, and a certain method of spelling it is in wide international use, then the model to be most closely copied is the visual appearance of the internationally shared portion of that name as spelled (or respelled) in the Latin alphabet. Take the names of the planets in the five major European languages:

English	Russian	Spanish	French	German
Mercury	Merkurie	Mercurio	Mercure	Merkur
Venus	Venera	Venus	Venus	Venus
Mars	Mars	Marte	Mars	Mars
Jupiter	Yupiter	Jupiter	Jupiter	Jupiter
Saturn	Saturn	Saturno	Saturne	Saturn
Neptune	Neptun	Neptuno	Neptune	Neptun
Uranus	Uran	Urano	Uranie	Uranus

## Notebook 3 – Morphology

From the appearances of these words, the choices all seem obvious. The only question would appear to be: Should we add /s/ or subtract the few final vowels when there are any? Clearly we should do the latter. The greatest variation in these natural names is precisely in these final vowels when they exist at all. This decision gives us either **?Mercur** or **?Merkur** for 'Mercury'; **Venus** for 'Venus'; **Mars** for 'Mars'; **Jupiter** for 'Jupiter'; **Saturn** for 'Saturn'; **Neptun** for 'Neptune'; and **?Uran** or perhaps **?Uranus** for 'Uranus'. Only the choices between **?Mercur** and **?Merkur** and between **?Uran** and **?Uranus** remain. **?Mercur** would appear to imitate the slightly more widespread spelling. On the other hand, the pronunciation of [c] in the [Mercur-] sequence in all the languages that use it is Loglan /k/. This consideration tips the balance, in my opinion, in favor of **Merkur**. As for Uranus, **Uran** is clearly the best choice; it is the internationally-shared portion of these five written names, and its sequelae are extremely variable.

But now let us consider the pronunciation of these visually good copies of these low-variance portions of these six international words. The copies are so good, in fact, that in the astronomic or mythological contexts in which they are most likely to appear in print, they will surely remind any scholarly reader of what they mean. But how do they sound? Quite Loglandical, it turns out; (MEHR-koor), (VEN-oos)—not (VEEN-uhs) or (VEIGHN-oos)—(mars)—that's an /s/, not a /z/—(zhoo-PEET-ehr), (SAAT-oorn), (NEP-toon), and (OOR-aan). Only one, **Neptun**, sounds like the English word. But that is just as it should be. These words are international words. They have, like all the tools of shared scholarship, only international looks, not international sounds. If we in the Loglan community pronounce them in our special way, we are only doing what the rest of the international community of scholars has been doing all along.

**2.13 The Linnaean Polynomials:** The Linnaean "binomials"—actually, polynomials because some of them have three parts—are a huge body of zoological and botanical terminology—probably numbering well into the early millions by this time—that have been made standard in spelling by international covenant among biological scientists. The covenant is very simple. Provided a proposed new Linnaean name is unique, then however it is spelled on its first appearance in print (including any adventitious errors) is the way it will then be spelled by all who use it afterwards. Even Japanese and Chinese scholars, including those who use the Latin alphabet in no other way, use Latin letters and Linnaean spellings to represent species and/or genera in the text of their own scientific writing...even if that writing is intended only for domestic consumption.

This stern system is called "Linnaean" after the great Swedish naturalist Carl von Linne who proposed this remarkably enduring system of biological nomenclature several centuries ago. Loglan has no choice but to follow it exactly. It is clear that the names of **Homo sapiens neanderthalensis**, **Escherichia coli** and **Australopithecus afarensis** must be at least written identically in all languages...even if they are pronounced in a great variety of ways. Phonological diversity doesn't matter; what matters is that readers will continue to be presented with the same words. So here the job of The Institute is to find a way of pronouncing these letters that is (1) consistent with the phonology of Loglan, (2) modelled on some natural language way of pronouncing these polynomials that is imitable by speakers of the widest possible group of other languages, and (3) does not—as far as can be foretold by examining reasonably large samples—render any two Linnaean words identically in speech. Once that is done, then a way must be devised for transforming the binomials spoken by loglanists into Loglan text that will not cause us to deviate from the audio-visual isomorphism of the language.

As far as pronunciation is concerned, the habits of Romance-speaking scholars, especially Italian and Spanish ones, commend themselves on all three counts. Except for the pronunciation of [C c], Romance pronunciation is consistent with Loglan phonology—which is a phonological close relative—and it employs, like Loglan, a minimum set of widely used phonemes. Moreover, the Spanish and Italian communities of biologists have been pronouncing Linnaean words for several centuries; and it is unlikely that the pronunciation habits they have developed for that purpose produce homonyms with detectable frequency.

First, we will estimate this Romance-based pronunciation pattern; then in Sec. 14, we will consider how isomorphism in the neighborhood of Linnaean words may best be preserved.

**2.14 Pronunciation of the Linnaean Polynomials:** Here is our current estimate of how an Italian or a Spanish biologist, with some knowledge of how the French, German, English or Slavic proper names that are



## Notebook 3 – Morphology

often celebrated in the polynomials are "really pronounced", as well as some feeling for the spelling peculiarities of Loglan, would try to regularize the pronunciation of the Linnaean vocabulary:-

We suppose that such a scholar will pronounce every letter and letter combination as in standard Loglan except these:-

[ae] (pronounced as Loglan)	/e/	(eh)
[au]	/o/	(oh)
[aw]	/o/	(aw)
[ay]	/ei/	(ay)
[c] before [a o u] or C not h	/k/	
[ch] initially & finally	/tc/	(ch)
[ch] medially except in [sch]	/k/	
[ea] except final & in [eau]	/i/	(ee)
[eau]	/o/	(oh)
[ee]	/i/	(ee)
[eigh]	/e/	(eigh)
[eu]	/oi/	(oy)
[ew]	/u/	(oo)
[igh]	/ai/	(igh)
[j]	/dj/	(j/dg)
[h] before any C	/hy/	(huh)
[lh]	/l/	
[oe]	/o/	(oh)
[ou]	/u/	(oo)
[ow]	/ao/	(ow)
[ph]	/f/	
[p] before [n s t]	/py/	(puh)
[sh]	/c/	(sh)
[sch]	/c/	(sh)
[q]	/k/	
[rh]	/r/	
[th]	/t/	
[w] before V	/u/	(w)
[x] initially	/z/	
[x] non-initially	/ks/	
[y]	/i/	(ee)

Note that, in deference to our Loglan habits, our standard-setting Romance-speaking scholar will pronounce [C c] before [e i y] as Loglan /c/, that is, as written, and as /k/ elsewhere. Thus 'Cephalopoda' is to be pronounced (sheh-faa-loh-POHD-aa), while 'Coelenterata' is pronounced (koh-len-tehr-AAT-aa). This follows half the European custom of pronouncing [c] before "strong vowels" ([a o u]) like Loglan /k/, and before "weak vowels" ([e i y]) like /s/. Instead of /s/ we will ask our standard-setter to use our "native" Loglan /c/.

**Double consonants**, that is, pairs of identical consonant letters, will be pronounced as if they were a single instance of that consonant.

**2.15 Writing Linnaean Names:** Once Linnaean names are being pronounced in a standard way by loglanists, then the problem arises of how to furnish Loglan readers with clear indications of which names are Linnaean in a given text. Since some Linnaean names are vowel-final (**Homo**) and some are not (**sapiens**), their endings cannot be our clue; and while we are obliged like the rest of the scientific community to italicize, bold or underline Linnaean names in text, even these textual indications are not unique and are in any case difficult to represent unequivocally in speech. So in the interests of preserving isomorphism, a special Linnaean name operator is required. Instead of **la** we will use **laa** (LAA-aa) for announcing Linnaean names. We will further specify that **laa** always be followed by a pause in speech and that Linnaean names never be used without this operator, that is, as vocatives. (So you may not address your visitor as **\*Hoi Canis lupus** even though da is one.)

## Notebook 3 – Morphology

Moreover, the variable number of terms in these polynomials—sometimes monomials (**Australopithecus**), most frequently binomials (**Escherichia coli** or **E coli**), but sometimes trinomials (**Homo sapiens neanderthalensis**)—would mean that we could not tell when a V-final breathgroup following a Linnaean term was simply another Linnaean term or a string of regular words. The threat of this ambiguity forces us to treat all Linnaean names as monomials in speech, looking to a separate algorithm to partition these transitional monsters, e.g., **Homosapiensneandertalensis** (hoh-moh-saa-PEE-ens- neigh-aan-dehr-taa-LEN-sees), **Kanislupus** (kaa-nees-LOOP-oos), back into their proper, and properly-spelled, polynomial parts. Because of their length we anticipate that some of these transitional monomials will be multiply stressed; but it is doubtful that the existence of these occasional multiple stresses will be more than adventitiously useful in partitioning.

The algorithm that will complete the isomorphism has not been written. And when it is written it will probably involve repeated lookups in extremely long tables. But the principle of that algorithm is clear. Loglan can continue to be isomorphic, even in the region of the Linnaean polynomials, if a machine can be programmed to imitate the performance of the microbiologist (or the ichthyologist, or the botanist, or whatever, who hear Linnaeans from their specialties) who can listen to /eceRIKiaKOLi/ and know that da has heard a Linnaean name that da can confidently type as [**Escherichia coli**], including the bolding that is now conventional.

We know that such programs can be written; it is just a question of waiting for some competent loglanist with a biological bent to find the time.

**2.16 The Post-Nominal Pause:** Many loglanists have misunderstood the requirement that there be a pause after names to mean that, in order for Loglan speech to be understood at all, these "obligatory pauses" must be carefully produced every time. Not so. Sloppy Loglan has about as good a chance of being correctly understood—at least by a sympathetic human listener—as sloppy speech in any natural tongue...which is to say that the chances are pretty good. The difference is that non-sloppy Loglan speech can be impeccably understood by any properly instructed listener...even a machine.

**2.17 Resolving Names:** The resolution of a Linnaean name from the speechstream is trivial. One comes upon the Linnaean operator **laa** followed by its pause; one scans right for the next pause or the end of the utterance; all that lies between the two pauses, or between the pause and silence, will be the Linnaean name. Partitioning the monomial thus generated into its proper Linnaean parts, and respelling those parts in ways that recapture the original spellings exactly, are non-trivial steps but beyond the scope of this book.

The resolution of non-Linnaean names is also non-trivial but quite easy to follow. Before the resolver can resolve names from the speechstream it must know how to resolve regular words from it; for a crucial step in the resolution of most names is the resolver's attempt to resolve as regular words the prequel of some copy of a name-marker that has been found in it; see Sec. 2.6. Thus the algorithm described in this section presupposes the existence of algorithms like those that will be described later for little words and predicates; Secs. 2.33 and 2.59-60. However, our discussion of the name-resolving algorithm can be followed without reference to those later algorithms. All that is required is the suspension of disbelief that they can be written.

Given that a regular-word resolver exists, then, the first step in the resolution of a name is to:

1. Find the first (or next) breathgroup in the utterance that is C-final. If there are none, then there are no names. (Being followed by obligate pauses, all non-Linnaean names are either right portions of their breathgroups or occupy them entirely. So all and only consonant-final breathgroups have resident names.) Suppose we locate a C-final breathgroup.
2. Search the first N-2 phonemes of the breathgroup for all copies of the sequences /i la hoi loa loi sia siu/. If none are found, the entire breathgroup is a name and may be so resolved. (The last 2 phonemes in the breathgroup do not have to be searched; they will always be either part or all of the resident name.)
3. If one or more copies of the specified sequences are found, select the rightmost and discover whether its prequel will resolve into regular words. If its prequel does resolve, the name commences just to the right of that copy and stretches to the end of the breathgroup and may be so resolved. If the

### Notebook 3 – Morphology

prequel of some copy does not resolve, select the next rightmost copy if any remain and repeat this step until no copies remain.

4. If no copy with a prequel that resolves into regular words can be found, then the entire C-final breathgroup is a name. Return to step (1) for further breathgroups until all C-final breathgroups have been processed in this way.

## B. THE MORPHOLOGY OF STRUCTURE WORDS

**2.18 The Functions of Structure Words:** The shortest words of any language are its logical, numerical or grammatical words. These are words like 'or' 'of' 'is' 'a' 'if' 'one' 'two' 'the' and endings like '-s' '-ed' '-ing' and '-hood' in English. In Loglan short morphs of this kind include all the grammatical particles (tense words, articles, etc.), the connectives (logical and causal), all the prepositions that are not predicates, the case-markers, the pronouns, the other variables, and all the number- and letter-words.

Clearly, structure words are the words that shape the **structure** of an utterance, and into which the less frequently-used and typically longer **content** words, the predicates and names, are fitted as pictures into frames. Thus, **Le \_\_\_\_\_ pa \_\_\_\_\_** is a sentence-frame. Its nature is completely determined by two structure words. But any two predicates we care to choose can complete the picture so-framed. For example, let us use **mrenu** and **fumna**. If we drop these two predicates into the empty places in the frame in both possible orders, we will make two sentences out of them: **Le mrenu pa fumna** = 'The man was a woman' and **Le fumna pa mrenu** = 'The woman was a man'. If there are 10,000 predicates in the language, then there are 100,000,000 ways of filling this one frame. Each structural frame is therefore a set of possible sentences.

This is the function of structure words: to build the structural frames which the content words then fill out, thus creating the utterances of the language.

**2.19 The Four Little-Word Forms:** All **simple structure words** in Loglan are short; the longest have only three letters. Since the shortest predicates are four letters long, there is a complete visual and audible separation of these two kinds of regular words in Loglan. Therefore we will occasionally call the simple structure words **little words**, for in Loglan they are genuinely little. It is the **compound little words** described in the next section that attain substantial lengths.

The four forms of little words are .V VV CV and CVV. Using a typical member to represent each form-class, we can call them the **A-form** words, the **Ia-form** words, the **Da-form** words, and the **Tai-form** words.

The complete set of **A-form** words are **a e i o u**. In ordinary contexts, these are the five simple logical connectives 'or' 'and' 'if and only if' and 'whether' etc., although they are also used as letter words in spelling contexts; see Sec. 2.25. These five tiny words, as well as any compounds which commence with them, must always be preceded by pauses in speech and commas in text. For example, in **Da, a de pa kamla** /DA.aDEpaKAMla/ ('X and Y came') the pause-comma is obligatory.

There are 25 **Ia-form** words and all of them are attitude indicators. Examples are **ia** and **io** ('Certainly' and 'Probably'). No pause need accompany these **Ia-form** words or their compounds: **Da ia, e de io pa kamla** /daIA.edeIOpaKAMla/ (daa-YAA . eh-deh-YOH-paa-KAAM-laa) = 'X certainly, and Y probably came'.

Next are the 85 **Da-form** words formed of the 5 primary vowels combined with the 17 regular consonants. All possible **Da-form** words have currently assigned meanings. Indeed, the CV form is perhaps the hardest-working morphological form in the language, in the sense that more CV morphs occur in Loglan utterances than morphs of any other form. The best examples of **Da-words** are **da** itself and its four companions, **de di do du**. These are often well-translated into a sort of "mathematized" English by the gender-less, number-less, case-less mathematical variables 'X' 'Y' 'Z' 'W' 'Q'. For example, **Da kepti de di do du** means 'X is a ticket to Y from Z on carrier W for price Q'.

Finally, there are the more numerous but less frequently used **Tai-form** words. There is morphological space for more than 400 of these CVV-form morphs, and only about half of them have been assigned; so we still have plenty of room for growth.

[It is a widely believed myth among loglanists that we are "running out of CVV space." This is not true. The last 12 years of active lexical development have added only a few dozen **Tai-form** words to the 148 words of this form that had meanings in 1975. More than a few dozen empty places, namely 85, have been added morphologically to the CVV word-space since that time. Five new phonemes /h y x q w/ were added to the language—even though four of them, /y x q w/, added only letter-words to the CVV space, and then promptly occupied it—and the previously unused vowel-pairs /aa ee oo/ have been discovered to have uses and so have augmented the rows of the CVV table. Since the number of **Tai-**

## Notebook 3 – Morphology

form words that have actually been added to the language, including the new letter-words, is far less than 85, we have even more open CVV-space now than we had in 1975.]

The best examples of **Tai**-words are the letter-words **Bai Cai Dai** etc. [B C D etc.] of which **Tai** [T] itself is one; but many of the prepositions and adverbs of the language are also **Tai**-form. For example, **dio** is a "case tag" that means 'to/toward' and **sui** is a "discursive" that means 'also'.

As we shall see presently, the 3 irregular sounds /q w x/ and /y/ are fully represented morphologically only among the letter-words. In fact, CVV is the only little-word form in which /q w x y/ may occur. All other little words are made from the 5 primary vowels and the 17 regular consonants.

**2.20 Compound Little Words:** A compound little word is a string of simple little words concatenated pauselessly in speech and printed without spaces in text. Thus, **pacenoina** is a compound little word, or simply a **compound**. It is composed of **pa + ce + noi + na**. Most combinations of little words are permitted in compounds. One exception is that VV-form words may be compounded only with each other. For example, **uau** expresses a satisfied kind of happiness and **uiua** expresses a happy kind of satisfaction, and both are compound attitudinals. Compound attitudinals may be of any length; but they are the only kind of compound in which VV words may occur.

There is one more restriction. While V, CV and CVV words may be mixed together in a compound, two orders are proscribed: V words may not follow other V words (the result would look like a VV word), and they may not follow CV words (which result would simulate a CVV word). So V words may in fact only follow CVV words, which they do only rarely and then only when the **Tai**-word is of Cvv-form, that is, only when it contains a monosyllabic vv-pair like **Tai** itself. When V words do follow Cvv words in a compound, they lose their leading pause. That pause is retained when a .V word is used initially in a compound. Thus .V + CVV puts the obligatory pause in front of the compound where it belongs. For example, **anoi** = 'if' is such a word, and the leading pause is preserved in use: **Da, anoi de** /DA.anoiDE/ = 'X if Y'. \*Cvv + .V, in contrast, would put the pause in the middle of the word if the pause were retained. So it is not retained. (Words may have no medial pauses, of course; see Sec. 2.28.) An example of a Cvv + V compound is **MaiA** (MIGH-aa), the acronymic word for [MA], the US Postal Service's acronym for 'Massachusetts'. Note that /MAIa/ (MIGH-aa) is audibly distinct from /MAia/ (MAA-yah). This phonological distinction will be employed in the resolution of little words; Sec. 2.33.

There are also some "false compounds" that we must look at. Derivations of the CV + V kind do exist semantically even though morphologically they are proscribed. For example, the CVV word **noa** has a meaning which is derivable from **no** followed by **a** ('not and/or' or 'only if'). But the resulting **noa** is not a compound. It is a simple CVV word chosen with that semantic derivation in mind. This kind of derivation applies to the **Tai**-form letter-words as well. Thus **Tai** itself is semantically derived from the letteral [T] plus the suffix [ai]. But **Tai** is nevertheless a simple CVV word, and so is morphologically not subject to further resolution.

**2.21 Letter-Words:** These are a kind of structure word which have a very special morphology in Loglan. By a **letter-word** is meant a word like English 'em', 'eff', or 'dee' by which letters are spoken or read aloud from text in a given language. That they are words in the spoken forms of languages which have written forms, and therefore have characters of some kind—and of course, not all languages do—is clear. We can say 'There are nineteen effs on this page' in spoken English. But words like 'eff' seldom appear in English text. Instead, in writing such a sentence we would probably use the letteral [f] and type [There are nineteen f's on this page.]. In Loglan, letter-words appear as frequently in the written language as in the spoken one. There are also a great many Loglan letter-words, since there are separate words for the upper- and lower-case versions of each letter in the Greek and Latin alphabets. That makes 100 letter-words in all, as there are 26 letters in the Latin alphabet and 24 in the Greek one.

Each Loglan letter-word is formed by combining the Loglan phoneme associated with that letter with a suffix. If the letter is Latin, the association is automatic. The phoneme associated with a given letter is the sound that letter is given in reading Loglan text aloud. If the letter is Greek, however, some of the associations between characters and sounds are obvious and some are arbitrary. So the entire list of associations between Greek characters and Loglan sounds will be given presently.

## Notebook 3 – Morphology

**2.22 Suffixes for the 52 Latin Letter-Words:** Some of these words are **Tai**-form; some of them are **Ama**-form. The four suffixes required to generate all 52 Latin letter-words are as follows:-

For the	7	lower case	Latin vowels,	add <b>-si</b> ;	thus	<b>asi, esi, ..., ysi.</b>
		upper		add <b>-ma</b> ;	thus	<b>Ama, Ema, ..., Yma.</b>
	19	lower	consonants,	add <b>-ei</b> ;	thus	<b>bei, cei, ..., ze.</b>
		upper		add <b>-ai</b> ;	thus	<b>Bai, Cai, ..., Zai.</b>

In addition, there are the 7 single-letter abbreviations of the vowel letter-words provided by the 7 vowels **a e i o u w y** used as one-letter words. These, too, are letter-words. But unlike the **Ama**- and **asi**-form vowel-words, to which they are alternatives (allomorphs), the single-letter letter-words are ambiguous with respect to both language and case. As we shall see later, the single phoneme /a/ may stand for the upper case Latin letter **A**, for the lower case Latin **a**, for the upper case Greek alpha (we cannot display the Greek letterals on our font, so will be content to name them), or for the lower case Greek alpha. Which letteral the vowel /a/ is representing in any given case will depend entirely on the context in which we find it; see Sec. 2.24.

The reason there is a single-letter abbreviation for each of the Latin vowel-words is that, in many contexts in which the Latin letter-words are used, the preferred vowel-word is the vowel itself. Thus in spelling the Loglan word **ba** we wish to name the characters; so we say **bei, a** (bay . aa). We could say **bei asi** (bay . AA-see) if we chose, thus specifying the lower case Latin [a]; but that much univocality is not required in the spelling context and is seldom used. See Sec. 2.25 for more on spelling practice. In fact, the 7 single-letter vowel-words, **a, e, i**, etc., are used either in or as letter-words wherever the loss of case and language information is unimportant; for example, in making acronymic words; see Sec. 2.29.

Consonants cannot, of course, be spoken alone. Therefore they always require a vocalic suffix.

**2.23 Suffixes for the 48 Greek Letter-Words:** These words are also of either **Tai**-form or **Ama**-form, but of course four different suffixes are used. The suffixes required to generate the 48 Greek letter-words are as follows:-

For the	6	lower case	Greek vowels,	add <b>-fi</b> ;	thus	<b>afi</b> , etc.
		upper		add <b>-mo</b> ;	thus	<b>Amo</b> , etc.
	18	lower	consonants,	add <b>-eo</b> ;	thus	<b>beo</b> , etc.
		upper		add <b>-ao</b> ;	thus	<b>Bao</b> , etc.

Two phonemes in the Loglan phoneme set have no corresponding letters in the Greek alphabet. These are /c/ and /w/, (sh) and (eu). The remaining 24 Loglan phonemes have been tentatively assigned to the 24 letters of the Greek alphabet as follows. There are four with arbitrary associations and these are marked with a pound-sign [#]:

a	alpha	h#	eta	n	nu	t	tau
b	beta	i	iota	o	omega	u	upsilon
d	delta	j#	xi (ksi)	p	pi	v#	psi
e	epsilon	k	kappa	q	theta	x	chi
f	phi	l	lambda	r	rho	y#	omicron
g	gamma	m	mu	s	sigma	z	zeta

As I say, these assignments are tentative. The Institute would be pleased to consider any proposal based on a better understanding of Greek phonemics than this one displays.

To use these tables to build a Greek letter-word, proceed as follows.

- (1) Suppose we want the word for lower case Greek gamma.
- (2) The suffix for l.c. Greek consonants is **-eo**.

## Notebook 3 – Morphology

- (3) Gamma is associated with the Loglan phoneme /g/.
- (4) So the required letter-word is **g + eo = geo**, a **Tai**-form word.

**2.24 Uses of Letter-Words:** Letter-words are currently being used in five contexts: (1) In spelling, see next section. (2) In making acronymic words like **CaiIzA** (shai-EEZ-aa) for 'CIA' (see-igh-EIGH), see Sec. 2.29. (3) In forming dimensioned numbers like **nenimei** (neh-NEE-may) [10m] for '10 meters', see Sec. 2.31. (4) As letter-variables both in mathematics (**toXai** [2X]) and in ordinary discourse (**Bai groda Cai** = 'B is bigger than C', which is often abbreviated in text to [B groda C]). (5) To form scientific predicates, for example, **geoykreni** (geigh-oh-uh-KREH-nee) for 'gamma-ray', which is made from the letter-word for lower case gamma, **geo** (GEIGH-oh), the hyphen /y/ (uh) (see Secs. 2.48 and 2.55), and the predicate **kreni** (KREH-nee), which means 'ray'.

**2.25 Spelling Aloud:** To spell a word aloud in Loglan, one uses **Tai**-words for the consonants and either **A**-words or **Ama** type words for the vowels at the speller's option. Normally da will use **A**-words for the vowel-letters. But if capitalization is to be reported, or there is any other source of confusion in the context, da may choose to use **Ama** words for greater explicitness. Thus the string of utterances (for so the grammar will perceive it) **Tai A I** /tai.a.i/ (tigh . aa . ee) will be taken by the Loglan auditor to spell the word [Tai]. In English, we would say 'Capital tee. Eigh. Eye.' More explicitly da might wish to say **Liu Tai nu leasri li, Tai A I lu** = 'The word 'Tai' is spelled (character-written) 'Tai. A. I.' A guide for this string is (lee-oo-TIGH-noo-leigh-AAS-ree-lee . tigh . aa . ee). There is more on spelling in the grammar.

**2.26 Little Word Predicates:** There are three series of words in Loglan that considered morphologically are compound little words but which are treated by the grammar as predicates. These are (1) **mathematical predicates**, (2) **acronymic predicates**, and (3) **identity predicates**. Identity predicates are **bi** and its analogs and compounds (see Lexeme BI in the Lexicon) and require no special morphological treatment. The other two series of little word predicates do require special treatment and will be discussed in the next section and in Sec. 2.29.

**2.27 Mathematical Predicates:** There are the two series of these words: the **cardinals** and the **ordinals**. They are generated by attaching either the cardinal suffix /-ra/ or the ordinal suffix /-ri/ to any number word or other quantifier; see Lexeme PREDA for the complete list of non-numerical quantifiers. Examples are **tora** = 'is a dyad/a twosome' and **rari** = 'is the "all-th" or final member of sequence...'. Like other predicates, mathematical ones are stressed penultimately. So they must be separated from any preceding number-word or quantifier by a pause in speech or a comma in text. (Such commas are not strictly necessary in text, but it is considered good writing style to use them because they contribute to the isomorphism of the language.) Thus **Kambei leva fe, fefera galno veslo mi** /KAMbeilevaFE.feFEraGALnoVESlomi/ = 'Bring those five fifty-five gallon containers to me' must be partitioned into (at least) two breathgroups at the juncture between the quantifier and the cardinal predicate to prevent /KAMbeilevafefeFEraGALnoVESlomi/ from being heard. The pauseless production would mean 'Bring that five-hundred-and-fifty-five gallon container to me'.

[This stress rule is new since 1983. It was decided to bring the stress in mathematical predicates—once initially stressed—into line with that of all other predicates. The cost of this rectifying move is phonologically a modest one: the occasional use of a quite naturally-occurring pause. Morphologically it eliminates an exception. It allows us to say that all predicates are stressed penultimately.]

**2.28 The "No Pausing Inside Words" Rule:** The preceding resolution demonstrates the necessity of not pausing inside a word...especially not inside a compound one, which will often have some other resolution if

## Notebook 3 – Morphology

the speaker does inadvertently pause. This is because no Loglan word legitimately contains a pause and so the resolver makes good use of whatever pauses it hears. (This means that the tiny stops that do occur acoustically in some vowel disyllables, e.g. in /a,o/, will be measured by the machine—and by the human auditor, for that matter, who is usually unaware of them—as "too brief to be a pause".) New loglanists frequently object to this rule: 'But sometimes I don't know the word, and I have to hesitate until I remember how it goes!' That is true; and human auditors—your teacher, for example—will understand this...and probably you. But the **machine** will not understand you until your Loglan speech becomes "fluent". That, in fact, is what any increase in fluency largely is. It is the elimination of just these morphologically unnecessary (and sometimes downright misleading) pauses from the hesitant speech with which you and every other learner will inevitably begin.

**2.29 Acronymic Predicates:** Morphologically, acronymic words are compound little words; but grammatically they are predicates. They are one of several classes of little word predicates (see Sec. 23) which have this slightly misleading morphology.

But what are "acronymic words"? Just as a letteral, let's say [T], is a visual abbreviation of its letter-word **Tai**, so a Loglan acronym is a visual abbreviation of its acronymic word. Thus wherever [CCC] occurs in Loglan text it is simply an optional, written abbreviation of the word **CaiCaiCai**; and both are pronounced (shigh-SHIGH-shigh), for, like all predicates, acronyms are stressed penultimately. This is not quite like the handling of acronyms in English. The English acronym [CCC], for example, is a representation in writing of the "spelling-form" 'See-see-see'; but this form never occurs in written English. So a closer parallel with Loglan acronyms is the use of compound numerals in both languages. Thus, [123] is a shorthand written notation for the spoken expression 'one-hundred-twenty-three' in English just as [123] is shorthand for **netote** in Loglan. Both "longhand" forms may occur in writing as well as in speech. Just so, [CCC] is shorthand for **CaiCaiCai**; and **CaiCaiCai** is not only the "spelling aloud" of [CCC] but it is a written word as well.

Conventionally, the acronyms of Loglan are restricted to Latin letterals, both upper and lower case. The variety of internationally-used acronyms that can be "spelled out" as acronymic words in Loglan includes not only common sequences of Latin upper-case letterals, like [USA UN DNA], but also alphanumeric sequences like [U234] and chemical formulas like [H2SO4]. Also, chemical symbols containing both lower- and upper-case letterals, such as [Fe] (Iron) and [As] (Arsenic), are also uniquely spellable as Loglan acronymic words once certain decoding conventions are taken into account. But first let us look at the rules for turning such acronyms into Loglan words. These "spelling-out" rules are as follows :-

1. All consonant-letterals in an acronym are represented in the acronymic word by their full 3-letter words. Thus **peicei** is the word of which [pc] is the written abbreviation; and both expressions are pronounced /PEIcei/ (PAY-shay).
2. With a very few exceptions to be noted later, each Latin vowel-letteral in an acronym is represented by that single vowel phoneme in the acronymic word. Thus **DaiNaiA** = /dai + nai + a/ is the reading-aloud, or spelling out, of **DNA**; and both expressions are pronounced /daiNAIa/ (digh-NIGH-aa).
3. If a pair of Latin vowel-letterals are adjacent in an acronym, the vowel phonemes by which they are read aloud are hyphenated with /z/ in the acronymic word. Thus **CaiIzA** = /cai + i + z + a/ is [CIA] read aloud. Both are pronounced /caiIZa/ (shigh-EEZ-aa). Similarly, **AzAzA** (aa-ZAA-zaa) is [AAA] read aloud.
4. If there is a 2-letter element symbol in a chemical acronym which, like [Ca] (Calcium), is composed of an upper-case consonant-letteral plus a lower-case Latin vowel-letteral, the two are hyphenated with /z/ in the acronymic word. Thus, [Ca] itself is read as **Caiza** (SHIGH-zaa) while [CaCO3] (calcium carbonate) is read as **CaizaCaiOte** (shigh-zaa-shigh-OH-teh). Note that [O] is not hyphenated to the second [C]. Thus [O] is not a lower-case appendage of that [C].



### Notebook 3 – Morphology

5. To read an acronym containing Greek vowel-letterals as a Loglan word, those vowel letters must be read aloud by their full three-letter words, e.g. /AMo/ or /AFi/. Any Greek consonant-letter in an acronym must be hyphenated with /z/ to any immediately following vowel of either nationality.
6. All non-initial numerals in an acronym are pronounced as number-words in the acronymic word. Thus **Utotefo** = /u + to + te + fo/ is the word of which [U234] is the abbreviation. Both are pronounced /utoTEfo/ (oo-toh-TEH- foh).
7. Acronyms with initial numerals—rare forms at best—are not allowed to be transformed into Loglan acronymic predicates. If they were, they would be taken for dimensioned numbers, e.g., **temei** = [3m] (TEH-may) 'three meters'.
8. Dimensioned numbers may have acronyms as their right-hand parts. Thus [100USD] might be the written form for 'one hundred US dollars', in which [USD] is an acronym. The whole expression, then, would be pronounced /nema,uSAIdai/ (neh-maa-oo-SIGH-digh) in which the resolver would detect the /u/ as part of the acronym [USD] and therefore not pair it with the preceding /a/. Written out, the compound word for [100USD] is **nemaUSaiDai**.
9. Acronymic words whose acronyms would imitate existing or even possible Loglan words are not permitted. This **\*peia** is not permitted because its acronym is [pa] which imitates **pa**; see Sec. 28 for the acronym recovery rules. In speech /PEIa/ and /pa/ would be quite distinct, but in written text the acronym [pa] would be indistinguishable from the word [pa].

**2.30 Pause and Stress Around Acronymic Words and Letter-Words:** Stress is always penultimate in acronymic predicates, that is, stress falls on the syllable that is second from the last, as indeed it does in all predicate words.

Given the penultimate stress rule, Rule 5 in the preceding section requires that the juncture between single-vowel connectives and a quantifier, as in **U totefo le mrenu** ('Whether 234 of the men'), be protected by either a pause or by final stress on the quantifier: thus either /u.totefoleMREnu/ or /utoteFOleMREnu/ would be morphologically distinct from the production with the penultimately stressed acronymic word, namely /utoTEfoleMREnu/ ('Be U234 (to/at/about) the men(?)' whatever that would mean). My prediction is that usage will follow the second, more economical route. But let us see.

The fact that acronymic words, being predicates, are always penultimately stressed may also be used to force the resolution of a string of separate letter-words (which otherwise might imitate an acronym) by stressing their final member. Thus, in /VEDmabaicaiDAI/ the sequence /baicaiDAI/ cannot resolve as the acronym **BaiCaiDai** [BCD], for that would have to be penultimately stressed: /baiCAIdai/. Thus the resolver can write **Vedma Bai Cai Dai** - 'Sell B to C for D', or even [Vedma B C D], from the production /VEDmabaicaiDAI/ without benefit of pauses. Pauses may of course be used to separate acronymic words from each other and from the number- and/or letter-words they might otherwise absorb, or optionally (never obligatorily) to separate letter-words from one another. Thus **Bai, DaiNaiA Cai** = 'B is the DNA of C (that is, part of C's genome)' may be univocally produced with only that one pause: /BAI.daiNAIacai/ or, just as effectively, /BAI.daiNAIaCAI/. However, if we drop the pause, the resulting production /BAIdaiNAIacai/ resolves as **BaiDai NaiA Cai** ('Be a BD type of NA to/of/about, etc. C'), and if at the same time the stress is dropped from /BAI/, then the resulting production, /baidaiNAIaeai/ resolves as **BaiDaiNaiA Cai** = 'Be a BDNA to/of/about, etc. C'. So stress is an important feature of the speechstream in the neighborhood of letter-words.

As we've just seen, /VEDmabaicaiDAI/ will resolve as **Vedma Bai Cai Dai** without benefit of pauses. But if we put the stress on the middle term in such a set, we get something that appears at first sight—that is, before we rule on it—to be resolutionally ambiguous: /VEDmabaiCAIdai/. There's an acronym here, alright, but it might be either **BaiCaiDai** (/baiCAIdai/), or **CaiDai** (/CAIdai/) with the letter-word **Bai** coming before it. Clearly we must rule on this case. The best morphological ruling is to let the resolver take such a pauseless, penultimately stressed string of letter-words as the longest acronym it can be; and then use pauses

## Notebook 3 – Morphology

to mark off other cases. Under this rule, the sense of 'Sell B to CD' is given by **Vedma Bai, Caidai**; and this will be uniquely resolved from either /VEDmabai.CAI dai/ or /VEDmaBAI.CAI dai/; but /VEDmabaiCAI dai/ uniquely resolves as **Vedma BaiCaiDai** and means 'Sell BCD'. Similarly, 'This is a ticket from B to C on DF for (price) G' is **Ti ketpi Bai Cai, DaiFai Gai**, which is uniquely given by /tiKETpibaiCAI.DAI faigai/ (and some other variations). But again, only a single pause is obligatory.

**2.31 Pause and Stress Around Dimensioned Numbers:** There is a final problem involving acronyms, and that is how to protect the junctures around the acronym-bearing dimensioned numbers of Rule 8, Sec. 29. Our example was [100USD]. This is spelled out in text as [nemaUSaiDai], and partially spelled out as [100USaiDai]; but in any case it is produced in speech as /nema,uSAI dai/, the close-comma indicating that the /a,u/ is disyllabic. The problem is how to prevent these objects from being misheard as "indefinite descriptions", that is, with the same grammar as **Ne mre nu** = 'One man'. Like **mre nu**, **USaiDai** is a predicate. How is it that the resolver does not hear **Nema USaiDai**—which would mean 'Some one hundred instances of U.S. dollars (i.e., things worth one U.S. dollar, eg., the bills or coins themselves)'—in this production? Again, the default rule is to let the resolver hear the longest dimensioned number it can hear, and use pauses to mark off other cases. So the resolver is instructed not to resolve pauseless productions like /nema,uSAI dai/ as two words when it can resolve it as one. Indeed, the two-word phrase **Nema USaiDai** would be parsed as an indefinite description. So when indefinite descriptions involving acronymic predicates are intended, the speaker must pause between the quantifier and that predicate. Thus **Nema USaiDai** /NEma.uSAI dai/ gets the now-intended meaning of 'Some one-hundred U.S. dollars' exactly. Even consonant-initial acronyms require this protection. Thus, **Nema NaiZaiDai** = 'Some 100 New Zealand dollars' must also be pause-bearing (/Nema.naiZAI dai/) in order to distinguish it from the dimensioned number **nemaNAIZAI DAI** /nemanaiZAI dai/. This last expression is a single word, a quantifier, and might be the measure of some quantity, say '100 New Zealand dollars- worth of wool' - **NemaNaiZaiDai lunli**. No pause is needed in this indefinite description. The two stressed syllables in the pauseless production /nemanaiZAI daiLUNli/ will effect the desired resolution.

Notice that /NEma.uSAI dai/ differs from /nema,uSAI dai/ in two respects: there is a pause and a stressed /NE/ in the first production but not in the second. The extra stress in the first production does help the human auditor resolve these two productions correctly; but it may not be relied on by the machine's resolver. For example, if the quantifier is monosyllabic, the natural tendency to stress one syllable relative to another vanishes. So /ne.uSAI dai/ requires the pause to distinguish it from /ne,uSAI dai/. Between this last pair of productions the phonological difference is now minimal; but it works.

One final note about stress in the neighborhood of dimensioned numbers. When such a number is of minimal length, consisting of only one letter-word, say, with a default **ne** assumed, and is being used in a numerical description—as in **Ti langa ta lio mei** = 'This is longer than that by (one) meter'—then the descriptor **lio** ('the number...') must be pronounced disyllabically and its second syllable must be stressed. This pattern will always give the desired resolution. Thus /liOMEi/ will resolve as **lio mei**, while the ambiguous production \*/liOMEI/ is in danger of picking up any following letter-word as the final syllable of the number: /liOMEI mei/ => **lio meimei** = 'the number (one) millimeter'. True; such a sequel is rare. But pronouncing **lio** disyllabically and stressing its final syllable offers complete protection against this accident; for then the one-syllable number is in effect the ultimate syllable of a quasi-compound, the phrase **lio mei**, which is in effect being stressed penultimately. This determines the right juncture of **mei** and prevents it from being heard as part of anything else.

**2.32 Acronym Recovery Rules:** The word-formation rules of Sec. 26 are sufficient to transform any (allowable) acronym into an acronymic word. But they are not sufficient to allow the recovery of every (allowable) acronym from its sound in speech. The consonant-words will decode uniquely, of course; but the single vowel-phonemes to which the vowel-words have been (nearly always) reduced will not. So to enable unique decoding of the single vowel phonemes in spoken acronymic words, certain already widely-used international conventions have been adopted. These are:-

## Notebook 3 – Morphology

- a. **"Nationality"**: The default assumption is that all the vowel-letterals in an acronym are Latin. If Greek vowels are used, they must be spelled out. If all the letter-words are single vowels, it is assumed that the entire acronym is Latin and upper case. Thus /aZAza/ is [AAA]; and the written word is **AzAZA**.
- b. **Case**: It can be inferred from the word-formation rules that every single letter vowel-word which is joined by a hyphen to a preceding upper-case consonant-word represents a lower-case Latin letteral. Thus /FAIze/ is [Fe] and the acronymic word is **Faize**. But let's go further. Let's also assume that every single-letter vowel-word that is not hyphenated to a preceding Latin consonant-word has the same case as that word, or as any following consonant-word if the vowel-word in question happens to be initial. Then /CAIo/ will decode as [CO] and /Unai/ as [UN]; and the two words will be spelled out as **CaiO** and **UNai**, respectively. (It is quite a different matter whether [UN] will regularly refer to the United Nations in Loglan, or to the (impossible) diatomic compound of Uranium and Nitrogen; we trust the former...although discussion of the latter is not of course impossible.) /FAIe/ of course will then be [FE]; /FAIze/ would still be [Fe]; and /FEIze/ (FAY-zeh) might well be taken to be [fE] if such a curious acronym is ever needed. (At the rate at which acronyms are proliferating in the modern world, it may soon be.) But /FEIe/ (FAY-eh) would always have to be written out as **feie**, for its acronym \*[fe] would imitate the number-word **fe**.

The case convention will not allow the symbols for the two V-initial element words **Au** (Gold) and **Eu** (Europium) to be read aloud in the usual way. The problem is that both the letterals in just these two chemical symbols are vowels. So the above conventions would apply the default rule wrongly and decode the spoken words /AZu/ and /EZu/ as [AU] and [EU]. (One of these, [AU], as it happens, predicates another impossible diatomic compound, this one involving Argon. The other, [EU], is not a possible chemical acronym for [E] is not an element symbol.) So these two exceptional chemical acronyms will have to have their lower-case second vowels "spelled out" when spoken aloud, namely as /aZUusi/ and /eZUusi/, respectively. They can then be written [Au] and [Eu] as required. Elements whose symbols are single letterals are, of course, referenceable by the corresponding letter-words: thus Nitrogen by **Nai** and [N] and Oxygen by **Oma** and [O].

Interestingly enough, the symbols composed of single letter-words have the grammatical status of "arguments", that is, they function as designations; while acronymic words composed of two or more symbols—not necessarily all alphabetic—have, as we have noted, the grammatical status of predicates. Thus **Ta U235** /ta.utoTEfe/ means 'That's U-235'; while it takes "predification" (with Lexeme ME) to turn the argument **Uma** into a similar predicate: **Ta meUma** = 'That's U (in the sense of Uranium)'. The reason for this lack of parallelism between letter-words and acronyms is explained under Lexeme TAI in the Lexicon.

**2.33 Resolving Structure Words:** All structure words, whether simple little words or compounds, are resolved in a two-pass operation. On the first pass, the resolver **reduces** some part of some production that contains only little words to a string of simple little words, preserving whatever pauses and stresses it may find among them for use on the second pass. On the second pass, the resolver acts as a **compounder**. It places junctures (word-boundaries) around certain substrings of the string of little words which the reducer has identified (in this way, it "compounds" them), thus isolating those which are left as simple little words. It then turns both kinds of "lexes" (words) over to the lexer for classification as "lexemes". (The terms 'lex' and 'lexeme' are defined in the next chapter; for the moment you can think of a lex as a word, and its lexeme as the part of speech to which it belongs.) Well-formed text has had all this morphological work done for it by the writer; so text is turned over to the lexer directly.

[The "compounder" that will perform the second pass has not yet been written...although the lexer that will use its output has been. The latter is part of the preparer of the machine grammar which has, so far, been tested on textual input only. There is no doubt in my mind that a compounding algorithm to work with acoustic productions can be written, even if it might require a somewhat more elaborate array of pauses and stresses than are now thought to be sufficient. It is even possible that, to make the compounder work, some additional usage restraints on grammatical productions may also be found

## Notebook 3 – Morphology

necessary; but these will almost certainly be minor adjustments in the usage rules which will have no large effects on the grammar.]

The **reduction pass** works as follows. For **unconditional resolution** the little-word resolver requires that it be given a segment of a production that is known by other resolutions of the resolver to contain only little words. Such a segment or segments could be (a) all of a V-final breathgroup which contains no CC's and hence no predicates; (b) those parts of a V-final breathgroup which does contain one or more CC's but in which the region or regions occupied by predicates have been marked (this will have been done for it by the predicate resolver in Sec. 2.60); or (c) a C-final breathgroup in which the regions occupied by predicates, if any, and by the resident name have all been marked, and in which there is one or more residual segments known to contain only little words. Thus in unconditional resolution, the little-word resolver is the last of the three resolvers to go to work on some breathgroups, namely on those parts of it which have been set aside for it by the other resolvers.

But there is a complication. The little word resolver is also required to perform **conditional resolutions**. The reader may recall that before the name-resolver of Sec. 2.17 could locate the left edge of the name that it knows is always resident in a C-final breathgroup, it had to have the "regular word resolver" (evidently a joint effort of the little word and the predicate resolvers) attempt to perform a conditional resolution of the prequel of any apparent name-marker that it found. This resolution was conditional because it could fail. Only when it found a name-marker whose prequel did resolve could the name-resolver mark off the region of the C-final breathgroup that was not occupied by the resident name. But not all prequels of apparent name-markers resolve. (The few that don't are, as we will see, easily identified.)

This seems circular but is not. The segment of a production which the name resolver calls a "prequel", and which is the string of phonemes that is given to the little word and predicate resolvers for conditional resolution, is definitely marked. In particular, it is the string that lies between the copy of a name-marker whose prequel it is and the left edge of that breathgroup. In fact the only difference between these prequel strings that the resolver is asked to resolve conditionally, and those more confidently marked portions of a breathgroup which are known to contain only regular words, is that the attempt to resolve the former may fail. The latter will resolve...if the utterance from which it is taken is well-formed. But the answer to the conditional resolution question may occasionally be no. The resolver must be able to provide that sort of information as well as give the results of a successful resolution which it knows beforehand will succeed.

Let us commence with **unconditional resolution**. We will assume that the name- and predicate-resolvers have done their work, and that we now have a pauseless segment of a breathgroup that is known to consist entirely of little words. Such segments may be initial or non-initial in their breathgroups. We recall that strings of little words will consist only of V VV CV CVV elements; that V elements can only be initial in breathgroups or follow CVV elements; and that if a V element does follow a CVV element pauselessly the latter must be of Cvv form (as for example in the acronymic word **MaiA**); see Sec. 2.29. This detail is used in Step (3c) of the resolution procedure given below.

We proceed as follows:

- 1) We ask first if the segment is breathgroup-initial, that is, immediately preceded by a pause. If it is, continue. If it is not, go to Step (4).
- 2) If the first sound in the pause-preceded segment is V, the first word is either a V or a VV. To find out which, count the consecutive V's. If the number is odd, the first V is a word, and the remainder of the string of V's, if any, is composed of one or more VV words and are so resolved. If the number is even, the entire string of V's is composed of VV words and are so resolved. In this way, we arrive at the first C, if one exists in the segment, or at the end of the segment.
- 3) If the first sound (or the sound now being examined) in the segment is C, the next sound is a V; for, by construction, there are no CC's in this segment. So examine the next two sounds. Of the three possibilities, namely CVCV, CVVC, CVVV—CVCC is again impossible by construction—the first two resolve immediately:

## Notebook 3 – Morphology

- 3a) If CVCV, the first CV is a CV-word, and it is followed by the unresolved sequence CV... This returns us to (3).
- 3b) If CVVC, it can only resolve as CVV plus sequence C... (Recall that \*CV + V is proscribed.) This too returns us to (3).
- 3c) If it is CVVV, we must listen to the first two Vs and proceed as follows.
  - 3c1) If they form a monosyllabic pair, giving CvvV, then Cvv is resolved as a word and we go with a V-initial sequence to (4), which handles V-initial segments that are not initial in their breathgroups.
  - 3c2) If they are disyllabic, giving CV,VV, then this sequence can resolve in two ways. First, as CV + VV; second as a disyllabic CV,V + a V-initial sequence. To discover which, count the consecutive V's if any exist beyond CV,VV. If the number is zero or even, then the sequence resolves as a CV word followed by one or more VV words. If the number is odd, it resolves as a disyllabic CV,V word which is also followed by one or more VV words. (There can be no V-form words in the string of consecutive V's.) This brings us to the end of the segment or to the next C, which returns us to (3).
- 4) We know that the segment is not initial in its breathgroup. Therefore if it is V-initial, then it can only commence with a string of one or more VV elements; and these are so resolved. This takes us either to the end of the segment or to the first C; and the latter returns us to (3). If the non-initial segment is C-initial, we also return to (3).

In this way, all little words in well-formed utterances are unconditionally resolved.

We now consider the **conditional resolution** of a segment which is the prequel of a possible name-marker.

- a) We ask first if the segment is C-final. If it is, it will not resolve into regular words and is so reported. If the segment is V-final, continue.
- b) Is it breathgroup-initial? That is, is the segment preceded by a pause? Then any number of V's may lie between its C's, or before the first C, or after the last C. So resolve it unconditionally; for it will resolve into regular words.
- c) Is it C-initial? If it is, every possible C-V pattern will resolve, so resolve it unconditionally. If it is V-initial, count the V's before the first C, if there is one, or in the segment if it contains no C. If there is an even number of V's, it will resolve unconditionally. It either is, or its head is, a string of VV-form elements. But if the number of V's is odd, it will not resolve and is so reported.

In sum, only if a segment is (i) C-final, or (ii) breathgroup-medial and commences with an odd number of V's will it fail to resolve. All other prequels of apparent name-markers will resolve unconditionally.

## C. THE RESOLUTION OF PREDICATES

**2.34 The Functions of Predicates:** Predicates are the words with which claims are made and objects described in any language. Thus **Da mre<sup>nu</sup>** = 'X is a man' and **Ta blan<sup>u</sup> bak<sup>so</sup>** = 'That's a blue box' as well

## Notebook 3 – Morphology

as **Le mrenu** = 'The man' and **Leva blanu bakso** = 'That blue box' essentially involve the predicates **blanu**, **bakso** and **mrenu** and 'is blue', 'is a box' and 'is a man' in each language. Predicates are also the most numerous words in any language, and in Loglan they belong to a single part of speech. But this is a grammatical matter which will be discussed more fully in Chapter 4. Morphologically, Loglan predicates are of two kinds, (a) the **little word predicates**—like the cardinals, ordinals, acronyms and identity words that have already been described as compound structure words—and (b) the **morphological predicates** that remain to be described. These are the words which are immediately detectable morphologically as predicates and are the subject of the present chapter. In this chapter the term 'predicate' will refer only to words which are immediately detectable as such unless it is specifically qualified to include little word predicates.

**2.35 A Temporary Stress-Marking Convention:** As we have already seen, stress is an important datum in resolving all varieties of predicates. So in this chapter, I will conventionally use an apostrophe ['] to mark the stressed vowel in most predicate words. Thus, **mre'nu** is merely a swifter way of writing **mrenu** /MREnu/. My use of this convention does not mean that stress is now to be marked with apostrophes in written Loglan. Quite the contrary. Stress in predicates is still regularly inferable from the structure of the word, and so is omissible from the written form. It is to familiarize the reader with the inferable pattern of stress among predicates that I shall be marking it here.

**2.36 The Post-Emphatic and Intervocalic Pauses:** Two occasions require that the speaker pause before a predicate. One is after emphasis; the second is when the predicate itself commences with a vowel. Neither is marked by a comma in text.

Occasionally an emphatic or stressed little word or syllable comes just before a predicate: /le**TO**.MREnu/ = 'The two men!'. If a pause were not used at this juncture, the production /le**TOM**REnu/ would resolve as **le to'mre nu**. At the present writing (1987) **to'mre** is not a word; but it is a legitimate word-form and would be heard as a new word by a novice or a computer.

Occasionally some non-initial predicate in the speechstream commences with a vowel. [None could in 1975; but some borrowed predicates now do.] Unless the preceding word is a name, which requires its own post-nominal pause, the preceding word will be regular and so end with a vowel. So in all such cases the required pause will be intervocalic. Intervocalic pauses are very brief, a mere glottal stop. Nevertheless, the pauses in **lo a'tmo** /lo.ATmo/ = 'atoms' and **le albumi!** /le.alBUMi/ (leh . aal-BOOM-ee) = 'the albumin' are distinguishable from the same pairs of vowels pronounced pauselessly but disyllabically: /loATmo/ and /le,alBUMi/ (leigh-aal-BOOM-ee). These last productions contain audible glides between the adjacent vowels and would resolve not as phrases but as the single predicate words **loatmo** and **lealbumi**. Since these are also permissible borrowings they would be heard as such.

So it is important for the accurate resolution of predicates that both the post-emphatic and the intervocalic pauses be observed. Fortunately, they are both easy to observe, being quite common in natural speech.

**2.37 Stress in Predicates:** Stress is on the penultimate syllable in all predicate words including the numerical ones and acronyms. Some complex predicates like **saa'dja** have disyllabic terms; so the penultimate stress rule does not mean that stress will always fall on—or even in—the penultimate term. In **saa'dja** (saa-AAD-jaa) it does, or at least in it; in **matso'e** (maat-SOH-eh) = 'mother's-sister' it doesn't. Similarly, if **rie** in **casrie** is taken to be monosyllabic (as it may be; see Sec. 1.14), stress will fall on the first, or penultimate, term: (SHAAS-ryeh). If it is pronounced disyllabically—and currently it may be, at the speaker's option—then stress will fall on the first half of the last term: (shaas-REE-eh).

Neither hyphens nor buffering vowels are counted in fixing stress. Thus **mekykiu** is (MECK-uh-kyoo), and **matyma** is (MAAT-uh-maa).

Stress in other compounds and disyllabic little words has not yet been fully regularized...and indeed may never be, since as we have seen word-resolution does not depend on uniform stress production in these regions of the Loglan speechstream. (Will this remain a free area in Loglan morphology? Let's watch and

## Notebook 3 – Morphology

see.) Thus, indicators when disyllabic and initial in an utterance tend to be stressed on their second syllables (eigh-AA-moo-GOHT-soh) but need not be, and on their first syllables when alone or final: **Eo** = "Please!" is "EH-oh"; and **Mu gotso eo** = 'We go, please!' is "moo-GOHT-soh-EH-oh". But again, these patterns are not involved in resolution, and so are strictly optional.

**2.38 The Forms of Predicates:** Predicates are distinguishable from strings of little words by one major feature and two minor ones. The major feature is that, among the regular words of the language, that is, among non-names, all and only predicates contain adjacent consonants. One of the minor features is that, unlike little words, predicates may contain the sound /y/; where /y/ does occur in the predicates of unbuffered dialects (see Sec. 1.8) it is a hyphen. The other is that all predicates but not all structure words are stressed penultimately. Predicates share vowel-finality with the other regular words, the structure words.

Predicates may be of any length greater than three phonemes. Indeed any 4-letter or longer word is a "resolvable predicate"—although not always an allowable one—if it meets the following five requirements:

1. It must be uniquely resolvable as a single word.
2. It must be non-monosyllabic.
3. It must have at least one consonant-pair.
4. It must be penultimately stressed.
5. It must be vowel-final.

It will repay us to consider the first two requirements carefully.

First, nothing is a predicate that resolves as a phrase. Thus **\*tokli'ni** is not a predicate because it is the phrase **to kli'ni**. Similarly, **\*tosma'bru** is not a predicate because it breaks up into an operator, **to**, and something that does resolve as a predicate, namely **sma'bru**.

The second requirement is related to the first. Consider the monosyllabic pseudo-word **\*spai**, once thought to be a permissible borrowing. But if **\*spai** were a predicate, whenever it happened to be stressed relative to its sequel, it would join with the first V- or vv-final syllable of that sequel to create something that would always be another predicate. For example, suppose someone wanting to use **\*spai** as a predicate said /SPAIlēMREnu/ intending to say **\*Spai le mrenu** = 'Spy on the man'. But what da's auditors will actually hear is **Spaile mrenu**; for **spai'le** meets all the requirements of a resolvable predicate. (Check them out.) Thus no monosyllabic predicate will always resolve uniquely. Whenever it was non-final and stressed, it would absorb the next syllable and create a word that would run beyond itself. Therefore there can be no monosyllabic resolvable predicates.

The three final requirements guarantee that we are talking about predicates and not about little words or names. That there should be at least one consonant-pair, and that stress be penultimate are the two discriminators of predicates in Loglan; and that they be vowel-final is a property of all non-names. But because of the intervocalic pause rule, that first CC need not be early in the word. Indeed, it may be among the last three letters of a predicate and preceded by an indefinite number of vowels. For example, **aoaoaoao'sti** (ow-ow-ow-ow-OWS-tee), though improbable, is a resolvable predicate. (Check it out.)

Suppose the speaker didn't say what da meant. Suppose one of those **ao**'s was intended as an attitudinal, say the third one. Then all the **ao**'s to its left would also resolve as attitudinals, and the predicate would start immediately to its right, thus requiring that intervocalic pause: /aoaoao.aoAOsti/. Once da corrects da's error and the pause is spoken in this way, the new utterance resolves as **Ao ao ao aoaosti**; and **aoao'sti** is the resolvable predicate. Notice that no comma is used in text to announce this intervocalic pause.

The portion of a resolvable predicate that comes before its first CC (its prequel) must be composed entirely of vowels except that its first element may be a consonant. Thus **taoao'sti** is also a resolvable predicate.

There is one category of resolvable predicates which are not allowable. These are the borrowings that fail the "Slinkui Test"; of these **\*sli'nkui** itself is an instance; see Sec. 2.58.

## Notebook 3 – Morphology

[The domain of possible predicate-forms has been hugely expanded over the limited 2 mod 3 sequences of 1975. Most of the new word-space is reserved for borrowings. Not only will this space be sparsely occupied, but as we will see later (Sec. 25), only some of that space is allowable.]

**2.39 Three Kinds of Predicates:** Predicates may be partitioned morphologically into three main types: **primitives**, which are of the two 5-letter forms [CV'C/CCV'] + CV; **complexes**, which have—unremarkably—a fairly complex pattern of forms which we will discuss presently; and **borrowings**, which are the residual of the other two. That is to say, anything is a borrowing that is an allowable predicate and not a primitive or a complex.

**2.40 Primitives:** These are the 5-letter words like **ma'tma** and **bru'di** ('mother' and 'brother'). There is no other form of primitive predicate. Some words that are morphologically primitives are derivationally borrowings. For example, **simba** (SEEM- baa) 'lion' is primitive in shape but was borrowed unchanged from Swahili, where it also means 'lion'. Similarly, the international word for 'meter' is **metro** in Loglan. This word too has been borrowed, but this time from the international scientific community where both the concept and some form of the word are all but universal.

Thus, derivationally considered there are **S-Prims** (taken from the international vocabulary of science), **N-Prims** (native or local words for things or concepts that originate in that place), **I-Prims** (international words, like **telfo** for 'telephone', that have spread internationally from a single origin), and finally, **C-Prims**, or those composite primitives which have been derived from as many of the eight target languages as possible and which are the "semantic primes" of the language.

**2.41 Complexes:** These are words like **furve'a** (foor-VEIGH-aa) = 'buy', derived from the phrase **fu ve'dma** = 'the one who is sold something'; **me'kykiu** (MEK-uh-kyoo) = 'ophthalmologist', derived from **menki kicmu** = 'eye doctor', and **geoykre'ni** (goy-uh-KREH-nee) = 'gamma-ray'. Like compounds, complexes are composed of separately identifiable, meaning-bearing parts. These parts are called affixes. Unlike the parts of compound little words, however, all of which are intact little words, the affixes of which complexes are made are often, but not always, foreshortened versions of their source words. Thus **-kreni** is not foreshortened; but **mek-**, **-kiu** and **-vea** are. The sources of the foreshortened words or affixes found in complexes are usually, but not always, primitive predicates. The sources of **mek-**, **-kiu** and **-vea** are primitives, namely **menki**, **kicmu** and **vedma**; but some are not. The sources of **geoy-** and **fur-** are the little words **geo** 'lower-case gamma' and **fu**, which is usually a sort of generalized passive. **Fu** has to be lengthened to become an affix; so the lengthener /r/ is used. Unlike /y/, which is a buffering hyphen—in **me'kykiu**, for example, /y/ buffers the two /k/s, which if allowed to come together would shorten the production to /MEKiu/, which would resolve as the little word string **me kiu**—/r/ is a kind of "morphological glue". It keeps things from falling apart. Without /r/ /fuVEa/ would resolve as **fu vea**, a meaningless little word phrase.

There is a **defining metaphor** standing behind each complex predicate. That metaphor or phrase constitutes the derivational source of the complex word. Thus, the metaphor behind **furvea** is the phrase **fu vedma**, which means 'the 3rd place of the predicate **vedma**'. **Vedma** in turn means '...sells...to...for price...'; so its 3rd place designates the buyer. From this we may infer that **furve'a** means '(to) buy'; and indeed it does. Behind **me'kykiu**, as we have already learned, stands the phrase **menki kicmu** 'eye doctor' and behind **geoykreni** stands the standard international scientific metaphor 'gamma-ray'. Complexes thus come in a wide variety of shapes and sizes. All of them, however, resolve into a unique string of affixes with or without linking hyphens; see Sec. 2.46 for the contexts of hyphenation.

A final characteristic of every well-made complex of **\*tosmabru**-shape is that it has passed the Tosmabru Test; see 2.56. (**\*Tosmabru** didn't; therefore we are obliged to star it.) There is more on complexes in Secs. 2.48-56.



## Notebook 3 – Morphology

**2.42 Borrowings:** This is the residual class composed of all the words that are resolvable as predicates—that is, that have the properties of predicates listed in Sec. 2.38—but which are not primitives and not complexes, and which have also passed the Slinkui (SLEENG-kwee) Test given in Sec. 2.58. It happens that all such words, or nearly all of them, have been borrowed from some natural language or group of languages. Thus **i'glu** is identifiable as a borrowing because it is a legitimate predicate that is neither a primitive nor a complex. Derivationally we know it comes from the Inuit (Eskimo) word of identical pronunciation. But even if we didn't know that, if we just looked at its sounds and sound-sequences alone, we would know that **i'glu** was a predicate of some kind. It has the required consonant-pair in /gl/; it is vowel-final; it is not a monosyllable; it is penultimately stressed (/IGlu/); and it cannot break up as a string of words or absorb parts of its environment to make a larger word (which is what words that fail the Slinkui Test unfortunately do). Furthermore, we know that **i'glu** is not a primitive because it is not like either **matma** or **brudi** in form. And finally, we know it is not a complex because it cannot be "parsed" into a string of affixes with or without hyphens; see Sec. 2.61. Some of these claims are certainly not obvious. But they are all true of **i'glu**; and your own developing morphological sense of the language will soon assure you that all these claims are true whenever they are true of some borrowing that you are considering putting into the language.

Derivationally, borrowings are as diverse as the languages and cultures from which they come. Thus **i'glu** comes directly from the languages of the Inuit, **alkurcu'fa** (aal-koor-SHOOF-aa) from the Arabic for artichoke; anthropology buffs will certainly be able to guess that **atla'tli** (aat-LAAT-lee) comes from the same Nahuatl word 'atlatl' (aht-ll-AHT-ll) for 'spear-thrower' that has been adopted by anthropologists for spear-throwers the world over; **pro'a** (PROH-aa) comes through English from the Malay 'perahu'; and **proto'ni** (proh-TOH-nee) imitates the Italian pronunciation of that word exactly but is really derived from all the European words for 'proton', including Italian 'protoni', inasmuch as they are all cognate borrowings of this Greek word for 'first'. Indeed the largest stock of borrowed words in Loglan comes from the international vocabulary of science. Thus words like **i'glu** and **atla'tli** are borrowed from natural language words in Loglan because cognate forms are widely used in an international scholarly discipline, namely anthropology; the Loglan word for 'fluorine' is **fluori'ni** (floo-aw-REE-nee) because of the nearly universal use of this word, or something very like it, in international chemistry; and **proto'ni** (proh-TOH-nee) is our word for 'proton' because, by what amounts to an international agreement, Greek 'proton' has been borrowed for this concept by all the cultures which do physics except those of China and Japan, in which a local metaphor of similar meaning is used.

There is more on borrowings in Secs. 2.57-58 and 2.62-63.

**2.43 Consonant-Pairs:** Any predicate word must have at least one CC. Long predicates often have several such pairs, and in some borrowings there are strings of 3 or even 4 adjacent consonants. Any adjacent pair of such consonants may be regarded as either an initial pair, like /br/ in **brudi**, or a medial pair, like /tm/ in **matma**. A different set of pairs is permissible in each position. The **permissible initials** are a proper subset of the **permissible medials**.

In an initial triple (e.g., /str/) both pairs (/st tr/) must be permissible initials. In a medial triple in a complex (e.g., /ntr/) the first pair (/nt/) need only be a permissible medial, but the second pair (/tr/) must be a permissible initial. In borrowings, both pairs in a medial triple (e.g., in /ntl/ both /nt/ and /tl/ need only be permissible medials.

**2.44 Permissible Medials:** Intelligibility studies of the consonant-joints have shown that the list of permissible medials that limited the predicate-forms of the 1975 language can be considerably lengthened. Any  $C_1C_2$  is now permissible in mid-word unless:

- |   |                  |
|---|------------------|
| (a) $C_1 = C_2$   | E.g., <b>*kk</b> |
| (b) $C_1$ is h  | <b>*ht</b>       |
| (c) $C_1$ is the unvoiced variant of $C_2$                              | <b>*pb</b>       |
| (d) both are <b>c s j</b> or <b>z</b> ,                                 | <b>*cs</b>       |
| (e) $C_1$ is <b>p t k</b> or <b>f</b> and $C_2$ is <b>j</b> or <b>z</b> | <b>*pj</b>       |



## Notebook 3 – Morphology

In 1975 permissibility far exceeded use. The only no-longer-permissible initial CC which was actually used in 1975 is **nr**. Its elimination was required to make an earlier version of hyphen **r** work, but with the addition of hyphen **y**, **nr** is possible again and could be added.

The restrictions on both initial and medial consonant-pairs apply only to predicates. In names, anything even vaguely imitative of the natural word is permissible: e.g., **Xrucof**, **Strndl**.

New permissible initials may be added from time to time; for example, **sf** is now being considered.

**2.48 The Decipherability of Complexes:** Deciphering a Loglan complex predicate like **mekykiu** is always possible because the affixes and hyphens of which complexes are made are always resolvable; that is, their joints can always be found out. Moreover, once an affix has been resolved in this way it can always be associated with exactly one predicate or little word, which means that the predicate can always be uniquely deciphered by the listener or reader who knows the meanings of all the affixes. Thus, while some predicates have been assigned more than one affix, no affix has been assigned to more than one predicate. Moreover letter-words, which imitate some affixes, never appear in the same contexts as the predicate-derived affixes. So if **mek** means **menki** 'eye' in one word, as it does in **mekykiu** 'eye-doctor', it will always mean **menki** in whatever other words it may appear. In this way, the defining metaphor underlying every complex can always be recovered.

The potentially ambiguous cases are those predicate affixes which happen to duplicate letter-words, mostly Greek ones, in current scientific usage. Thus **beo** is the CVV-form affix of the predicate **begco** = 'beg' or 'request'. But **beo** is also lower case beta, and is occasionally used in such technical terms as **beoytse'ro** (beigh-oh-uht-SEHR-oh) = 'beta-error'. By convention such letter-word affixes are always prefixes and always attached by hyphen /y/, not /r/, to words which are recognizable as predicates without them (as **tse'ro**, for example, is). In a similarly initial position a predicate affix would be attached by hyphen /r/. Thus, **beordo'u** (beigh-awr-DOH-oo) means 'beg-give' or '(to) grant a request'. CVV-form predicate affixes may also be final in complexes, where they bear no hyphen. Thus **gadbe'o** (gaad-BEIGH-oh) means '(to) pray', being derived from **gandi begco** = 'god-beg'. In sum, the letter-word or "beta" sense of **beo** is always expressed by **beoy-** (beigh-oh-uh), while the "beg" sense of **beo** is expressed in complexes either by **beor-** (beigh-awr) or by **-beo** alone. By marking them with a distinctive hyphen, the ambiguity that might otherwise arise from using letter-words in complexes is avoided.

The affixes currently assigned to predicates are shown in List 3. The affixes currently assigned to little words are given in Sec. 2.54. Any letter-words is a potential affix provided it is used prefixively and hyphenated with /y/.

**2.49 Affix-Length and Frequency of Use:** Two types of affixes are used in the construction of complexes: "long" ones and "short" ones. **Long affixes** are either 4 or 5 letters long; **short affixes** are all 3 letters long. By judicious selection of the lengths of the affixes the word-builder chooses, da can adjust the length of the resulting complex to be (roughly) inversely proportional to its expected frequency of use, thus satisfying the Zipf principle of bio-linguistical efficiency; see the reference to Zipf's work in **Loglan 1**.

**2.50 Term-Reduction:** When a complex like **ni'lboi** ('child') is composed entirely of short affixes (**nil** + **boi**) it is called **fully reduced**. The words in the defining metaphor of a complex (in this case **nilca botci**) are called its **terms**. To replace a term in a metaphor with a short affix is to **reduce** it; to replace it with a long affix is to leave it **unreduced**. The set of short affixes has been designed to be an efficient "reduction machine". Currently, the efficiency of this machine is 95% in the exact sense that 95% of the terms of complex predicates in the 1975 dictionary could be reduced to short affixes if we wished them to be. The efficiency of the Loglan short-affix set—which is probably the largest in any language—will probably diminish as the semantic range and variety of complex predicates in the language increases. What this means is that complexes which are late additions to the language are less likely to be reducible than early ones... which is, of course, a fitting outcome since the early ones are those which are used more frequently.

## Notebook 3 – Morphology

**2.51 Long Affixes:** Every primitive has two long affixes. The one used in final positions is the 5-letter primitive itself, e.g., **-kre'ni** in **geoykre'ni**. The one used in non-final positions is the unique 4-letter affix that may always be obtained by removing the final vowel of a primitive. That is to say, primitives have been constructed in such a way that no two of them differ in only their final vowels unless they are members of the language/nationality/culture trios such as **spana/spani/spano** between which differentiation in a complex is seldom necessary. The place of that final vowel is then occupied by hyphen /y/. Take the metaphor **ni'lca bo'tci**, which means 'girl-boy', for the concept of "child" (not in the sense of "offspring", but the creatures who, along with puppies, play in village streets). Stripping the (non-distinctive) /a/ from **nilca** and replacing it with /y/ yields **nilcybo'tci**. So both or all affixes in a complex may be long. Or the first may be short and the second long: **nilbo'tci**. Or the other way round: **ni'lcyboi**. Or, as we have seen, both may be reduced. (**Ni'lboi** is in fact the dictionary entry for this sense of 'child'; it is derived from a Chinese-type metaphor, the logical sum of girls or boys.) Thus, every Loglan complex is in principle polymorphic, so its length may usually be adjusted to changing frequency of use.

**2.52 Short Affixes & Their Derivations:** The short affixes are all three letters long. There are four forms: CCV, CVC, Cvv with a monosyllabic vowel-pair, and CV,V with disyllabic vowels. For most purposes, but not for all, Cvv's and CV,V's may be considered variants of the same form. Examples of the four affix forms, showing all possible derivations from primitives, are as follows:

- CCV: **cli** from **clika**, **fra** from **farfu**; rarely, **pre** from **papre**, **klu** from **kutla**.
- CVC: **mat** from **matma**, **nil** from **nirli**; rarely, **rim** from **trime**, **tan** from **trana**.
- Cvv: **boi** from **botci**; rarely, **lia** from **clina**.
- CV,V: **cea** from **cenja**; rarely, **keo** from **ckemo**.

Affixes with the "best", i.e., the most memorable derivations—like **cli**, **mat** and **boi**—have been assigned to the most **productive** primitives, that is, the ones that appear in the greatest number of metaphors in the 1975 dictionary; and so their incidence in speech will hugely outnumber affixes derived by less obvious routes... like **kiu**, **tan** and **lia**.

This fact may be usefully employed in constructing an optimal order for learning the short affixes in the affix-set. For example, if one has learned the most productive CCV-assignments first, and so has learned that **cli** belongs to the very productive **clika**, then, when one comes upon the less productive **clina**, one not only knows that its most obvious affix has been pre-empted one can easily understand why. This makes psychological room, so to speak, for learning **lia**.

This optimized order is the one in which the affixes are presented to the learner in **MacTeach 2**.

**2.53 Affix-Assignment & Coverage:** Not all primitives have short affixes. Currently the most productive 600-odd of the 800-odd C-primis have been assigned at least one short affix; and between 800 and 850 short affixes have been assigned. Thus, some productive primitives have two or more short affixes for use in different contexts. For example, **madzo** has both **mad** and **mao**; **cenja** both **cen** and **cea**; and **godzi** has all three possible short affixes, namely **god goz** and **goi**.

Short affixes were assigned to primitives in such a way that their **coverage** of the metaphors in the 1975 lexicon was maximized, that is to say, the number of terms in that pool of metaphors that was reducible by those affixes was maximized. Currently that coverage figure is 95%; see Sec. 2.50.

Short-affix space is extremely limited: 1,965 possibilities exist, of which the most useful 42% have already been assigned. Often what remains in the empty spaces of the assignment tables are forms like ZVA, ZVE and ZVU which are unassignable. When this "unusable" space is removed—unusable that is, by the existing set of primitives, or anything like it—it turns out that about 75% of the usable space has already been used.

## Notebook 3 – Morphology

Moreover, there is often strong competition between primitives for the remaining affix space. At the present, mature stage of affix assignment, such conflicts can only be resolved by remaking more primitives. Affix assignment virtually stopped in 1982 when (1) coverage reached 95%, and (2) the cost to old loglanists of relearning remade primitives reached approximately 1 in 7 of the original 1975 list. This may have been premature. Perhaps 40 or 50 additional primitives should be remade, which would lift term-reduction efficiency by another 2% or 3%.

**2.54 Pre-empted CVr Affixes:** A short list of r-final CVC affixes was pre-empted for the 10 digits and 8 other operators found frequently in complexes. Currently these are: **fer for fur, jur, ner nir nor nur, pir, rar rer ror, ser sor, ter tor, ver vor**. Apart from the digits they come, of course, from the non-numerical quantifiers **ra re ro**, the negative **no**, the decimal point **pi**, and the three simple conversion operators, **nu fu ju**.

**2.55 Making Complexes:** To make a complex predicate, one should first gauge the expected frequency of use of the concept to be rendered. Then a metaphor must be contrived that allows the degree of term reduction necessary to produce a word of the length appropriate to that frequency rank. [The Institute will eventually publish length/frequency tables for the guidance of word-makers.] Suppose a complex is to be fully-reduced. The rules for combining short affixes are as follows:-

1. CCV's may go anywhere in a CPX; they never require a following hyphen.
2. CVCs may be only non-final; if the consonant-joint with the next affix is impermissible, they must be hyphenated with /y/. E.g., **mekykiu** and **keryrou**.
3. Cvv's and CV,V's may go anywhere, but if initial in anything but a 2-term word with final CCV, e.g., **boicli**, they must be hyphenated with /r n/.

As an illustration of Rule 3, if it is desired to combine **boi + cli + mre** ('boy-like-man'), an /r n/-hyphen must be introduced between the first 2 terms: **boircli'mre** (boy-rr-SHLEEM-reh). Otherwise, **boi** would "fall off". Without the hyphen, /boiCLImre/ would resolve as the phrase **boi climre** = 'beta is a "similar-man" (whatever that means)'; and it doesn't matter that the sequence **cli'mre** is currently unassigned. To the computer's ear—and that of the novice—this is an unknown complex with a decipherable metaphor, "similar man".

Similarly, in a 2-term word composed entirely of Cvv's and/or CV,V's, e.g., **hai + hoi**, the joint between them must also be hyphenated with /r n/: **hairhoi**. Otherwise the production /HAIhoi/ is just a pair of little words. There is more on hyphenation in. Sec. 2.46.

Long affixes may always be substituted for short ones. If the final term is to be left unreduced—for example, if greater length is required, or if there is no usable short affix for that primitive in that position—then its 5-letter affix may be used. If a non-final term is to be left unreduced, then its 4-letter affix plus hyphen /y/ may be used.

Each trial word must then be tested for resolvability.

**2.56 The "Tosmabru Test":** Only one type of complex word has known resolvability problems. These are the "Tosmabru-type" words in which the last term is a long affix of **ma'bru**-form in which the medial consonant-pair is (like **br**) a permissible initial. All non-final terms (there may be several) must be CVC's (like **tos**) for resolution problems to arise. Tosmabru-type words can break up into phrases...as indeed /toSMABru/ does: **to smabru**.

To test your trial word for "Tosmabru failure", examine all the C/C joints between the CVC affixes, and between the last CVC and the **mabru**-term. If all those C/C joints are **bridged** by permissible initials, then the word will break up. But if any C/C joint is **unbridged**, i.e., is impermissible as an initial CC, your word

## Notebook 3 – Morphology

will not break up. It has passed the Tosmabru Test. Interestingly enough, only one unbridged joint per word is sufficient to ensure resolvability.

Suppose your trial word is **?gusnilbo'tci**. Its intended terms are **?gus/nil/botci**. S/N is bridged; L/B is not bridged. One unbridged joint is enough. Therefore it passes. The production /gusnilBOTci/ cannot resolve as the phrase **gu \*snilbotci** because **\*snilbotci** fails the "Slinkui Test", which is a test for borrowings; see Sec. 2.58. (Before we've tested it **?snilbotci** is an apparent borrowing.) Suppose our trial complex is **?gusnitca'pri**. Its intended terms are **gus/nit/capri**. Both S/N and T/C are now bridged. Therefore **\*Gusnitcapri** fails. It is in fact the phrase **gu snitcapri**. **Sni/tca/pri** is clearly quite a different complex.

Tosmabru failures may be repaired in several ways. One is to try different affixes, or even different metaphors, until a joint appears that is not bridged. Another is to hyphenate. Hyphenating the first C/C joint in **\*gusnitcapri** (but not the second) would repair it. Thus **gusynitcapri** (goos-uh-neet-SHAAP-ree) resolves. There is also a simpler solution in this case. **Gusto**, the obvious source of **gus**, has another short affix, namely **gut**. Using **gut** instead of **gus** in **\*gusnitcapri** demolishes the first bridge, and so **gutnitcapri** resolves.

**2.57 Allowable Borrowings:** There are four tests which must be passed by a usable borrowing. Any word which is (1) a resolvable predicate (Sec. 2.38), (2) neither a primitive nor a complex, (3) free of "imitation hyphens" (it obviously can have no real ones if it is not a complex), and (4) neither a **Slinkui-type** word nor a **left-extension** of such a word, e.g., **\*aslinkui**, is an allowable borrowing.

**\*Slinkui** is a disallowed word, typical of its class, that on the surface seems perfectly permissible as a borrowing. It is certainly a resolvable predicate. It is certainly not either primitive or complex. But then one notices that if **\*slinkui** were allowed in the language, a certain perfectly regular complex, namely **paslinkui** (which means 'ancestor', from 'past-linear-kin'), would not uniquely resolve. For if **\*slinkui** were allowable, **paslinkui** would break up as the phrase **pa \*slinkui**.

On the other hand, if we disallow **\*slinkui**, and all words of its type, then **paslinkui**, and any word of its type, will not break up. The resolver would first obtain **\*slinkui**; note that it was disallowed; look at the prequel; find **pa**; and then it would say, Right! **Paslinkui** must be it, then. That's the only thing /pasLINKui/ can be, since **\*slinkui** is not allowed. (This, incidentally, is how **Aslinkui-type** words would also get resolved...if we allowed them. But these left-extensions of the basic Slinkui-form are disallowed for a slightly different reason which I will take up presently.)

The principle underlying the exclusion of Slinkui-type words from the domain of borrowings (some apparent Slinkui-words, by the way, e.g., **smabru**, are quite acceptable as complexes) is that complex predicates are so much more important in Loglan semantics than borrowings are—it is only the imitativeness of borrowings, after all, and not their number, that we intend to enhance by giving them a large word-space—that whenever there is a competition for a given portion of the word-space between these two varieties of predicate, it is the complexes that must always win out.

Thus, Slinkui-type borrowings (but not complexes) are not only excluded from the class of allowable borrowings, they are excluded from the language. That is why the "perfect complement set" I once thought we could achieve for borrowings is not perfect. It has a big empty swath cut right down the middle of it, namely the space occupied by all those resolvable predicates that are neither primitives nor complexes yet fail to pass the "Slinkui Test"; see below.

A second, similarly large emptiness in the domain of resolvable predicates is caused by our disallowing **Aslinkui-type** words as well. While these words are definitely resolvable once Slinkui's have been disallowed (but not before), the great variety of those possible leftward extensions of any Slinkui-form makes them difficult to see or hear. Indeed, they impose a "double double-take" on the resolver, a twisted loop in the resolution algorithm. So for these reasons—or this reason, if it is the same one—**Aslinkui's** too have been disallowed.

Let us finally consider the reasons behind the third test for allowable borrowings. We said it must contain no **imitation hyphens**, that is, no **y-** or **r-**involving sequences that imitate the hyphens of a complex. This exclusion is not always strictly necessary for resolution. It is not necessary, for example, when the word containing the apparent hyphen will not resolve as a complex. But in anticipation of the day when borrowings may also be involved in complexes—when some Inuit-speaking loglanist invents the word

## Notebook 3 – Morphology

**i'glymao** for 'igloo-maker', for example—in which case hyphenated segments in a complex may be exceedingly irregular, let us now exclude all those imitation hyphens from "simple" borrowings that might confuse such constructions later. What this means is that no /y/, no /r/ in the context /CVVrC/, and no /n/ in /CVVnr/ is permitted in a borrowing.

**2.58 The "Slinkui Test":** Any CC-initial trial borrowing must be given this test; no other need be. It is simply performed. Put a test prefix, say **pa**, on your trial word. Let's say it's actually **?slinkui**. Then try to resolve **?paslinkui** as a complex; see Sec. 29. If the effort succeeds, as it does with **pas/lin/kui**, then the trial-word fails. If the term-resolving effort fails, as it would, for example, with **pas/lin/kuti**, then no coalescence with a preceding CV operator is possible; so the trial word passes.

Suppose your trial word is **?spe'a**. **?Spea** has a form that was once accepted as an allowable borrowing. But **?spe'a** obviously fails the Slinkui Test. **Paspe'a** immediately resolves as complex and so **\*spea** fails as a borrowing. [Thus a whole set of words, once accepted as legitimate borrowings—albeit in a narrower, modularly defined lineage—now vanishes from the domain of allowable borrowings. And so from the language.] But trial words that fail the Slinkui Test can usually be easily repaired. Any disturbance in the smooth structure of the imitation complex that takes over after the first letter—e.g., the **linkui**-part of **\*slinkui** or the **pea**-part of **\*spea**—will also suffice to make it pass the test. For example, **\*sli'nkui** could be repaired as **slie'nkui**, **slinkui'a**, **slini'kui**, or, as noted, **slinku'ti**. Indeed, dozens of variations on each trial word are usually possible; one of them might imitate its natural model almost as well as the word that failed.

Some Slinkui-form sequences discovered by the resolver in the course of word-resolution are actually complexes. For example, if one presented **?sma'bru** to the Slinkui Test, one would find out it was a Slinkui. No question of "failing" as a borrowing is involved here...merely the identification by the resolver of certain legitimate sequences as having Slinkui-form. There is more on this diagnostic use of the Slinkui Test in the next two sections.

**2.59 The Resolution & Partial Classification of Predicates:** Under the new morphology, the resolver proceeds very much as before, except that it must now deal with the possibility of vowel-initial predicates and it must perform the Slinkui Test on any CC-initial predicate-form sequence it encounters. Let us see why.

Suppose some such sequence "fails the test", that is, turns out to be a Slinkui. Since Slinkui-type borrowings are not allowed, and therefore do not exist in any well-formed string the resolver will encounter, the discovery that it has found something that looks like one—that this particular CC-initial sequence is indeed of Slinkui-form—can only mean (1) that it's a complex (e.g., **sma'bru**), or (2) that the sequence is part of some larger predicate. Slinkui-form sequences can only exist as words if they are complexes; so if some sequences exist that are not complexes, they must be parts of words. The kind of word of which they are parts must obviously be predicates. Moreover, the part they are is a right-hand part since the right juncture of the Slinkui-sequence will already have been fixed; see next section. Therefore there must be a left-hand part out there somewhere.

What kind of a left-hand part could it be? Well, if weren't for the Aslinkui exclusion, it could be anything that could precede an impermissible initial consonant pair that turned up as the first CC of a predicate. For in truth, the **sl** of a **\*slinkui** that is not complex cannot begin a word. Since Aslinkui-type borrowings have been excluded from the language as well (see above), there is only one allowable form for the left-hand part of that larger word, and that is the CV-form that will by definition turn any Slinkui into a complex. There will always be one, of course, or the non-complex Slinkui would not be there. So the resolver looks for the CV ahead of each such Slinkui; always finds one (in any well-formed utterance); and puts left-juncture to the left of that CV and hence the word.

The word that results from this manoeuvre will always be a well-formed complex; the Slinkui Test will have seen to that. In a sense, the resolver short-circuits the normal recognition procedure in these Slinkui-cases. For normally the recognition of a complex is subsequent to its resolution as a predicate.

## Notebook 3 – Morphology

**2.60 The Predicate Resolution:** The new algorithm for resolving predicate words is in some ways simpler than the 1975 one. The resolution of a predicate still commences, of course, with the discovery of a CC in the left-to-right scan of some breathgroup. That first CC—there may be others, of course—may be either a permissible initial (**sp**) or impermissible one (**rk**). If the FIRST CC is IMPERMISSIBLE, the algorithm records that fact as a prevailing condition and then performs the following steps:-

1. Is there a stressed V' or vv' immediately ahead of the first CC? If so (e.g., **i'sp, ia'rk**), mark it and record STRESS FIRST as a prevailing condition. If not, find the first instance of a V' or vv' that follows the first CC (e.g. **spe', spudei', rida', rkoti'**) and mark it.
2. Find the first instance of a V or vv after the marked V' or vv' (**i'spai, ia'rko, spe'a, spudei'tai, rkia'mpa**, etc.). Put right juncture to the right of that V or vv (**i'spai|, ia'rko|, spe'a|, spudei'tai|, rkia'mpa|**, etc.).
3. If the condition STRESS FIRST and/or FIRST CC IMPERMISSIBLE obtains, find the prequel by locating the first instance of a C or ./ to the left of the first CC (**.aoi'spai|, tia'rko|, morkia'mpa|**). Put left-juncture at that ./ or before that C (**|aoi'spai|, |tia'rko|, |morkia'mpa|**) and exit.
4. Perform the Slinkui Test on any sequence that remains. (Stress will follow its first CC, and it will be a permissible initial.) If it is a Slinkui (e.g., **spe'a|, spa'rgu|, sma'bru|**), send it to the term resolver to discover whether it resolves as a complex. If its terms do resolve (e.g., **sma'/bru|**), put left-juncture immediately to its left (**|sma'bru|**). If it doesn't resolve as a complex, find the CV that will always be to its left and put left-juncture to the left of that CV (**|tospe'a|, |mispa'rgu|**). In either case, tag it as a recognized complex and exit. If it is not a Slinkui (e.g., **spudei'tai|, spu'ta|**), put left juncture immediately ahead of the first CC (**|spudei'tai|, |spu'ta|**) and exit.

Clearly term-resolution is involved in at least some cases of word-resolution. We must now consider how the resolver will "parse" the terms of a complex when the predicate is one.

**2.61 Term-Resolution:** Once a predicate has resolved as a word, the recognition of its primitiveness, if it is primitive, is trivial: it is either of **matma-** or **brudi-**form, or it is not. If the resolved word is not primitive, or if a Slinkui-form sequence has been identified by the resolver, an attempt is made to parse the word or sequence as a complex, that is to say, to resolve its terms if it has terms. If it has no hyphenation sequences and is 0 mod 3, then, if it is complex, its terms will just "fall out" in 3's, and each triple will be a legitimate affix: **sma/bru, nil/boi, tar/ses/mao, roj/mad/ses/mao**, and so on. If they don't or aren't (e.g., **tai/rko, mor/kia/mpa**), it is not complex. If it has no hyphens and is not 0 mod 3, then it will have a 5-letter final affix if it is complex; and the rest will fall out in 3's. If they don't and/or there is no long final affix, it is not complex. If it has hyphens or hyphen-like sequences, e.g., /y/, /CVVrC/ or /CVVnr/, it is complex. (Borrowings, remember, may contain no imitation hyphens.) So the algorithm then examines the regions between the hyphens, and between any hyphen and the ends of the word, and the lengths of these regions, together with their positions in the word, will always uniquely determine what lengths of affixes will be found in them if the complex is well-formed. (It will also identify whatever irregularly formed sequences it finds among the affixes as derived from borrowings, when this development is upon us.) In this way, every predicate word that is a well-formed complex will uniquely resolve into its terms; and every word that resolves into terms will be a well-formed complex.

**2.62 The Recognition of Borrowings:** The recognition of a borrowing thus always follows on the failure to hear or see it as a complex once it has resolved as a word. Thus, one knows one has a borrowing only by failing to resolve it as a complex. Thus **tia'cro** is complex because its two affixes, **tia** and **kro**, are well-formed. **Tia'rko, morkia'mpa** and **fainstoi'a** are not complex because their partition as complexes fails.



## Notebook 3 – Morphology

**Rko** and **mpa** are not well-formed CCV-affixes; and **fainstoi'a**, while 0 mod 3 and hyphenless, does not fall out into legitimate 3's: **fai/nst/oia**. And so all these words are borrowings.

The construction of a borrowing thus often proceeds by putting impermissible initials (**rk mp ns**) in judicious places, or by making sure that its affix-length pieces do not have affix-form shapes (**nst oia**). Thus, we may say that whatever is an allowable predicate word and not a primitive or a complex is a borrowing.

Observe, by the way, the operation of the pair-from-the-left rule of Sec. 1.15 in determining the character of the stressed syllable of **fainstoi'a**: /fainSTOIa/ (fighn-STOY-aa), which is quite a different sound from (fighn-STOH-yaa).

**2.63 Making Borrowed Predicates:** The procedure for making a predicate based on some natural language model involves four distinct steps:

- 1) Build a preliminary trial word that imitates the model word as well as possible...either its sounds or its letters, or rarely, both. When the model predicate is a scientific word, the word-maker may be guided by the transcription system given in Sec. 2.14 for pronouncing Linnaean names. Thus the predicate 'cephalopod'—which is either derived from or the source of the Linnaean name **Cephalopoda**—might well be initially transcribed as **?cefalopo'd**; for this is our recommended pronunciation, and hence a rewriting in Loglan phonemes, of the portion it shares with the related Linnaean name. Even if there is no related Linnaean term, transcribing it as if it were Linnaean is a good procedure to follow when borrowing scientific words.
- 2) Make sure the trial word is a resolvable Loglan predicate, repairing it as necessary to give it the necessary properties; see Sec. 2.38. **?cefalopo'd** (sheh-faa-loh-POHD) lacks two of these properties. It is not V-final and it has no CC. Let us repair the first defect by giving it the final **-a** of the corresponding Linnaean name, and the second by inserting **r** before the second C. /rC/ is always an impermissible initial, and having one sufficiently early in the word is crucial for preventing the first C and its following one or more V's from "falling off the word". **?cerfalopo'da** (shehr-faa-loh-POHD-aa) is the word that results from these moves.
- 3) Make sure that the developing word is not primitive or complex, and that it doesn't resolve as a phrase. This means checking its form against the two forms reserved for primitives—which require a more cautious procedure of derivation than you may wish to get involved with here—and attempting to resolve it first as a phrase, and when that fails, as a string of affixes. Both attempts fail with **?cerfalopo'da**. Nothing falls off; and term resolution starts out with **cer/fal/..** but then encounters a sequence which is impossible in a complex, namely **...opoda**. So the word passes the third test.
- 4) Check it for false hyphens. **?cerfalopo'da** has no /y/s, no /CVVnr/s, and the /cerf/ sequence just misses being an instance of /CVVrC/; so the developing word passes this test as well.
- 5) Test it for Slinkui Failure and the Aslinkui Condition. The first test is relevant only if the trial word commences with CC, as this one doesn't. The second condition can exist only in a trial word whose first CC is non-initial, is a permissible initial, comes before the stress, and has a prequel that resolves as words. Only in such words will the prequel be loosely-attached. The first CC, namely /rf/, is not a permissible initial; so the trial word avoids the Aslinkui Condition as well.

**Cerfalopo'da** passes all tests and is thus an allowable borrowing. Eventually, for at least the borrowed portion of the Loglan scientific vocabulary, this word-borrowing procedure can easily be made algorithmic. The borrowing algorithm is under development now. But much of the Loglan scientific vocabulary will not be borrowed. Large portions of it may be captured Chinese-style by local metaphor. For example, rather than borrow the sounds of 'xenophobe' one might prefer to "borrow" its idea, the idea of the "stranger-fearing" person. That would produce a scientific complex that would be decipherable by any Loglan reader, not just those learned in the vocabulary of Western science. If this were our policy, **gu'rfia** might well become the

### Notebook 3 – Morphology

word for 'xenophobe/-ic' in Loglan. It is made from **gutra firpa** = 'stranger-fear' and is thus a literal translation from the Greek. Whence **lopo gu'rfia** would be 'xenophobia', and so on. Still, the borrowing **zernofo'bi**—made by much the same procedure as **cerfalopo'da** was made above—is also possible, and has the advantage of being immediately recognizable to those who know the psychiatric vocabulary. But even in English—one might say especially in English—such Graeco-Latin words exclude from understanding precisely those whose educations have not permitted them to know.

To borrow or not to borrow. Lehnwort vs. lehn übersetzung. There are many good arguments on both sides of this question. As yet, The Institute has taken no position. We prefer to wait until the loglaphone community has extended itself internationally—and also intellectually, one would hope—beyond the narrow confines of anglophone computer science and those related Western disciplines of logic, linguistics and anthropology where interest in Loglan began some decades ago.

## CHAPTER 3

### LEXICON (WORDS & SPEECH PARTS)

**3.1. Definitions and Conventions:** In this chapter" the vocabulary of Loglan is partitioned into its 69 "parts of speech", or sets of grammatically interchangeable words. These sets are called **lexemes**, and we will consider them one at a time in the alphabetic order of their simplest or most representative members. These representative members, written in upper case letters, serve as **lexeme names**. We commence with Lexeme A, the Afterthought Connectives, and end with Lexeme ZO, the Quantity Abstractor.

If a lexeme has more than one member, these are called its **allolexes**. The allolexes of a lexeme are its equally permissible alternative expressions; they are the "interchangeable elements" of which that lexeme is composed. Lexemes which have only one member are called **monolexic**. In the section devoted to each lexeme, our object will be to show how its various allolexes are formed morphologically, how it is used by the grammar, and the range of meanings of its allolexes. Whenever possible, the list of allolexes is complete.

The allolexes of all the lexemes of a language are the **lexes** of that language. They are the words or phrases which, like English 'nevertheless', the grammar treats as a single word. In Loglan all lexes are words; that is, they are spoken pauselessly and written without internal spaces. The **lexer** is that part of the human central system, or of some computer program designed to process language, which identifies lexes as such, and which assigns them to their lexemes. In the system of computer programs that constitute the **machine grammar** of Loglan, the lexer is part of the **preparser** subsystem.

From the point of view of the grammar, each allolex of a non-monolexic lexeme is an equally legitimate occupant of whatever place any of them occupies in any utterance. Thus it makes no difference to the grammar which allolex of a lexeme a speaker chooses; for the grammatical structure of an utterance does not change when one allolex of a lexeme replaces another. So in a certain sense, if a learner has learned how to use one allolex of a lexeme, da has learned how to use them all. For they all have the same grammatical privileges. Thus, there are only semantical differences between allolexes of the same lexeme, not grammatical ones. It is in that sense that allolexes of the same lexeme are **grammatically interchangeable**.

The lexemes of Loglan are **exclusive**. That is, if a lex is a member of a certain lexeme, then it is a member of no other. This is not true of English, where the word 'bank', for example, belongs to at least three lexemes, each with distinct grammatical roles. The uniqueness of lexemic assignment in Loglan removes a major source of linguistic ambiguity from the language. This is the **lexemic ambiguity** ('They are flying planes') which seems to be both massive and universal in natural languages.

Lexemes whose names are 'M' followed by a numeral are called **M-lexemes** or **machine lexemes**. These and certain other machine-oriented lexemes do not appear in human Loglan and are visible and audible only to the machine. The human user of the grammar need not concern daself with them. The role of machine lexemes in the machine grammar is explained in the next chapter. The names of lexemes used only by the machine are \*-ed in the section headings. A total of 17 lexemes are occasioned only by the machine's needs. So from the human user's point of view, there are only 52 lexemes in Loglan.

Nine lexemes, or about one out of six in the human lexicon, differentiate connective words from one another. This is an extraordinarily high proportion compared to any natural language. As a logical language, Loglan makes great use of the principle of **connectivity**, the principle by which the truth-values of several or many sentences may be related to one another in a single sentence.

The PREDA lexeme is the largest lexeme in the lexicon, containing about 80% of all dictionary entries. PREDA's are all the predicates of the language, that is, its noun-, verb- or adjective-like words as well as many of its adverbs and prepositions.

Many allolexes of the non-PREDA lexemes are compound structure words. The **compounding formulas** which show how these lexes are generated are composed of lexeme names, the sign of alternation [/], of concatenation [+], and of identity [=]. Thus 'NI/TAI + FI = UI' means that members of the NI (nee) or TAI (tigh) lexemes may be joined together with members of the FI (fee) lexeme to produce compound structure words which will be lexed as members of the UI (wee) lexeme.

The words 'operand', 'modificand' and 'connectand' are originally Latin words meaning 'that which is operated upon', 'that which is modified' and 'that which is connected', respectively; they are used freely throughout the text.

## Notebook 3 – Lexicon

'R' followed by a numeral, e.g., 'R12', refers to a particular rule of grammar that will be found under that number in the next chapter.

The **parse** of an utterance is a specification of how the grammar generates or understands it. A parse may be partially shown by a full parenthesization of the utterance once it is stripped of any other punctuation. The nested parentheses show the order in which the lexes in the utterance are to be grouped. Thus the parse of **Da, a de, e di** = 'X or Y and Z' is partially given by **((da a de) e di)**. This shows that the expression is to be understood by first grouping **da a de** together, and then grouping the result with **di**. Note that this structure is independent of the semantic values of **a** and **e**, which are in fact allolexes of the same lexeme (Lexeme A, the first lexeme in the lexicon). In this chapter, the production sign [=>] will sometimes mean 'is parsed as'.

Reference will be made throughout this chapter to various grammatical structures whose origins and functions will not be completely understood until they are studied in the grammar. In general, the lexicon acquires its structure from the grammar while the grammar acquires its semantic variety from the lexicon. In truth they are inseparable; neither can be deeply studied without reference to the other.

### Lexeme A: Afterthought Connectives (Eks)

These are the afterthought connectives used between predicates, arguments, linked arguments, argument modifiers, and sentence modifiers. They are said to be "afterthought" because they are left-grouping, e.g., **da, a de, e di** => **((da a de) e di)**. So connectands may always be added to an ekked string without disturbing previously understood meanings.

When used between predicates, eks are prefixed with M11 by the preparser; when connecting linked arguments, the preparser inserts M1; when connecting argument modifiers, M6; and when connecting sentence modifiers—called simply "modifiers" in the grammar—M5. Only when used to connect arguments are eks unmarked.

The complete list of simple eks is:

a	...and/or...
e	...and...
o	...if and only if...
u	...whether...
nuu	whether..., ...
noa	...only if...
noe	not...but...
nou	not...whether...
anoi	...if...
enoi	...but not...
onoi	...or...but not both
nuanoi	whether..., not...
noanoi	not both...and...
noenoi	neither...nor...

In addition, any member of the PA-Lexeme may be appended to an ek, producing, for example, tensed, located or even "motivated" connectives: **efa** = 'and then'; **evi** = 'and here (at this place)'; **emoi** = 'and with this intention'; and so on. All such compounds are treated grammatically as members of A. Not many of these A + PA compounds have been explored, however; their semantic domain seems to transcend any found in the natural languages. The loglanist is invited to explore this new country. See Lexeme PA for a list of the PA components of A + PA compounds.

### Lexeme ACI: Hyphenating Eks

## Notebook 3 – Lexicon

These are any member of the A-Lexeme suffixed by **-ci**: thus **aci**, **eci**, **apaci**, **noanoici**, and so on, are all hyphenating eks. ACI connectives have an effect analogous to that of hyphen **ci** in a predicate string (see Lexeme CI); that is, they consolidate the two adjacent elements into a single connectand. At the moment, ACI is confined grammatically to ekking arguments and predicates; among the latter, the preparser will have prefixed M9 to the ACI word. The use of hyphenating eks could, of course, be extended to other ekked structures in the grammar, for example, to ekked modifiers. It would be grammatically costly to do so, however, since introducing ACI to an ekking structure usually involves the addition of 2 to 4 rules, and a new M-lexeme will always be required for every structure so accommodated. So for the present, and until a use for them in other contexts can be demonstrated, hyphenating eks are confined to connecting the two main ingredients of a sentence, arguments and predicates.

### Lexeme AGE: Right-Grouping Eks

These are any member of the A-Lexeme suffixed by **-ge**: thus **age**, **ege**, **apage**, **noanoige**, and so on, are all right-grouping eks. AGE connectives have an effect analogous to the effect of the grouping operator **ge** in a predicate string; that is, they consolidate the entire right portion of a string of ekked elements into a single connectand. At the moment AGE words are confined grammatically to ekking arguments and predicates; among the latter, the preparser will have prefixed M11 to the AGE word. The use of right-grouping eks could, of course, be extended to other ekked structures in the grammar, for example, to ekked modifiers. It would be grammatically costly to do so, however, since introducing AGE to a new ekking structure usually involves the addition of 2 to 4 rules, and always involves the addition of another M-lexeme. So for the present, and until a use for them in other contexts can be demonstrated, right-grouping eks are confined to arguments and predicates.

### Lexeme BI: Identity Operators

These are the "little word predicates" which must be kept grammatically separate from the PREDA-Lexeme because their compounds are recognized by the preparser. If it weren't for this morphological function, BI and kin would be members of PREDA. The current list of BI words is

bi	is equal/identical to
bie	is a member of
biu	is part of
cie	is less than
cio	is greater than

and the compounds

ciebi	is less than or equal to
ciobi	is greater than or equal to

Other compounds may be made by prefixing **no-** or **nu-** to any of these with obvious meanings. Thus **nocio** makes the same claim as **ciebi**. No doubt mathematicians will have other allolexes to add to BI.

BI is an open lexeme; new members may added at any time.

### \*Lexeme BAD

## Notebook 3 – Lexicon

This is special lexeme used by the lexer as a bin for words it doesn't recognize. Human users may also have such a bin for unlexed words; but unlike the machine, humans usually guess what lexeme unknown words belong to.

### Lexeme CA: Predicate Word Connectives (Sheks)

These are the connectives used to join individual predicate words in an afterthought, i.e., left-grouped, mode. Morphologically, each shek is the result of inserting **c** before the characteristic vowel of an ek (see Lexeme A). The current list of sheks is

ca	...and/or...
ce	...and...
co	...if and only if...
cu	...whether...
nucu	whether..., ...
noca	...only if...
noce	not...but...
nocu	not...whether...
canoi	...if...
cenoi	...but not...
conoi	...or...but not both
nucunoi	whether..., not...
nocanoi	not both...and...
nocenoi	neither...nor...

### Lexeme CI: The Interverbal Hyphen

CI is monolexic; its one allolex, **ci**, therefore has a unique role in the language. **Ci** is used between pairs of predicate units—usually, individual predicate words—to make larger predicate units inside predicate strings. Such "hyphenated" units then function as single modifiers or modificands in the predicate string.

### Lexeme CUI: The Shek Left-Parenthesis

CUI is also monolexic; its one member, **cui**, is the left-parenthesis inside predicate strings. It is used with sheks (see Lexeme CA) to increase their leftward scope beyond the single predicate word which is the normal left-connectand of a shek. No right -parenthesis is necessary in predicate strings.

### Lexeme DA: Variables

The allolexes of DA are found in six phonemically distinct series: (1) the five replacing, or 3rd person, variables **da de di do du** ('X' 'Y' 'H' 'W' 'Q') linguists call the use of such short expressions to replace longer ones "anaphora", hence these are the anaphoric variables); (2) the four non-designating variables **ba be bo bu** ('x' 'y' 'w' 'q'), the "bound variables" of the logician); (3) the three 1st and 2nd person variables **mi tu mu** ('I/me', 'you', 'we/us', the latter in the sense of 'you and I/me jointly'); (4) the mixed 1st and 3rd person variables **mua mue mui muo muu** (the other sense of 'we/us': 'I/me and X/Y/H/W/Q jointly'); (5) the mixed 2nd and 3rd person variables **tua tue tui tuo tuu** ('you and X/Y/H/W/Q jointly'); and (6) the three pairs of demonstrative variables, namely **ti ta** ('this' vs. 'that' as applied to non-linguistic objects), **toi toa** ('this (most recent) utterance' vs. 'that (earlier) utterance'), and **tiu tau** ('the situation referred to by this (most recent)

## Notebook 3 – Lexicon

utterance' vs. 'the situation referred to by that (earlier) utterance'). The last two pairs of demonstratives do not, so far as we know, have parallels in natural language. But the distinctions they make will be found to be indispensable to the user of a logical language.

### Lexeme DIO: Argument Tags

Argument tags are optional prepositions that can be selectively used before any argument, but are especially useful in three situations: when (1) the argument so tagged is out of its usual order, as may be required during translation, for example, to match some natural word-order; (2) the tagged argument is a **sufori** argument of its predicate (experience has shown that the meanings of the first 3 places of a predicate are easy to remember, but that if a predicate has **sufo** (four or more) places, the meanings of the **sufori** places are difficult to remember); or (3) the auditor is suspected of not knowing the place structure of that predicate, as is often the case, for example, when the speaker is a teacher and the auditor is a learner.

Loglan argument tags are of two types: (a) the 11 optional case tags, and (b) the five ordinal tags once called "HB-tags" [for "Hixson-Bonewits", the nom-de-plume of the loglanist who first proposed them]. The two systems are discrete; they will seldom if ever be used together in the same utterance.

The 11 optional case tags, together with their source words and meanings, are as follows:

Dio	(dirco)	'to/for'	Destinations or Beneficiaries
Duo	(durzo)	'by'	Actors, Agents, Doers
Foa	(folma)	'in/of'	Wholes, Sets, Collectivities
Fua	(fruta)	'-'	Outputs, Purposes, Products
Koe	(konte)	'by/for'	Quantities, Amounts, Values
Mau	(cmalo)	'than'	Lessers in greater/lesser than relations
Neu	(nerbi)	'under'	Conditions, Fields, Circumstances
Piu	(pisku)	'in/on'	Parts, Passives, Properties
Roa	(groda)	'than'	Greater in greater/lesser than relations
Sau	(satci)	'from'	Sources, Origins, Reasons, Causes
Veü	(vetci)	'by/via'	Deeds, Means, Routes, Events, States

Studies have shown that these eleven cases accommodate all the places of the 800-odd primitive predicates; see List 4, Cases.

Morphologically, each case tag is derived from a Loglan primitive predicate meant to be of mnemonic value in learning its meaning, e.g., **groda** for "greater". Tags are derived by a procedure similar to that used to derive CVV-form affixes from their primitives, but with this difference: if the result of applying the affix-deriving procedure produces a letter word (all forms ending in **-ei -ai -oi** or **-ao** are letter words; see Lexeme TAI), the letter **-u** must be substituted for the final **-i** or **-o** of that letter word. Four tags on the current list show the **u**-substitution: **man**, **neu**, **sau** and **veü**. All case tag assignments are still provisional. The Institute would welcome suggestions before GPA (The Institute's acronym for 'Going Public Again') for alternative tags that have plainer derivations, or which have been drawn from primitives with more useful mnemonics.

It is believed that the eleven cases that these tags represent constitute a (nearly) complete set in the sense that the places of (nearly) any future Loglan predicate will be capable of being assigned some ordered subset of these 11 cases in such a way that no two places of that predicate will be assigned the same case unless the occupants of those places can be exchanged without altering the truth-value of any claim made with that predicate. This criterion is more complex to read about than it is to use. For example, the first two places of **ciktu** = '...is equal to...in dimension/feature...' are currently assigned the **Foa** Case, the entire string of case assignments being **foa foa piu** (Whole Whole Part). The dual assignment of the **Foa** Case to two places of this predicate reflects the fact that any two occupants of these places in a true sentence may be exchanged without rendering the resulting sentence false. Thus if X is equal to Y on dimension W, then Y is equal to X on that dimension.

The English prepositions in the third column above are meant only to suggest how these tags may occasionally be translated into English. They are not definitions of the case tags.

## Notebook 3 – Lexicon

The second system of argument-tagging—an historically earlier one, in fact—uses five ordinal tags to identify nothing more than the number of the tagged argument's normal position in the place structure of its predicate; thus **zua zue zui zuo zuu** mean 'the normally 1st', 'the normally 2nd', '3rd', '4th', or '5th argument of its predicate', respectively. [Since they were proposed and adopted in 1977, the ordinal tags had been assigned the morphemes **pua pui pue puo puu**; they were shifted to their present /zuV/ values in 1987 to accommodate a useful development of the tense system; see PA.]

It is very likely that we will not need both tagging systems. One will almost certainly be found to be enough. We have installed them both in this still-experimental version of the language in order to discover which tagging system will be used preferentially. As usual, we will let usage decide.

### Lexeme DJAN: Name Words

These are all the C-final words found by the lexer. Names are used in four places in the grammar: they are formed into strings by R68-9 (**Djan Pol Djonz**); they are made into designations by R80, which precedes them with **la** (**La Djan Pol Djonz, pa pasko merki mursi ci kapta** = 'John Paul Jones was an early American sea-captain'); they are made into vocatives by R177, which puts **Hoi** in front of them (**Hoi Djan Pol** = 'O John Paul!'); and they are used as unmarked vocatives by R179 (**Pol, gotsi** = 'Paul, go!').

### \*Lexeme END

This machine-oriented lexeme has exactly one allolex, the period or full stop [.]. This is a special, and probably temporary, concatenation sign used by the machine to prepare a specimen composed of 2 or more utterances for sequential parsing.

### \*Lexeme FI: The Utterance Ordinal Suffix

This monolexic lexeme is used only by the preparser to recognize the NI/TAI + FI compounds. These are the "utterance ordinals" which are members of Lexeme UI, for example, **Nefi** and **Rafi** ('Firstly' and 'Finally'). The suffix **-fi** is incidentally also used to form the lower-case Greek vowel letter-words **afi efi ifi**, etc.; but the recognition of letter-words is handled by the lexer directly and does not require **-fi** as a separate lexeme; see TAI for other letter-words.

### Lexeme GE: The Grouping Operator

**Ge** is the only member of its lexeme; it serves to make right-extending groups within predicate strings. Like all strings of similar elements in Loglan, predicate strings group left whenever they are unmarked: (i) **((preda preda) preda) preda**. **Ge** functions as a left-parenthesis, sometimes matched with an optional right-parenthesis **gue**, and interrupts that normal pattern by turning all that follows **ge**, or all that lies between **ge** and **gue** in case there is a **gue**, into a single modificand; see Lexeme GUE. Thus (ii) **(preda ge ((preda preda) preda))** is a **ge**-marked string. Notice that the normal right-grouping pattern resumes in the portion of the string that follows **ge**. **Ge** is often well-translated by the English phrase 'for a'; for example, 'That's big for a little woman's suitcase' = **Ta groda ge cmalo fumna racbao**. The predicate in this utterance has the same structure as (ii).

### Lexeme GI: The Fronting Operator



## Notebook 3 – Lexicon

GI has two allolexes, **gi** and **goi**. GE allows the "fronting" of strings of possibly arguments and/or modifiers, that is to say, it permits the speaker to move all or some of these items to the front of the sentence from their Loglan-normal position at the rear. **Gi** thus permits Loglan to be spoken in the O-S-V (Object-Subject-Verb) word-order by appearing between Object and Subject (or between Object and Verb, in the case of imperatives). There is usually no English word or phrase corresponding to **gi** in the O-S-V sentences which it helps in this way to translate. For example, in the sentence 'In May to Europe, John travels' - **Na la Femen, dio la Europas, gi la Djan, traci** Loglan **gi** corresponds to nothing in the English sentence except the comma. **Gi** is in this sense one of Loglan's "spoken punctuation marks". This same sentence in Loglan-normal order is **La Djan, traci la Europas, na la Femen**. Note that the case tag (preposition) **dio** (see Lexeme DIO) is left out in this minimally marked normal word order.

It is not necessary to front the entire string of normally trailing arguments and/or modifiers when using **gi**; only some of them need be shifted. But if only one is shifted, it must be marked by **gi**. Thus **Na la Femen, gi la Djan, traci la Europas** - 'In May, John travels to Europe' also requires **gi**, and corresponds to a much more common English word order.

GI's second allolex, **goi**, permits the objects called "prenex quantifiers" in the argot of logicians to be identified as such. Prenex, or sentence, quantifiers are the strings of quantified variables that precede a sentence in which these same variables appear in unquantified form; e.g., 'For every x there is a y such that y is faster than x' = **Baba be goi be kukra ba**. Thus **goi** may often be translated by English 'such that'.

### Lexeme GO: The Inversion Operator

GO is monolexic; it is the predicate inversion operator. **Go** is used in a predicate string to exchange the positions of the right and left segments formed by some partition of that string. The segmentation of the original string must coincide with some modifier-modificand joint in it. Thus (i) **Da mutce groda mrenu** may be segmented between **groda** and **mrenu**, and inverted as (ii) **Da mrenu go mutce groda** = 'X is a man who is very big'; but (i) may not be segmented between **mutce** and **groda** and inverted as (iii) **Da groda mrenu go mutce** = '\*X is big man who is very', (iii) is a legitimate inversion, but of another Loglan sentence: (iv) **Da mutce ge groda mrenu** = 'X is extreme for a big man'. Thus inversion with **go** sometimes renders some of the punctuation of the original sentence redundant, in this case **ge**, and so omissible. Thus (v) **\*Da ge groda mrenu go mutce** means the same thing as (iii)—and as (iv), for that matter—and parses as essentially the same structure; but the redundant **ge** is not considered good usage, and so I have \*-ed the utterance in which it appears.

### Lexeme GU: The Optional Right Boundary Marker ("Comma")

GU is also monolexic and is the optional comma-word. It is used to mark the right boundaries of structures whose boundaries would not be clear without it. **Gu** is never used redundantly, however, i.e., when the boundary with which it coincides is already clear. **Gu** is sometimes, but not always, replaceable by a literal comma [,] in text or by a pause in speech. One must know the particular grammar rules involved in this and other uses of **gu** to know whether such substitutions are admissible. In general, the cases in which **gu** may be replaced by pause/comma far outnumber the few cases in which it may not.

### Lexeme GUE: The GE-Group Optional Terminator

GUE is monolexic and, like GU, optional: that is, **gue** may always be omitted when it isn't needed. **Gue** has only one function: to terminate a **ge**-group when that group does not run to the end of the predicate string. Thus, **gue** is always matchable with a **ge** but not all **ge**'s have matching **gue**'s; see Lexeme GE.

### Lexeme HOI: The Vocative Marker

Also monolexic, HOI is the left or leading mark of some vocative expressions. If the vocative is a name word (Lexeme DJAN) or a string of name words (**Djan Pol Djonz**), leading **Hoi** is optional unless it is needed to separate a vocatively used name from a preceding name...perhaps one used designatively: **Donsu ti la Djan Pol Djonz, Hoi Pit** = 'Give this to John Paul Jones, O Pete!' Here **Hoi**—in which the **H** is always capitalized in text—serves, like English 'O', to separate the vocative name from the designative one. Without it, **Donsu ti la Djan Pol Djonz Pit** says 'Give this to John Paul Jones Pete', which designates a person with a tetradic name. So **Hoi Pit** is obligatory if Pete is to be successfully instructed to be the donor. **Hoi** is optional in front of names which need no separation from the prequel: **Donsu ti le ditca, Pit** = 'Give this to the teacher, Pete'; for in such contexts it is clear where the vocative name begins.

A second use of **Hoi** is to mark descriptions used vocatively, such as **Nenkaa Hoi Ganbra** = 'Come in, O Noble One!' Here the **Hoi** is obligatory. Without it the invitation becomes **Nenkaa ganbra** = 'Be in-comingly noble!' In text, the initial letters of all the predicate words in vocative descriptions are capitalized.

### Lexeme HU: The Interrogative Argument

HU is monolexic; its one member, **hu**, is the interrogative argument. That is to say, **hu** makes possible the kinds of questions to which arguments are answers. **Hu** may be spoken alone or take the position of some argument in an utterance. For example, **Hu hijra** asks 'Who is present?' Just **Mi**, or **La Djan**, or **Mi hijra** are all possible answers...some, admittedly, more useful than others.

HU is also used by the compound-lexer to identify **Nahu**-type compounds. The most common of these are **Nahu** = 'When?', **Vihu** = 'Where?' and the four kinds of Loglan 'Why?': **Kouhu** = 'Because of what cause?', **Moihu** = 'Because of what motive, or with what intent?', **Rauhu** = 'Because of what reason, or with what justification?' and **Soahu** = 'From what premise(s)?' or 'With what logical justification?'. All these PA + HU compounds are members of the Lexeme UI; see UI for the complete list. If it weren't for this use of HU by the parser, **hu** would be just another allolex of Lexeme DA; q.v.

### Lexeme I: Sentence Connectives (Eesheks)

**I** is the principal sentence connective, but there are many others. Almost any compound which has **I**- as its leading element (except **ICI** and **IGE** words; see below) is an eeshek. Thus, there are **I** itself and the 14 eesheks formed by preceding a shek with **I**-:

<b>I</b>	... . . .
<b>Ica</b>	...; or..., and possibly both
<b>Ice</b>	...; and...
<b>Ico</b>	...if and only if...
<b>Icu</b>	...; whether or not...
<b>Inucu</b>	whether...; ...
<b>Inoca</b>	...only if...
<b>Inoce</b>	not...but...
<b>Inocu</b>	not...whether...
<b>Icanoi</b>	...if...
<b>Icenoi</b>	...; but not...
<b>Iconoi</b>	...; or..., but not both
<b>Inucunoi</b>	whether..., not...
<b>Inocanoi</b>	not both...; and...
<b>Inocenoi</b>	neither ...nor...

## Notebook 3 – Lexicon

In addition, any of these 15 eesheks may be compounded with a following PA word, just like any ek may be; and these, too, are members of the I-Lexeme. But the eeshek that is usually used in such compounds is **I** itself. If the PA word is a temporal, the result is a tensed connective (e.g., **Ifa** = 'And later...'); if the PA is a spatial, the result is a located connective (**Ivu** = 'And far away...'); and if the PA word is a causal operator—these are a well-defined subset of the PA-Lexeme—the result is one of the afterthought causal connectives. These are important for translating from, or to minds trained within, the Indo-European languages, so I give the complete set of 16 afterthought causal connectives here. Grammatically, the important thing to remember is that each of these I + PA words is a member of the I-Lexeme and treated just like any other eeshek:

Ikou	Effect	E because of cause C.
Inukou		C; therefore E.
Inokou		E although C.
Inunokou		C; nevertheless E.
Imoi	Act	A because of motive M.
Inumoi		M; therefore A.
Inomoi		A although M.
Inunomoi		M; nevertheless A.
Irau	Decision	D because of reason R.
Inurau		R; therefore D.
Inorau		D although R.
Inunorau		R; nevertheless D.
Isoa	Conclusion	C because of premise(s) P.
Inusoa		P; therefore C.
Inosoa		C although P.
Inunosoa		P; nevertheless C.

Thus, the category of sentence connectives in Loglan is extremely varied; for example, **Inocenoipacenoina** (**Inoceno** + **pacenoina**) is one of its potential members...although it is mind-boggling to even think of what it might mean. Clearly this is a domain of logical possibilities that far exceeds the naturally sayable. Loglanists are invited to explore this vast semantic domain and make what use of it they can or will.

### Lexeme ICI: Hyphenating Eesheks

These **Ici** and **Icaci**-type words are recognized by the compound-lexer. They are eesheks in all their wonderful variety but adorned with a trailing **-ci** that makes them hyphenating. The hyphenating eesheks are a special variety of sentence connectives which have the same effect on a string of connected sentences as ACI-words have on a string of ekked arguments or predicates, namely, they "hyphenate" or close-bind a pair utterances in a string of connected utterances, making a single connectand out of them.

### Lexeme IE: The Identity Interrogative

**Ie**, the single allolex of this lexeme, is the interrogative particle by which identity questions are raised. Given any sort of designation—a name, for example, or a description, or a variable—**ie** may be prefixed to it; and this converts the argument, or actually the sentence in which the argument is embedded, into a question which asks about the identity of the designatum of that argument. Thus, **Ie la Djan** means 'Which John?', **Ie le mrengu jia pa kamla lepo foatci** = 'Which man that came to dinner (i.e., to the "formal eating")?', or **Ie tu pa godzi go trena** = 'Which (of) you went by train (took the train)?' Thus **ie** has roughly the sense of English 'which...' or 'which of...' as applied to some set of possible identities.

### Lexeme IGE: Right-Grouping Eesheks

This variety of sentence connectives is formed by attaching to any eeshek a trailing **-ge**; thus **Ige** and **Icage** are right-grouping eesheks, and are recognized as such by the compound-lexer. IGE is a special variety of **I**-connective which has the same effect as AGE has among connected predicates or arguments. In a string of connected utterances, it groups all that follow it into a single connectand.

### Lexeme JE: The First Linking Operator

**Je**, which is grammatically unique, is the first linking operator. It is used to attach 2nd arguments to predicate words: thus, **Le farfu je le botci** = 'The father of the boy' and **Ta kukra je lo litla, grobou** = 'That's a faster-than-light ship'. See JUE for links to *suteri* arguments.

### Lexeme JI: Argument Modification Links

At present there are five of these operators: **ji ja jii jie** and **pe**. All of them attach either modifiers or other arguments to arguments. In this way they accomplish "local modification" as opposed to the sentence-wide or "adverbial" role that unattached modifiers perform.

**Ji** is the identifying link (**Le mrenu ji vi le hasfa** = 'The man in the house'); **ja** is the predicating link, i.e., it gives incidental information about the designatum (**Le mrenu ja le ditea** = 'The man, who is incidentally the teacher'); **jii** is the identifying membership link (**Le mrenu jii le merka** = 'The man who is one of the Americans'); **jie** is the predicating membership link (**Le mrenu jie le brudi** = 'The man, who is incidentally one of the brothers'); and finally **pe** which is the postfixing genitive operator (**Le bukcu pe le ditca** = 'The book of the teacher'). **Pe** provides an alternative to the prefixed genitive (**Le le ditca, bukcu** = 'The teacher's book') just as in English, which is also a language in which both forms of the possessive exist.

JI is an open lexeme; other allolexes may be added from time to time.

### Lexeme JIO: Subordinate Clause Links

There are currently two of these words, **jia** and **jio**; both are used to attach subordinate clauses to arguments. The two JIO words differ grammatically from JI words in that the operands of JIO words are sentences— although such "sentences" may be single predicate words, i.e., "imperatives"— while the operands of JI are either arguments or modifiers. Thus, **Da jio prano** means 'The X who runs'. As in the case of JI words, the distinction between the identifying and predicating senses of subordinate clauses is scrupulously maintained. Thus **Jio**-clauses identify; they correspond to the restrictive clauses which are usually not set off by commas in English text: **La Djan, jio prano ga blonda** = 'The John who runs is blond'. In contrast, **jia**-clauses predicate; they correspond to the non-restrictive clauses which are usually set off by commas in thoughtfully composed English text: **La Djan, jia prano ga blonda** = 'John, who incidentally runs, is blonde.' Notice how, with restrictive clauses, the translation into English actually assigns a word to Loglan **la**: 'The John who runs is blonde'. With non-restrictive clauses made with **jia**, Loglan **la** has, as usual, no translation: 'John, who incidentally runs, is blonde.' The English word 'incidentally' is one of the few unequivocal ways of showing that an English clause is non-restrictive. In Loglan the linking word itself carries that meaning.

### Lexeme JO: Metaphorizers

## Notebook 3 – Lexicon

**Jo** and **kin** are the metalinguistical operators by which a loglanist, if speaking or writing carefully, can call attention to whatever portion of da's speech is non-literal. **Jo** itself signifies that the preceding word was used metaphorically. **Tojo** means that the two preceding words were used metaphorically; **tejo**, that the three preceding words were so used, and so on. **Raja** anywhere in an utterance means that the utterance as a whole is not to be taken literally.

### Lexeme JUE: The Second Linking Operator

**Jue**, the sole member of its lexeme, is used in conjunction with **je** to link 3rd and subsequent arguments to predicate words; see JE. For example, if one wanted to designate a ship by alluding to the fact that it was not only faster than light (**kukra je lo litla**) but faster than light by 20 kilometers a second, one could say **Le kukra je lo litla jue lio tonikeimeikuasei, grobou** ('The faster-than-light by twenty-kilometers-per-second ship'). Somewhat less fancifully, **Le farfu je to botci jue la Meris** designates the father of two boys by (or through) Mary.

It is grammatical but bad usage to link sutori arguments to the predicate of a sentence with **je** and **jue**: **\*Da farfu je to botci jue la Meris**. Equally understandable and better usage is the simpler preposition-free form **Da pa farfu to botci la Meris** = 'X fathered two boys through Mary'.

### Lexeme KA: Prefix Members of Forethought Connectives (Keks)

Keks are separated pairs of connective words, like English 'Either...or...', which are used to make forethought connections between a wide variety of elements. Connections of this kind are said to be "forethought" in that the speaker must decide what kind of connection da is going to use before mentioning the elements which are to be connected.

KA and the next lexeme, KI, are two of the most widely distributed lexemes in Loglan grammar. They are used with M2 in R13 and R19 to kek links and linkargs; with M3 in R37 and R150 to kek predicates; and without any advance marking at all to kek modifiers in R63, arguments in R106, termsets in R125 and sentences in R160.

Each kek is composed of a prefix member chosen from the KA Lexeme and an infix member chosen from the KI lexeme, q.v. Thus **Ka da ki de farfu** is a sentence with a kekked 1st argument and means 'Either X or Y, and possibly both, are fathers'. Like the English expressions 'Either...or...', 'If...then...', 'Both...and...' and 'Neither...nor...', the earliest or "prefix" member of each pair—in this case **Ka**—announces that a connection is about to be made. **Ka** also specifies the kind of connection it is going to be, namely a logical alternation. The prefix element is then followed by the left-connectand, which is in turn followed by the infix member of the connective pair—in this case **ki**—and the connection is then completed by the right connectand. Thus it takes a minimum of 4 elements to make a kekked connection...5 if the KA element must be marked by either M2 or M3 for the machine.

There are two main series of keks. The first are the forethought logical connections like English 'Either...or...'; the second are the forethought causal connectives like English 'Because..., (therefore)...'. The logical series is semantically more fundamental and will be described first.

**The Logical Keks:** Morphologically, the prefix portion of a logical kek is formed by inserting a /k/ before the characteristic vowel of an ek; see the A Lexeme. The infix portion is then either **ki** or **kinoi** depending on whether the corresponding ek does not or does end with **-noi**. To show this relationship between eks and logical keks, here is the complete list of logical keks shown alongside eks of the same meaning:

a	ka...ki...	either...or...and possibly both
e	ke...ki...	both...and...

## Notebook 3 – Lexicon

o	ko...ki...	if and only if...then...
u	Nuku...la...	#...whether...
nuu	ku...ki...	whether..., ...
noa	kanoi...ki...	if ...then...
noe	kenoi...ki...	both not...and...
nou	nukunoi...ki...	#not...whether...
anoi	ka...kinoi...	#...if...
enoi	ke...kinoi...	both...and not...
onoi	ko...kinoi...	either...or...but not both
nuunoi	ku...kinoi...	whether..., not...
noanoi	kanoi...kinoi...	either not...or not...and possibly neither
noenoi	kenoi...kinoi...	neither ...nor...

Notice that the transformation of ek into kek is irregular in the case of the four independence connections: **u nuu nou nuunoi**. Also, I have marked with '#' as possibly misleading the English translations of a different four Loglan connections which have no forethought renderings in English.

**The Causal Keks:** Morphologically, the prefix element of a causal kek may be formed by removing the leading /i/ from a causal eeshek (see Lexeme I) and adding a final /ki/: thus **kouki** is obtained from **Ikou**, an I-word. A more direct way of describing these compound members of KA is to say that they are PA + KI when the PA word is a causal. (Temporal and spatial keks have not yet been defined; but they are certainly possible.) Here is the list of the causal keks paired with eesheks of the same meaning. Notice, however, that the eeshek used to derive the kek is the one that precedes that kind of causal element (i.e., cause or effect) in an afterthought construction. In the following lists, E = Effect where C = its Cause; A = Action where M = its Motive; D = Decision where R = the justifying Reason; and C = Conclusion where P = the supporting Premise.

E Ikou C	Nukouki E ki C	*Therefore E because of C.
C Inukou E	Kouki C ki E	Because of C, E.
E Inokou C	Nunokouki E ki C	*Nevertheless E although C.
C Inunokou E	Nokouki C ki E	Although C, E.

A Imoi M	Numoiki A ki M	*Therefore A in order to M.
M Inumoi A	Moiki M ki A	In order to M, A.
A Inomoi M	Nunomoiki A ki M	*Nevertheless A although M.
M Inunomoi A	Nomoiki M ki A	Although M, A.

D Irau R	Nurauki D ki R	*Therefore D because of R.
R Inurau D	Rauki R ki D	Because of R, D.
D Inorau R	Nunorauki D ki R	*Nevertheless D although R.
R Inunorau D	Norauki R ki D	Although R, D.

C Isoa P	Nusoaki C ki P	*Therefore C given P.
P Inusoa C	Soaki P ki C	Given P, C.
C Inosoa P	Nunosoaki C ki P	*Nevertheless C although P.
P Inunosoa P	Nosoaki P ki C	Although P, C.

The \*-ed entries are those for which no grammatical English translation seems possible.

Again we notice that in these logically sophisticated regions of the vocabulary, the fineness of grain of the Loglan lexicon far exceeds that found in any natural language. It is likely that Loglan speakers trained—as, for some time, all will be—in the coarser distinctions of the natural languages will find most of this domain of forethought causal connectives too fastidious to be usable. But it is also possible that the

## Notebook 3 – Lexicon

availability of these extremely precise logico-causal distinctions will lead some loglanists down linguistic pathways which will eventually take them to some powerful reflections not easily formulated in the natural languages, and that others will then follow in their linguistic footsteps in pursuit of these same or equally powerful new insights.

### Lexeme KI: Infixes for Forethought Connectives (Keks)

There are just two of these, **ki** and **kinoi**; and they are used as infixes with the prefix elements of the forethought connectives described in Lexeme KA.

### Lexeme KIE: The Left-Parenthesis

KIE is monolexic, its sole member **kie** being the left or leading parenthesis. **Kie** is often represented in text by the punctuation mark [(], but is always pronounced [kyeh] when read aloud. Used with **kiu** below.

### Lexeme KIU: The Right-Parenthesis

KIU is also monolexic, its one member **kiu** being often represented in text by the mark [)] but pronounced [kyoo] when read aloud. Used with **kie** above.

LA has two allollexes, **la** and **laa**, the latter being a special operator for Linnaean names. **La** has the sole function of generating designations based on ordinary, non-Linnaean names. **La** may be used to precede either a sequence of one or more name-words (**La Djan Pol Djonz**) or a string of one or more predicate units (**La Redro Nu Herfa** = 'The Red-Haired One'). The initial letters of name-words are always capitalized; but the words in **la**-marked predicate strings are also given capital initials in text.

Names are seldom unique; but they are always used as unique designations in the contexts in which they appear. Thus, there will normally be only one person answering to the call 'John Paul Jones!' in the situation in which **La Djan Pol Djonz** is used, and only one red-haired person in the context in which **La Redro Nu Herfa** is successfully used.

To use names vocatively, the vocative operator HOI (q.v.) is used in place of **la** in either type of expression.

**Laa** is used only with Linnaean polynomials; see Sec. 2.13.

### Lexeme LAE: Indirect Designation Operators

LAE has two members, **lae** and **sae**. **Lae** is used to designate something by operating on a sign or address of that something. Generally, **lae** precedes a designation of some linguistic entity, i.e., a quotation, although it need not; any object that can serve as a sign of some other object will do. But if it precedes a quotation, **lae** enables us to designate the referent-of-the-referent of that quotation. For example, (i) **Liegai, War and Peace, gai** designates the English expression 'War and Peace'. That English expression, however, is often a sign of—a kind of label or address of—a certain volume in many English-speaking persons' libraries, namely an English translation of that Russian novel known in English as 'War and Peace'. If **lae** were used to precede (i), the resulting Loglan expression (ii) **Laeliegai, War and Peace, gai**, would no longer designate this English expression, but the volume in my library, say, which has this English title

## Notebook 3 – Lexicon

printed on its back. So when I ask you (iii) **Eo kambei mi laeligai, War and Peace, gai**, what I am asking you to bring me is not that English expression, but the volume whose address it is. Similarly, if I don't know the name of the person who lives at 123 Main street, but I would like you to take this copy of "War and Peace" to him or her (and we were speaking Loglan) I could say (iv) **Eo kambei laeligai, 123 Main Street, gai laeligai, War and Peace, gai**, which is equivalent to saying in exceedingly careful English 'Please bring to the person whose address is '123 Main Street' the object with the title 'War and Peace'.'

**Sae** allows us to perform the inverse of this indirect addressing manoeuvre. By prepositioning **sae** before any argument, we may use the resulting expression to designate the sign or signs of which the designatum of that argument is the referent. For example, **Sae levi bukcu** might be used to designate the title or titles of "this book" that I am holding in my hand, say in whatever languages it has been translated, or to designate its "addresses" conceived in some other way, say some particular library's shelving code for it. Thus **sae** undoes what **lae** does; and so both **sae lae** and **lae sae** do nothing.

### Lexeme LE: Descriptors

Descriptors are words that make designations out of predicate expressions. LE is an open lexeme presently composed of a group of 1-initial words, all of which have a 'the'-like quality. In the following lists, **preda** stands for any predicate expression used descriptively...a "descriptive predicate", in the terminology of the grammar. Here is the entire current list of simple LE-words as they might operate on any **preda**:-

<b>Le preda</b>	The one thing I mean, or each of the set of things I mean, which I believe appears or appear to you to be a preda, or to be predas. E.g., <b>Le mrenu pa fumna</b> = 'The man was a woman/the men were women'. (Explicit plurals are managed with quantifiers; see NI.)
<b>Lo preda</b>	The mass individual composed of all the instances of preda there are. E.g., <b>Lo cutri ga djipo lo clivi</b> = 'Water is important to life.'
<b>lea preda</b>	Each of the set of all things which are predas. E.g., <b>Lea humni ga razdou</b> = 'All humans are rational (give reasons).'
<b>loe preda</b>	The characteristic or normal individual which best exemplifies the predas in the present context. E.g., <b>Loe panzi ga fotli loe humni</b> = 'The typical chimpanzee is stronger than the typical human.'
<b>lue preda</b>	The particular set of predas, or apparent predas, which I have in mind. E.g., <b>Lue monca gorla ga numcmalo</b> = 'Mountain gorillas are few in number.'
<b>lua preda</b>	The set composed of all the predas there are. <b>Lua ficli ga mutce laldo</b> = 'Fishes are very old.'

Compound LE-words may be formed by postfixing to one of these simple LE-words either (a) a DA-word (any variable) or (b) a TAI-word (any letter variable), and/or (c) one or more PA-words. Usually the PA-words are spatial or temporal. The possessive adjectives **lemi** ('my'), **lemu** ('our'), **leda** ('X's') and **leTai** ('T's') are formed in this manner. So also are the demonstrative descriptors **levi** and **leva** ('this...' and 'that...') and the tensed descriptors **lefa**, **lena** and **lepa** ('the future...', 'the present...' and 'the former...') as in **Lefa bragai** = 'The future king'. Some of these LE + PA words are very long, e.g., **Lepacenoincefa bragai je la Frans** = 'The-once-and-future king of France'; but all of them can be plainly deciphered. Thus, **le + pa + ce + noi + na + ce + fa** means nothing more mysterious than 'the-before-and -not-now-and-after (king of France)'. All these words behave grammatically just like **le**.



### Lexeme LEPO: Event Operators

These are the LE + PO and LE + ZO (see PO and ZO) compounds that make arguments out of sentences. They are recognized as LEPO words by the preparser. LEPO words are known as event operators, although they designate properties (for example, with **lepu**) and quantities (with **lezo**) as well. Their grammar is quite different from that of LE-words, which they resemble morphologically, in that LEPO-words but not LE-words take whole sentences as operands. Sometimes it looks as if the operand of an event operator is nothing more than a predicate expression—e.g., **lepo sucmi** = 'the swim'—but this is because the operand is here the simplest of all sentences, a one-word imperative. The event of swimming could be fully specified: **lepo da sucmi de di** = 'X's swim to Y from W'. So unlike the operands of ordinary descriptors, the operands of event operators may be embellished without limit: **Lepo lemi brudi ji la Djan, pa kamla lepo fomitci, e stolo mu ne nirne** = 'The time my brother John came to dinner and stayed with us for a year.'

Notice that the events, properties or quantities described with event operators may be particular (**lezo da pa sucmi** = 'the amount X swam') or general (**lopo sucmi** = 'swimming'), or general but constrained in various ways: e.g., **lovipo sucmi** = 'The here-swimming (the swimming here)' and **lenapo penso** = 'The now-thinking (the thinking now)'.

Loglan event descriptions are philosophically of some interest. Their designata are the objects of thought, for example; and so event descriptions accomplish what in the Indo-European languages is accomplished by so-called "indirect discourse". Thus, 'He thought that it was going to rain' is translated into Loglan with an event description: **Da pa jupni lepo ti fa crina** = 'X opined the event of this (place) being later rained-on.'

### Lexeme LI: The Left Quotation Operator

LI is monolexic and is the opening mark for ordinary, or "weak" quotation. **Li** is used with the close-quote **lu** to quote grammatical Loglan, usually in strings of two or more words; see LIE and LIU for other varieties of quotation.

### Lexeme LIE: The Strong Quotation Operator

LIE is also monolexic; it is the leading mark of "strong" quotation. That is, it is the mark by which non-Loglan strings or even strings of malformed Loglan may be unequivocally quoted. **Lie** is used with two instances of any arbitrarily chosen word X which are placed at each end of the string to be quoted. The boundary marker X is usually a letter-word (see TAI) and each instance of it must be separated from the quoted string by a pause: **lieX, <quoted string>, X**. /X/ must not of course appear adventitiously within the string being quoted.

**Lie**-quotation is most often used to quote foreign language strings. When it is, the upper-case Loglan letter-word for the Latin letter which is initial in the Loglan predicate for that particular foreign language is used conventionally as the boundary marker X: **LieSai, No hablo Español, Sai**. **Sai** is the sign of **spana**, the predicate meaning 'is Spanish/an element/expression of the Spanish language'.

### Lexeme LIO: The Number Designator

LIO is the monolexic number designator. **Lio** is used before mathematical expressions to designate numbers. For example, **Lio ne natra numcu** = 'The number one is a natural number' and **Ti langa ta lio topifemei** = 'This is longer than that by 2.5 meters.'

## Notebook 3 – Lexicon

### Lexeme LIU: The Single-Word Quotation Operator

LIU is used to quote single words. It is monolexic and used only before single, well-formed Loglan words to designate them. For example, **Liu liu logla purda** = 'The word 'liu' is a Loglan word'. See LI and LIE for other varieties of quotation.

### Lexeme LU: The Right Quotation Operator

LU is monolexic and is the closing mark for ordinary, or "weak" quotation. **Lu** is used with the open-quote **li** to quote grammatical Loglan, usually in strings of two or more words; see LIE and LIU for other varieties of quotation.

### Lexemes M1\* through M11\*

M1 is a machine lexeme inserted by the preparser before any instance of A that is connecting linked arguments.

M2 is a machine lexeme inserted by the preparser before any instance of KA that is connecting linked arguments.

M3 is a machine lexeme inserted by the preparser before any instance of KA that is connecting predicates.

M4 is a machine lexeme inserted by the preparser before any NO that it is negating a modifier.

M5 is a machine lexeme inserted by the preparser before any A that is connecting sentence modifiers.

M6 is a machine lexeme inserted by the preparser before any A that is connecting argument modifiers.

M7 is a machine lexeme inserted by the preparser before any PA that is inflecting a predicate.

M8 is a machine lexeme inserted by the preparser before any NO that is negating a marked predicate, as shown by PO, ZO or M7.

M9 is a machine lexeme inserted by the preparser before any ACI that is connecting predicates.

M10 is a machine lexeme inserted by the preparser before any A that is connecting predicates.

M11 is a machine lexeme inserted by the preparser before any AGE that is connecting predicates.

### Lexeme ME: The Predifying Operator

ME is the monolexic operator that turns any designation into a predicate ("predifies" it). Prefixed to the first word in that designation, **me** creates a predicate with a meaning, often vague, that can be associated with that designation. Some examples: **Ta metu cadre** = 'That's a you-ish dress'; **Plizo le mela Kraislr** = 'Use the Chrysler'; **Da mela Nuiork** = 'X is a New Yorker'; **De mele sitci** = 'Y is of the city (citified? cityish?)'.

Such usages create "false compounds", i.e., forms like **metu**, **mela** and **mele** which are recognized as not being members of single lexemes by the preparser, as true compounds always are, and so are decomposed. Thus, **metu** is lexed as the two-lexeme string ME DA before it goes to the parser; see DA.

### Lexeme NI: Quantifiers

The quantifier lexeme NI has in principle a limitless number of allolexes, since it contains not only the digits **ni ne to te fo fe so se vo ve** ('none' 'one' 'two' 'three' 'four' 'five' 'six' 'seven' 'eight' 'nine'), but all the

## Notebook 3 – Lexicon

compounds of whatever length that might ever be made from them. Thus, **netotefofesosevove** as well as **neni** ('ten') is an allolex of NI.

NI also includes **ho**, the interrogative quantifier ('How many?' or 'How many of...?'); thus **Ho da pa kamla** = 'How many of X came?' It includes three of the non-numerical quantifiers as well, namely **re ro ru** ('most' 'many' 'enough'); the other two r-words, **ra ri** ('all' 'few'), having certain compounding duties to perform—they make the cardinal and ordinal compounds that will be classified as PREDA—must occupy a separate lexeme, namely RA.

Other allolexes of NI are **sa si su** ('approximately (all)' 'at most (one)' 'at least (one)'), which may either be prefixed to other NI-words (**savo** = 'approximately eight' **sivo** = 'at most eight' **suvo** = 'at least eight') or used alone with their default values understood (**sa** = 'nearly all', **si** = 'at most one', **su** = 'at least one'). Finally, there is a series of mathematical expression-building operators, not yet complete, consisting of the decimal point **pi** [.] , the double-zero **ma** [00], the triple-zero **mo** [000], the division sign **kua** [/] (read 'over'), the inverse division sign **kue** (not available on this font, but read 'under'), the addition sign **poi** [+], the subtraction sign **nio** [-], the negation sign **niu** (not available), the multiplication sign **tia** [\*], the power sign **sua** [] (read 'to the ...-th'), the left and right parentheses **kie** and **kiu** [( )], which in this context, may often be read as 'the quantity', and for more complex nesting, the left and right brackets **gie** and **giu** [[ ]], also often readable as 'the quantity'. Mathematical expressions formed with these operators—for example, **tokuafo** [2/4] and **tosuanifuo** [2|-4] (read '2 to the power minus 4')—are also NI words.

NI has another group of allolexes which are formed of some numerical expression plus one or more letter-words of the TAI-Lexeme, e.g., **nevepinifomeimei** [19.04mm]. These are the dimensioned numbers; they, too, as well as the TAI words themselves, may be components of mathematical expressions. For example, consider **tobeisuato** [ 2b|2 ] = 'two-bee-squared' vs. **tobeikiusuato** [2b)|2] = 'two-bee-the-quantity-squared'. The "grammar" of such expressions, i.e., the set of compounding rules for NI compounds, has not yet been written.

A final type of NI compound may be formed by prefixing **ie** (see Lexeme IE) to any other NI word, simple or compound, e.g., **ieto** or **ietobeigiusuato**. Such quantifiers have the sense of 'which...of...?', as in **Ieto le mrenu** = 'Which two of the men?'...or, indeed, as in **Ietobeigiusuato le protoni** = 'Which two-bee-the-quantity-squared of the protons?' The internal structure of such compounds may be as complex as the speaker or writer likes; but grammatically, such expressions count as just another quantifier.

There is one usage convention involving NI and TAI words that resolves an ambiguity that cannot be solved by preparsing. This is the apparent ambiguity of the expression **nei preda**. Is (i) **nei** a quantifier, in which case this is an indefinite description meaning 'n predas'? Or is (ii) **nei** a letter-variable, in which case it means 'n is a preda'? Consider that the speaker has already established that the letter word **nei** [n] in some mathematical expression is to represent a certain large unknown number. May da then lift it out of that expression and use **nei** as a quantifier, saying, for example, **\*Nei mrenu pa kamla** (interpretation (i))? Da may not; the parser would declare this expression to be unparsable (and so I have starred it). What is happening here is that **Nei** has been lexed as a TAI word (interpretation (ii), which happens to be the correct one); but so-lexed, the expression will not parse. Standing alone, as it does, **Nei** will not be parsed as a NI word, i.e., a quantifier. Whatever da's intent, da's auditor will hear **Nei** as a letter variable, presumably as a replacement of some longer argument in which 'n' figures; because that is grammatically what **Nei** is. So what da has said would translate into the equally ungrammatical English non-sentence '\*Enn is a man came'. In other words, da would like to use **nei** as a quantifier, but may not because any well-behaved Loglan parser will treat it as an argument.

The way out of this dilemma is to adopt a certain usage convention. We may use the dimensioned number **nenei** [1n] whenever we would like to use the number represented by **nei** [n] as a quantifier. **Nei** is not a quantifier; but **nenei** [1n] is. Moreover, this usage is mathematically correct. The number one is the identity operator in this group; so  $1X = X$  for any X. So it doesn't matter mathematically whether we say **nei** or **nenei**. But it does matter grammatically; for only by using **nenei** will the parser behave in the way that we intend. To English-speakers, this usage has a certain cost. It will be a trifle awkward, at first, to have to say **Nenei mrenu pa kamla** ('One-enn men came') when one feels, following English usage, that one ought to be able to say more economically **\*Nei mrenu pa kamla**. But later on I trust we loglanists will discover the satisfaction of speaking precisely and so, understandably. Besides, by observing this convention we preserve the much more important use of TAI words as anaphoric variables; see TAI below.

### Lexeme NO: The Negation Operator

There is just one negative in Loglan and it is used in all of the ways that the various negatives of the natural languages are used. **No** is very widely distributed in Loglan grammar, occurring in 13 rules. In particular, **no** occurs (i) in R28 where it is used before predicate words and other predicate units (**Da bilti no nirli ckela** = 'X is a beautiful non-girls' school'); this is the close-bound negative that we often express with 'non-' in English. It occurs (ii) in R34 where it is used to negate kekked predicate units (**Da bilti no ke botci, e nirli ckela** = 'X is a beautiful not both boys and girls' school', that is, a school for beautiful boys or beautiful girls or beautiful things that are neither boys nor girls, but not for both beautiful boys and beautiful girls; the scope of this **no** is confined to the kekked pair of predunits **ke botci, e nirli**). **No** is used (iii) in R62 before modifiers (**Titci no na** = 'Eat (but) not now') where the preparser will have marked it with a leading M4. It is used (iv) in R64 where it negates kekked modifiers (**Titci no ke vi, e na** = 'Eat (but) not both here and now'); the preparser will also have marked this **no** with M4. It is used (v) and (vi) in R134 and R143 to negate tensed or otherwise marked predicates (**Da no pa mutce gudbi takna** = 'X was not a very good talker'); this is long-scope **no**, the tense operator having spread the scope of **no** over the entire predicate expression. The preparser will have looked ahead and found the **pa** or other scope-extending mark on the other side of **no**, and then inserted M8 before the **no** to tell the parser that this is a long-scope negative of a predicate it is dealing with. **No** is also used (vii) in R149 to negate identity predicates (**Da no bi de** = 'X is not equal to Y'); and (viii) in R151 to negate kekked predicates (**Da no ke gudbi takna, ki bilti ckela** = 'X is not both a good talker and a beautiful school'; here the scope of **no** covers the entire kekked predicate). It is also used (ix) in R159 to negate statements (**No da bilti ckela** = 'It is not the case that X is a beautiful school'); as suggested, this **no** has the sense of the logician's 'It is not the case that...'. **No** is used (x) in R161 to negate kekked sentences (**No ke da bilti ckela ki de gudbi takna** = 'It is not the case that both X is a beautiful school and Y is a good talker'. It is used (xi) alone in R168 where **No** is a fragmentary utterance, perhaps an answer. **No** is used (xii) in R183, where it is used with a "gap" (a comma or a **gu**) before any utterance, including fragmentary ones (**No, na la Nemen** = 'No; in January'); in such uses **no** often has the sense of being one of several fragmentary answers being made in the same utterance. Finally, **no** occurs (xiii) in R184 where such leading negatives may be preceded by free modifiers (**la no, na la Nemen** = 'Certainly not; in January').

**NO** also occurs in /noV/ compounds (**noa, noe**, etc.) that are some but not all of the A words. In this use it is rather like **NOI** (see next lexeme) which also combines with vowels and other elemental lexemes to produce A and other kinds of connective words.

#### \*Lexeme NOI: The Negation Suffix

This lexeme does not occur as a separate word. It is used only by the preparser to detect some sorts of A, CA, ICA, KA and KI words, namely those that end with the suffix **-noi**. For example, the A words **anoi** and **enoi**, the CA words **canoi** and **cenoi**, the I words **icano** and **iceno**, the KA words **kanoi** and **kenoi**, and the KI word **kino** all require that **NOI** be recognized by the preparser, in particular, by that portion of it that analyzes compounds, in order that they be assigned to their proper lexemes.

#### Lexeme NU: Conversion Operators

The simple allolexes of this lexeme are **nu fu ju**, the three conversion operators which in Loglan effect the passive voice and a great many other "place-shifting" constructions as well. Thus, **nu** exchanges the meanings of the 1st and 2nd places of a predicate (**Da nu bloda de** = 'X is hit by Y'); **fu**, of the 1st and 3rd places (**Da fu ketpi de** = 'X is the departure-point from which ticket Y is valid'); and **ju**, of the 1st and 4th places (**Da ju ketpi de** = 'X is the carrier on which ticket Y is valid'). **NU** also includes compounds like **nufu** and **nufuju**, and the series of numerically marked conversion operators **nute** (= **fu**), **nufu** (= **ju**), **nufe**, **nuso**, etc., an alternative series of conversion operators that accommodate extremely long place-structures. All these words are recognized by the CPD-lexer.

## Notebook 3 – Lexicon

NU words may also be used in incomplete utterances: **Da nu bloda** = 'X is hit', **Da fu ketpi** = 'X is a departure-point of some journey for which tickets are required', and **Da ju ketpi** = 'X is a carrier on which tickets are required.'

### Lexeme PA: Inflectors/Adverbs/Prepositions

PA is a "portmanteau" lexeme in that its allolexes belong to 5 semantically distinct series: tense operators, location operators, modal operators, causal operators, and the predicate marker **ga**. These very different kinds of words are members of the same lexeme not because their meanings are similar but because their grammar, though not their usages, happens to be identical.

As the title of this lexeme suggests, any of the PA words may be used in three grammatical contexts: (i) as an inflector of some predicate (**Da pa titci ta** = 'X ate that'); (ii) as an adverbial modifier of the main predicate expression of the sentence (**Da titci ta pa** = 'X eats that earlier') or, when linked to an argument by a JI word (q.v.), as a local modifier of that immediately preceding argument (**Da titci ta ji pa** = 'X eats that earlier one, i.e., that one that was before'); and (iii) as the prepositional head of a phrase or a clause which modifies the main predicate expression (**Da titci ta pa la Yen** = 'X eats that before nine' or **Da titci ta lia lo horma** = 'X eats that like a horse') or, if linked to an argument, then as a local modifier of that argument (**Da titci ta ji lia lo horma** = 'X eats that (thing) that is like a horse').

If the prepositional meaning of a PA word is known, then its adverbial meaning can generally be inferred by regarding the adverb as an ellipsis of a phrase or clause headed by that same word used as a preposition or conjunction. Thus, we may infer that **Da titci ta vi** = 'X eats that here' is short for **Da titci ta vi ti** = 'X eats that in/at this (place)'; hence that the sentence without **ti** means 'X eats that here'. Similarly, the inflectional meaning of a PA word may usually be inferred by constructing a phrase using that word as a preposition and some designation of the place or time of speech as the object of that preposition. Thus **Da pa titci** may be regarded as shorthand for **Da pa ti titci** - 'X, before this occasion on which we are speaking, eats'; hence 'X ate'.

So much for the three contexts in which PA words may appear. We may now describe the five semantic types of PA words:-

**Tense Operators:** These are **pa na fa**, the simple past, present and future tenses, respectively, and their numerous derivatives and compounds. For example, there are the continuous tenses **pia nia fia** ('was ...-ing' 'is ...-ing' 'will be ...-ing', which yield 'since...' and 'ever since' for the prepositional and adverbial senses of **pia**, respectively; I shall let the reader work out the corresponding senses of **nia** and **fia**. There are the habitual tenses **pua nua fua** ('habitually ...-ed' 'habitually ...s' 'will habitually ...'). There are the intervalized tenses **pazi paza pazu** (the immediate, intermediate and remote past, for example) as well as **fazi faza fazu** and even **nazi naza nazu** with their analogous meanings. There are the nine narrative or compound tenses, the first element of each compound relating the point of speech to the reference point in the narration...often called the "narrative present", the second relating the reference point to the event. The first three of the narrative tenses are the so-called "perfect" tenses of concluded action **papa napa fapa** ('had ...-ed' 'has ...-ed' 'will have ...-ed'); the next three are the "progressive" tenses of predicted action **pafa nafa fafa** ('was going to...' 'is now going to...' 'will be going to...'); the last three are the less often used but equally interesting tenses of "coincident" action—coincident because the reference point and the event coincide in these tenses—**pana nana fana** ('was then ...-ing' 'is just now ...-ing' 'will be ...-ing'). Beyond the narrative tenses are the quantified tenses, of which **rana rona rina suna tona nena nina** ('always' 'frequently' 'usually' 'rarely' 'sometimes' 'twice' 'once' 'never') are just a few of the many quantified possibilities which the loglanist is invited to explore. Various combinations of the above are also possible, e.g., **papia** ('had been ...-ing') and **pazufazi** ('was long ago going to ... immediately'). There are the connected tense words such as **pacenoina** ('no longer' or 'once', literally 'before-and-not-now'). Finally, any PA word may, of course, be incorporated into a descriptor via the LE + PA compounding formula; for example, one of the narrative tenses **pafa** might be so incorporated, as in **Lepafa ditca** = 'The-one-who-was-going-to-be a teacher'; see Lexeme LE. Some of these tensed descriptors are very long, but even the longest so far built are easy to decipher. Take **Lepacenoincefa bragai**. The **pacenoincefa**-part of this descriptor unravels easily as

## Notebook 3 – Lexicon

'before-and-not-now-and-after'; whence the whole expression might be elegantly translated as 'The once-and-future king.')

Even this long list of PA words is not complete. Loglanists are invited to explore their many possibilities.

**Location Operators:** These are **vi va vu**, 'at' or 'in', 'nearby' and 'far away', and their many derivatives, e.g., via **vii viu**, and so on. This list is incomplete.

**Modal Operators:** Unlike the tense and location operators, which are inherently systematic, the modal operators are a disordered set. We can do no better than list the ones in current use; no doubt others will be added.

ciu	(ciktu)	as much as/to the same degree as...
dia	(dilri)	for/on behalf of...
biu	(bivdu)	in manner/mode.../by method...
kae	(kanse)	re/concerning/as for/with regard to...
lia	(clika)	like/as/in the way that...
lui	(pluci)	for/in order to please...
mou	(mordu)	as well as/in addition to...
hea	(helba)	with... 's help/through agent...
sea	(setfa)	instead of/in place of...
tie	(trime)	with..., a tool or means

Notice how the modals differ from the case tags of the DIO Lexeme. In a certain sense, modal phrases may be used to extend the place-structure of nearly any predicate; so they may be thought of as "itinerant cases". But because such phrases can turn up nearly anywhere, they may never be distinctive features of any predicate. Therefore, the modal preposition may never be omitted from its argument, as a case tag may be, relying on the auditor to infer the modality of that argument from its context. So the modality of an argument must always be explicitly announced by some attendant modal preposition.

Loglanists are just beginning to explore the uses of modals as inflectors and adjectives. We expect adventurous speakers to make many interesting discoveries in these uncharted waters. With a little thought, almost any of these strange new usages may be sensibly interpreted. For example, what does **Da durzo de hea** mean? In particular, what is the adverbial sense of **hea**? Well; if you think about it, **hea** used non-inflectionally and without an argument must at the very least mean that X did Y with the help of someone else, that is, "helpedly". So **Da durzo de hea** must mean 'X does Y with help'. What about **hea** as an inflector? Reasoning analogically from, let us say, the sense of **pa** itself in these three positions—as a preposition **pa** means 'before (this designated time)', as an adverb 'before some undesignated time (presumably inferable from context)', and as an inflector 'before this particular time, namely the point of speech'—**Da hea durzo de** must mean that X's doing Y was helped by someone or something current, perhaps a person present at the time of speech. Could we translate it as 'X does Y with your help', you the listener? Extending the currency principle to another case, what does **Da sea durzo de** mean? As a preposition **sea** means 'instead of'. So I would assume that the specimen means that X did Y instead of someone or something else, someone or something which was in some sense present at the time of speech. Could it mean that X does Y instead of your doing Y? I.e., in your place? But note that the speaker, too, is present at the point of speech...but perhaps less interestingly so, since **da** is always present. As I say, loglanists are just beginning to explore this vast new domain of meanings that has been opened up by the machine grammar work.

**Causal Operators:** These are the PA words which, when prefixed with /i/, form the causal connectives which are members of the I-Lexeme; q.v. Here the same sixteen causal relationships are invoked by causal prepositions; and these same sixteen words may, if interpreted by ellipsis, be used adverbially and even as inflectors. The most common use of these elemental causal operators is as prepositions, and that is the sense of the English translations given here. Some of these meanings do not exist in the natural languages, so their Loglan meanings are occasionally hard to think out.

## Notebook 3 – Lexicon

kou C	because of cause C
nukou E	therefore/with effect E
nokou C	despite cause C
nunokou E	nevertheless (unexpected) effect E
moi M	because of motive M
numoi A	so action A
nomoi M	despite motive M
nunomoi A	nevertheless (unexpected) action A
rau R	because of reason R
nurau D	thus decision D
norau R	despite reason R
nunorau D	nevertheless (unjustified) decision D
soa P	because of premise(s) P
nusoa C	thus consequence(s) C
nosoa P	despite premise(s) P
nunosoa C	nevertheless (unentailed) consequence C

When the prepositional sense is plain, the adverbial and inflecting senses are easily inferred.

**Ga:** **Ga** is the fifth kind of PA word, and there is just one of its kind. **Ga** is a boundary-marker; it is used as the left boundary marker of any predicate whose left boundary would otherwise be unclear. **Ga** is therefore one of the "spoken punctuation marks" of Loglan as is suggested by its g-initial form. (The other boundary markers are **ge gi go goa goi gu gue**, q.v.) The left boundary of a predicate happens to coincide with the inflecting position of the PA words; so **ga** has the grammar of PA words, and is therefore lexemically a PA word even though it is semantically quite unlike the rest of them. Moreover, unlike other PA words **ga** is used only in the inflecting position, and then only when the immediately preceding word—exclusive of the "free modifiers" which, for these and other purposes, constitute a kind of "grammatical noise"; see Lexeme UI—is a predicate. Thus **ga** is used only in contexts like **Le sadji ga fumna** ("The wise one is a woman") from which its omission or removal would cause an unintended description to be heard: **Le sadji fumna** = 'The wise woman'.

It is considered bad usage to use **ga** where its removal would not cause this ambiguity. Thus **\*La Selis, ga fumna** and **\*Tu ga fumna**, although grammatically understandable, do not occur in well-formed Loglan speech. Here, then, is another grammatical superset, only some of whose members are actually used.

### Lexeme PAUSE: The Pause-Comma

This is the lexemic element, represented by a pause in speech or a comma in writing, which has survived the preparer's inspection of its context and has been found to be genuinely lexemic, that is, necessary for the parser to be told about. The two written expressions of PAUSE are as a comma [,] and as a pound-sign [#]. Pauses come to the preparer as [,]s. The preparer then removes all instances of [,] which are not lexemic, i.e., which, like the morphemic pauses after name-words or before vowel-initial connectives, do not actually

figure in the grammar rules, and rewrites the others as [#]s. Only [#]s occur in the preparsed strings which are delivered to the parser.

### Lexeme PO: Abstraction Operators

There are exactly two of these, the event-abstractor **po**, and the property abstractor **pu**. There is a third abstractor, **zo**, which abstracts quantities; but because pseudo-instances of **zo** are occasionally generated while making acronymic PREDAs, the preparser, which must identify these kinds of PREDAs, requires that **zo** have its own lexeme; see ZO. This is a consequence of the fact that /z/ is the acronymic hyphen. So **CIO** is **CaiIzO**, which the preparser will read as a compound formed of the lexemes TAI + I + ZO and then classify it as a PREDA. Thus the preparser uses ZO in its lexing operations. If it weren't for this, **zo** would be a member of PO, which it is like in all grammatical respects.

The two PO words and ZO occur in 3 contexts: (1) Close-bound or "short-scope" PO/ZO occurs just before a predicate word or predunit, and in this context the scope of the abstraction is just that single word or predunit; e.g., **Da po sucmi ditca** = 'X is an act-of-swimming teacher', presumably one who teaches that activity. It should not be inferred that abstraction is necessary here. This same claim can be made with the shorter concrete metaphor in **Da sucmi ditca** = 'X is a swim(mer) teacher', and loglanists generally prefer these concrete forms. Where precision is needed, however, the **Da po sucmi ditca** metaphor is available to the loglanist just as the gerund with '-ing' is available to the anglophone...and with perhaps less ambiguity about what is intended. (2) Long-scope PO/ZO occurs when it is followed by either (2a) a pause-comma or a **gu**, in which case the scope of abstraction is the entire ensuing predicate expression, including any arguments, e.g., **Da po, sucmi ditca lo frasa** = 'X is a case of being a swimmer teacher of French' (perhaps one who teaches French while swimming?), or (2b) either an argument or a modifier, in which case no other mark is needed to tell the parser that the PO has long scope; e.g., **Da po de sucmi ditca lo frasa** = 'X is a case of Y's being a swimming teacher of French'. In both these contexts, the operand of PO/ZO is a sentence, even though it may be a sentence without arguments—i.e., an unspecified imperative—in context (2a). (3) PO/ZO also occurs in LEPO compounds (q.v.); in this descriptive context the abstraction is always long-scope; e.g., **Lepo de sucmi ditca lo frasa** = 'The event of Y's swim-teaching French'.

The semantic distinction between **po** and **pu**, between what we call in English "events" and what we call "properties", is perhaps the most difficult to understand in Loglan. The difference between these two operations is clearest in the descriptive context, that is, between the designate of event-descriptions and those of property-descriptions as abstracted, let us say, from the same predicated relationships. Any differences found here should apply, of course, to the other contexts. Let us examine, then, both an event-description and the property-description abstracted from the same relationship.

**Lepo** evidently abstracts a case, state, condition or event of any length—as long as an epoch or as short as a sneeze—from some predicated relationship. **Lepu**, in its turn, may abstract a property or quality from the same relationship. Suppose the predicated relationship is a motherhood between a mammalian mother X, an offspring Y and a father Z. We note first that the event-descriptions abstracted from such relationships are time-bound; like stories, mammalian motherhoods have beginnings, middles, and ends, and can apparently involve substantial segments of the "life-lines" of several or even many individuals. On a common understanding of motherhood, this one will involve those segments of X's and Y's lifetimes from the moment of her conception of Y and either her death or Y's, whichever is earlier. Of Z's lifetime, at least in many mammalian species, only a brief snippet need be considered to be part of X's motherhood, namely that snippet in which his sperm was delivered to her in some fashion. Of course more may be involved; in species in which pair-bonding takes place, more is involved. But this is sufficient to describe the event-state-condition of motherhood in mice, for example. We note that the thing that **lepo** designates about these 3 individuals is laid out in space-time. It has a beginning, a duration and an end. And the relevant segments of each of the three life-lines describe paths through space as well.

The property description of this same relationship, in contrast, is essentially time-free. It is the least set of facts which would establish the truth of that predicated relationship if it were true. That Y developed from an ovum produced by X's ovarian tissue is one of them; that that ovum was fertilized by a spermatozoan produced by Z's testicular tissue is another one. Are there any others? It would seem not. These two properties of these three individuals would, if known, be sufficient to establish X's maternity of Y by Z in any court of law...or, more likely, in any laboratory. And that, it would appear, is what a property is: it is



## Notebook 3 – Lexicon

that lean abstraction from the richness of the world that is just sufficient to enable us to know the truth of some matter.

The event-description is, in contrast, a "fat" concept. The designata of descriptions like **lepo da mamla matma de di** ('The event-state-condition of X's mammalian motherhood of Y by father Z') are often rich and complicated segments of the world; and they have a richer dimensionality than properties do. They have duration, for example, as properties do not. The designatum of **lepo da mamla matma de di** ('The property of X's mammalian maternity of Y by father Z') consists, in contrast, of two very simple but sometimes difficult to establish facts; and those facts are strangely free of time and other complications.

In sum, designata of **po**-abstractions are rich, multi-dimensional objects distinguished by duration; the designata of **pu**-abstractions are least sets of sufficient facts. The designata of **zo**-abstractions are the leanest of all, of course, being simply numbers...and sometimes uninteresting numbers at that. While the amount of heat in this room, and the amount of blue in that painting may well be interesting numbers, **lezo da mamla matma de di** is not. How shall we measure the quantity of mammalian motherhood, as it relates to three specified individuals? Give it one if it obtains between them, zero otherwise? Not a fruitful enterprise. If one or more of the three required participants remains undesignated, however, then the number **lezo da mamla matma**, the amount of mammalian motherhood in which X is involved, for example, might have some interest.

In short, the **zo**-abstraction is the least widely applicable of the three abstractions. **Zo** is useful only with those properties or relations which science has managed to quantify in some way. Thus, until we know how to measure the blueness of a painting or the motherness of an animal, **lezo ta blanu** and **lezo ta matma** will have designata we will not know how to find.

### Lexeme PREDA: Predicate Words

Predicate words, while the most numerous items in any lexicon—they comprise about 80% of the Loglan dictionary at the present time—are called by only two grammar rules, R21-2. PREDA words enter the grammar through this narrow window in the rule group called "predicate units", and as predunits they are then variously elaborated as the noun-like structures of Loglan, the arguments, and the verb-like ones, the predicates. But because they all enter through the same narrow window, any predicate word may end up being either.

Morphologically, PREDA words come in five varieties:-

(1) The CC-bearing, V-final words of the lexicon like **preda** itself. These range from short borrowed words like **iglu** to long complex predicates like **rojmaosenmao** ('agronomist'). All PREDAs of this kind are identified by the lexer. They are by far the largest group of PREDA words that have actually been built so far; they include noun-like predicates like **mrenu** ('is a man'), verb-like ones like **godzi** ('goes from...to...by route...'), adjective-like ones like **corta** ('is shorter than...by amount...'), preposition-like ones like **bitsa** ('is between...and...'), and adverb-like ones like **mutce** ('is extreme in dimension...' which is more frequently used as an adverbial modifier of other predicates, as in **Da mutce sadji** = 'X is very wise'). Loglan makes no grammatical distinction whatever between these various kinds of predicate words...as is attested by their being members of the same lexeme.

(2) The numerical predicates made with the suffixes **-ri** and **-ra** which are members of the RA Lexeme; the exact compounding formula for these kinds of PREDAs is NI/RA (+ ...) + RA. This formula generates the two infinite sets of the ordinal and cardinal numbers, respectively, e.g., **neri** = 'is first in sequence...' and **nera** = 'is a monad', as well as certain irregular numericals like **sutori** ('is second or subsequent in sequence...') which are so much more efficient than the English circumlocutions which they translate that they have already crept into the English speech of some loglanists.

(3) The 6 predicate variables, which are two sorts: **bua bue buo buu**, the non-designating predicate variables used in logic (**Raba rabe rabua goi ko ba bua be ki be nu bua ba** = 'For every x, y, and predicate P, x P y if and only if y nu P x'); and **dui dua**, the demonstrative predicate variables, or "proverbs" as they

## Notebook 3 – Lexicon

are sometimes called in English grammar, which have meanings much like English 'do'. For example, 'He meant to go; and he did'. This sentence could be rendered into Loglan as **Da moi godzi, ice da dui**. Like all demonstratives, the **i**-final member of the pair is used for proximate items, whether in time or space, the **a**-final one, for distant items.

(4) Acronymic PREDAs such as **CaiIzO** [CIO]; these may be made or borrowed at will by the user. Some Loglan acronyms, like DNA and USA, have simply been borrowed from existing acronyms in the natural languages...but of course they are then repronounced as Loglan words. Thus DNA in Loglan text is the abbreviation of the compound Loglan word **DaiNaiA**, and USA is short for **USaiA**. Acronymic predicates may also be derived internally, that is, from commonly used Loglan phrases whose frequency of use may have increased so much that the Zipf principle (that the length of a linguistic expression be inversely proportional to its frequency of use) will insist that a short Loglan expression be found for it. The increasing frequency of use of long, usually technical expressions (e.g., 'deoxyribonucleic acid') is, of course, the chief cause of acronym-formation in both the natural languages and in Loglan.

It is worth noting in passing that, while the part of speech of a compound letter-word, such as **MaiTai** [MT], is that of an acronymic PREDA, a single letter-word, such as **Mai** or **Tai**, is classified as an instance of Lexeme TAI, the letter-variables. Letter-variables, in turn, are one of the elemental forms of the Loglan argument. So a large grammatical shift takes place in the movement of **Mai** to **MaiTai**; see Lexeme TAI for further discussion of this point.

Finally, (5), there is the predicate interrogative **he**, which is of course also a member of the PREDA lexeme. **He** may be used in all places where any other PREDA may be used:- As a modifier: **Da he forli** = 'X is how strong?'; as a modificand: **Da mutce he** = 'X is very what?'; as the entire predicate expression: **Tu he** = 'You're what?' or 'How are you?'; as the entire utterance: **He** = 'What?', often used in the sense of 'What did you say?' A more polite form of the last expression is **Eo he** = 'Please, what?'

In conclusion, the predicate words of Loglan are of an immense number and variety; but all of them have exactly the same grammar. To know how to use one of them is to know how to use them all.

### \*Lexeme RA: Numerical Predicate Suffixes

There are two of these, **ra ri**, and both are used both as ordinary NI words and to form the numerical predicates among the PREDAs. If it weren't for the use of Lexeme RA by the preparser, which must lex these numerical predicates by analyzing them as little word compounds, **ra ri** would, like their companions **re ro ru**, be allolexes of NI.

### Lexeme TAI: Letter Variables

There are exactly 100 of these letter variables in Loglan, 52 of them are words for the Latin characters, 26 upper-, and 26 lower-case; and 48 of them are words for the Greek characters, 24 upper-, and 24 lower-case. In the sequel, the word 'letteral', which is an analog of 'numeral', will stand for the phrase 'letter character'. Thus, just as [1] is the numeral for the number-words 'one' and **ne**, so [t] is the letteral for the letter-words 'tee' and **tei**. For their construction, see Secs. 2.21-23.

All letter-words are members of the TAI Lexeme. TAI enters the grammar at just two points, in rules R79 and R81. In R79 TAI words enter as simple arguments, as in such utterances as **Tai mrenu** = 'Tee is a man'; in R81, they enter as the operands of the numerical descriptor **lio**, as in **Lio Tai numcu** = '(The number) Tee is a number', for this, too, is a possible use of a letter-variable. As arguments, the uses of TAI words are very similar to those of DA words, the replacing and personal variables, etc. Like DA words, TAI words may be used as possessives in LE + TAI compounds, e.g., **Letei kapma** = 't's hat'. Other compounds made with TAI are the dimensioned numbers formed by NI + TAI, e.g., **Sai corta Tai lio nemakeimei** = 'S

## Notebook 3 – Lexicon

is shorter than T by 100km (that is, by 100 kilometers)'. The word **nemakeime** is, of course, a member of NI. A final use of TAI words in compounds is to make acronymic PREDAs; these are often strings of TAI or A words interspersed with NI words, the latter being confined to non-initial positions: e.g., **HaitoSaiOfo** [H2S04]. Sometimes the morphological rules call for one or more instances of interstitial **-z-**, the acronymic hyphen, e.g., as in **CaIzO**; see Sec. 2.29 for a fuller account of the construction of acronymic PREDAs.

Since simple TAI words are arguments and compound ones are predicates, this leads to a grammatical problem for the chemical abbreviations, some of which are simple TAI words, e.g., **Sai** or [S] for Sulphur, and others compound, e.g., **Caiza** or [Ca] for Calcium. This means that **Da Caiza** is a legitimate sentence (meaning 'X is Ca (i.e., Calcium)') but that **Da Sai** is not. In fact, **Da Sai** is a pair of arguments, perhaps an ordered couple, best translated 'X, S'. This lack of grammatical parallelism between the simple and compound letter-words is unfortunate, but it appears to be necessary and is easily accommodated. 'X is Sulphur' can be easily said symbolically in Loglan by the addition of another syllable: **Da meSai**. Thus the predifier ME (q.v.) turns the argument **Sai** into a predicate form.

Assigning the simple TAI words to one lexeme and the compound ones to another is justified by the fact that doing so leads to two very considerable conveniences, surrendering either of which would make the language poorer. One convenience is that if simple TAI words are arguments, then a rich and virtually limitless mechanism of "anaphora" (replacement of long designations by short ones) is made available to the loglanist, with the result that even ordinary Loglan speech may very well become "mathematized": **Ama pa donsui Bai, Cai** = 'A gave B (gift) C', or **A pa donsui B, C** in the letteralized written form. We may expect that eventually such usages will exploit the full set of 100 Loglan letter variables. To give up this elegance by making simple TAI words predicates (whence **\*Le Ama pa donsui le Bai le Cai**) would be to abandon this potentially powerful Whorfian experiment before it had been tried. So this convenience argues that at least the TAI words should be arguments.

A second and opposite convenience is assured if acronyms, which are nearly always compound letter-words, are treated grammatically as predicates. If this is done, then those acronyms which arise to fill the Zipfean need to shorten originally long predicate expressions—for example, as 'DNA' shortens 'deoxyribonucleic acid'—may be used in exactly the same ways as the longer expressions they replace. Thus, 'That's DNA' will replace 'That's deoxyribonucleic acid'; and in Loglan the same abbreviation will be **Ta DaiNaiA**. To make acronyms arguments, would be to undo this elegance. It would require that acronyms be prefixed by the predifier **me-** whenever their naturally predicative role was required—which would be very frequently. Thus, **\*Da meDaiNaiA** would be the required form for the predicate 'is an instance of DNA'; and again a useful elegance would be lost.

The solution adopted preserves both elegances. But it requires that when acronyms are used in designations, they be treated as the operands of some descriptor. Sometimes **lo** is the best descriptor, as in **Donsui mi leva batpi je lo DaiNaiA** = 'Give me that bottle of DNA'. But sometimes **la** is best, as in **La USaiA groda gunti** = 'The USA is a big country'. It depends entirely on whether, in the given context, the designation so generated is meant to be unique or not. The English description 'the USA', like its parent phrase 'the United States of America', is nearly always meant as a unique designation; therefore its translation into Loglan should probably be graced with **la**. Loglan makes such intentions explicit.

As mentioned, when it is desired to use a letter-word as a predicate, it must first be prefixed by **me**; see Lexeme ME.

### Lexeme UI: Free Modifiers

These are the words which, unless initial in an utterance, are counted by the preparser as "grammatical noise". So UI words (along with several other items discussed at the end of this section) are removed by the preparser from the string to be parsed. They are restored to it by the postparser after the string has been parsed. We are justified in effecting this temporary removal of noisy elements from the string given to the parser because free modifiers are literally free to appear anywhere in any utterance, and so their appearance somewhere in a certain utterance gives no grammatical information about it. Free modifiers are therefore almost purely semantic devices. Their restoration after parsing assures that their semantic information will not be lost.

## Notebook 3 – Lexicon

The entire grammatical significance of a UI word therefore lies in whether it appears at the head of its utterance or within it. When a UI word appears at the head of an utterance, it is taken to modify that utterance as a whole (**Ia mi ditca** = 'Certainly, I am a teacher'). When one appears non-initially, it is taken to modify the individual word it follows: **Mi ia ditca** = 'I, certainly, am a teacher (I don't know about the others)'; **Mi ditca ia** = 'I teach, certainly (I'm not sure what else I do)'. To follow a word with a free modifier is often to emphasize it. Thus the last two utterances might well be spoken /**MI**iaDITca/ and /mi**DIT**caia/, respectively, while the first, with its free modifier in the Loglan-normal initial position, might receive no emphatic stress at all: /iamiDITca/.

UI words are of five semantic types: First, there are the **attitudinals**, which are used to express rather than report the speaker's attitude toward what da is saying, or toward some aspect of the world that is evoked by what da says. Second, there are the **discursives**. These are the words that call attention to how the units of the speaker's discourse are related to one another, or to the utterances of another speaker. Third, there are the **relative interrogatives**, the 'When?' 'Where?' 'How?' and 'Why?' series made by compounding any PA word with HU. Fourth, there are the **utterance ordinals** made by compounding any NI with FI. Finally, there are the **salutations**. These are the greetings, farewells, and expressions of gratitude and acknowledgement. All five kinds of UI words may be placed anywhere in an utterance, although unless special emphasis is intended, the initial position is stylistically preferred.

In general, UI words are extralogical; that is to say, they do not alter the truth-values of the sentences which they adorn. For a speaker to be convinced that da came, or to ask why da came, has no bearing on whether or not da came.

Now, in detail: -

**Attitudinals.** All attitudinals are VV-form words; all the VV-form words in Loglan are attitudinals except **ie**, the identity interrogative. The exceptional **ie** is more grammatically constrained than the other VV-form words but is related to them semantically, as the act of interrogating expresses an attitude; see Lexeme IE. The current list of simple attitudinals is:-

ae	Yes, I nope so/that/to...	
ai	Yes, I will/intend to...	(-ai is also L. u.c. cons. suff.)
ao	Yes, I want to...	(-ao is also Gk. u.c. cons. suff.)
au	I don't care whether...	
ea	Let's.../I suggest that...	
ei	Is that so?/Is it true that...?	(-ei is also L. l.c. cons. suff.)
eo	Please...	(-eo is also Gk. l.c. cons. suff.)
eu	I suppose/Let's suppose that...	
ia	Yes, that's certainly true/Certainly/I agree that...	
ii	Perhaps...	
io	Probably.../I think so/that...	
iu	I don't know whether...	
oa	I/you must...	
oe	I/you should...	
oi	I/you may...	
ou	It doesn't matter ethically whether I/you...	
ua	(completion) There!/At last!	
ue	(surprise) Well!/Oh!	
ui	(pleasure) Gladly/Happily...	
uo	(anger/annoyance) What!	
uu	(sorrow/regret) Alas.../Sorry that...	

Note that several VV-forms are still unassigned: **aa ee oo**. These may yet be given meaning as loglanists learn how to handle these disyllables which to anglophones still seem odd.

## Notebook 3 – Lexicon

To expand the domain of attitudinal expression, any number of compound attitudinals may yet be constructed. For example, **uiua** has a clear meaning: pleasure + completion = satisfaction; **ueui** also has a clear interpretation...and a usage; for example, it might be used to greet a long-lost friend standing at one's door. But few of these attitudinal compounds have been made. Again, the loglanist is invited to explore a new semantical domain.

**Discursives.** All the discursive UI words are CVV in form; but by no means all the CVV-form words are members of the UI Lexeme. For example, all of the case-tags (DIO), most of the letter variables (TAI), and many of the modifying prepositions (PA) are also of CVV-form. So there is some potential for confusing CVV-forms that seem to have similar functions. The way to distinguish the discursive UI words from the prepositional PA and DIO words is to ask whether the word you're uncertain about may be used prepositionally. If it may be, it's probably either a PA word or a DIO word. If not, it's almost certainly one of the discursives.

But what about the many words with vaguely "adverbial" meanings? If the word in question changes the claim of the predicate of the sentence in which it occurs, it's probably a PA word; a further test is whether you can turn it into a preposition. But if the word in question seems to relate the utterance as a whole to some other utterance, stated or implied, it's probably a discursive.

Here is the current list of discursive UI words. Note that none of them may be used prepositionally, that the meanings are generally adverbial but that each such adverbial meaning seems to appeal either explicitly (**cia**) or implicitly (**coa**) to another utterance or, like **feu**, to allude to outside information:-

bea	(bleka)	For example/By way of illustration (cf. <b>gea</b> )
buo	(bufpo)	However/In contrast/On the other hand
cea	(cenja)	Changing topics/(New paragraph)
cia	(clika)	Similarly/Like the foregoing
coa	(corta)	Summarizing/In short
dau	(dakli)	Hence it is probable that
dou	(donsu)	Given/By hypothesis/Per assumption
fae	(fanve)	And vice versa (reverses the order of terms in a previous claim to form a new one)
fau	(fando)	Finally/In conclusion
feu	(fekto)	In fact/Actually/According to the facts
gea	(genza)	In particular/As an instance (cf. <b>bea</b> )
kii	(kliri)	Clearly/Obviously/Plainly/Of course
kuu	(kumtu)	Generally/Generalizing from the above
leu	(clesi)	Anyway/Anyhow/In any case/event
nie	(snire)	In detail/Closely examined
saa	(sapla)	Loosely/Roughly/Simply speaking
sui	(sumji)	Also/Too/Moreover/Furthermore/In addition
voi	(volti)	Skipping details/Without going into details

Notice that, like the DIO words, when a natural derivation of the discursive from its primitive mnemonic yields a TAI word, then the final vowel of that first derivative is changed to /u/. **Dau** and **leu** illustrate the /u/-transformation in the above list. What this means is that TAI words have derivational precedence over other CVV-form words.

**Relative Interrogatives.** These are the compounds made from any PA word plus the argument interrogative HU ('Who?'). The meanings of such compounds are always readily decipherable, since they are nothing more than contractions of prepositional phrases in which the object of the preposition, its operand, is always **hu**. So **nahu** means 'At what time?' or 'When?'; **vihu** means 'In or at what place?' or 'Where?'; **biuhu** means 'In what manner?' which is one of the many varieties of 'How?' (nearly any modal operator will fit here); and **kouhu** means 'Because of what?' or one of the numerous Loglan 'Why?'s'. All the PA words—and there are many hundreds of them—are capable of generating such question words, and with much greater precision than is apparently ever available in the interrogative forms of natural language.

## Notebook 3 – Lexicon

The surprising thing about these interrogatives is their very simple grammar. Like the true-false interrogative **ei**, they can turn any utterance into a question. The whole utterance is its operand. Thus **ei** turns **Toi tradu** ('This is true') into the question **Ei toi tradu** = 'Is this true?'. Similarly, **kouhu** turns it into the causal question **Kouhu toi tradu** = 'Why (because of what cause) is this true?' And three other varieties of 'Why?', namely **Moihu Rauhu** and **Soahu**, are available to the loglanist to lend surgical precision to da's inquiries.

**Utterance Ordinals.** This is the series of compounds formed by attaching the suffix **-fi** to any NI or TAI word. The results, **Nefi Tofi Tefi** and finally **Rafi**, allow the loglanist to number da's utterances—or, for that matter, to number any sequence of clauses or terms within an utterance—with effects like English 'Firstly' 'Secondly' 'Thirdly' and 'Finally' or 'Lastly'. Letter-ordinals may also be employed: **Amafi Baifi Caifi** and so on; and for the purposes of outlining, the lower-case Latin and the two Greek series of utterance letter-ordinals may also be used: **asifi beifi ceifi**, **Amofi Baofi Caofi** and **afifi beofi ceofi**, etc. These words, too, are grammatical noise.

**Salutations.** There are four such currently assigned UI words. They are **loi** ('Hello') **loa** ('Goodbye') **sia** ('Thank you') and **siu** ('You're welcome'). Like the attitudinals the salutations are expressions of the speaker's feelings or desires, but this time, toward da's interlocutor or someone da wishes to speak with. No doubt there will be more such words in time. Their grammar is identical to that of the free modifiers. They may be used anywhere; or, if no special internal emphasis is desired, they are usually placed, like the other UI words, at the head of the utterance: **Loi mi bi la Djan** = 'Hello! I'm John'.

**A Note on Other Free Modifiers.** In addition to the UI words, there are three other structures that are handled as free modifiers in Loglan grammar. These are (1) the vocative marker HOI (q.v.) together with any string of one or more name-words that may optionally follow it (**Hoi nenkaa** = 'Hey, come in! vs. **Hoi Djan, nenkaa** = 'O John, come in!'); (2) any unmarked string of one or more name-words, that is, unmarked by either LA or HOI (**Djan Pol Djonz, nenkaa** = 'John Paul Jones, come in!'); and (3) the parenthetic marker KIE together with the parenthesized utterance and the closing parenthesis KIU which will always follow KIE in well-formed speech (**Kie Rafi kiu Djan, pa nenkaa** = '(Finally) John came in').

These three elements together with Lexeme UI constitute the complete list of free modifiers in Loglan.

### Lexeme ZE: The Joining Operator

ZE is monolexic; it contains only **ze**, the "joining" or "mixing" operator which has the sense of 'and, jointly'. **Ze** is not a logical connective; it does not allow a number of connected claims to be made simultaneously. A claim made with **ze** is always a single claim. Thus, **Da redro, e cmalo** makes two claims about X, namely that it is red, and that it is small. **De redro ze nigro** makes only one claim about Y, namely that it is red-and-black mixed together. Perhaps it has red stripes alternating with black ones; perhaps it has red dots on a black field. But if a thing is **redro ze nigro**, it is not true that it is either red or black separately.

ZE may be used to join arguments or, as above, predicates. The designata of **ze**-joined arguments are like teams. If **Da ze de pa berti leva felstaga** ('X and Y, jointly, carried that log (fallen trunk)' then what is being asserted is that the team of them did. We may gather that it is most unlikely that either of them could have carried it separately.

The grammar of ZE is very similar to that of the afterthought connectives; see Lexeme A. However, since **ze**-binding is "tighter" than shek-binding, that is, since strings like **Da, e de ze di** = 'X, and Y and Z, jointly' will always parse as **(da e (de ze di))**, grammatically "earlier" rules are required to effect **ze**-joints than shek-joints. This is what requires **ze** to be in its own lexeme.

ZE is also used by the preparser to recognize acronymic PREDA's. This is because pseudo-**ze**'s are sometimes generated by the acronymic hyphen **-z-**.

## Notebook 3 – Lexicon

### **\*Lexeme ZI: Magnitude Suffixes**

There are three of these words, **zi za zu**, indicating small, intermediate, and large magnitudes, respectively, in their role as suffixes in tense (**fazi** = 'right away') and location (**vizu** = 'in this region') compounds; see PA. A separate lexeme is required to enable the prepaser to recognize such compounds and also to identify some acronymic PREDA's, because, just as with ZE (q.v.), pseudo-instances of **za zi zu** may be generated in these words by the acronymic hyphen **-z-**.

### **\*Lexeme ZO: The Quantity Abstractor**

ZO is monolexic; it's sole member is **zo**, the quantitative abstraction operator. It has a grammar parallel to that of PO, q.v. The only reason **zo** occupies a separate lexeme is because, just as with ZE and ZI words, pseudo-**zo** syllables may be generated by the use of hyphen **-z-** in some acronymic PREDA's. If it weren't for this mechanical use of ZO by the prepaser, **zo** would be a member of PO.

## CHAPTER 4

### GRAMMAR (UTTERANCE FORMS)

**4.1 Design Objectives:** The objectives that have controlled the design of Loglan grammar have been, first, to accommodate the rich variety of claims and designations found in natural language but, second, to do so with a grammar modeled on the predicate calculus, thus facilitating both the logical manipulation of its utterances by its users and the design of a language with parsimonious metaphysical assumptions. Third, it was to make the operations of that grammar natural enough to fit the human central nervous system so as to make the language speakable, and fourth, it was to be flexible enough to imitate the natural word orders of the planet's major tongues when desired...for example, in translation. But it was also, fifth, to be expressible in a rule set small enough to be easily learned by adults, possibly even under experimental conditions. Finally, sixth, the grammar was to be syntactically unambiguous for two reasons. The most obvious one was to make interaction with machines possible. Somewhat less obvious is a reason that is the same reason logicians and mathematicians require unambiguous codes, namely to make it possible rigorously to entertain implausible ideas...a prime requirement of a logical language.

Syntactic unambiguity was achieved heuristically, and therefore doubtably, as early as 1963, and was not to be formally demonstrated until 1982. A few years before that achievement a constructive-proof algorithm for demonstrating conflict-freeness in certain classes of computer languages had become available and was soon adapted to disambiguating human grammars. Syntactic unambiguity was lost again in 1984 when The Institute's small computer proved inadequate to the task of servicing the growing grammar. It was recently redemonstrated for the enlarged language when The Institute acquired more capacious computing machinery in 1986. During the Winter and Spring of 1987 a large backlog of planned new grammatical features were installed in the language one by one; and I was gratified to discover that the condition of conflict freeness was easily returned to each time. In short, given our present tools for disambiguating human grammars, there is no reason to believe that Loglan grammar will not remain syntactically unambiguous for the indefinite future.

**4.2 Definitions and Conventions:** Let us first look at a typical grammar rule and provide a terminology for discussing its parts:

150: kekpred	=> M3 KA predicate KI predicate	Da <b>ke ckano ki briga</b> = X is <b>both kind and brave.</b>
151:	=> NO kekpred	De <b>no ke ckano ki bunbo</b> = Y is <b>not both kind and a fool.</b>
Local.		

This is the grammar rule by which "kekpred", or forethoughtfully connected, predicates are formed. The specimens on the right illustrate each line of the grammar rule. The number on the left gives its position in the grammar. The remark 'Local' tells us that the rules which use 'kekpred' are all nearby.

Each numbered line in a grammar rule is called a **rule**. The sign [=>] in each rule may be read 'may be developed as' or 'may produce' and is called the **production sign**. The expression to the left of the production sign in a rule is called its **left-half**. If a rule has no explicit left-half, it is assumed to have the same left-half as the first preceding rule that has one.

Rules 150-1 are all the rules in this grammar that have 'kekpred' as their left-halves. A **grammar rule** is the complete set of rules in a given grammar that have a given left-half. Within a grammar rule, the order of rules is arbitrary.

This is a **context-free grammar**. In such grammars the left-halves of all grammar rules are single elements. **Right-halves** may either be single elements, in which case the rule is called a **replacement**, or they may be strings of two or more elements (up to about five in number), in which case, it's an **expansion**.



## Notebook 3 – Grammar

Elements which are written entirely in upper-case letters, like 'KA' and 'NO', or in a mixture of upper-case letters and numerals, like 'M3', are called lexemes. Recall that lexemes are sets of grammatically interchangeable words, roughly corresponding to the "parts-of speech" of conventional grammar.

Elements whose names are written at least partly in lower-case letters—for example, 'kekpred' and 'predicate'—are known as **gramemes**. By definition, each grameme in a grammar appears in the left-half of exactly one of its grammar rules. At present, there are 88 gramemes in Loglan grammar; so there are 88 grammar rules.

The remark 'Local' appended to the 'kekpred' grammar rule tells us that the rules that use this grameme are all nearby (actually, in the same "grammar group"; see below). Sometimes a series of one or more numerals occupies the space occupied here by 'Local'; these are references to the non-local rules which use the grameme defined by that grammar rule. The forward references are given first; the backward references, if any, are separated from the forward ones by a semicolon.

The right-halves of a grammar rule are called the **allograms** of its grameme. 'kekpred' has two allograms. When either gramemes or their allograms are referred to in text, their names will be shown in angle brackets. Thus we can say that <kekpred> has the allograms <M3 KA predicate KI predicate> and <NO kekpred>. Alternatively, we can write the Kekpred Grammar Rule as <kekpred => NO kekpred | M3 KA predicate KI predicate>. In this second formulation the bar [|] is used to separate the names of allograms and is read 'or'.

A grameme that appears in the allograms of another grameme, as <predicate> appears in one of the allograms of <kekpred>, is said to **be used** by that other grameme. A grameme which is used in one of its own allograms is said to be **recursive**. In general, we will find it advantageous to develop early in the grammar the gramemes that will be used by later gramemes.

A sequence of grammar rules which has been ordered by the above principle, and which terminates in a widely used structure, is called a **grammar group**, or sometimes just a **group**. Ideally, all the gramemes except the final one in a group are used only within that group. When this is true of a grameme in a group, that grameme is said to be a **local** to that group and so is marked 'Local' as above. Groups are usually given the name of the major non-local grameme with which they terminate. For example, <kekpred> is part of the Predicates Group (Rules 128-54) and is local to that group. The last grameme developed in the Predicates Group is <predicate>.

As shown in the example, each rule will be illustrated by a **specimen** of Loglan that it helped produce. When we need to refer to a rule or specimen, we will use the ordinal of the rule but prefix it with 'R' or 'S'; thus S150-1 are the specimens of R150-1. The part of a specimen that is in boldface corresponds to the part produced by its rule. In the specimens given for R150-1 above, all but the Loglan words [Da] and [De] were produced by the two <kekpred> rules; so all but those parts of the two specimens are shown in boldface. A corresponding pattern of bolding is shown in the English translations of the Loglan specimens.

Some rules use M- or machine lexemes. For example R150 uses M3. Machine lexemes are inaudible to the human ear, but, as we will see later, they are indispensable for machine parsing. They are in fact lexemes in **machine Loglan**, which may be thought of as that "dialect" of Loglan that machines will be able to read and hear. By convention, the name of a machine lexeme is formed of the letter 'M' followed by a numeral. There are at present 11 machine lexemes in Loglan: M1-M11. All are involved in one way or another with extending the limited 1-element **lookahead** of the **LRI parser**. This Yacc-generated parser is one of the three components of our **machine grammar**. Another component is the **preparser** which, among other **mechanizing** functions, inserts machine lexemes into the strings to be parsed. These machine lexemes, together with certain other lexemes (like ERROR in the first grammar group), are of concern only to the machine and so are not part of **human Loglan**. These and other computational features are removed from the parsed string by the third component of the machine grammar, the **postparser**. It is the function of the postparser to **humanize** the parse by making it intelligible to humans. We retain these machine-oriented features in our exposition of the grammar in the current notebook only because some of our readers may be interested in how machines may be said to "understand Loglan". We assume that most readers, however, will wish to ignore all grammatical embellishments that have been put there solely for machines. So we will glance only briefly at these computational devices. Readers interested primarily in the computability of the language must go to other sources for detailed information about how this has been accomplished; e.g., **Notebook 1**, 1982.

**4.3 The Structure of Loglan Grammars:** The grammar rules of Loglan may be conveniently divided into twelve functional groups. We will discuss those groups here in the order in which the listener is likely to make use of them (the so-called "top-down" order). In that order they are the groups of rules that govern the formation of (1) Optional Punctuators, (2) Linked Arguments, (3) Predicate Units, (4) Descriptive Predicates, (5) Sentence Predicates, (6) Modifiers, (7) Argument Modifiers, (8) Arguments, (9) Term Sets, (10) Predicates, (11) Sentences, and (12) Utterances.

In general, later structures on the above list involve earlier ones. Thus, looking at the grammar from the "bottom-up", we would find that

Utterances require Sentences;

Sentences require Modifiers, Predicates and Arguments;

Predicates require Sentence Predicates and Term Sets;

Term Sets require Arguments and Modifiers;

Arguments require Descriptive Predicates and Argument Modifiers;

Argument Modifiers require Modifiers;

Modifiers require Arguments;

Sentence Predicates require Predicate Units and Descriptive Predicates;

Descriptive Predicates require Predicate Units;

Predicate Units require Linked Arguments;

Linked Arguments require Arguments again, which are thus the most "circular" structures in the language, in that they are used by nearly everything which they in turn use; and that

Optional Punctuators are used by nearly every other grammar group.

Therefore we will start with Group A, the Optional Punctuators, for these are at the "top" of the parse tree. We will then work "down" to Group K, which develops Utterances, for these are near its "root". The grammar is thus presented in its "top-down" order; it goes from the "leaves" (the lexemes) toward the "root" (the <utterance> grameme itself) of the parse tree. This is the same order, by the way, in which the rules appear in the formal (machine) grammar which was given to Yacc to build the LR1 parser. While this may not be the best order in which to teach grammar rules to human learners, it is the most economical one in that it minimizes forward references. It is thus most suitable for a work whose most enduring value to its users will probably be its utility as a reference work.

### Group A. The Optional Punctuators

#### Rules 1-7

The three "optional punctuators" **gu**, **gue** and PAUSE, the latter being represented by a pause in speech and a comma [,] in text, are optional only in the sense that they may be omitted when they are not necessary. But in fact, usage requires that these punctuators be omitted when the intended grouping is clear without them; so they are seldom optional in any real sense. Oddly enough, the computer sometimes treats these punctuators as "present in spirit" even when they are omitted. S1 is an example of this.

- |        |                    |  |
|--------|--------------------|--|
| 1: err | => ERROR<br>Local. | Mu titci fa ( <b>ERROR</b> ) = We eat later. (The machine sees this as an error because it expects a <b>gu</b> . But if it supplies the dummy lexeme ERROR at the end of this utterance, it can complete the parse. We arrange for it to do this by making <err> an allogram of <gu> in R2, and of <gue> in R4.) |
| 2: gu  | => err             | Mu titci vi ( <b>ERROR</b> ) = We eat here.  |

## Notebook 3 – Grammar

- 3:           => GU                   Mu titci vi **gu** le supta = We eat here the soup. (Without **gu** the utterance would be heard as 'We eat in the soup'; so **gu** is necessary here.)  
121-2.

Both **Mu vi titci le supta** = 'We here eat the soup' and **Mu titci le supta vi** = 'We eat the soup here' avoid the need for punctuation. The unpunctuated word orders tend to be stylistically preferred in Loglan but are certainly not obligatory. In fact, the judicious use of punctuation makes almost any conceivable word-order possible in Loglan, a feature which is very useful in translating whenever one wishes to reproduce the flavor of a natural language text as closely as possible.

There is also an invisible '(ERROR)' at the end of S3 as of nearly all specimens. But we don't show it here because the <err> allogram of <gu> is not part of R3.

- 4: gue       => err                   Da bilti ge cmalo nirli ckela (**ERROR**) = X is beautiful for a small-girls school. (Punctuation at the end of an utterance may always be omitted. Again, the computer will compensate for such "errors" by inserting the dummy ERROR Lexeme.)
- 5:           => GUE                   Da bilti ge cmalo nirli **gue** ckela = X is a beautiful small-girls [pause] school, i.e., a school for small girls who are beautiful. (Here the **gue** serves as a right parenthesis matched with **ge**.)  
22-3.
- 6: gap       => gu                   Mu titci fa **gu** le mitro = We eat later the meat. (Again, an explicit mark is called for; to omit it is to generate the phrase **fa le mitro** = 'after the meat'.)
- 7:           => PAUSE                 Mu titci fa, le mitro = We eat later, the meat. (When <gap> is invoked, the required mark may be a comma in writing or a pause in speech.)  
10, 15-6, 24, 59-60, 83, 87, 90, 92, 94, 98, 100, 130-1, 183-4.

As may be seen from the references, **gu** alone is required in only two rules, R121-2; **gue** is used in similarly few places, R22-3; but the triple option presented by <gap>, which may be executed by a **gu**, pause/comma or frequently by nothing at all, is extremely widely used. In fact <gap> is the most widely-used grameme in Loglan grammar. It provides an opportunity to mark the right boundary of phrases or clauses whenever such a boundary would be unclear without it. Loglanists tend to use pauses and commas at such points when dealing with human readers and interlocutors in conditions of low noise, and to use **gu**'s in high noise or when addressing computers or other unforgiving auditors. See the numerous rules listed above for the occasions on which these punctuators may be used.

Optional punctuators are not the only "punctuation marks" in Loglan. There are also six punctuators which are essential to the structures they mark and so may never be omitted. These are the "grouping operator" **ge**, which may be found in R22-3 along with the optional **gue**; the "inversion operator" **go**, found in R48; **goa**, which marks the V-O-S word-order in R155; and the "fronting operators" **gi** and **goi**, found in R165-66. In addition, there is a punctuator-like member of the PA-Lexeme, **ga**, which is used exclusively for marking the left boundaries of otherwise unmarked predicate expressions when these would be absorbed by just-preceding descriptions; see R139. For example, **ga** in **Le mrenu ga sadji** = 'The man is wise' prevents **Le mrenu sadji** = 'The man (sort of) wise one' from being heard. <Gap> would work here—for example, **\*Le mrenu, sadji** also parses in the required way—but is regarded as bad usage (and so is \*-ed) because there are some kinds of descriptive arguments after which <gap> fails to produce the required separation. **Ga** always works and so is preferred.

This is the entire punctuation system of Loglan. See the rules cited for the details.

## Group B. Linked Arguments Rules 8-19

## Notebook 3 – Grammar

**Je** and **jue** are the two preposition-like words that attach strings of one or more arguments to predicate words. Sometimes the predicate word involved is buried in a predicate expression (**Da kukra je lo litla, grobou** = 'It's a faster-than-light ship'); sometimes it is the last word in a description (**Le selrispe farfu je la Djek** = 'The proud father of Jack'). In either case, the linking words have the effect of binding a string of elements into a single unit. Thus if the **je** and **jue** of S8 were removed, the string remaining would be composed of three distinct arguments: **To ketpi da de** = 'Two tickets, X, Y'. The two linking operators bind this triad into a single argument: 'Two tickets to destination X from point-of-departure Y'. The present group of rules shows how to construct the right part of such expressions, the "linked arguments". The predicate units, like **ketpi**, to which they are attached are constructed in the next group; see R26.

8: juelink	=> JUE argument Local.	To ketpi je da <b>jue de</b> = Two tickets to <b>X from Y</b> .
9: links1	=> juelink	<b>Jue de = From Y.</b>
10:	=> juelink linksl gap Local.	<b>Jue de jue di = From Y on W.</b>
11: links	=> linksl	<b>Jue de = From Y.</b>
12:	=> links M1 A linksl	<b>Jue de, a jue di = From Y or from W.</b> (The pause after <b>de</b> is a "morphemic pause", necessary for the resolution of <b>a</b> .)
13:	=> M2 KA links KI linksl 174.	<b>Ke jue de ki jue di = Both from Y and from W.</b>
14: jelink	=> JE argument Local.	To ketpi <b>je da</b> jue de = Two tickets to <b>X from Y</b> .
15: linkargsl	=> jelink gap	<b>Je da = To X.</b>
16:	=> jelink linksl gap Local.	<b>Je da jue de = To X from Y.</b>
17: linkargs	=> linkargsl	<b>Je da = To X.</b>
18:	=> linkargs M1 A linkargsl	<b>Je da, e je de = To X and to Y.</b>
19:	=> M2 KA linkargs KI linkargsl 26, 173.	<b>Ke je da ki je de = Both to X and to Y.</b>

When links are connected, as they are for example in S18, the linking words need not be repeated. Thus, Loglan **Je da, e de** conveys the same notion as S18 does just as English 'To X and Y' conveys the same notion as 'To X and to Y'. Sometimes in both languages, however, this degree of explicitness is desired.

Links and linkargs are used in only three non-local rules. Links are used in R172, and linkargs in R26 and R173. In R172-3 both links and linkargs appear as fragmentary utterances, such as might be answers to **Jue hu** and **Je hu** ('To/from/by whom?') questions. Linked arguments enter the main stream of the grammar at just one place, however, namely in R26 of the next group, where they are attached to predicate words.

Notice that two machine lexemes have been introduced in Group B. They are M1 which marks the "eks" (the **a**-form afterthought connectives) which connect both links and linkargs; and M2 which marks the "keks" (the **ka**-form forethought connectives) which connect these objects. The reader will recall that machine lexemes are put in place by the preparser, one of whose tasks is to extend the limited 1-element lookahead of the machine. Human brains do not need these warnings since our capacity to remember and inspect long strings is so much greater than that of any parsing algorithm used currently by machines.

**Group C. Predicate Units**

**Rules 20-33**

Predicate units, or "predunits" as we will sometimes call them, are either single predicate words, with or without certain inflecting operators such as **nu**, **no** or **po**; or they are strings of such possibly inflected predicate words which have been made into a predicate unit by **ge**, or by a **ge** matched with a subsequent **gue** (R22-3); or they are arguments that have been turned into predunits by the "predification" operator **me** (R24); or they are predicate words which have been augmented by the attachment of one or more linked arguments (R26).

20: predunitl	=> PREDA	Ba <b>sucmi</b> = Something <b>swims</b> .
21:	=> NU PREDA	Ba <b>nu sucmi</b> = Something <b>is swum to</b> .
22:	=> GE despredE gue	Da briga <b>ge musmu janto</b> = X is brave <b>for a mouse hunter</b> . (The <gue> is not activated here; it would be redundant if it were.)

**Ge** is the grouping operator; it is only meaningful if it precedes a string of two or more predunits. Thus **\*briga ge musmu**, while grammatical, is proscribed as bad usage, for it can mean nothing other than what **briga musmu** ('brave mouse') already means. Again, redundant marks are avoided. <despredE>, which is the operand of **ge**, is made in the next group; see R45-6.

23:	=> NU GE despredE gue	Da <b>nu ge briga janto</b> = X is a <b>quarry of brave hunters</b> . (Nu is one of the conversion operators; and <b>ge</b> extends the scope of <b>nu</b> over the entire <despredE> string, in this case <b>briga janto</b> , which then takes its place-structure from its last term, in this case <b>janto</b> . Thus X is a kind of quarry.)
24:	=> ME argument gap Local.	Ba <b>mela Ford</b> = Something <b>is a Ford</b> . ( <b>Me</b> turns any argument into a predicate unit with a meaning associated with that argument.)
25: predunit2	=> predunit1	Da <b>kukra grobou</b> = X is a <b>fast ship</b> .
26:	=> predunitl linkargs Local.	Da <b>kukra je lo litla</b> , grobou = X is a <b>faster-than-light ship</b> .

R26 is the primary use of the linked arguments made in the preceding group. Their use in descriptions also passes through this rule. Usage imposes a certain restraint on the use of R26, however. The <predunitl> to which linkargs is attached by it must not, by convention, be the last unit in a string unless that string is a descriptive predicate, i.e., one used in making arguments. Thus **Le mutce kukra je lo litla** = 'The thing which is very much faster than light' is an acceptable use of linked arguments. In fact the link word is necessary if this description is to be heard as one argument rather than two. The link word is also necessary in **Da kukra je lo litla, grobou** in which the argument is linked to a non-final predicate unit in a string. (We will call this the "internal specification" of a predicate.) But arguments are not allowed to be linked to final units in predicate strings when these are being used as sentence predicates. Thus while **\*Da mutce kukra je**

## Notebook 3 – Grammar

**lo litla** (presumably to render 'X is very much faster than light') is perfectly grammatical, it is a proscribed usage because there exists a preposition-free form that does the same job more elegantly: **Da mutce kukra lo litla**. Again, usage in Loglan avoids redundant markings, such as this quite unnecessary **je**. The same link words are not of course redundant inside predicate strings, or to link a descriptive argument to other arguments. But they are redundant if used to link a sentence predicate to its argument set, or the arguments in that set to one another.

This is the first instance we have encountered of the "grammatical superset" phenomenon. By this I mean that we have written a rule of grammar that generates a domain of grammatical utterances which is larger than the domain of "good utterances" allowed by usage. Another way of saying this is that not all grammatical utterances are acceptable as good usage. Any economically written set of grammar rules will occasionally generate such effects. They are perfectly harmless. We need only add a set of usage rules to such a grammar to ensure that any grammatical utterance that meets the usage rules will also be interpretable. We have at present no way of interpreting the difference between **\*Da kukra je de** and **Da kukra de**; so we do not use the former. There are other supersets in Loglan grammar; I will call the reader's attention to them as we encounter them.

27: predunit3 => predunit2  
28: => NO predunit3  
Local.

Da **sadja** = X is **wise**.  
Da **no bimbo** = X is **no fool**. (This is short-scope negation; it applies to a single predunit. Long-scope negation is accomplished elsewhere.)

29: predunit4 => predunit3  
30: => predunit4 ZE predunit3  
Local.

Ti **nigro** = This is **black**.  
Ti **nigro ze babe** = This is **black-and-white (mixed)**. (**Ze** is a special connective which has the sense of mixing properties.)

31: predunit => predunit4  
32: => PO predunit4

Ti **blanu** = This is **blue**.  
Ti **po blanu** = This is a state of **being blue**. (This is short-scope abstraction; other uses of **po** take whole sentences as operands; see R130-1.)

33: => ZO predunit4  
34, 36, 49-50.

Ti **zo blanu** = This is an amount by which something is blue.

**Zo** is kept out of the PO-Lexeme only because the preparer needs to recognize quasi-ZO in lexing acronyms. This is because **-z-** is the acronymic hyphen, and when it precedes /o/, as it does in **CaiIzO** [CIO], the result looks like a compound of TAI + I + ZO to the preparer. If **zo** were part of PO, it could not lex this acronym correctly. For this entirely mechanical reason—which probably has no meaning for human lexers—PO and ZO must be kept lexemically separate.

Predunits are the building blocks with which "descriptive predicates" (predicates used as "nouns") and "sentence predicates" (predicates used as "verbs") are made. These will be constructed in the next two groups. No machine lexemes have been necessary in the construction of predicate units.

### Group D. Descriptive Predicates

#### Rules 34-48

These are the predicate expressions which, when used as operands of any of the descriptive operators—**le** and **kin**, for example, or the name operator **la**, or any quantifier—create that special kind of argument

## Notebook 3 – Grammar

known as nouns or noun-phrases in Indo-European grammar. We call such arguments "descriptions" because they employ a predicate expression to "describe" some feature of the designated thing.

The following sequence of rules differs from that in the next group, in which sentence predicates are made, primarily in one feature: descriptive predicates may have kekked head units, e.g., the **ke forli ki sadji** part of S35; sentence predicates may not. It turns out that ambiguities are generated if sentence predicates are permitted to have kekked head units, while the same constructions are quite unambiguous in descriptions. So kekked head units are provided in this rule group—the <kekpredunit> of R35 and 37-8—and absent from the next.

34: despredA	=> predunit	Le <b>forli</b> = The <b>strong one</b> .
35:	=> kekpredunit	Le <b>ke forli ki sadji</b> = The <b>both strong and wise one</b> .
36:	=> predunit CI despredA Local.	Le <b>denro simba ci janto</b> = The <b>dangerous lion-hunter</b> . (Without <b>ci</b> the predunits in the string would group left, thus 'The dangerous-lion hunter'.)
37: kekpredunit	=> M3 KA descpred KI despredA	Le <b>ke forli ki sadji</b> = The one who is <b>both strong and wise</b> .
38:	=> NO kekpredunit Local.	Le <b>no ke forli ki sadji</b> = The one who is <b>not both strong and wise</b> .

Another machine lexeme is introduced here, M3 which extends the parser's lookahead over KA again, this time to announce to the parser that a predicate is being kekked. Unmarked keks connect arguments.

39: despredB	=> despredA	Le <b>sadji</b> = The <b>wise one</b> .
40:	=> CUI despredC CA despredB 52, 54.	Le mutce <b>cui fizdi forli ce sadji</b> = The one who is very <b>physically-strong and (very) wise</b> .

**Cui** is the left boundary mark of left connectands formed of more than one predunit. Without **cui** the single units on either side of the "shek" (the CA-connective) are taken as its connectands. No similar mark for right connectands is necessary.

41: despredC	=> despredB	Le <b>mrenu</b> = The <b>man</b> .
42:	=> despredC despredB 52.	Le <b>mutce bunbo mrenu</b> = The <b>very foolish man</b> .

This is the rule by which predicate strings are generated. Because it is left-recursive, all unmarked predicate strings are left-grouping.

43: despredD	=> despredB	Le <b>sadji</b> = The <b>wise one</b> .
44:	=> despredD CA despredB 56.	Le <b>forli canoi sadji</b> = The <b>strong if wise one</b> .
45: despredE	=> despredD	Le <b>mrenu</b> = The <b>man</b> .
46:	=> despredE despredD 22-3.	Le <b>forli canoi sadji mrenu</b> = The <b>strong if wise man</b> .
47: descpred	=> despredE	Le <b>troku hasfa</b> = The <b>stone house</b> .

## Notebook 3 – Grammar

48:                   => despredE GO despred                   Le **hasfa go troku** = The **house of stone**.  
58, 72-4, 76, 97

Notice that several grammar rules in this group besides the last one are non-local. All these other non-local rules are employed at similar points in the next group, where sentence predicates are made. But as all descriptive predicates may have kekked head units, care is taken to ensure that these structures borrowed from the descriptive sequence do not end up as the heads of sentence predicates; for there they would generate ambiguities.

The descriptive predicate grameme itself, <despred>, is used as the right part of R58, the final rule in the construction of sentence predicates, and at various points (R72-4, 76 and 97) in the construction of arguments, which are made in Group G.

### Group E. Sentence Predicates

#### Rules 49-58

This group of grammar rules is structurally parallel to the preceding one, the only difference being the one already noted, namely that sentence predicates may not have kekked head units. Note that some "despred" gramemes figure in these "senpred" rules, but that they are never initial in an allogram; this renders them harmless. Thus **mutce ke briga ki ckano** is a permissible sentence predicate while **ke briga ki ckano mrenu**, with its kekked head unit, is not. The reason for this proscription is not hard to find. If **ke briga ki ckano mrenu** were permitted as a sentence predicate, as in **Da ke briga ki ckano mrenu**, presumably to mean 'X is both a brave and a kind man', then the parser could not distinguish between this use of keks to connect predicate units—it is **briga** and **ckano** that are being connected here and not **briga** and **ckano mrenu**—and the later use of the same keks in Group I to connect whole predicate expressions. This would generate an ambiguity between the two rules, for they would both be capable of generating the same string. So the auditor could not tell whether the speaker meant his keks to connect just preddunits, as above, or whole predicate strings, as in Group I. The latter possibility would give an entirely different parse tree, one with the interpretation 'X is both a brave person and a kind man'. As things stand, the second interpretation is the only legitimate one...precisely because kekked head units are not allowed in sentence predicates.

49: senpred1	=> predunit	Da <b>mrenu</b> = X is a <b>man</b> .
50:	=> predunit CI senpred1 Local.	Da <b>simba ci janto mrenu</b> = X is a <b>lion-hunting man</b> .
51: senpred2	=> senpred1	Da <b>sadji</b> = X is <b>wise</b> .
52:	=> CUI despredC CA despredB Local.	Da grada <b>cui simba janto ce sadji</b> = X is a great <b>lion-hunter and</b> (a great) <b>wise one</b> .
53: senpred3	=> senpred2	Da <b>prano</b> = X <b>runs</b> .
54:	=> senpred3 CA despredB Local.	Da <b>prano a fleti</b> = X <b>runs and/or flies</b> .
55: senpred4	=> senpred3	Da <b>mrenu</b> = X is a <b>man</b> .
56:	=> senpred4 despredD Local.	Da <b>simba mrenu</b> = X is a <b>lion man</b> .
57: sentpred	=> senpred4	Da <b>simba fumna</b> = X is a <b>lion woman</b> .
58:	=> senpred4 GO despred 128.	Da <b>fumna go simba</b> = X is a <b>woman of lions</b> .



## Notebook 3 – Grammar

All the gramemes in this tiny group are local except the last one, and that one, <sentpred>, figures in just one later rule: R128 in Group I, which makes the objects called simply "predicates". It is in this later group that the sentence predicates made here will be supplied with their "termsets": the strings of modifiers and/or arguments that complete them as predicates. Before that can be done the termsets must be constructed. The modifiers and arguments of which termsets are composed will be prepared in the next two groups.

The making of sentence predicates has involved no machine lexemes.

### Group F. Modifiers

#### Rules 59-67

Setting aside the "free modifiers" of R176-9, which are quite differently distributed and composed, the (unfree) sentence modifiers which occur in Loglan utterances are of two sorts. They are either (1) phrases, such as **pa da** ('Before X'), in which some PA-word functions as a preposition and some argument, in this case a very simple one, functions as its "object" or operand; or they are (2) argument-less PA-words, such as **pa** itself, which function in the sentence like English "adverbs". Thus **pa** without an argument is an ellipsis for 'Before some assumed time, say the moment of speech' and thus means simply 'Earlier'. Group F is exclusively concerned with the construction of these two kinds of modifiers. Its first grammar rule, <mod1>, gives the basic architecture of Loglan modifiers, which is exceedingly simple. Notice that both the **pa** and the **pa da** varieties of modifiers may be optionally followed by the <gap> triplet, the choice of **gu**, pause/comma or nothing at all, which is one of the optional punctuators developed in Group A.

59: mod1       => PA gap  
60:             => PA argument gap  
                Local.

Mu titci **pa** = We eat **earlier**.  
Mu titci **pa la Ven** = We eat **before nine**. (When a modifier comes at the end of an utterance an explicit gap would be redundant and so is not used.)

61: mod       => mod1  
62:             => M4 NO mod  
                Local.

**Fa** = **Afterwards**. (Later/then/etc.)  
**No fa** = **Not afterwards**.

The parser needs to know what kind of **no** this is; so the preparer has looked ahead of this **no** (as the parser can't) and found that it is negating a modifier. So the preparer erects the signpost M4 before the **no** to inform the parser of its discovery. (The preparer is thus a kind of forward scout for the machine's parser... which snuffles along blindly—but very mentally—with its nose to the ground.)

63: kekmod     => KA modifier KI mod  
64:             => M4 NO kekmod  
                Local.

**Ke fa ki pa da** = **Both afterwards and before X**.  
**No ke fa ki pa da** = **Not both afterwards and before X**.

65: modifier   => mod  
66:             => kekmod  
67:             => modifier M5 A mod  
                89, 118.

**Pa da** = **Before X**.  
**Kanoi fa ki pa** = **If later then earlier**.  
**Fa noa na** = **Later only if now**.

Two machine lexemes are introduced in this short group, M4 which marks those **no**'s that negate modifiers, and M5 which marks the eks that connect them. The keks of R63 do not need to be marked.

## Notebook 3 – Grammar

### Group G. Arguments Rules 68-116

This is the longest rule group in the grammar, comprising nearly a quarter of all its rules. But the argument rules nevertheless constitute a coherent group, with few gramemes used elsewhere until we get to <argument> itself. Once developed, arguments are then used very widely throughout the grammar...for example, in the preceding group to make modifying phrases.

The question may well arise: Why should not arguments precede modifiers in their development? The answer is that arguments use modifiers in their development as well, and that arguments unlike modifiers have a great variety of backward references. They are, recall, used as early as Group B, Linked Arguments. It seemed best to put them in a position where all they will use has already been built; and then let them have all the backward references they require. In short, arguments are the most circular structures in Loglan grammar in that they are more frequently embedded in themselves.

An argument is a designation, and in Loglan as in the natural languages there are many kinds of designations. A preview of what is to be found in this large and important rule group may therefore be in order. We will commence with "names" (R68-9, 80), and move on through "definite descriptions" (R72-6, 83) and the various kinds of "quotations" (R84-6). Then we will provide for "event descriptions"—called "indirect quotations" in European grammar—whose handling in Loglan (see R87) expresses one of the fundamental philosophical orientations of the language (namely that objects of thought and imagination are best treated as a subset of definite descriptions). Then we provide for the "indefinite descriptions" (**Ne mrenu** vs. **Le mrenu**) in R97-100 and 102; and finally the various ways of connecting arguments are given toward the end of the group; R103-16.

68: name	=> DJAN	La <b>Djan</b> , ditca = <b>John</b> is a teacher.
69:	=> name DJAN 179.	La <b>Djan Djonz</b> , ditca = <b>John Jones</b> is a teacher.

The commas in S68-9 represent "morphemic pauses", that is, the pauses required after names which permit their resolution. But this is a morphological matter, and the grammar pays no attention to such pauses. In fact the preparser eliminates all morphemic pauses before presenting a string to the parser. The morphemic pauses between the parts of a multiple name are necessary in careful speech—otherwise the lexer would hear them as single names, e.g., **Djandjonz**—but such internal pauses are not represented by commas in text. It would distract the eye to do so.

70: mex	=> RA	<b>Ra</b> da ditca = <b>All of</b> them are teachers.
71:	=> NI 170.	<b>Ne</b> da ditca = <b>One of</b> them is a teacher.

The <mex> grammar rule will be much more elaborate once the mathematicians set to work on it. But I have thought it best to keep <mex> rudimentary until the rest of the grammar is functioning smoothly.

72: descriptn	=> LE descpred	<b>Le ditca</b> pa gudbi = <b>The teacher</b> was good.
73:	=> LE mex descpred	<b>Le te ditca</b> = <b>The three teachers.</b>
74:	=> LE arg1 descpred	<b>Le le ditca gu bukcu</b> = <b>The teacher's book.</b>
75:	=> LE mex arg1	<b>Le to da ditca</b> = <b>The two of them</b> teach.
76:	=> LA descpred Local.	<b>La Ditca</b> ga sadji = <b>Teacher</b> is wise.

## Notebook 3 – Grammar

Note the "possessive" provided in R74. An alternative possessive form will be provided in R88 in which the order will be possessor-possession, as in the Romance languages: 'El libro de la maestra'. Like English ('The teacher's book' and 'The book of the teacher') Loglan has both possessive orders. As usual our policy is to provide ways of speaking every common natural word-order and let usage decide between them.

Note also, from R76, that descriptions may be used as names. When this is done the predicate words in the **la**-marked description are all capitalized in text, and the expression has, like other names, the sense of being a unique designation in that context. Thus there is only one **la Farfu** in the Smith family, just as there is only one Rat in **The Wind in the Willows**. Names may not be used as predicates, however. Thus **\*Da Djan** cannot be used to mean 'X is a John' because this expression is not parsed as a sentence. (It is \*-ed because it lacks a pause-comma before the name.) The **Djan** in **Da, Djan** as parsed by our grammar is neither a name nor a predicate but one of the free modifiers. It is a vocative, in fact, which modifies **Da**. So the proper rendering of **Da, Djan** into English is 'X, John' as if one were answering a question asked by John about the identity of something. The proper rendering of of English 'X is a John' into Loglan is **Da mela Djan**. In this expression the predifier **me** turns **la Djan** into a predicate (see R24); and **Da** is the first argument, or "subject", of that predicate.

77: arg1	=> DA	<b>Da ditca</b> = X teaches (is a teacher).
78:	=> HU	<b>Hu ditca</b> = <b>Who</b> teaches?
79:	=> TAI	<b>Tai ditca</b> = <b>T</b> teaches.
80:	=> LA name	<b>La Tam</b> , bi ba = <b>Tom</b> is (that) someone.
81:	=> LIO TAI	<b>Lio nei</b> cmalo = <b>The number n</b> is small.
82:	=> LIO mex	<b>Lio te</b> cmalo <b>lio nei</b> = <b>The number three</b> is smaller than <b>the number n</b> .
83:	=> descriptn gap	<b>Le ditca</b> ga gudbi = <b>The teacher</b> is good.
84:	=> LI (utterance LU)	<b>Li</b> , <b>Le ditca ga gudbi</b> , <b>lu steti</b> = <b>Quote Le ditci ga gudbi close-quote</b> , is a sentence. (The commas here are stylistic only, i.e., neither morphemic nor lexemic.)
85:	=> LIU (word)	<b>Liu lu</b> purda = <b>The word lu</b> is a word.
86:	=> LIE (X, string, X)	<b>Lie Dai</b> , <b>Deutschland</b> , <b>Dai</b> dotca purda = <b>The string 'Deutschland'</b> is a German word. (These commas are obligatory. 'X' is any Loglan word that does not appear in the quoted string. I usually use the initial letter of the Loglan word for the language in which string is written, in this case <b>D</b> for <b>dotca</b> .)
87:	=> LEPO sentence gap Local.	<b>Lepo da pa cluva de</b> , viekli = <b>That X loved Y</b> is clear.

The parenthetic portion of R84 is carried invisibly through the first parsing of the utterance and then parsed "re-entrantly" as an independent utterance. The parenthetic portions of R85-6 are not even seen by the parser. It is assumed that they either do not need to be (R85) or cannot be (R86) parsed.

## Notebook 3 – Grammar

88: argmod1	=> JI argument	Da <b>ji la Djan</b> , ditca = The X <b>who is John</b> , teaches. ( <b>Ji</b> identifies.)
89:	=> JI modifier	Da <b>ja vi le hasfa</b> ga kicmu = X, <b>who is in the house</b> , is a doctor. ( <b>Ja</b> predicates.)
90:	=> JIO sentence gap Local.	Da <b>jio brudi de</b> , murpeu = The X <b>who is a brother of Y</b> , is a seaman (seaperson). ( <b>Jio</b> identifies.)

**Pe** is also an allolex of JI and, like the 'de' of Romance-style possessives, puts the possessor last. Thus the phrase **pe le ditca** in **Le bukcu pe le ditca** ('El libro de la maestra') is also a production of R88.

91: argmod	=> argmod1	Da <b>jia sucmi</b> = X, <b>who (incidentally) swims</b> . ( <b>Jia</b> predicates.)
92:	=> argmod M6 A argmod1 gap Local.	Da <b>jia sucmi, e jia prano</b> - X, <b>who (incidentally) swims and runs</b> .
93: arg2	=> arg1	<b>Da = X</b> .
94:	=> arg1 argmod gap Local.	<b>Da ji de = The X who is Y</b> .
95: arg3	=> arg2	<b>Sai = S</b> .
96:	=> mex arg2 Local.	<b>Te Sai = Three of the S's</b> .
97: indef1	=> mex descpred Local.	<b>Te solda</b> pa kamla = <b>Three soldiers</b> came.
98: indef2	=> indef1 gap Local.	<b>Te solda</b> ga ditca = <b>Three soldiers</b> are teachers.

Either the comma or the **gu** of <gap> can be used here in place of **ga** to separate the first argument from its predicate. In fact **\*Te solda, ditca** parses in substantially the same way as **Te solda ga ditca** does. But I have starred it because it is probably not a good idea to adopt this usage, despite its tempting economy. The reason is that if the first argument ends in a linked argument, <gap> will not successfully separate it from an upcoming predicate; instead <gap> will allow the intended predicate to be absorbed into its first argument. For example, **Le farfu je le botci, talna** does not say that the father of the boy is an Italian, as a careless speaker might have intended; instead it is a designation of some boy-fathering Italian: 'The father-of-the-boy (type of) Italian'. Replacing the comma with **ga** conveys the intended meaning unequivocally: **Le farfu je le botci ga talna** - 'The father of the boy is Italian'. Thus the stronger marker **ga** is necessary in some cases. Since **ga** will work in all cases and <gap> only in some, it seems wiser for loglanists to habituate themselves to the always-successful marking move...even though **ga** is less economical than the pause/comma permitted by <gap> that would often do the job.

The frowned-upon uses of <gap> form another unused portion of a superset of grammatical possibilities.

99: indefinite	=> indef2	<b>Te solda = Three soldiers</b> .
100:	=> indef2 argmod gap Local.	<b>Te solda jia nigro = Three soldiers, who are (incidentally) black</b> .

## Notebook 3 – Grammar

101: arg4	=> arg3	<b>Bai = B.</b>
102:	=> indefinite	<b>Ne fumna = A (one) woman.</b>
103:	=> arg4 ZE arg3	<b>Bai ze Cai = B and C jointly.</b>
104:	=> arg4 ZE indefinite Local.	<b>Bai ze ne fumna = B and a woman jointly.</b> (Ze among arguments has the effect of forging single, teamlike entities.)
105: arg5	=> arg4	<b>Bai = B.</b>
106:	=> KA argument KI arg6 Local.	<b>Kanoi Bai ki Cai = If B then C.</b>
107: arg6	=> arg5	<b>bei groda cei = b is bigger than c.</b>
108:	=> DIO arg6	<b>Mau cei gi bei groda = Than c, b is bigger.</b> (Mau is one of the optioned case-tags of the DIO-Lexeme. Derived from <b>cmalo</b> it marks the lessers in "greater-lesser than" relationships. For the uses of the fronting operator <b>gi</b> , see R165-6.)
109:	=> IE arg6	Mau <b>ie cei gi bei groda</b> = Than <b>which c</b> is b bigger? (DIO-words are always omissible; <b>Ie cei gi bei groda</b> means the same thing. Yet <b>mau</b> is not redundant. It suggests in advance the kind of predicate that is coming up.)
110:	=> LAE arg6 Local.	Donsu da <b>lae bei</b> = Give X to <b>whoever has address b</b> (or of whom b is a sign in some other sense).
111: arg7	=> arg6	bei groda <b>cei</b> = b is bigger than c.
112:	=> arg6 ACI arg7 Local.	bei groda <b>cei, e dei, onoi fei</b> = b is bigger than c, and than <b>d or f but not both.</b> (Onoi is "exclusive or" and <b>-ci</b> turns its connectands into the right connectand of <b>e</b> , i.e., it hyphenates them.)
113: arg8	=> arg7	<b>bei groda cei = b is bigger than c.</b>
114:	=> arg8 A arg7 Local.	bei groda <b>cei, e dei, onoi fei</b> = b is bigger than c <b>and d, or f but not all three.</b> (Without <b>-ci</b> these "afterthought" connectives group left.)
115: argument	=> arg8	<b>bei groda cei = b is bigger than c.</b>
116:	=> arg8 AGE argument 117; 60, 24, 14, 8.	bei groda <b>cei, ege dei, onoi fei</b> = b is bigger than c, <b>and d or f but not both.</b> (The <b>-ge</b> suffix groups all that follows it to the right, and thus has the same effect as a

## Notebook 3 – Grammar

-**ci** suffix on the 2nd connective;  
cf. S112.)

Notice that <argument> has only one forward reference, and that is the first rule in the next group. Most references to <argument> are backward, which only means that arguments have already been widely used. Indeed, we may recall that they have been used to construct one class of modifiers, namely prepositional phrases (R60); they have been used to build one kind of predunits, namely those composed of arguments preceded by the predifier **me** (R24); and of course arguments figured in linked arguments (R8 and 14).

The single forward reference to R117 leads of course to the most common use of arguments in Loglan utterances, their contribution to the "terms" and "termsets" which are major constituents of nearly every Loglan utterance.

### Group H. Terms & Term Sets

#### Rules 117-127

A term is either a modifier or an argument. So terms are strings of one or more arguments and/or modifiers in any order. Term sets are strings of none or more terms and include the connections of such strings with one another. Group H is really two groups, both of them very small. One of them, R117-20, constructs <terms>; the other, R121-7, constructs <termset>. The uses of these two gramemes are as follows:-

Terms are used as the first arguments ("subjects") of sentences, and include any immediately following modifiers. Thus the argument-modifier pair **La Djan, na la Fomen** in **La Djan, na la Fomen, traci ti** = 'John, in May, travels here' is an instance of <terms>. But so also are the strings of fronted modifiers and/or arguments which come before the subject of a sentence. The only difference is that terms that come before a subject must be set off from it by **gi**. For example, **Ti na la Fomen** in **Ti na la Fomen, gi da traci** = 'Here in May, X travels' is also an instance of <terms>. **Gi** is the fronting operator; it announces that a string of such fronted terms has been concluded.

Termsets, in contrast, are always deployed after their predicates. For example, **ti** in **Da traci ti** is a <termset>. Indeed, all sets of "sutori" (second-and-subsequent) arguments, including any sentence modifiers with which they may be mixed (**ti fa** in **Da traci ti fa** - 'X travels here later'), or any strings of modifiers, or of arguments and modifiers mixed together, when they follow their predicates—for example, as **fa, na la Fomen** in **Da traci fa, na la Fomen** ('X travels later, in May') follows **traci**—are instances of <termset>.

As we will see, termsets are made of terms. But there are two differences in the resulting structures. First, termsets, but not terms, may be null, i.e., represented by the right boundary marker **gu** or by nothing at all; see R121. Second, termsets (but again not terms) may be connected to each other. Just why these two manoeuvres are useful can probably only be discovered by using them. But I can give you a preliminary glimpse of their utility for termsets by pointing out that they (and not terms) are grammatically attached to sentence predicates; see R128. These predicate-containing constructions then, together with their possibly null termsets, become the "predicates" of Loglan sentences; and such predicates may be connected to each other in every possible way...including ways which allow them to have "joint termsets". For example, in the English sentence 'John loved and hated Mary', 'Mary' is probably intended by the speaker to be a joint termset of the connected pair of predicates 'both loved and hated', that is, to be the "direct object" of both "verbs" (to use an older grammatical terminology). Of course we cannot be sure of this in English; but in Loglan we can be. It is just this kind of potentially ambiguous construction that in Loglan necessitates either nullifying or truncating individual termsets in order to make room for joint ones. For example, the two possible renderings of 'John loved and hated Mary' into Loglan are (i) **La Djan, pa cluva, e tsodi la Meris**, in which **la Meris** is an individual termset, and (ii) **La Djan, pa cluva, e tsodi gu la Meris**, in which it is a joint one. What has happened in (ii) is that in it **e** and **gu** have nullified the individual termsets of **cluva** and **tsodi**, respectively, and made room for a joint one.

Such manoeuvres are never required of terms. But if this structural distinction seems too intricate for easy recall, you may prefer to remember a simpler, positional one: terms occur ahead of their predicates; termsets occur after them. This surface property will take you a long way toward managing their uses properly.

## Notebook 3 – Grammar

117: term	=> argument	Da pa cluva <b>de</b> = X loved <b>Y</b> .
118:	=> modifier Local.	Da pa cluva <b>fa</b> = X loved <b>later</b> .
119: terms	=> term	Da pa cluva <b>de</b> = X loved <b>Y</b> .
120:	=> terms term 155, 157-8, 165-66, 171.	Da pa cluva <b>de fa la Ven</b> = X loved <b>Y after nine</b> . (The order of terms is syntactically unimportant. Thus <b>Da pa cluva fa la Yen, de</b> means approximately the same thing.)

Notice that all these references are forward. In effect, they list the many uses of <terms>. R155 deploys it as the delayed subject of "goa sentences" (sentences in V-O-S word order); R157 uses it as the fronted modifiers of a goa-sentence; R158 uses it as the subject, accompanied by any following modifiers, of a declarative sentence in Loglan-normal (S-V-O) word order; R165-6 arranges for its use as the fronted arguments, with accompanying modifiers, of sentences in O-S-V word order, but also as the fronted modifiers or prenex quantifiers of any utterance; and in R171 <terms> accommodates fragmentary utterances, for example, answers to **Hu**, **Vihu** and **Nahu** ('Who?', 'Where?' and 'When?') questions.

We now use <terms> to develop <termset>.

121: termset1	=> gu	Da pa cluva de, e tsodi <b>gu</b> fa la Ven = X loved Y and hates (some- one), (both) after nine. ( <b>Gu</b> represents the null termset and turns <b>fa la Ven</b> into a joint termset. A comma is insufficient here; therefore this allogram is not <gap> but <gu>.)
122:	=> terms gu Local.	Da pa cluva de fa la Ven, e tsodi <b>di gu</b> va do = X loved Y after nine, and hates W, (both) near Q. (Here <b>gu</b> truncates the 2nd termset and prepares for a joint termset, <b>va do</b> , a modifier of both predicates, to be spoken.)

Notice that it is <gu> and not <gap> that is used to truncate or nullify an individual termset. This is because the use of a comma or a pause at such points can lead to ambiguities. So it is either **gu** itself or, when it is not required, nothing at all that terminates a termset.

123: termset2	=> termset1	Da pa cluva <b>de fa</b> = X loved <b>Y later</b> .
124:	=> termset2 A termset1	Da pa cluva <b>de fa, e di pa</b> = X loved <b>Y later and W earlier</b> .
125:	=> KA termset2 KI termset1 Local.	Da pa cluva <b>ke de fa ki di pa</b> = X loved <b>both Y later and W earlier</b> .
126: termset	=> termset2	Da sanpa <b>de di vi do</b> = X is a sign of <b>Y to W at Q</b> .
127:	=> PAUSE termset2 128, 139, 141, 148, 154.	Da sanpa, <b>de di vi do</b> = X is a sign, of <b>Y to W at Q</b> . (This dis-

## Notebook 3 – Grammar

cretionary pause-comma in a place where <gap> does not occur must be explicitly provided for.)

Predicates are equipped with individual termsets in R128 and 148, and with joint termsets in R139, 141 and 154. All these applications of <termset> take place in the next group, which constructs predicates. Termsets are used for no other purpose in the grammar than to construct predicates.

### Group I. Predicates

#### Rules 128-154

In this group, the predicate expressions that make the claims of sentences are constructed. Basically they consist of a sentence predicate from Group E coupled with a termset, possibly null, from Group H; this coupling is accomplished in R128. In the sequel we'll call such expressions simply "predicates". Predicates may be "marked" or "bare", that is, prefixed by tense or abstraction operators (which includes the left-marker **ga**) or not so prefixed; see R129-31. This distinction is fundamental; it involved building two tracks through the grammar of predicates, one for bare forms, one for marked ones. For example, in a long sequence of rules, R132-42, a distinction between "backpreds" and the two kinds of "front" predicates, "barefronts" and "markfronts", is gradually developed; and in R137 and 140, this distinction is finally used. It turns out that the forms called backpreds are going to be used as the right connectands of both kinds of ekked connections, the "bareekpreds" and "markeekpreds". These are the bare and marked versions, respectively, of ekked predicates, that is, of connected predicates in which the connections are made with afterthought or A-form connectives. Barefronts and markfronts, of course, are destined to be the left or leading connectands of these same ekked predicates; and the ekked connections themselves will of course reflect the "bareness" type of their leading elements.

By R143-54, the bare vs. marked distinction is no longer important. Identity predicates ("identpreds") are now developed—BI and kin may not be treated as "just another PREDA" because they enter into little word compounds and must be recognized by the lexer—and finally kekked, or forethoughtfully connected, predicates ("kekpred") are developed. Along the way, in R135-6 and 146-7, the ACI- and AGE-forms of afterthought connections are developed in the same way that they were developed for arguments.

128: barepred	=> sentpred termset Local.	Da <b>fumna</b> = X is a woman. (Here we are using the null termset <gu>, which, because it is final, may be expressed by <err>, i.e., omitted altogether.)
129: markpred 130:	=> M7 PA barepred => PO gap sentence gap	Da <b>pa fumna</b> = X was a woman. Da <b>po de fumna</b> = X is a case of Y's being a woman. (Both <gap>s may be null here, the 2nd because it is final, the 1st because there is no following <predunit> for the normally close-binding 'po' to bind to. <b>Po</b> will not "stick" to <b>de</b> , and so does not need to be separated from it.)
131:	=> ZO gap sentence gap Local.	Da <b>zo de hatro</b> = X is the amount of heat in Y (by which Y is hot). <gap>s as in S130.)
132: backpred1	=> barepred	Da <b>fumna, e ditca</b> = X is a woman



Notebook 3 – Grammar

133:	=> markpred	and <b>a teacher</b> . (Both these connectands are barepreds.) Da fumna, e <b>pa ditca</b> = X is a woman and <b>was a teacher</b> .
134:	=> M8 NO backpred1 Local.	Da fumna, e <b>no ga blanu marpi</b> = X is a woman and <b>not a blue snake</b> . (Here <b>ga</b> extends the scope of 'no' over the whole backpred. <b>Ge</b> would also work but is bad usage.)
135: backpred	=> backpred1	Da fumna, e <b>ditca</b> = X is a woman and <b>a teacher</b> . (Backpreds are always right connectands.)
136:	=> backpred1 M9 ACI backpred Local.	Da fumna, e <b>ditca, aci stude</b> = X is a woman, and <b>a teacher or student</b> . ( <b>Aci</b> binds <b>ditca aci stude</b> into the right connectand of <b>e</b> , thus altering the normal left-grouping of after-thought connection.)
137: bareekpred	=> barefront M10 A backpred Local.	Da <b>fumna, e ditca, aci stude</b> = X is a <b>woman, and a teacher or student</b> . (This is a bare "ekpred" because its leading element is bare.)
138: barefront	=> barepred	Da <b>ditca lo dotca</b> = X is a <b>teacher of German</b> .
139:	=> bareekpred termset Local.	Da <b>ditca, e stude gu lo dotca</b> = X is a <b>teacher of, and a student of, German</b> . (Again <b>gu</b> prepares for the joint termset.)
140: markepred	=> markfront M10 A backpred Local.	Da <b>pa ditca, e cluva</b> = X <b>was a teacher and a lover</b> . (This is a marked "ekpred" because its leading element is marked.)
141: markfront	=> markpred	Da <b>pa ditca lo frasa</b> = X <b>was a teacher of French</b> .
142:	=> markepred termset Local.	Da <b>pa ditca, e cluva gu lo frasa</b> = X <b>was a teacher of, and a lover of, French</b> .
143: predicate2	=> barefront	Da <b>ditca</b> = X <b>is a teacher</b> .
144:	=> markfront	Da <b>pa ditca</b> = X <b>was a teacher</b> .
145:	=> M8 NO predicate2 Local.	Da <b>no ga blanu tcaro</b> = X <b>is not a blue car</b> . (This is long-scope negation. X may not be a car of any kind.)
146: predicate1	=> predicate2	<b>Ditca = Be a teacher!</b>
147:	=> predicate2 M11 AGE predicate1 Local.	<b>Gudbi, ege ckano, a briga = Be good, and be kind or brave!</b> (If <b>e</b> replaced <b>ege</b> , this would read 'Be good and kind, or be brave!')

## Notebook 3 – Grammar

148: identpred	=> BI termset	<b>Da bi le mrenu jio pa godzi la Frans</b> = X is the man who went to France.
149:	=> NO identpred Local.	<b>Da no bi le mrenu jio pa godzi la Frans</b> = X is not the man who went to France.
150: kekpred	=> M3 KA predicate KI predicate	<b>Ke ckano ki briga</b> = Be both kind and brave!
151:	=> NO kekpred Local.	<b>No ke ckano ki bunbo</b> = Don't be both kind and a fool!
152: predicate	=> predicate1	<b>Da mrenu</b> = X is a man.
153:	=> identpred	<b>Da bi le mrenu</b> = X is the man.
154:	=> kekpred termset 155, 158, 162.	<b>Da ke briga ki bunbo gu raba</b> = X is both brave (about everything) and a fool about everything.

Predicates, once developed, are used in constructing sentences. All three of these references are to the next group, in which sentences are made. In R155 <predicate> is used to make the Goa-sentence; in R158 it is part of the declarative sentence in normal word order; and in R162 it is the whole of the Loglan imperative.

### Group J. Sentences Rules 155-164

This small group of rules catalogues the kinds of Loglan sentences, and gives their formulas. Starting with Goa-sentences (R155), it moves on to the various forms of statements (R156-9), and finally to kekked sentences (R160-1). Finally in R162-4, imperatives, statements and kekked sentences—a variety of what are called "compound sentences" in English grammar—are bundled together into the <sentence> grameme.

155: goasent	=> GOA predicate GOA terms Local.	<b>Goa seidjo le so fumna goa to mrenu</b> = Among the six women are two men. (This is the V-O-S word-order which is standard in some languages. It is a heavily marked form in Loglan.)
156: statement	=> goasent	<b>Goa seidjo Bai goa Cai</b> = Included in B is C.
157:	=> terms goasent	<b>Vi levi kruma goa seidjo Bai goa Cai</b> = In this room, included in B is C. (A variety of the Goa-sentence with fronted modifiers.)
158:	=> terms predicate	<b>Da fumna</b> = X is a woman. (This is the Loglan-normal word-order, S-V-O. Here usage restricts the expression of <terms> to a single argument, the subject, followed by none or more modifiers.)

## Notebook 3 – Grammar

159:	=> NO statement Local.	<b>No da fumna = It is not the case that X is a woman.</b>
160: keksent	=> KA sentence KI sentence	<b>Ko da bi lio ne ki sumduo de - If and only if X is 1, add Y.</b>
161:	=> NO keksent Local.	<b>No da ko bi lio ne ki sumduo de = It is not the case that if and only if X is 1, (you should) add Y.</b>
162: sentence	=> predicate	<b>Fumna = Be a woman!</b>
163:	=> statement	<b>Da fumna = X is a woman.</b>
164:	=> keksent 174-5; 130-1, 90, 87.	<b>Kanoi ti bi la Fomen, ki fumna = If this is April, be a woman!</b>

The two forward references to <sentence> in R174-5 carry sentences into the next group, where they are only one among the many kinds of utterances. But the back references of <sentence> are more interesting. R130-1 show how sentences may be nested inside predicates by means of the abstraction operators **po pu** and **zo**; R90 shows how a sentence may be nested in an argument modifier by using the operators **jia** and **jio** (which produce the Loglan equivalents of 'John, who was the man who came to dinner' and similar constructions); and R87, which permits sentences to become the basis of event descriptions ('Before John came to dinner' and similar clauses). In these three ways, sentences may be embedded in sentences, the embedding sentences embedded in further sentence, and so on, ad infinitum.

Before moving on to utterances it may be worth a moment to consider how it came about that the two instances of <terms> in R157-8 are allowed by usage to have such different expressions. In R157 it is not good usage—in fact, it would not even be interpretable—to include any arguments at all in the optional string of fronted modifiers which usage allows <terms> to be as a possible embellishment of the Goasentence. In contrast, usage expects to find something very different as an expression of <terms> in R158. Here, <terms> is the subject of the Loglan-normal sentence; and so it must consist of exactly one argument accompanied by any number of, including zero, following modifiers. Thus, R157 and R158 generate two massive supersets of possibilities only some of which are good usage. Indeed a very different set in each domain is allowed by usage.

The reader may be interested to learn the reason why these supersets appear at just this point in the grammar. They were fashioned as an alternative to writing an extremely long and awkward, and probably always imperfect, lookahead-extending procedure into the preparser ...or rather, to abandon one that had already been written. It was first thought that such a procedure could detect without parsing whether a set of early terms was fronted, or fragmentary, or the subject of a declarative sentence plus some early modifiers, or so on; and that the M-Lexemes inserted by that procedure would then permit a grammar of sentences to be written which was devoid of these supersets. But on further study of this matter the Academy came to the conclusion that the hope of writing such a lookahead-extending algorithm was largely illusory; that exceptional cases could probably always be found that could not be handled by it; and that therefore it was preferable from the standpoint of computer science to adopt the superset solution to the "fragment handling" problem rather than attempt to solve it by elaborating the preparser algorithm in new and dubious directions.

Besides, there are good theoretical reasons to believe that just such supersets are part of every natural grammar, and that it is the job of an entirely different kind of rule, the usage rule, to declare which subdomains in the domains of the parsible are in this higher sense allowable. If so, then what we are doing with Loglan at this point is an ancient linguistical move which may be performed in all languages in approximately the same way. What is most important about this solution is that it solves the problem of parsing fragmentary utterances within the grammar, thus preserving the syntactic unambiguity of the language without dependence on the presumed completeness of a preparser algorithm.

### Group K. Utterances Rules 165-194

## Notebook 3 – Grammar

Group K is the last group of rules in Loglan grammar. In it are shown all the varieties of utterances, including sentence fragments like answers and monosyllabic responses (which occasion the supersets of Group J) as well as sentences themselves. In addition, the two major elaborations of sentences are provided in this group, namely (1) leading modifiers like **Na la Fomen, gi** in **Na la Fomen, gi mia pa godzi la Europas** and/or fronted arguments like **Dio da gi** in **Dio da gi de pa takna** - 'To X, Y spoke', and (2) the so-called "prenex quantifiers" with which logicians embellish certain kinds of claims: **Ba rabe goi** in **Ba rabe goi ba kunci be** = "There is an x such that for every y, x is related to y" ('Someone is related to everyone'). Both kinds of elaborations involve the <terms> grameme from Group H, and are carried out in R165-6 and R175.

165: headterms => terms GI

**Dio da piu de gi di pa takna = To X about Y, W spoke.** (**Dio** and **piu** are 2 of the 11 optional case tags; **dio** is derived from **dirco** and means that X is a "beneficiary or destination" of some kind; **piu** is from **pisku** and means that Y is a kind of "part, passive or topic".)

166: => headterms terms GI

**Ba rabe goi dio be gi ba takna = There is an x such that for every y, to y, x talks.** (**Goi** is a 2nd allolex of GI and fronts the "prenex quantifiers" of logicians' talk. This is an unusual word-order, but may be used.)

Usage requires that instances of <terms> used with the prenex fronting operator **goi** be limited to strings of arguments, that is, that there be no modifiers mixed in. Moreover, such arguments must obviously be matchable, except for their quantifiers, with arguments occurring in the body of the sentence to which such headterms are attached; see R175. Thus, the **ba** in **raba** is matchable with the second instance of **ba** in S166. In other words, when **goi** is the fronting operator, the grammatical possibilities opened up by Rules 165-6 form a superset of the forms accepted by usage. Usage and grammar coincide, however, when **gi** does the fronting. <Terms> used with the fronting operator **gi** in these rules may be strings of arguments and/or modifiers in any mixture.

167: uttA => A

**Enoi = And-not** (the first but not the second). (An answer to a HA-question. The monosyllabic allolexes of A, namely, **a e o u**, are also used in spelling. See R191 for **i**.)

168: => NO

**No = No.** (Answer to an **Ei**-question.)

169: => IE

**Ie = Which?**

170: => mex

**Ne. To. Te. Fo = One. Two.**

**Three. Four.** (Used either to answer **Ho**-questions—'How many?'—or in counting.)

171: => terms

**La Djan, na la Formen = John, in April.** (A pair of answers to two **Hu**-questions: 'Who? When?' It is the utterance fragments produced by this rule that

### Notebook 3 – Grammar

		creates the lookahead problem solved by the "superset solution" in R157-8. Calling subjects, fronted arguments, and answers <terms> relieves the LR1 parser of the necessity of making finer-grained decisions until it is deeper into the parse.)
172:	=> links	<b>Jue lemi sorme = Through my sister.</b> (An answer to <b>Jue hu-</b> questions.)
173:	=> linkargs	<b>Je le fotpa botci jue lemi sorme = By the fat boy through my sister.</b> (An answer to <b>Je hu-</b> questions.)
174:	=> sentence	<b>La Pit, pa takna = Pete talked.</b>
175:	=> headterms sentence Local.	<b>Dio la Djan, na lepo la Ruprt, pa kamla gi la Pit, pa takna = To John, when Rupert came, Pete talked.</b>
176: headmod	=> UI	<b>Ue nenkaa = Well, come in!</b>
177:	=> HOI (name/descriptn,)	<b>Hoi Redro Nu Herfa, nenkaa = O Red Haired One, come in!</b> (The name or description is parsed re-entrantly after the primary parse is concluded.)
178:	=> KIE (utterance KIU)	<b>Kie pazu kiu ti crina = (At last) it rained here.</b>
179:	=> name Local.	<b>Djan, nenkaa = John, come in!</b>

The four allograms of <headmod> are also the "free modifiers" of Loglan. These are the elements that may occur anywhere in a Loglan utterance, and are taken to "modify"—if that interpretation is at all meaningful—the immediately preceding word, or, if initial, the utterance as a whole. Precisely because these objects may occur anywhere, their occurrence at any given location gives the parser no information. Therefore, their incidence constitutes a kind of "grammatical noise". So except for these headmods—which are the first elements in any utterance-initial strings of them—they are removed by the preparser before the utterance is "handed to" the parser. Thus, the parser never sees attitudinal words like **ia** ('certainly') or discursives like **pou** ('however') unless one occurs at the head of some utterance. Moreover, it never sees names used vocatively, such as **Djan** without a preceding **la**, unless they too are initial in some utterance. This is the preferred position for a free modifier, by the way, because then it is taken as modifying the utterance as a whole...which is usually what is intended. Finally, except for parsing them reentrantly, i.e., when the parse of the main utterance is concluded, the parser never sees descriptive expressions used vocatively with **hoi**, or parenthetic expressions marked by **kie** and **kiu**, unless these, too, are initial in some utterance. For reentrant parsing, see the comment after the last rule in the grammar, R194.

180: uttB	=> uttA	<b>Nenkaa = Come in!</b>
181:	=> headmod	<b>Djan = John!</b>
182:	=> headmod uttA Local.	<b>Pidr, zvosto = Peter, stay out!</b>
183: neghead	=> NO gap	<b>No, na la Fomen, gi da pa kamla = It is not the case that in April, X came.</b>
184:	=> headmod NO gap	<b>Ui no, na la Fomen, gi da pa kamla</b>

Notebook 3 – Grammar

	Local.	= <b>Happily it is not the case that</b> in April, X came.
185: uttC	=> uttB	<b>Na la Fomen, gi da pa kamla = In April, X came.</b>
186:	=> neghead uttC Local.	<b>Ui no, na la Fomen, gi da pa kamla = Happily it is not the case that in April, X came.</b>
187: uttD	=> uttC	<b>De prano = Y runs.</b>
188:	=> uttC ICI UttD Local.	<b>Da sucmi, ice de prano, icanoici di fleti = X swims, and Y runs if W flies.</b> (Again the <b>-ci</b> suffix binds the 2 final sentences into the right connectand of <b>ice</b> .)
189: uttE	=> uttD	<b>Di fleti = W flies.</b>
190:	=> uttE I uttD Local.	<b>Da sucmi, ice de prano, icanoi di fleti = X swims and Y runs, if W flies.</b> (This is the normal left-grouping of afterthought connection without the special grouping effects of <b>-ci</b> or <b>-ge</b> .)
191: utterance	=> I	Fei. Lei. E. Tei. <b>I</b> = Ef. El. Ee. Tee. <b>Eye.</b> (This rule is used only to parse and generate the letter <b>i</b> when used in spelling.)
192:	=> uttE	<b>De prano = Y runs.</b>
193:	=> I uttE	<b>Icanoi di fleti = If W flies.</b>
194:	=> uttE IGE utterance	<b>Da sucmi, ige de prano, icanoi di fleti = X swims; and Y runs if W flies.</b> (Again the <b>-ge</b> suffix groups all subsequent connected utterances into a single right connectand of the connective so-altered.)

The <utterance> grameme is also used covertly—i.e., carried invisibly through the parse—by KIE in R180 and by LI in R84. These are parenthetic expressions and bits of quoted Loglan, respectively, and like the names and descriptive vocatives of R179, such embedded utterances are parsed "reentrantly". This means that after the parse of an embedding utterance is complete the parser will return to parse the embedded expression, and will continue to do so until all embedded expressions have been parsed.

## LIST 5. THE TEACHING CORPUS

(LWs listed where 1st used)

### A. Imperatives & Responses (18)

(eo ao ai ae ti tu mi no)

- A01 Hapci  
hapci  
Be Happy!
- A02 Eo skitu  
eo skitu  
Please sit!
- A03 Stali eo  
stali eo  
Stand up, please!
- A04 Hompi ti  
hompi ti  
Drink this!
- A05 Ao kerju tu  
ao (kerju tu)  
I want you to take care of yourself!
- A06 Ai mi kerju mi  
ai (mi [kerju mi])  
Yes, I'll take care of myself.
- A07 Eo helba mi  
eo (helba mi)  
Please help me!
- A08 Ao no mi helba tu  
ao (no [mi (helba tu)])  
I don't want to help you.
- A09 No takna  
no takna  
Don't talk!
- A10 Ae no takna mi  
ae ([no takna] mi)  
I wish you wouldn't talk to me.

- A11 Eo no godzi  
eo (no godzi)  
Please don't go!
- A12 No totco mi  
(no totco) mi  
Don't touch me!
- A13 No mi totco tu  
no (mi [totco tu])  
I'm not touching you.
- A14 Ae no mi puncko tu  
ae (no [mi (puncko tu)])  
I hope I'm not hurting you.
- A15 Eo no nordri mi  
eo ([no nordri] mi)  
Please don't forget me!
- A16 No eo nordri mi  
([no eo] nordri) mi  
Don't, please, forget me!
- A17 No nordri eo mi  
(no [nordri eo]) mi  
Don't forget, please, me!
- A18 No nordri mi eo  
(no nordri) (mi eo)  
Don't forget me, please!

### B. Address & Response; Offers & Replies (20)

(loi loa sia siu ea oi mu)

- B01 Djim  
djim  
Jim.
- B02 Loi  
loi  
Hello!
- B03 Eo nenkaa

## Notebook 3 – Teaching Corpus

	eo nenkaa Please enter.	B15	Ao ao I want to. (OK./Alright./etc.)
B04	Sia ai sia ai Thanks, I will.	B16	Eo mi titci ti eo (mi [titci ti]) Please, I eat this? (May I eat this?)
B05	Loi, Fred loi fred Hello, Fred.	B17	Oi oi You may. (OK./Alright./etc.)
B06	Djin, loi djin loi Jean, hello!	B18	Eo no mi titci ti eo (no [mi (titci ti)]) Please, I not eat this?
B07	Sia, Ted sia ted Thanks, Ted.	B19	Oi no oi no You may not. (It's ok if you don't.)
B08	Oi resto oi resto You may lie down.	B20	Loa, Tcet loa tcet Goodbye, Chet.
B09	Sia ao no (sia ao) no Thanks, I don't want to.		
			<b>C. Addressing vs. Naming (23)</b> <b>(la ta e hoi)</b>
B10	Oi titci ti oi (titci ti) You may eat this.	C01	Djan, santi djan santi John, be quiet!
B11	Sia ao sia ao Thanks, I want to.	C02	La Djan, santi (la djan) santi John is quiet.
B12	Siu, Selis siu selis You're welcome, Sally.	C03	Prano, Mel prano mel Run, Mel!
B13	Sia ae no (sia ae) no Thanks, I hope not to. (I'd rather not.)	C04	Prano la Mel prano (la mel) Run to Mel!
B14	Ea mu gotso ea (mu gotso) Let's go. (I suggest we go.)	C05	Prano ta, Mel



## Notebook 3 – Teaching Corpus

	prano (ta mel) Run to that, Mel!		Steven is a son of Ruth.
C06	Prano ta, Mel, e ti, Pol prano ([ta mel] e [ti pol]) Run to that, Mel, and to this, Paul!	C17	La An, takna la Ritcrd, mi (la an) (takna [(la ritcrd) mi]) Anne talks to Richard about me.
C07	Keit, skesa mi keit (skesa mi) Kate, kiss me!	C18	An, takna la Ritcrd, mi an (takna [(la ritcrd) mi]) Ann, talk to Richard about me.
C08	Skesa, Keit, mi (skesa keit) mi Kiss, Kate, me!	C19	La Djeik, farfu, Djan, mi la Suzn (la djeik) ([farfu djan] [mi (la suzn)]) Jake is the father, John, of me through Susan.
C09	Skesa mi, Keit skesa (mi keit) Kiss me, Kate!	C20	La Djeik, farfu la Djan, mi, Suzn (la djeik) (farfu [(la djan) (mi suzn)]) Jake is the father of John by me, Susan.
C10	Skesa mi Hoi Keit skesa (mi [hoi keit]) Kiss me, O Kate!	C21	Mel, prano la Djan Pol Djonz mel (prano [la <djan pol> djonz]) Mel, run to John Paul Jones!
C11	Skesa la Bab, Hoi Keit skesa (la [bab (hoi keit)]) Kiss Bob, O Kate!	C22	Prano la Djan Pol Djonz Mel prano (la [<djan pol> djonz] mel) Run to John Paul Jones Mel!
C12	Skesa la Bab Keit skesa (la [bab keit]) Kiss Bob Kate (a person by the name of 'Bob Kate').	C23	Prano la Djan Pol Djonz, Hoi Mel prano (la [(djan pol) (djonz <hoi mel>)]) Run to John Paul Jones, O Mel!
C13	Karl, kukra la Dik karl (kukra [la dik]) Carl, be faster than Dick.	<b>D. Descriptions (11)</b> <b>(le ne su gu)</b>	
C14	La Karl, kukra la Dik (la karl) (kukra [la dik]) Carl is faster than Dick.	D01	Bleka le nirda bleka (le nirda) Look at the bird!
C15	La Stivn, sunho, Ruf (la stivn) (sunho ruf) Steven is a son, Ruth.	D02	Eo penso le nable eo (penso [le nable]) Please think about the problem.
C16	La Stivn, sunho la Ruf (la stivn) (sunho [la ruf])	D03	Eo penso ne nable eo (penso [ne nable])

Notebook 3 – Teaching Corpus

	Please think about one problem (i.e., exactly one problem).		You're what? (How are you?)
D04	Eo penso su nable eo (penso [su nable]) Please think about a problem (i.e., at least one problem).	E02	Mi djela mi djela I'm well.
D05	Ai mi ckano le bunbo ai (mi [ckano (le bunbo)]) I will be kind to the fool.	E03	Ta he ta he That's what? (What's that?)
D06	Mi driki le purda mi (driki [le purda]) I remember the word.	E04	Da muzgi da muzgi It's music.
D07	La Celdn, godzi le sitci (la celdn) (godzi [le sitci]) Sheldon goes to the city.	E05	Le cersi ga he (le cersi) (ga he) The chair is what? (How's the chair?)
D08	Le rodlu ga gudbi le fitrua (le rodlu) (ga [gudbi (le fitrua)]) The road is better than the path.	E06	Da komfu da komfu It's comfortable.
D09	Le ditca ga fundi mi le groda (le ditca) (ga [fundi (mi <le groda>)]) The teacher likes me better than (he does) the big one.	E07	Da he komfu da (he komfu) It's how comfortable? (How comfortable is it?)
D10	Le troli ga plizo la Frenk, ne mipli (le troli) (ga [plizo <la frenk> <ne mipli>]) The controller uses Frank as one example.	E08	Da nurmue komfu da (nurmue komfu) It's moderately comfortable.
D11	Le tcaro ga djipo le ponsu le kolro (le tcaro) (ga [djipo (<le ponsu> <le kolro>)]) The car is important to the owner for the color.	E09	Le nurci he le (nurtci he) The edible what?
		E10	Le nurtci herba le (nurtci herba) The edible plant.
		E11	Le botci ga he sucmi (le botci) (ga [he sucmi]) The boy is what kind of swimmer?
		E12	Le botci ga tarle sucmi (le botci) (ga [tarle sucmi]) The boy is a tired swimmer.
	<b>E. Questions with he; Demonstratives &amp; Plurals; Replacement with da (18) (he da na ri ro levi leva)</b>		
E01	Tu he tu he		

### Notebook 3 – Teaching Corpus

E13	Leva he tcaro leva (he tcaro) That what-kind-of car?	F05	La Bab, bi le hapci (la bab) (bi [le hapci]) Bob is the happy one.
E14	Leva tcaro he leva (tcaro he) That car-type-of what?	F06	Da hapci hu da (hapci hu) He's happy about what?
E15	Levi ri herba ga he (levi ri herba) (ga he) These few plants are what?	F07	Le ckela le ckela The school.
E16	Da vendu da vendu They're poisonous.	F08	Ie le ckela ie (le skela) What school? (Which "the school"?)
E17	Leva ro juntii na he (leva ro juntii) (na he) Those many young ones are now (doing/being) what?	F09	Le cninu ckela le (cninu ckela) The new school.
E18	Da na takna da (na takna) They're now talking.	F10	Ta bi le cninu ckela ta (bi [le (cninu ckela)]) That is the new school.
<p><b>F. Identity Questions &amp; Sentences; Replacement with de &amp; dui (25) (ie bi hu i de dui)</b></p>		F11	La Selis, cluva (la selis) cluva Sally loves. (Sally is in love.)
F01	Da mrenu da mrenu He's a man.	F12	Da cluva hu da (cluva hu) She loves whom?
F02	Ie da ie da Which he?	F13	Da cluva la Pit da (cluva [la pit]) She loves Pete.
F03	La Bab la bab Bob.	F14	I de cluva hu i (de [cluva hu]) And he loves whom?
F04	Ie la Bab ie (la bab) Which Bob?	F15	La Alis la alis Alice.
		F16	Ie la Alis

Notebook 3 – Teaching Corpus

	ie (la alis) Which Alice?	G01	Ei ti breba ei (ti breba) Is this bread?
F17	La Muhamed Alis la (muhamed alis) Muhamed Ali.	G02	Ia da breba ia (da breba) Yes, it's bread.
F18	Da godzi la Italias da (godzi [la italias]) He goes to Italy.	G03	Ei ti ckela ei (ti ckela) Is this a school?
F19	Da godzi de hu da (godzi [de hu]) He goes to it from what? (He goes there from where?)	G04	No. I da hasfa no . i (da hasfa) No; it's a house.
F20	Da godzi de la Danmark da (godzi [de (la danmark)]) He goes to it from Denmark. (He goes there from Denmark.)	G05	Ei ta hasfa ei (ta hasfa) Is that a house?
F21	Hu matma leva botci hu (matma [leva botci]) Who is the mother of that boy?	G06	Ia no. I da vemsia ia no . i (da vemsia) Certainly not; it's a store.
F22	Ti dui ti dui This (person) is.	G07	Ei tu cidja ei (tu cidja) Are you awake?
F23	Mi gudbi tu mi (gudbi tu) I'm better than you.	G08	No. I mi sonli no . i (mi sonli) No; I'm asleep.
F24	Tu gudbi mi hu tu (gudbi [mi hu]) You're better than me at/for what?	G09	Ei tu cluva mi ei (tu [cluva mi]) Do you love me?
F25	Mi gudbi tu ti mi (gudbi [tu ti]) I'm better than you are at/for this.	G10	Ia mi cluva tu ia (mi [cluva tu]) Yes, I love you.
		G11	Ei toi tradu ei (toi tradu) Is that true? (Toi refers to the last remark.)
		G12	Ei toa logla

**G. Yes/No Questions & Answers; Utterance  
Demonstratives (12)  
(ei ia toi toa)**

Notebook 3 – Teaching Corpus

ei (toa logla)

Was that Loglan? (Toa refers to an earlier remark.)

H10 Mi penso toi fa  
mi (penso [toi fa])  
I will think about that later.

**H. Tenses; Time Questions & Answers; Local Modification; Punctuation (22)  
(pa fa ji ipou nahu)**

H01 Ei tu na bleka mi  
ei (tu [na (bleka mi)])  
Are you now looking at me?

H11 Na gi mi penso da  
(na gi) (mi [penso da])  
Now I'm thinking about it.

H02 No. Ipou mi fa bleka tu  
no . ipou (mi [fa (bleka tu)])  
No. But I will look at you.

H12 Na mi, penso da  
(na mi) (penso da)  
At me, think about it. (When I do, think about it.)

H03 Ei tu pa tidjo mi  
ei (tu [pa (tidjo mi)])  
Were you heavier than me?

H13 Na mi penso da  
(na mi) (penso da)  
At me, think about it. (When I do, think about it.)

H04 Ia mi pa mutce tidjo tu  
ia (mi [pa (<mutce tidjo> tu)])  
Yes, I was much heavier than you.

H14 Na gi la Djan, totco tu  
(na gi) ([la djan] [totco tu])  
Now John is touching you.

H05 Eo tu fa helba mi  
eo (tu [fa (helba mi)])  
Please, will you (later) help me?

H15 Na la Djan, totco tu  
(na [la djan]) (totco tu)  
When John (does), touch yourself.

H06 Ai mi fa helba tu  
ai (mi [fa (helba tu)])  
Yes, I will (later) help you. (Both senses of 'will' are involved.)

H16 Mi na totco tu  
mi (na [totco tu])  
I'm now touching you.

H07 Mi groda tu  
mi (groda tu)  
I'm bigger than you.

H17 Mi ji na, totco tu  
(mi [ji na]) (totco tu)  
I, as of now, touch you. (I.e., the present-I touches you.)

H08 Tu nahu groda mi  
(tu nahu) (groda mi)  
You as of when are, were, or will be, bigger than me?

H18 Mi na fundi tu ji pa  
mi (na [fundi (tu <ji pa>)])  
I now like you as you were before.

H09 Mi groda tu fa ti  
mi (groda [tu (fa ti)])  
I will be bigger than you hereafter (after this).

H19 Mi no fundi tu pa  
mi ([no fundi] [tu pa])  
I didn't like you then.

H20 Ei tu na saadja toi  
ei (tu [na (saadja toi)])

## Notebook 3 – Teaching Corpus

	Do you now understand this?	I09	Fa le cimra fa (le cimra) After the summer.
H21	Ia mi dui ia (mi dui) Yes I do.	I10	Fazi fazi Right away.
H22	No mi dui no (mi dui) No I don't.	I11	Fa la Nevesonin fa (la nevesonin) After 1960.
<b>I. Time Phrases (21)</b>			
<b>(pahu fahu fazi pazu tiu)</b>			
I01	Ai mi ditca ai (mi ditca) I intend to teach. (I'm going to teach.)	I12	Na tiu gi mi pa clafo ([na tiu] gi) (mi [pa clafo]) At that (reported event), I laughed.
I02	Nahu nahu At what time/date/period?	I13	Fa ta gi ti rodja ([fa ta] gi) (ti rodja) After that, this (will) grow.
I03	Fa la Netomen fa (la netomen) After December.	I14	Fa ta, rodja (fa ta) rodja After that, grow!
I04	Pa la Torin pa (la torin) Before the Second.	I15	Fa gi ti rodja (fa gi) (ti rodja) Later this (will) grow.
I05	Na le monza na (le monza) In the morning.	I16	Ti fa rodja ti (fa rodja) This will grow.
I06	Pahu pahu Before when?	I17	Mi kicmu pazu tu mi (kicmu [pazu tu]) I was a doctor long before you (were).
I07	Pa le natli pa (le natli) Before (the) night.	I18	Ai mi fa traci na la Nenimen ai (mi [fa (traci <na (la nenimen)>])) I intend to travel in October.
I08	Fahu fahu After when?	I19	Na la nenimen gi mi fa traci la Frans ([na (la nenimen)] gi) (mi [fa (traci <la frans>)]) In October I will travel to France.
		I20	Mi fa traci da pa la Netomen

Notebook 3 – Teaching Corpus

mi (fa [traci (da <pa (la netomen)>))]  
I will travel to it before December.

vi (le hasfa)  
In the house.

I21 Mi fa traci pa la Netomen da  
mi (fa [traci <pa (la netomen)> da])  
I will travel before December to it.

J10 Vahu  
vahu  
Near where?

**J. Space Questions & Answers; Space Phrases  
(17)  
(vi va vu vihu vahu vuhu)**

J01 La Djan, titci vi  
(la djan) (titci vi)  
John eats here.

J11 Va la Rain  
va (la rain)  
Near the Rhine.

J02 Ei tu stolo vi  
ei (tu [stolo vi])  
Do you stay here?

J12 Vuhu  
vuhu  
Far from where?

J03 No. I mi stolo va  
no . i (rai [stolo va])  
No; I stay there.

J13 Vu le mursi  
vu (le mursi)  
Far from the sea.

J04 Ei tu stude vi  
ei (tu [stude vi])  
Are you a student here?

J14 Vi la Ditroit, gi da pa malbi  
([vi (la ditroit)] gi) (da [pa malbi])  
In Detroit he was sick.

J05 No. I mi stude vu  
no . i (mi [stude vu])  
No; I am a student far away.

J15 Da fa zbuma va le vrici  
da (fa [zbuma (va <le vrici>))]  
It will explode near the

J06 Ei tu ji vi, stude  
ei ([tu (ji vi)] stude)  
Are you, as are here, a student? (I.e., the  
part of you that is here?)

J16 Da na genbalci de vu le  
monca  
da (na [genbalci (de <vu (le monca)>))]  
J17 Da na genbalci de ji vu le monca  
da (na [genbalci (de <ji (vu [le  
monca])>))]  
He's now rebuilding the one (that is) far  
from the mountain.

J07 Vihu tu sonli  
vihu (tu sonli)  
Where do you sleep?

**K. Existentials & Universals; Completion (34)  
(ba be bo bu raba rabe rabo nibe ifeu  
inusoa vina uu)**

J08 Va tu  
va tu  
Near you.

K01 Ba najda  
ba najda  
Something is a knife. (There are knives.)

J09 Vi le hasfa

Notebook 3 – Teaching Corpus

<p>K02 Ei ba smano ei (ba smano) Is something smoke? (Is there smoke?)</p>	<p>K12 Ba nilboi na la Nevevonin ba (nilboi [na (la nevevonin)]) There are children in 1980. (There were /will be children in 1980.)</p>
<p>K03 Ia ba smano ia (ba smano) Yes, something is smoke. (Yes, there is smoke.)</p>	<p>K13 Ei raba cninu ei ([ra ba] cninu) Is everything new?</p>
<p>K04 Ei ba breba ei (ba breba) Is something bread? (Is there any bread?)</p>	<p>K14 No. I ba no cninu no . i (ba [no cninu]) No; something is not new. (No; some things are not new.)</p>
<p>K05 Ia. I ba breba vi ia . i (ba [breba vi]) Yes; something is bread here. (There's bread here.)</p>	<p>K15 Ei raba cluva rabe ei ([ra ba] [cluva (ra be)]) Does everybody love everybody else?</p>
<p>K06 Ei ba bukcu vi ei (ba [bukcu vi]) Is something a book here? (Are there any books here?)</p>	<p>K16 No. I ba no cluva rabe no . i (ba [(no cluva) (ra be)]) No; there are some who don't love everyone else.</p>
<p>K07 Uu no ba bukcu vi uu (no [ba (bukcu vi)]) Sorry, nothing is a book here. (There are no books here.)</p>	<p>K17 Ifeu ba no cluva be ifeu (ba [(no cluva) be]) In fact, there are some who don't love anyone else.</p>
<p>K08 No ba cutri vina no (ba [cutri vina]) Nothing is water here now. (There is no water here now.)</p>	<p>K18 Ei raba kunci be ei ([ra ba] [kunci be]) Is everyone related to someone else?</p>
<p>K09 Vihu ba najda vihu (ba najda) Where is something a knife? (Where are there knives?)</p>	<p>K19 No. I ba kunci nibe no . i (ba [kunci (ni be)]) No; some are related to no one else (to zero somethings).</p>
<p>K10 Ba najda vi levi drara ba (najda [vi (levi drara)]) Something is a knife in this drawer. (There are knives in this drawer.)</p>	<p>K20 Raba gudbi be bo (ra ba) (gudbi [be bo]) Everyone is better than someone at/for something.</p>
<p>K11 Nahu ba nilboi nahu (ba nilboi) When is something a child? (When were there/will there be children?)</p>	<p>K21 Ba paslinkui rabe bo ba (paslinkui [(ra be) bo]) Someone is an ancestor of everyone through someone.</p>



## Notebook 3 – Teaching Corpus

K22	Ba vegri be rabo ba (vegri [be (ra bo)]) Something is greener than something else to everyone.		Therefore it is a ticket to somewhere from somewhere on some carrier for some price.
K23	La Djan, godzi la Frans (la djan) (godzi [la frans]) John goes to France.	K33	Mi no vizka mi (no vizka) I do not (i.e., cannot) see.
K24	Inusoa de godzi da ba inuso (de [godzi (da ba)]) Therefore he goes there from somewhere.	K34	Inusoa mi no vizka raba rabe inuso (mi [(no vizka) «ra ba» <ra be>]) Therefore I do not see anything against any background.
K25	Ti groda ti groda This is big.	<b>L. Predicate Strings; Grouping, Hyphenation, Connection &amp; Inversion (22)</b> <b>(ge go ci ce ke ki)</b>	
K26	Inusoa da groda ba be inuso (da [groda (ba be)]) Therefore it's bigger than something by some amount.	L01	La Frans, grada gunti (la frans) (grada gunti) France is a great country.
K27	Ti bitsa ti bitsa This is in between.	L02	Da he grada gunti da ([he grada] gunti) It's how great a country? (I.e., what kind or degree of greatness is involved?)
K28	Inusoa da bitsa ba be inuso (da [bitsa (ba be)]) Therefore it's between something and something else.	L03	Da bilca grada gunti da ([bilca grada] gunti) It's a militarily great country.
K29	Ti vedma ti vedma This (person) is a seller.	L04	La Spat, simba janto kangu (la spat) ([simba janto] kangu) Spot is a lion-hunting dog.
K30	Inusoa da vedma ba be bo inuso (da [vedma (<ba be> bo)]) Therefore he sells something to someone at some price.	L05	Da he simba janto kangu da ([he simba] janto] kangu) He's a what-kind-of-lion hunting dog? (Used for what kind of lions?)
K31	Ti racketpi ti racketpi This is a travel-ticket.	L06	Da frika simba janto kangu da ([frika simba] janto] kangu) He's an African-lion-hunting dog.
K32	Inusoa da ketpi ba be bo bu inuso (da [ketpi (<(ba be) bo> bu)])	L07	Da cmalo ge janto kangu da (cmalo [ge (janto kangu)]) He's small for a hunting dog.

## Notebook 3 – Teaching Corpus

- |   |  |
|---|--|
| <p>L08 Da cmalo janto ci kangu<br/>da (cmalo [janto ci kangu])<br/>He's a small hunting-dog.</p>  | <p>L18 Da kubra kanpi ci tcaro rodlu<br/>da ([kubra (kanpi ci tcaro)] rodlu)<br/>It's a wide race-car road. (I.e., a road for<br/>wide race-cars.)</p> |
| <p>L09 La Mbelis, dorja cefli<br/>(la mbelis) (dorja cefli)<br/>Mbeli is a war chief.</p>   | <p>L19 La Djan, prano go kukra la Djek<br/>(la djan) ([prano go kukra] [la djek])<br/>John runs faster than Jack.</p>                                  |
| <p>L10 Da he dorja cefli<br/>da ([he dorja] cefli)<br/>He's what-kind-of-war chief? (A chief for<br/>what kind of wars?)</p>  | <p>L20 Da botsu go vedji groda<br/>da (botsu go [vedji groda])<br/>It's a boat that is very large.</p>   |
| <p>L11 Da he ge dorja cefli<br/>da (he [ge (dorja cefli)])<br/>He's what kind of a war-chief? (What's he<br/>like as a war-chief?)</p>  | <p>L21 Da slano takna go staryfoa<br/>da ([slano takna] go staryfoa)<br/>He's a slow talker who is full of surprises.</p>                              |
| <p>L12 Da ckano ge dorja cefli<br/>da (ckano [ge (dorja cefli)])<br/>He's kind for a war-chief.</p>   | <p>L22 Da ge slano takna go staryfoa<br/>da ([ge (slano takna)] go staryfoa)<br/>He's, for a slow talker, full of surprises.</p>                       |
| <p><b>M. More Connections &amp; Groupings in<br/>Predicate Strings (26)<br/>(gue cui canoi ka kanoi)</b></p>  |  |
| <p>L13 Da junti ge terla famva bilra<br/>da (junti [ge &lt;terla famva&gt; bilra])<br/>He's young for a world-famous billiard-<br/>player.</p>                                    | <p>M01 Da bilti cmalo ce nirli<br/>da (bilti [cmalo ce nirli])<br/>She's (a) beautifully small (thing) and a<br/>beautiful girl.</p>                   |
| <p>L14 Da kusti ge slano torkrilu ci dampa<br/>da (kusti [ge (slano &lt;torkrilu ci dampa&gt;)])<br/>It's expensive for a slow bicycle- pump.<br/>(I.e., a slow-acting pump.)</p> | <p>M02 Da bilti ke cmalo ki nirli<br/>da (bilti [ke cmalo ki nirli])<br/>She's beautifully both small and a girl.</p>                                  |
| <p>L15 Da junti famva ce terla ci sadji<br/>da (junti [famva ce (terla ci sadji)])<br/>He's young to be famous and worldly-<br/>wise.</p>   | <p>M03 Da bilti ge cmalo nirli gue ckela<br/>da ([bilti (ge &lt;cmalo nirli&gt; gue)] ckela)<br/>It's a beautiful small-girls' school.</p>             |
| <p>L16 Da junti ke famva ki terla ci sadji<br/>da (junti [ke famva ki (terla ci sadji)])<br/>He's young to be both famous and<br/>worldly-wise.</p>                               | <p>M04 Da bilti cmalo ge nirli ckela<br/>da ([bilti cmalo] [ge (nirli ckela)])<br/>It's beautifully small for a girls' school.</p>                     |
| <p>L17 Da junti ce terla ci famva sadji<br/>da ([junti ce (terla ci famva)] sadji)<br/>He's a young (and) world-famous sage.</p>  | <p>M05 Da bilti cmalo nirli ci ckela<br/>da ([bilti cmalo] [nirli ci ckela])<br/>It's a beautifully small girls-school.</p>                            |

### Notebook 3 – Teaching Corpus

- |  |  |
|--|--|
| <p>M06 Da bilti ge cmalo ge nirli ckela<br/>da (bilti [ge (cmalo &lt;ge (nirli ckela)&gt;)])<br/>It's beautiful for a small type of girls' school, i.e., for a girls' school which is small.</p>           | <p>M15 Da bilti ci cmalo ce nirli ckela<br/>da ([bilti ci cmalo) ce nirli] ckela)<br/>It's a beautifully-small and a girls' school.</p>                                    |
| <p>M07 Da bilti cmalo ci nirli ci ckela<br/>da (bilti [cmalo ci (nirli ci ckela)])<br/>It's a beautiful small-girls-school.</p>  | <p>M16 Da bilti canoi cmalo ci nirli ckela<br/>da ([bilti canoi (cmalo ci nirli)] ckela)<br/>It's a beautiful if for small-girls school.</p>                               |
| <p>M08 Da bilti ce cmalo nirli ckela<br/>da ([bilti ce cmalo) nirli] ckela)<br/>Its a beautiful-and-small girls' school, i.e., for girls who are both beautiful and small.</p>                             | <p>M17 Da bilti ce cmalo ge nirli ckela<br/>da ([bilti ce cmalo] [ge (nirli ckela)])<br/>It's beautiful and small for a girls' school.</p>                                 |
| <p>M09 Da bilti cmalo ca nirli ckela<br/>da ([bilti (cmalo ca nirli)] ckela)<br/>It's a beautifully small and/or a (beautiful) girls', school.</p>   | <p>M18 Da bilti ce cmalo nirli ci ckela<br/>da ([bilti ce cmalo] [nirli ci ckela])<br/>It's a beautiful, and a small, girls-school.</p>                                    |
| <p>M10 Da bilti ka cmalo ki nirli ckela<br/>da ([bilti (ka cmalo ki nirli)] ckela)<br/>It's a beautifully either small and/or a girls', school.</p>  | <p>M19 Da bilti cui cmalo nirli ce ckela<br/>da (bilti [cui (cmalo nirli) ce ckela])<br/>It's a beautiful small-girl and (a beautiful) school.</p>                         |
| <p>M11 Da bilti cmalo nirli ce ckela<br/>da ([bilti cmalo] [nirli ce ckela])<br/>It's a beautifully small girl-and-school, i.e, it's both a beautifully small girl and a beautifully small school (!).</p> | <p>M20 Da bilti cmalo ci nirli ce ckela<br/>da (bilti [(cmalo ci nirli) ce ckela])<br/>It's a beautiful small-girl and school.</p>   |
| <p>M12 Da bilti ge cmalo ce nirli ckela<br/>da (bilti [ge «cmalo ce nirli» ckela])<br/>It's beautiful for a small, and for a girls', school, i.e., for a school which is small and for girls.</p>          | <p>M21 Da bilti cmalo ce ge nirli ckela<br/>da (bilti [cmalo ce (ge &lt;nirli ckela&gt;)])<br/>It's (a) beautifully small (thing) and (beautiful for a) girls' school.</p> |
| <p>M13 Da bilti ge cmalo nirli ce ckela<br/>da (bilti [ge (cmalo &lt;nirli ce ckela&gt;)])<br/>It's beautiful for a small girl and for a small school.</p>   | <p>M22 Da bilti ce cmalo ca nirli ckela<br/>da ([bilti ce cmalo) ca nirli] ckela)<br/>It's a beautiful and a small, or it's a girls', school.</p>                          |
| <p>M14 Da cui bilti cmalo ce nirli ckela<br/>da ([cui (bilti cmalo) ce nirli] ckela)<br/>It's a beautifully small, and a girls', school.</p>   | <p>M23 Da bilti ce kanoi cmalo ki nirli ckela<br/>da ([bilti ce (kanoi cmalo ki nirli)] ckela)<br/>It's a beautiful, and, if a small then a girls', school.</p>            |
|  | <p>M24 Da bilti ce cmalo nirli ce ckela<br/>da ([bilti ce cmalo] [nirli ce ckela])<br/>It's a beautiful-and-small girl-and-school.</p>                                     |
|  | <p>M25 Da bilti cmalo ce nirli ca nirli<br/>da (bilti [(cmalo ce nirli) ca nirli])</p>   |

## Notebook 3 – Teaching Corpus

	It's a beautifully small thing and a beautiful girl, or it's a beautiful school.	N09	Da po, corta mrenu da (po [corta mrenu]) It's a case of being a short man.
M26	Da bilti cmalo ce ka nirli ki ckela da (bilti [cmalo ce (ka nirli ki ckela)]) It's a beautifully small thing, and either a (beautiful) girl or a (beautiful) school.	N10	Da pa mutce po ge corta mrenu da (pa [mutce (po <ge (corta mrenu)>)]) It was an extreme case of being a short man.
<b>N. Event/State Predicates; Other Abstractions (16) (po pu zo di)</b>		N11	Da po mi corta la Djan da (po [mi (corta <la djan>)]) It's a case of my being shorter than John.
N01	Da zbuma da zbuma It explodes (is an explosive).	N12	Da po la Pit, mutce corta la Djan da (po [(la pit) <mutce corta> <la djan>)]) It's a case of Pete's being much shorter than John.
N02	Da po zbuma da (po zbuma) It's an explosion (an event of exploding).	N13	Da pu de fotli di da (pu [de (fotli di)]) X is a (the) property of Y being stronger than W.
N03	Da pa clado po zbuma da (pa [clado (po zbuma)]) It was a loud explosion.	N14	Da pu de fotli da (pu [de fotli]) X is an instance of Y's strength.
N04	Da po mrenu da (po mrenu) It's a manhood (a state of being a man).	N15	Da zo de blanu di da (zo [de (blanu di)]) X is an (the) amount by which Y is bluer than W.
N05	Da pa corta po mrenu da (pa [corta (po mrenu)]) It was a short manhood.	N16	Da zo de blanu da (zo [de blanu]) X is an (the) amount of blue in Y.
N06	Da pa corta po mrenu bukcu da (pa [(corta <po mrenu>) bukcu]) It was a short-manhood book. (A book about short manhoods?)	<b>O. Mass &amp; Event Descriptions; Mass Event Descriptions (27) (lo lovi lepo lopo)</b>	
N07	Da pa corta ge po mrenu bukcu da (pa [corta (ge <(po mrenu) bukcu>)]) It was short for a manhood book.	O01	Ei da fundi lo malna ei (da [fundi (lo malna)]) Is he fond of milk? (Does he like milk?)
N08	Da po ge corta mrenu da (po [ge (corta mrenu)]) It's an event, state, or case of being a short man.	O02	Ei tu pa titci lo nikri ei (tu [pa (titci <lo nikri>)])

Notebook 3 – Teaching Corpus

Did you eat cheese?

- O03 Ei tu fa janto lo simba  
ei (tu [fa (janto <lo simba>)])  
Will you (be) hunt(ing) lions? (Are you going to hunt lions?)
- O04 Ia mi fa janto da  
ia (mi [fa (janto da)])  
Yes, I'll (be) hunt(ing) them. (Yes, I'm going to hunt them.)
- O05 Ei lo nirda vi sucmi  
ei ([lo nirda] [vi sucmi])  
Do birds here-swim? (I.e., swim here, in this place?)
- O06 Ei lo nirda vi, sucmi  
ei ([lo nirda] [vi] sucmi)  
Do birds here swim? (More global; does the event of birds swimming take place here?)
- O07 Ei lo nirda ji vi, sucmi  
ei ([lo nirda] [ji vi] sucmi)  
Do (the) birds here swim? (I.e., the local birds?)
- O08 Ei lovi nirda ga sucmi  
ei ([lovi nirda] [ga sucmi])  
Do local birds (here-birds) swim?
- O09 No. I da fleti  
no . i (da fleti)  
No; they fly.
- O10 Leva sonda ga he  
(leva sonda) (ga he)  
That sound was what?
- O11 Lopo zbuma  
lopo zbuma  
Explosion(s).
- O12 Ei tu pa hirti lepo zbuma  
ei (tu [pa (hirti <lepo zbuma>)])  
Did you hear the explosion?
- O13 Ia mi pa hirti da. I lopo zbuma ga clado  
ia (mi [pa (hirti da)]) . i ([lopo zbuma] [ga clado])  
Yes, I heard it. Explosions are loud.
- O14 Ei tu pa hirti le clado po zbuma  
ei (tu [pa (hirti <le (clado [po zbuma]>)]))  
Did you hear the loud explosion?
- O15 Ei tu hirti lo clado po zbuma  
ei (tu [hirti (lo <clado (po zbuma) >)])  
Do you hear loud explosions?
- O16 Ei tu fundi lopo sucmi  
ei (tu [fundi (lopo sucmi)])  
Do you like swimming?
- O17 Ei lepo prano pa nardu  
ei ([lepo prano] [pa nardu])  
Was the run difficult?
- O18 Ei lopo nilboi ga treci  
ei ([lopo nilboi] [ga treci])  
Is childhood interesting?
- O19 Ei lopo mormao lo simba ga nardu  
ei ([lopo (mormao <lo simba>)] [ga nardu])  
Is killing lions difficult?
- O20 Lo ficli ga spalii lo mursi  
(lo ficli) (ga [spalii (lo mursi)])  
Fish live in the sea.
- O21 Lo tongu ga trime lopo takna  
(lo tongu) (ga [trime (lopo takna)])  
The tongue is the tool of speech.
- O22 Lo tcina ga hasfa lo tongu  
(lo tcina) (ga [hasfa (lo tongu)])  
The chin is the house of the tongue.
- O23 Lopo takna ga proju lo sonda  
(lopo takna) (ga [proju (lo sonda)])  
Speech produces sound.

Notebook 3 – Teaching Corpus

O24	Lopo dirlu ga ckozu lopo kecri (lopo dirlu) (ga [ckozu (lopo kecri)]) Loss is a cause of sadness.		mi (krido [lepo (<la paris> <garsitci (la frans)>)]) I believe that Paris is the capital of France.
O25	Lepo kanpi sucmi pa valna (lepo [kanpi sucmi]) (pa valna) The swimming-meet (event of competitive swimming) was violent.	P07	Mi djacue lepo la Marz, redro levi carta mi (djacue [lepo (<la marz> Credro (levi carta)>)]) I claim that Mars is redder than this shirt.
O26	Le, po kanpi sucmi pa valna (le [(po kanpi) sucmi]) (pa valna) The competition swimmer was violent.	P08	Le stude pa djadou le surva lepo da pluci de (le stude) (pa [djadou (<le surva> <lepo (da [pluci de])>)]) The student informed the servant that he/she (the servant) pleased him/her (the student).
O27	Lo, po kanpi sucmi ga valna (lo [(po kanpi) sucmi]) (ga valna) Competition swimmers are violent.		
<p><b>P. Specified &amp; Nested Event Descriptions (22)</b> <b>(No new LWs.)</b></p>		P09	La Ruprt, djacue lepo la Denvr, bitsa le vrici lepo kamda (la ruprt) (djacue [lepo <la denvr> <bitsa ([le vrici] [lepo kamda])>)]) Rupert claims that Denver is between the river and the fighting.
P01	Ei tu djano lepo mi stude ei (tu [djano (lepo <mi stude>)]) Do you know that I am a student?	P10	La Dag, furmoi lepo trati lepo skesa la Meris (la dag) (furmoi [lepo (trati <lepo (skesa [la meris])>)]) Doug intends to try to kiss Mary.
P02	Tu danza lepo jmite hu tu (danza [lepo (jmite hu)]) You want to meet whom?	P11	Ei la Degol, pa krido lepo kanmo lepo rilnamdou la Frans ei ([la degol] [pa (krido <lepo (kanmo [lepo (rilnamdou <la frans>)])>)]) Did de Gaulle believe that (he) could glorify France?
P03	Le nirli pa takna mi lopo da traci (le nirli) (pa [takna (mi <lopo (da traci)>)]) The girl talked to me about her travels.	P12	Lepo le mreanu pa trati ga fekto (lepo [(le mreanu) (pa trati)]) (ga fekto) That the man tried is a fact.
P04	Le farfu pa takna le detra lepo nu fatru (le farfu) (pa [takna (<le detra> <lepo (de [nu fatru])>)]) The father talked to the daughter about his being troubled.	P13	Lepo le botci pa fundi lopo sucmi ga gudcae (lepo [(le botci) (pa <fundi (lopo sucmi)>)]) (ga gudcae) That the boy was fond of swimming is lucky.
P05	Ei tu danza lepo mi tcihea tu ei (tu [danza (lepo <mi (tcihea tu)>)]) Do you want me to feed you? (I.e., help you eat?)		
P06	Mi krido lepo la Paris, garsitci la Frans		

### Notebook 3 – Teaching Corpus

- |  |   |   |
|--|---|---|
| P14  | <p>Lepo lepo le matma pa kamla pa tradu pa fatru<br/>(lepo [(lepo &lt;(le matma) (pa kamla)&gt;) (pa tradu)]) (pa fatru)<br/>The fact that the mother came was true was troubling.</p>  | <p>He knew it was going to rain before the first drops fell.</p>  |
| P15  | <p>Mi pa viadja lepo lepo prano pa kukra mi (pa [viadja (lepo &lt;(lepo prano) (pa kukra)&gt;)])<br/>I observed (saw-knew) that the running was fast.</p>   | <p>P22 Da pa djano lepo fa nicfea pa lepo le neri clife fa felda<br/>da (pa [djano (lepo &lt;fa (nicfea [pa (lepo &lt;(le [neri clife]) (fa felda)&gt;])&gt;])<br/>He knew it was going to snow before the first leaves fell.</p> |
| <p><b>Q. Attitude Indication; Conversion, Negation &amp; Superlatives (33)</b><br/><b>(uo ue ua uu ui nu fu)</b></p> |   |   |
| P16  | <p>Mi jupni lepo lepo mi danza lepo helba la Meris, pa gudcae mi (jupni [lepo (&lt;lepo (mi [danza (lepo &lt;helba (la meris)&gt;)])&gt; &lt;pa gudcae&gt;)])<br/>I think that my desire to help Mary was fortunate.</p>                | <p>Q01 Uo, la Keit, pa dupma mi uo ([la keit] [pa (dupma mi)])<br/>Damn, Kate deceived me! (Annoyance.)</p>   |
| P17  | <p>Mi penso lopo firpa lopo crina mi (penso [lopo (firpa &lt;lopo crina&gt;)])<br/>I think about fear of (the) rain.</p>  | <p>Q02 Ue tu nu dupma ue (tu [nu dupma])<br/>Well, so you're deceivable! (Surprise.)</p>  |
| P18  | <p>Mi pa rulkao lepo santi gu, lepo helba la Bab mi (pa [rulkao (&lt;lepo (santi gu)&gt; &lt;lepo (helba [la babj]&gt;)])<br/>I was obliged to be silent in order to help Bob.</p>  | <p>Q03 Ua tu no nu dupma ua (tu [no (nu dupma)])<br/>Thank goodness you're not deceivable! (Satisfaction.)</p>  |
| P19  | <p>Lo nimla ga cnida lopo cluva gu, lopo clivi (lo nimla) (ga [cnida &lt;lopo (cluva gu)&gt; &lt;lopo clivi&gt;)])<br/>Animals need to love in order to live.</p>   | <p>Q04 Da nu vegri de di da ([nu vegri] [de di])<br/>X is less green than Y to W.</p>   |
| P20  | <p>Lo humni ga cnida lopo nu cluva gu, lopo hapci clivi (lo humni) (ga [cnida &lt;lopo ([nu cluva] gu)&gt; &lt;lopo (hapci clivi)&gt;)])<br/>Humans need love (i.e., to be loved) in order to live happily.</p>                         | <p>Q05 Da gutra nu begco de da ([gutra (nu begco)] de)<br/>It's a strange request by him. (A strange thing for him to request.)</p>   |
| P21  | <p>Da pa djano lepo fa crina gu, pa lepo le neri drida fa felda da (pa [djano (&lt;lepo (fa [crina gu])&gt; &lt;pa (lepo [(le &lt;neri drida&gt;) (fa felda)])&gt;])<br/>Please don't be troubled by me! (Don't let me bother you!)</p> | <p>Q06 Uu tu nu fatru uu (tu [nu fatru])<br/>I'm sorry you're troubled.</p>   |
|  |   | <p>Q07 Eo no nu fatru mi eo ([no (nu fatru)] mi)<br/>Please don't be troubled by me! (Don't let me bother you!)</p>   |
|  |   | <p>Q08 Ui tu no nu fatru</p>  |

Notebook 3 – Teaching Corpus

- ui (tu [no (nu fatru)])  
I'm glad you're not troubled.
- Q09 Da nu gudbi de ba  
da ([nu gudbi] [de ba])  
X is bettered by (worse than) Y for some purpose x.
- Q10 Di no nu gudbi raba rabe  
di ([no (nu gudbi)] [(ra ba) (ra be)])  
W is not betterable by anything for any purpose, i.e., is best.
- Q11 No nu gudbi  
no (nu gudbi)  
Don't be betterable! (Be the best!)
- Q12 No ge kukra prano  
no (ge [kukra prano])  
Don't run fast!
- Q13 No, kukra prano  
no (kukra prano)  
Don't run fast!
- Q14 No kukra prano  
(no kukra) prano  
Not fast run! (Run but not fast!)
- Q15 Da fu ge briga stuci  
da (fu ge [briga stuci])  
X is a teller of brave stories.
- Q16 Da fu brigystuci  
da (fu brigystuci)  
X is a teller of bravestories.
- Q17 Ei da nu kamla de  
ei (da [(nu kamla) de])  
Is X where Y comes from? (Is X the departure-point of Y's coming?)
- Q18 Da fu lerci de  
da ([fu lerci] de)  
X sends letters to Y. (X corresponds with Y.)
- Q19 Ei da sanpa de di  
ei (da [sanpa (de di)])  
Does X mean Y to W? (Is X a sign of Y to W?)
- Q20 Ia de nu sanpa da di  
ia (de [(nu sanpa) (da di)])  
Yes, Y is the meaning of X to W.
- Q21 Ia di fu sanpa de da  
ia (di [(fu sanpa) (de da)])  
Yes, W assigns meaning Y to X. (Or: W sees significance Y in X.)
- Q22 No fu sanpa mi ta  
(no [fu sanpa]) (mi ta)  
Don't think I'm meant by that! (Don't interpret that sign as meaning me!)
- Q23 No nu takna  
no (nu takna)  
Don't be talked to! (Don't let yourself be talked to.)
- Q24 No fu takna  
no (fu takna)  
Don't be talked about. (Don't be someone people talk about.)
- Q25 Da fu gudbi  
da (fu gudbi)  
It's a purpose for which some things are better than others.
- Q26 Mi fu namci da  
mi ([fu namci] da)  
I can name him. (I know him by name.)
- Q27 Ta fu garti  
ta (fu garti)  
That's something to be grateful for.
- Q28 Nu garti mi ta  
(nu garti) (mi ta)  
Be the recipient of my gratitude for that!
- Q29 Uu no ba nu hompi vi



Notebook 3 – Teaching Corpus

uu (no [ba <nu hompi> vi])  
Alas, there's nothing potable here.

Q30 Da nu vizka ba be  
da ([nu vizka] [ba be])  
X is visible to someone or something  
against some background.

Q31 De no nu vizka raba rabe  
de ([no (nu vizka)] [(ra ba) (ra be)])  
Y is invisible to everything against every  
background.

Q32 Da no nu kukra rutma de ti raba  
da ([no <nu kukra>) rutma] [(de ti) (ra  
ba)])  
X is the fastest route to Y from here by  
any path.

Q33 Da no nu plizo raba lopo roadru  
da ([no (nu plizo)] [(ra ba) (lopo roadru)])  
X is unusable by anyone for hammering.  
(X is useless for hammering.)

**R. Counting, Quantifying & Numerical  
Questions (21)**

**(to te fo fe so se vo ve iesu iene ho hoba toba  
teba foba soba)**

R01 Konduo le batpi  
konduo (le batpi)  
Count the bottles!

R02 Ne. I to. I te. I fo. I foba batpi  
ne . i to . i te . i fo . i ([fo ba] batpi)  
One. Two. Three. Four. There are four  
bottles.

R03 Ne. To. Te. Fo. Foba batpi  
ne . to . te . fo . (fo ba) batpi  
One. Two. Three. Four. There are four  
bottles.

R04 Bleka le fo batpi  
bleka (le fo batpi)  
Look at the four bottles!

R05 Totco to le fo batpi  
totco (to [le fo batpi])  
Touch two of the four bottles! (Any two.)

R06 Kambei mi ne le to batpi  
kambei (mi [ne (le to batpi)])  
Bring me (either) one of the two bottles!

R07 Su levi fe batpi ga veslo lo viski  
(su [levi fe batpi]) (ga [veslo (lo viski)])  
At least one of these five bottles contains  
whiskey.

R08 Iesu de dui  
(iesu de) dui  
Which one (or more) of them do?

R09 Levi to de dui  
(levi to de) dui  
These two of them do. (Should this ,really  
be ternary?)

R10 Foba tugle levi tobme  
(fo ba) (tugle [levi tobme])  
There are four legs on this table. (This  
table has four legs.) Literally, 'Four  
somethings x are legs of this table.'

R11 Hoba tugle leva tobme  
(ho ba) (tugle [leva tobme])  
How many legs has that table?

R12 Leva tobme ga nu tugle soba  
(leva tobme) (ga [(nu tugle) (so ba)])  
That table is belegged with six. (...is six-  
legged.)

R13 Mi nu tugle ho ba  
mi ([nu tugle] [ho ba])  
I am belegged by how many? (I have how  
many legs?)

R14 Ne uu  
ne uu  
One, alas!

## Notebook 3 – Teaching Corpus

R15	<p>Ho mu nu tugle toba (ho mu) ([nu tugle] [to ba]) How many of us are belegged with two? (...are two-legged?)</p>		<p>At least one man came. (Some men came.)</p>
R16	<p>Ho le se tobme ga nu tugle soba (ho [le se tobme]) (ga [(nu tugle) (so ba)]) How many of the seven tables are six-legged?</p>	S03	<p>Ru mrenu pa kamla (ru mrenu) (pa kamla) Enough men came.</p>
R17	<p>Iene da iene da Which one of them?</p>	S04	<p>Le fo mrenu pa kamla (le fo mrenu) (pa kamla) The four men came.</p>
R18	<p>Ho le bekti ji vi levi kruma, ga nu (ho (le bekti) (ji &lt;vi (levi kruma)&gt;)) (ga [(nu tugle) (te ba)]) How many of the objects in this room are three-legged?</p>	S05	<p>Te le fo mrenu pa kamla (te [le fo mrenu]) (pa kamla) Three of the four men came.</p>
R19	<p>Ho le bekti ji vi, nu tugle (ho [(le bekti) (ji vi)]) (nu tugle) How many of the objects here are legged (have legs)?</p>	S06	<p>Ru le fo mrenu pa kamla (ru [le fo mrenu]) (pa kamla) Enough of the four men came.</p>
R20	<p>Ho le bekti vi nu tugle (ho [le bekti]) (vi [nu tugle]) How many of the objects are here legged (have legs here)?</p>	S07	<p>To le te le fo mrenu pa kamla (to [le te (le fo mrenu)]) (pa kamla) Two of the three of the four men came.</p>
R21	<p>Ho le bekti vi levi kruma ga nu tugle foba ([ho (le bekti)] [vi (levi kruma)]) (ga [(nu tugle) (fo ba)]) How many of the objects (while) in this room are four-legged?</p>	S08	<p>Le to le te le fo mrenu pa kamla (le to [le te (le fo mrenu)]) (pa kamla) The two of the three of the four men came.</p>
		S09	<p>Ho le te mrenu ho (le te mrenu) How many of the three men?</p>
		S10	<p>Le ho mrenu le ho mrenu The how many men? (How many men were there?)</p>
<p><b>S. Quantified Descriptions &amp; Questions (18)</b> <b>(iete iefo ieho)</b></p>			
S01	<p>Ne mrenu pa kamla (ne mrenu) (pa kamla) One man came.</p>	S11	<p>Te le ho mrenu te (le ho mrenu) Three of the how many men?</p>
S02	<p>Su mrenu pa kamla (su mrenu) (pa kamla)</p>	S12	<p>Iete le mrenu iete (le mrenu) Which three of the men?</p>
		S13	<p>Te le iefo mrenu</p>

## Notebook 3 – Teaching Corpus

	te (le iefo mrenu) Three of the which four men?		That measures in grams the number 300. (That weighs 300 grams.)
S14	Ie le ho mrenu ie (le ho mrenu) Which (set of) the how many men? (Two questions.)	T04	Ti langa ta lio tomeimei ti (langa [ta (lio tomeimei)]) This is longer than that by 2mm (by the number 2 millimeters).
S15	Iesu le ho mrenu iesu (le ho mrenu) Which (one or more) of the how many men?	T05	Ti skakubra lio mei ti (skakubra [lio mei]) This measures in width the number m. (By default, m = lm. So, this is one meter wide.)
S16	Ieho le ho mrenu ieho (le ho mrenu) Which how many of the how many men? (Three questions.)	T06	La Djan, pa donsu le botci le to dalri (la djan) (pa [donsu <le botci> <le to dalri>]) John gave the boy the two dollars (the two things worth a dollar, e.g., one-dollar bills or silver dollars).
S17	Se sorme pa mercea le se brudi (se sorme) (pa [mercea (le se brudi)]) (Each of some) Seven sisters married (each of) the seven brothers. (This doesn't say what the literal translations seems to say in English.)	T07	La Djan, pa penti lio totodai le mrenu le torkrilu (la djan) (pa [penti <(lio totodai) (le mrenu)> <le torkrilu>]) John paid 22 dollars (worth of money) to the man for the bicycle.
S18	Se sorme pa mercea ne le se brudi (se sorme) (pa [mercea (ne <le se brudi>)]) (Each of some) Seven sisters married (some) one of the seven brothers. (Better, but not quite.)	T08	Lepa ckemo pa sekmi lio to (lepa ckemo) (pa [sekmi (lio to)]) That (former) time-interval measured in seconds the number 2. (That was two seconds.)
<b>T. Measurement, Dimensioned Numbers &amp; Numerical Description (10)</b> <b>(lio lepa -ma -mei -dai)</b>			
T01	Ti dalri lio te ti (dalri [lio te]) This measures in dollars the number 3. (This is worth 3 dollars.)	T09	Lepo da clivi pa nime lio voto (lepo [da clivi]) (pa [nime (lio voto)]) His life measured in years the number 82. (He lived 82 years.)
T02	Ti dalri ti dalri This is a dollar. (I.e., a dollars-worth of something, e.g., money).	T10	Ta pa miksekmi lio to ta (pa [miksekmi (lio to)]) That measured in microseconds the number 2. (That was 2 microseconds.)
T03	Ta gramo lio tema ta (gramo [lio tema])	<b>U. Linked Description; Identity Clauses; Replacement with Letter-Words; Mixed Predicates and Arguments</b> <b>(24)</b>	

Notebook 3 – Teaching Corpus

	<b>(je jue ze sui -mo -ai -ei (dai/dei, etc.))</b>		
U01	Mi pa takna da ta mi (pa [takna (da ta)]) I talked to X about that.	U10	La Djonz, ji le farfu je la Meris, e je la Selis (la djonz) (ji [le (farfu <(je [la meris]) e (je [la selis])>))] The Jones who is the father of Mary and of Sally (independently).
U02	Hu bi da hu (bi da) Who is X?	U11	Dai bi le farfu sui je le matma je la Bab, jue le ditca dai (bi [le (<farfu sui> <(je [matma <(je [la bab]) <jue (le ditca)>])>))] D is also the father of the mother of Bob by the (male) teacher. (I.e., grandson Bob's father was a teacher.)
U03	Da bi le fremi je le botci da (bi [le (fremi <(je (le botci)>))]) X is the friend of the boy.	U12	Dai bi le farfu sui je le matma je la Bab gu, jue le ditca dai (bi [le (<farfu sui> <(je [le (matma <(je [la bab]) gu>)]) (jue [le ditca])>)]) D is also the father of the mother of Bob, through the (woman) teacher. (I.e., Bob's maternal grandmother was a teacher too.)
U04	Ie le botci ie (le botci) What (the) boy?	U13	Dai bi le farfu je Mai, jue le sorme je la Pidr dai (bi [le (farfu <(je mai) (jue [le (sorme <(je [la pidr])>])>))] D is the father of M through the sister of Peter.
U05	Le botci ji le brudi je le merfua je le furvea (le botci) (ji [le (brudi <(je (le [merfua (je <le furvea>))>))]) The boy who is the brother of the wife of the buyer.	U14	Le sorme je Pai jue hu le (sorme [(je pai) (jue hu)]) The sister of P through whom?
U06	Le furvea je hu le (furvea [je hu]) The buyer of what?	U15	Jue la Rabrt, ze la Celis jue ([la rabrt] ze [la celis]) Through (parents) Robert and Shelly jointly.
U07	Le furvea je le hasfa jue la Djonz, jue lio nemomodai le (furvea [(je <le hasfa>) (<jue (la djonz)> <jue (lio nemomodai)>)]) The buyer of the house from Jones for (the) one-million dollars.	U16	Inusoa Dai bi le farfu je Mai jue le sorme je Pai jue Rai ze Cai inusoa (dai [bi (le <farfu ([je mai] [jue (le <sorme ([je pai] [jue (rai ze cai)])>))])>)] So D is the father of M by thesister of P through R and C jointly.
U08	Ie la Djonz ie (la djonz) Which Jones?	U17	Inusoa Dai bi le farfu je Mai jue le sorme je Pai jue Rai e jue Cai
U09	La Djonz, ji le farfu je la Meris, ze la Selis (la djonz) (ji [le (farfu <(je ([la meris] ze [la selis])>))] The Jones who is the father of Mary and Sally jointly. (I.e., of the sib composed of Mary and Sally.)		

### Notebook 3 – Teaching Corpus

inusoa (dai [bi (le <farfu ([je mai] [jue (le <sorme ([je pai] [(jue cai) ]>))>))>))

So D is the father of M by the sister of P through R and through (Independent claims.)

It's a condition of being a mixture of black and non-black.

#### V. Identifying vs. Claiming Subordinate Clauses (7) (jio jia)

U18 Dai bi le farfu je Mai jue le sorme je Pai jue Rai, e la Celis

dai (bi [le (farfu <(je mai) (jue [le (sorme <(je pai) (jue [rai e (la celis)])>))>))>))

D is the father of M by the sister of P through R and Shelly.

V01 Le mrenu jio pa merji na la Somen, haimro de jio pa merji na la Nenimen  
([le mrenu] [jio (pa <merji (na [la somen])>)) (haimro [de (jio <pa (merji [na (la nenimen)])>))])

The man who was married in June is happier than the one who was married in November.

U19 Dai bi le farfu je Mai jue le sorme je Pai jue Rai gu, e je Cai

dai (bi [le (farfu <(je mai) (jue [le (sorme <([je pai] [jue rai] gu) e (je cai)>))>))>))

D is the father of M through the sister of P through (parent) R, and (she is also the sister) of C.

V02 Mi ji la Djan, merji la Meris, ji le detra je la Solomon  
(mi [ji (la djan)]) (merji [(la meris) (ji <le (detra [je (la solomon)])>])>

I, who am John, am married to (the) Mary who is the daughter of Solomon.

U20 Dai bi le brudi je Mai jue le sorme je Pai jue Rai, e jue Cai

dai (bi [le (brudi <(je mai) (jue [le (sorme <(je pai) ([jue rai] e [jue cai])>))>))>))

D is the brother of M through (one parent who is) the sister of P through (parent) R and through (parent) C. (This last jue is redundant.)

V03 Mi jia ditca ga merji la Meris, jia detra je la Solomon  
(mi [jia ditca]) (ga [merji (<la meris> <jia (detra [je (la solomon)])>])>

I, who am a teacher, am married to Mary, who (incidentally) is a daughter of Solomon.

U21 Eo vedma mi le ketpi je la Paris, jue la Romas, jue le kukra trena jue lio fenidai eo (vedma [mi (le <ketpi ([je (la paris)] [(jue <la romas>) (<jue (le [kukra trena])> <jue (lio fenidai)])>))>))>))

Please sell me the ticket to Paris from Rome by the fast train for (the price of) fifty dollars.

V04 La Meris, jia ditca la Franses, di (la meris) (jia [ditca (<la franses> di)])  
Mary, who teaches French to her.

V05 La Meris, jia La Pit, ditca la Franses, di (la meris) (jia [(la pit) (ditca (<la franses> di)])>))

Mary, such that Pete teaches French to her.

U22 Da nigro ze redro bakso  
da ([nigro ze redro] bakso)  
It's a black-and-red box.

U23 Da nigro ze no nigro  
da (nigro ze [no nigro])  
It's black-and-non-black.

V06 Raba jio katma ga titci be jio ratcu (ra [ba (jio katma)]) (ga [titci (be <jio ratcu>)])>

Everything which is a cat eats something else which is a rat.

U24 Da po nigro ze no nigro  
da (po [nigro ze (no nigro)])

## Notebook 3 – Teaching Corpus

V07 Raba jio redro nu herfa, e pa brana na la Nenemen Nevovoton, pa cenja be jio narmykoi  
 (ra [ba (jio <(redro [nu herfa]) e (pa [brana (na <la (nenemen nevovoton) >)]>))] (pa [cenja (be <jio narmykoi>)]))  
 Everyone who is red-haired and was born in November 1882 became a general.

### W. Sentence, Predicate & Argument Negation (18) (ni)

W01 La Djan, pa no gudbi prano  
 (la djan) (pa [(no gudbi) prano])  
 John was a non-good runner (i.e., a runner but not a good one).

W02 La Djan, no pa gudbi prano  
 (la djan) (no [pa (gudbi prano)])  
 John was not a good runner.

W03 La Djan, no pa godzi la Paris  
 (la djan) (no [pa (godzi <la paris>)])  
 John did not go to Paris.

W04 La Djan, no pa titci le pligo  
 (la djan) (no [pa (titci <le pligo>)])  
 John did not eat the apple.

W05 La Djan, pa titci ni le pligo  
 (la djan) (pa [titci (ni <le pligo>)])  
 John ate none of the apple(s). (I.e., zero (of the) apples.)

W06 La Djan, pa godzi ni la Paris  
 (la djan) (pa [godzi (ni <la paris>)])  
 John went to no Paris. (I.e., to zero of Paris.)

W07 Ni la Djan, pa godzi la Paris  
 (ni [la djan]) (pa [godzi (la paris)])  
 No (zero) John went to Paris.

W08 No la Djan, pa godzi la Paris  
 no ([la djan] [pa (godzi <la paris>)])

It is not the case that John went to Paris.

W09 La Djan, no pa kukra godzi la Paris (la djan) (no [pa <kukra godzi> <la paris>])  
 John didn't go quickly to Paris.

W10 La Djan, pa no kukra godzi la Paris (la djan) (pa [(<no kukra> godzi) (la paris)])  
 John went non-quickly to Paris. (I.e., he went, but slowly.)

W11 La Djan, pa kukra no godzi la Paris (la djan) (pa [(kukra <no godzi>) (la paris)])  
 John quickly didn't go to Paris. (But he did something else quickly!)

W12 La Djan, pa kukra godzi la Paris, ni la Romas (la djan) (pa [(kukra godzi) <la paris> <ni (la romas)>])  
 John went quickly to Paris from no Rome. (But from somewhere else?)

W13 Ni la Djan, pa kukra godzi la Paris, la Romas (ni [la djan]) (pa [(kukra godzi) (<la paris> <la romas>)])  
 No John went quickly to Paris from Rome. (But did someone else?)

W14 No la Djan, pa kukra godzi la Paris, la Romas no ([la djan] [pa (<kukra godzi> <(la paris) (la romas)>)])  
 It is not the case that John went quickly to Paris from Rome.

W15 La Djan, no pa kukra godzi la Paris, la Romas (la djan) (no [pa (<kukra godzi> <(la paris) (la romas)>)])  
 John didn't go quickly to Paris from Rome.

W16 La Djan, no ga kukra godzi la Paris, la Romas

## Notebook 3 – Teaching Corpus

	(la djan) (no [ga <kukra godzi > <(la paris) (la romas)>]) John doesn't go quickly to Paris from Rome.		About France, John talked to himself.
		X07	La Tam, la Frans, gi la Djan, pa takna ((la tam) (la frans)] gi) ([la djan] [pa takna]) To Tom about France, John talked.
W17	La Djan, no ga kukra godzi la Paris, la Romas (la djan) (no [ga <kukra godzi> <(la paris) (la romas)>]) John doesn't go quickly to Paris from Rome.	X08	Ti gi lo redro litla ga sanpa lo denro raba (ti gi) ([lo (redro litla)] [ga (sanpa <(1o denro) (ra ba)>]) Under these circumstances, red lights are a sign of danger to everyone (Here the 4th argument of a 5-place structure has been dropped, namely the behavior disposed by the sign.)
W18	La Djan, no kukra godzi la Paris, la Romas (la djan) ([no kukra] godzi] [(la paris) (la romas)]) John non-quickly goes to Paris from Rome.	X09	Lopo stise gu, ti gi lo redro litla ga sanpa ([(lopo <stise gu>) ti] gi) ([lo (redro litla)] [ga sanpa]) Stopping under these circumstances, red lights mean.

### X. Quotation of Loglan; Fronted Arguments

(9)

(li lu liu gi)

X01	La Djan, pa cutse li, Ai, lu la Tam (la djan) (pa [cutse (<li ai lu> <la tam>)]) John said 'Yes I will!' to Tom.
X02	La Tam, gi la Djan, pa cutse liu Ai ([la tam] gi) ([la djan] [pa (cutse <liu ai>)]) To Tom, John said 'Yes I will!'
X03	Liu Ai, la Tam, gi la Djan, pa cutse ([(Liu ai) (la tam)] gi) ([la djan] [pa cutse]) 'Yes I will!' to Tom, John said.
X04	La Tam, gi la Djan, pa cutse ([la tam] gi) ([la djan] [pa cutse]) To Tom, John said (something).
X05	La Frans, gi la Djan, pa takna ([la frans] gi) ([la djan] [pa takna]) About France, John talked (to someone).
X06	La Frans, gi la Djan, pa takna da ([la frans] gi) ([la djan] [pa (takna da)])

### Y. Predicates from Arguments and Prenex Quantifiers (11)

(me me- goi)

Y01	Liu tceму mela Djenis, rutma purda, e sackaa liu mutce (Liu tceму) ([(<me (la djenis)> rutma) purda] e [sackaa (liu mutce)]) 'Tceму' is a "Jenny-route" word and derived from 'mutce'.
Y02	Liu tceму jia sackaa liu mutce mela Djenis, rutma purda ([liu tceму] [jia (sackaa <liu mutce>)]) ([(me <la djenis>) rutma] purda) The word 'tceму' which derives from 'mutce' is a "Jenny-route" word.
Y03	Ba pa meliu me forma holdu le lengu ba (pa [<(me [liu me]) forma> holdu] (le lengu)) There was a 'me'-form hole in the language.
Y04	Raba ra bua goi, ko ba melo bua, ki ba bua

## Notebook 3 – Teaching Corpus

([(ra ba) (ra bua)] goi) (ko [ba (me <lo bua>)] ki [ba bua])

For every something x and every predicate P, if and only if x is a manifestation of P-hood, then x is P.

(le kangu) (pa [(bleka <je mi>) go (me <li mi danza lepo hasfa godzi lu>)])

The dog looked at me in an "I want to go home" sort of way.

Y05 Raba ra bua goi, ko ba melea bua, ki ba bua

([(ra ba) (ra bua)] goi) (ko [ba (me <lea bua>)] ki [ba bua])

For every something x and every predicate P, if and only if x is a manifestation of the class of all that are P, then x is P.

Y06 Raba ra bua goi, ko ba bie lea bua, ki ba bua

([(ra ba) (ra bua)] goi) (ko [ba (bie <lea bua>)] ki [ba bua])

For every something x and every predicate P, if and only if x is a member of the class of all that are P, then x is P.

Y07 Li, Da melo preda, lu durduo snola li, Da preda, lu

(li da me lo preda lu) ([durduo snola] [li da preda lu])

'Da melo preda' reciprocally entails 'Da preda'.

Y08 Ta meda po muvdo

ta ([me da] [po muvdo])

That's an X-type move. (Said of a chess-player, dancer, etc.)

Y09 Ta memi tcaro

ta ([me mi] tcaro)

That's a "me"-car. (Expresses me, belongs to me, etc.)

Y10 Le kangu pa meli, Mi danza lepo hasfa godzi, lu bleka mi

(le kangu) (pa [<me (li mi danza lepo hasfa godzi lu)> bleka mi]).

The dog "I want to go home"-looked at me. (I.e., gave me that "I want to go home" look.)

Y11 Le kangu pa bleka je mi go meli, mi danza lepo hasfa godzi, lu

### Z. Prenex Quantifiers (7)

#### (goi)

Z01 Raba be goi, be matmaa ba

([(ra ba) be] goi) (be [matmaa ba])

For every something x there is a y such that y is the maternal grandmother of x.

Z02 Be raba goi, be matmaa ba

([be (ra ba)] goi) (be [matmaa ba])

There is something y such that for every x, y is x's grandmother.

Z03 Ba lea humni goi, ba paslinkui da

([ba (lea humni)] goi) (ba [paslinkui da])

There is something x such that for every human, x is his or her ancestor.

Z04 Raba goi, cluva ba

([ra ba] goi) (cluva ba)

For every something x, love x!

Z05 Lea humni goi, cluva da

([lea humni] goi) (cluva da)

For every human, love him or her!

Z06 Lea humni go clivi goi, cluva da

([lea (humni go clivi)] goi) (cluva da)

For every human who is alive, love him or her!

Z07 Lea clivi humni goi, da gi cluva

([lea <clivi humni>) goi] da gi) cluva

For every living human, him or her love.

### AA. Connected Arguments & Predicates; Joint Argument Sets (32)

#### (a anoi onoi noa efa epa gugu do)

AA01 La Djan, e la Meris, fundi la Bab



### Notebook 3 – Teaching Corpus

- ([la djan] e [la meris]) (fundi [la bab])  
John and Mary like Bob. Paul is the father of Bob by Sally, or of Beth (but not of both).
- AA02 La Djan, efa la Meris, fundi la Bab  
([la djan] efa [la meris]) (fundi [la bab])  
John, and later Mary, liked Bob. AA12 La Pol, jia farfu la Bab, e la Pit  
(la pol) (jia [farfu (<la bab> e <la pit>)])  
Paul, who is the father of Bob and Pete.
- AA03 Ke la Djan, ki la Meris, fundi la Bab  
(ke [la djan] ki [la meris]) (fundi [la bab])  
Both John and Mary like Bob. AA13 La Pol, jia farfu la Bab, gu, e la Pit  
(la pol) (jia [farfu (<(la bab) gu> e <la pit >)])  
Paul, who is the father of Bob and Pete.  
(This gu is redundant and insufficient to alter the meaning of All.)
- AA04 Di, e de ke fundi da ki tsodi la pit  
(di e de) (ke [fundi da] ki [tsodi (la pit)])  
He and she both like him and hate Pete. AA14 La Pol, jia farfu la Bab, gugu, e la Pit  
([la pol] [jia (farfu (<(la bab) gu>) gu]) e  
(la pit)  
Paul, who is the father of Bob; and Pete.  
(It takes a double gu to say this. Perhaps we should introduce a guu for this? One that would dissolve into two gu's before the parse?)
- AA05 Di, e de fundi da, epa tsodi la Pit  
(di e de) ([fundi da] epa [tsodi (la pit)])  
He and she like him, and before that, hated Pete. AA15 La Pol, farfu, e perdia la Bab, e la Pit  
(la pol) (farfu e [perdia (<la bab> e <la pit>)])  
Paul is a father, and (he) tutors Bob and Pete.
- AA06 Di, e de fundi da, e tsodi gu, la Pit  
(di e de) ([fundi da] e [tsodi gu]) [la pit])  
He and she like him better than, and hate, Pete. AA16 La Pol, farfu, e perdia gu, la Bab, e la Pit  
(la pol) ([farfu e (perdia gu)] [(la bab) e (la pit)])  
Paul is the father of, and the tutor of, Bob and Pete.
- AA07 Di, e de fundi, e tsodi gu, la Pit  
(di e de) ([fundi e (tsodi gu)] [la pit])  
He and she (both) like and hate Pete. AA17 Da farfu ke la Bab, ki la Pit, la Meris  
da (farfu [(ke <la bab> ki <la pit>) (la meris)])  
He is the father of both Bob and Pete by Mary.
- AA08 Di, e de ke fundi ki tsodi gu, la Pit  
(di e de) ([ke fundi ki (tsodi gu)] [la pit])  
He and she both like and hate Pete. AA18 Da farfu kanoi la Bab gu, ki la Pit, la Meris  
da (farfu [kanoi <la bab> gu] ki (<la pit> <la meris>)])  
He is the father if of Bob (by someone) then of Pete by Mary.
- AA09 La Pol, farfu la Bab, efa la Pit  
(la pol) (farfu [(la bab) efa (la pit)])  
Paul is the father of Bob and later Pete. AA19 Da farfu kanoi la Bab, ki la Pit, la Meris
- AA10 La Pol, farfu la Bab, la Selis, onoi la Bet  
(la pol) (farfu [(la bab) <la selis> onoi <la bet>)])  
Paul is the father of Bob by Sally or by Beth. (But not by both.)
- AA11 La Pol, farfu la Bab, la Selis, gu, onoi la Bet  
(la pol) (farfu [(<(la bab) (la selis)> gu) onoi (la bet)])

### Notebook 3 – Teaching Corpus

- da (farfu [(kanoi <la bab> ki <la pit>) (la meris)])  
He is the father of if Bob then Pete by Mary.
- AA20 Da farfu la Bab gu, e la Pit, la Meris, a la Betis  
da (farfu [( <la bab> gu) e (<la pit> <(la meris) a (la betis)>)])  
He is the father of Bob (by someone), and of Pete by Mary or Betty.
- AA21 Da farfu la Bab, anoi la Pit, la Meris, a la Betis  
da (farfu [( <la bab> anoi <la pit>) <la meris> a <la betis>)])  
He is the father of Bob if of Pete, by Mary or Betty.
- AA22 Da farfu la Bab, e la Pit, la Meris, gu, a la Betis  
da (farfu [«([la bab] e [la pit]) (la meris)> gu) a (la betis)])  
He is the father of (both) Bob and Pete by Mary, or of Betty.
- AA23 Da farfu la Bab, e la Pit, ka la Meris, ki la Betis  
da (farfu [<la bab> e <la pit>) (ka <la meris> ki <la betis>)])  
He is the father of (both) Bob and Pete by either Mary or Betty.
- AA24 Da farfu ke la Bab, ki la Pit, ka la Meris, ki la Betis  
da (farfu [(ke <la bab> ki <la pit>) (ka <la meris > ki <la betis >)])  
He is the father of both Bob and Pete by either Mary or Betty.
- AA25 Da farfu la Bab, gu, e la Pit, la Meris, e lendia  
da ([farfu <(la bab) gu> e <(la pit) (la meris)>]) e lendia  
He is the father of Bob, and of Pete by Mary, and teaches languages.
- AA26 Da farfu la Bab, e la Bet, e ditca la Franses, di, e do
- da ([farfu (<la bab> e <la bet>)]) e [ditca (<la franses> <di e do>)])  
He is the father of Bob and Beth and teaches French to her and him.
- AA27 Da farfu, e ditca la Franses, gu, di, e do  
da ([farfu e (ditca <(la franses) gu>)]) [di e do])  
He is the father of, and teaches French to, her and him.
- AA28 Da farfu, e ditca la Franses, di, e do  
da (farfu e [ditca (<la franses> <di e do>)])  
He is a father, and teaches French to her and him.
- AA29 Da farfu, e ditca la Franses, di gu, e do  
da (farfu e [ditca <([la franses] di) gu> e do])  
He is a father, and teaches French to her, and (teaches) him (as a language).
- AA30 Da farfu ba be, noa ditca la Franses, ba, e be  
da (farfu (ba be) noa [ditca <la franses> <ba e be>)])  
He is the father of someone by somebody only if (he) teaches French to that someone and that somebody.
- AA31 Da kanoi farfu ba be, ki ditca la Franses, ba, e be  
da (kanoi [farfu (ba be)] ki [ditca (<la franses> <ba e be>)])  
He, if the father of someone by somebody, then teaches French to that someone and that somebody.
- AA32 Liu tceму kanoi sackaa liu mutce ki mela Djenis, rutma purda  
(liu tceму) (kanoi [sackaa (liu mutce)] ki [( <me (la djenis)> rutma) purda])  
'Tceму', if derived from 'mutce', is a "Jenny"-route word.
- BB. Causal Inflectors, Modifiers & Phrases**  
(31)

Notebook 3 – Teaching Corpus

**(moi soa kou moipa numoi kouhu moihu nukouhu)**

([moi ta] gi) (mi [pa (tokna ti)])

So that that (goal would be reached) I took this.

BB01 Mi pa godzi moi  
mi (pa [godzi moi])  
I went with a purpose, i.e., because of some motive.

BB12 Mi pa danza ta numoi ti  
mi (pa [danza (ta <numoi ti>)])  
I wanted that (goal) therefore this (act).

BB02 Mi moipa godzi  
mi (moipa godzi)  
I purposefully went.

BB13 Ti pa rodja kou lepo tu cuidru da  
ti (pa [rodja (kou <lepo (tu [cuidru da])>)])  
This grew because of the fact that you watered it.

BB03 Mi ji moi, pa godzi  
(mi [ji moi]) (pa godzi)  
I, as someone with a purpose, went.

BB14 Da pa rodja kou lepo cuidru  
da (pa [rodja (kou <lepo cuidru>)])  
It grew because of the watering.

BB04 Mi ji moipa, godzi  
(mi [ji moipa]) godzi  
I, as someone with a purpose then, go.

BB15 Da pa rodja kou lo cutri  
da (pa [rodja (kou <lo cutri>)])  
It grew because of water.

BB05 Moi mi, godzi  
(moi mi) godzi  
To get me, go. (I.e., using me as a goal or purpose.)

BB16 Ti pa rodja kouhu  
ti (pa [rodja kouhu])  
This grew because of what? (This grew why?)

BB06 Moi gi mi godzi  
(moi gi) (mi godzi)  
Purposefully, I go.

BB17 Tu pa cuidru da moihu  
tu (pa [cuidru (da moihu)])  
You watered it why? (For what purpose?)

BB07 Soa ta gi da bi de  
([soa ta] gi) (da [bi de])  
It follows from that that X = Y.

BB18 Ti pa rodja nukouhu  
ti (pa [rodja nukouhu])  
This grew with what consequences? (A converse 'why'.)

BB08 Soa lepo da bi de gi, da tsidru  
([soa (lepo <da (bi de)>)] gi) (da tsidru)  
Since X is the same as Y, X is guilty.

BB19 Kouhu da pa rodja  
kouhu (da [pa rodja])  
Because of what did it grow?

BB09 Da tsidru soa lepo da bi de  
da (tsidru [soa (lepo <da (bi de)>)])  
X is guilty because X is Y.

BB20 Nukouhu da pa rodja  
nukouhu (da [pa rodja])  
With what consequences did it grow?

BB10 Mi pa tokna ti moi ta  
mi (pa [tokna (ti <moi ta>)])  
I took this so that that (goal would be reached).

BB21 Moihu tu pa cuidru da  
moihu (tu [pa (cuidru da)])

BB11 Moi ta gi mi pa tokna ti

### Notebook 3 – Teaching Corpus

- Why did you water it? (For what purpose?)
- BB22 Moihu tu pa danza lepo da rodja  
moihu (tu [pa (danza <lepo (da rodja >))])  
Why did you want it to grow?
- BB23 Kou ta gi ti pa rodja  
([kou ta] gi) (ti [pa rodja])  
Because of that, this grew.
- BB24 Mi moi ditca  
mi (moi ditca)  
I purposefully teach.
- BB25 Moi gi mi ditca  
(moi gi) (mi ditca)  
Purposefully I teach.
- BB26 Moi mi, ditca  
(moi mi) ditca  
With me as a goal or prize, teach!
- BB27 Moi ba gi mi ditca  
([moi ba] gi) (mi ditca)  
With something as a goal, I teach.
- BB28 Mi ditca moi ba  
mi (ditca [moi ba])  
I teach in order to have, be or do something.
- BB29 Mi ditca moi lepo mi snire lo junt  
mi (ditca [moi (lepo <mi (snire [lo  
junti])>))])  
I teach in order to be near the young, i.e.,  
with that motive.
- BB30 Mi ditca kou lepo mi cluva lo junt  
mi (ditca [kou (lepo <mi (cluva [lo  
junti])>))])  
I teach because I love the young, i.e.,  
from that cause.
- BB31 Mi ditca rau lepo mi bremao lo junt  
lopo clivi
- mi (ditca [rau (lepo <mi (bremao [(lo  
junti) (lopo clivi)])>))])  
I teach in order to prepare the young for  
life, i.e., for that reason.
- CC. Compound Term Connectives (8)**  
**(enumoi enukou efa eva epa)**
- CC01 Mi, enumoi la Djan, pa godzi  
mi enumoi [la djan]) (pa godzi)  
I and therefore John went.
- CC02 Mi, efa la Djan, pa godzi  
(mi efa [la djan]) (pa godzi)  
I and then John went.
- CC03 Mi pa vizka la Meris, enukou la Pit  
mi (pa [vizka «la meris > enukou <la  
pit>])])  
I saw Mary and, as a consequence, Pete.
- CC04 Da farfu la Djek la Meris, enukou ni la  
Selis  
da (farfu [(la djek) <la meris> enukou <ni  
(la selis)>])])  
He is the father of Jack through Mary and,  
as a consequence, not through Sally.
- CC05 Mi pa godzi, enukou, pa vizka la Djan  
mi ([pa godzi] enukou [pa (vizka <la  
djan>)])  
I went and, as a consequence, saw John.
- CC06 Mi pa godzi, eva, pa vizka la Djan  
mi ([pa godzi] eva [pa (vizka <la djan>)])  
I went and there saw John.
- CC07 Mi pa godzi, epa, vizka la Djan  
mi ([pa godzi] epa [vizka (la djan)])  
I went, and, beforehand (i.e., before  
going), saw John.
- CC08 Mi pa godzi, enumoi, vizka la Djan  
mi ([pa godzi] enumoi [vizka (la djan)])  
I went and, as intended, saw John.

Notebook 3 – Teaching Corpus

**DD. Connective Questions (8)**  
**(ha enoi noenoi)**

- DD01 Tu danza lo skafi ha lo tcati  
tu (danza [(lo skafi) ha (lo tcati)])  
You want coffee how-connected to tea?  
(Do you want coffee, tea, both or  
neither?)
- DD02 Enoi eo  
enoi eo  
And-not, please. (Coffee but not tea,  
please.)
- DD03 Tu farfu la Djein, ha la Alis  
tu (farfu [(la djein) ha (la alis)])  
You are the father of Jane how-connected  
to Alice? (Are you the father of Alice or  
Jane?)
- DD04 E  
e  
And. (Both. I.e., of both Jane and Alice.)
- DD05 Tu farfu la Djein, e la Alis, la Meris, ha la  
Betis  
tu (farfu [<la djein> e <la alis>] <la  
meris> ha <la betis>)])  
You are the father of Jane and Alice,  
through Mary how-connected to Betty?
- DD06 Noenoi. I mi farfu la Djein, la Meris, gu, e  
la Alis, la Betis  
noenoi . i (mi [farfu (<[(la djein) [la  
meris]) gu> e <(la alis) (la betis) >)])  
Neither. I am the father of Jane through  
Mary, and of Alice through Betty.
- DD07 Inusoa tu farfu la Djein, a la Alis, la  
Meris, e la Betis  
inusoa (tu [farfu <(la djein) a (la alis)>  
<(la meris) e (la betis)>)])  
Therefore you are the father of Jane or  
Alice, by Mary and Betty.
- DD08 Ipou no tu farfu la Djein e la Alis la Meris  
a la Betis

ipou (no [tu (farfu <[(la djein) e [la alis])  
([la meris] a [la betis]) >)])

However, it is not the case that you are  
the father of Jane and Alice through  
Mary or Betty.

**EE. Internal Arguments (12)**  
**(No new LWs.)**

- EE01 Da grobou go kukra lo litla  
da ([grobou go kukra] [lo litla])  
It's a ship that is faster than light.
- EE02 Da kukra je lo litla gu, grobou  
da ([kukra (je <(lo litla) gu>)] grobou)  
It's a faster than light, ship.
- EE03 Da penso go kukra mi  
da ([penso go kukra] mi)  
He thinks faster than I do.
- EE04 Da kukra je mi penso  
da ([kukra (je mi)] penso)  
He's a faster-than-I thinker.
- EE05 Ta rutma je la Mineapolis, go kukra ti  
ta ([rutma <je (la mineapolis)>] go  
kukra] ti)  
That's a route to Minneapolis that is faster  
than this.
- EE06 Ta rutma je la Mineapolis, jue la Cikagos,  
go kukra ti  
ta ([rutma <(je [la mineapolis]) (jue [la  
cikagos])>] go kukra] ti)  
That's a route to Minneapolis from  
Chicago that is faster than this.
- EE07 Ta kukra je ti rutma la Mineapolis, la  
Cikagos  
ta ([kukra <je ti>] rutma] [(la mineapolis)  
(la cikagos)])  
That's a faster-than-this route to  
Minneapolis from Chicago.
- EE08 Ta no nu kukra rutma la Mineapolis, la  
Cikagos

## Notebook 3 – Teaching Corpus

- ta ([[no <nu kukra>) rutma] [(la mineapolis) (la cikagos)])  
That's an "unfasterable" route to Minneapolis from Chicago. (I.e., unbetterable in speed.)
- EE09 Da farbru je la Djan, go plumro la Pit, la Djan  
da ([[farbru <je (la djan)>) go plumro] [(la pit) (la djan)])  
He's a paternal uncle of John's who is more pleasing (as an uncle) than Pete to John.
- EE10 Da plumro je la Pit, jue la Djan, farbru la Djan  
da ([[plumro <(je [la pit]) (jue [la djan])>) farbru] [la djan])  
He's a more pleasing-than-Pete-to-John type of paternal uncle to John.
- EE11 Ba pa korfro je liu me holdu le lengu  
ba (pa [<korfro (je [liu me])> holdu] (le lengu)])  
There was shape-of-'me' hole in the language. (There was a 'me'- shaped hole in the language.)
- EE12 Ba pa holdu je le lengu gu, go korfro liu me  
ba (pa [<holdu (je [(le lengu) gu])> go korfro] (liu me)])  
There was a hole in the language which had the shape of 'me'.
- FF. Argument Ordinals (HB-tags) (25)**  
**(pua pue pui puo puu)**
- FF01 Da sanpa de di do du  
da (sanpa [<de di> do] du)  
X is a sign of Y to W disposing behavior/action H under circumstances Q.
- FF02 Pua da sanpa  
(pua da) sanpa  
X is a sign.
- FF03 Pue de sanpa  
(pue de) sanpa  
Y is a (referential) meaning.
- FF04 Pui di sanpa  
(pui di) sanpa  
W is a sign-interpreter, something that finds meaning in a sign.
- FF05 Puo do sanpa  
(puo do) sanpa  
H is a (behavioral) meaning, i.e., whatever a sign disposes some interpreter to do.
- FF06 Puu du sanpa  
(puu du) sanpa  
Q is a situation or set of circumstances which releases a disposition created by some sign.
- FF07 Puu lopo tcaro bapra ga sanpa pua lo redro litla pue lo humni  
(puu [lopo (tcaro bapra)]) (ga [sanpa (<pua (lo [redro litla]) > <pue (lo humni)>)])  
Operating a car is a circumstance making signs out of red lights to humans.
- FF08 Lopo tcaro bapra ga nufe sanpa pua lo redro litla pue lo humni  
(lopo [tcaro bapra]) (ga [(nufe sanpa) (<pua (lo [redro litla])> <pue (lo humni)>)])  
Operating a car is a circumstance making signs out of red lights to humans.
- FF09 Pue lo denro ga danri sanpa pui lo nimla  
(pue [lo denro]) (ga [(danri sanpa) (pui <lo nimla>)])  
Danger is a common sign-referent among animals.
- FF10 Lo denro ga danri nu sanpa pui lo nimla  
(lo denro) (ga [(danri <nu sanpa>) (pui <lo nimla>)])  
Danger is a common sign-referent among animals.

## Notebook 3 – Teaching Corpus

<p>FF11 Puo lopo prano ga danri sanpa pue lo denro (pue [lopo prano]) (ga [(danri sanpa) (pue &lt;lo denro&gt;)]) Running is a response commonly disposed by danger.</p>	<p>The father through (of offspring out of mother) Mary.</p>
<p>FF12 Lopo prano ga danri ju sanpa pue lo denro (lopo prano) (ga [(danri &lt;ju sanpa&gt;) (pue &lt;lo denro&gt;)]) Running is a response commonly disposed by danger.</p>	<p>FF20 Le farfu je ba jue la Meris le (farfu [(je ba) (jue &lt;la meris&gt;)]) The father of someone through Mary (of at least 1 of Mary's children).</p>
<p>FF13 Lo nimla ga fu sanpa puu lopo do clivi (lo nimla) (ga [(fu sanpa) (puu &lt;lopo (do clivi)&gt;)]) Animals exhibit behaviors that have been disposed by signs under circumstances of their being alive.</p>	<p>FF21 Le farfu je raba jio nu matma la Meris le (farfu [je (ra &lt;ba (jio [(nu matma) (la meris)])&gt;)]) The father of all who are mothered by Mary (all of Mary's children).</p>
<p>FF14 La Miniapolis, danri nu godzi pui la Seint Pol (la miniapolis) [(danri (nu godzi)) [pui (la &lt;seint pol&gt;)]) Minneapolis is a common destination in going from Saint Paul.</p>	<p>FF22 Da kukra je pui lio nenimeikuasei tcaro da ([kukra (je &lt;pui (lio nenimeikuasei)&gt;)] tcaro) It's a faster-(than-something)-by-10-meters-per-second type of car.</p>
<p>FF15 Ba godzi la Miniapolis, la Seint Pol ba (godzi [(la miniapolis) (la &lt;seint pol&gt;)]) Someone goes to Minneapolis from Saint Paul.</p>	<p>FF23 Pui lio nenimeikuasei kukra (pui [lio nenimeikuasei]) kukra Ten meters-per-second is a speed differential.</p>
<p>FF16 Puo ta vedma (pue ta) vedma That is a price.</p>	<p>FF24 Lio nenimeikuasei fu kukra (lio nenimeikuasei) (fu kukra) Ten meters-per-second is a speed differential.</p>
<p>FF17 Ta ju vedma ta (ju vedma) That is a price.</p>	<p>FF25 Ba kukra be lio nenimeikuasei ba (kukra [be (lio nenimeikuasei)]) Something is faster than something else by ten meters-per-second.</p>
<p>FF18 Ba vedma be bo ta ba (vedma [(be bo) ta]) Someone sells something to someone else for that.</p>	<p><b>GG. Compound &amp; Connected Tenses (11)</b> <b>(-fa- -pa- -na- ra- ne- ni- -noi-)</b></p>
<p>FF19 Le farfu je pui la Meris le (farfu [je (pui &lt;la meris&gt;)])</p>	<p>GG01 Mi fapa sucmi mi (fapa sucmi) I will have swum.</p>
	<p>GG02 Mi nepa sucmi mi (nepa sucmi) I swam once (i.e., one time).</p>

## Notebook 3 – Teaching Corpus

- |   |  |
|---|--|
| <p>GG03 Mi nipa sucmi<br/>mi (nipa sucmi)<br/>I never swam (i.e., zero times in the past).</p>  | <p>If you do not come, then I will not be there (i.e., be present there).</p>  |
| <p>GG04 Mi nina sucmi<br/>mi (nina sucmi)<br/>I never swim (i.e., zero times in any present time).</p>  | <p>HH02 No tu fa kamla, inoca no mi hijra va<br/>(no [tu (fa kamla)]) inoca (no [mi (hijra va)])<br/>You will not come only if I'm not there.</p>  |
| <p>GG05 Mi rana sucmi<br/>mi (rana sucmi)<br/>I always swim (i.e., at all present times).</p>   | <p>HH03 No mi fa hijra va, icanoi no tu kamla<br/>(no [mi (fa &lt;hijra va&gt;)]) icanoi (no [tu kamla])<br/>I won't be there if you don't come.</p>   |
| <p>GG06 Mi noina sucmi<br/>mi (noina sucmi)<br/>I don't now swim.</p>   | <p>HH04 Kanoi no tu fa kamla ki ke no mi hijra va<br/>ki tu fa kecdri<br/>kanoi (no [tu (fa kamla)]) ki (ke [no (mi &lt;hijra va&gt;)] ki [tu (fa kecdri)])<br/>If you don't come, then I'll not be there and you'll be sorry.</p> |
| <p>GG07 Mi pacenoina sucmi<br/>mi (pacenoina sucmi)<br/>I used to swim but don't any longer.</p>  | <p>HH05 No mi fa hijra va, icanoi no tu fapa kamla<br/>(no [mi (fa &lt;hijra va&gt;)]) icanoi (no [tu (fapa kamla)])<br/>I won't be there if you have not come.</p>  |
| <p>GG08 Mi noipacenoina sucmi<br/>mi (noipacenoina sucmi)<br/>I never have, and do not now, swim.</p>   | <p>HH06 Tu fapa kamla, icanoi mi hijra va<br/>(tu [fapa kamla]) icanoi (mi [hijra va])<br/>You will have come if I am there.</p>   |
| <p>GG09 Mi pacefa bragai<br/>mi (pacefa bragai)<br/>I was once, and will again be, a king.</p>  | <p>HH07 Mi fa hijra va, inoca tu fapa kamla<br/>(mi [fa (hijra va)]) inoca (tu [fapa kamla])<br/>I will be there only if you will have come.</p>   |
| <p>GG10 Mi noipacenoinacenoifa bragai<br/>mi (noipacenoinacenoifa bragai)<br/>I never have been, am not now, and never will be a king.</p>          | <p>HH08 Kanoi mi fa hijra va ki tu pa kamla<br/>kanoi (mi [fa (hijra va)]) ki (tu [pa kamla])<br/>If I am there, then you came.</p>  |
| <p>GG11 Facenoina mi, garni levi landi<br/>(facenoina mi) (garni [levi landi])<br/>After but not during me (i.e., my lifetime), rule this land.</p> | <p>HH09 Ka no mi fa hijra va ki tu pa kamla<br/>ka (no [mi (fa &lt;hijra va&gt;)]) ki (tu [pa kamla])<br/>Either I will not be there or you came (and possibly both).</p>  |
| <p><b>HH. Logically Connected Clauses (13)</b><br/><b>(inoca icanoi ica Ice)</b></p>  |  |
| <p>HH01 Kanoi no tu fa kamla, ki no mi hijra va<br/>kanoi (no [tu (fa kamla)]) ki (no [mi (hijra va)])</p>  | <p>HH10 No mi fa hijra va, ica tu pa kamla<br/>(no [mi (fa &lt;hijra va&gt;)]) ica (tu [pa kamla])<br/>I will not be there, and/or you came.</p>   |



## Notebook 3 – Teaching Corpus

HH11 Tu fapa kamla, ica no mi hijra va  
(tu [fapa kamla]) ica (no [mi (hijra va)])  
You will have come, and/or I won't be there.

HH12 Tu fapa kamla, ica ke no mi hijra va ki tu kecdri  
(tu [fapa kamla]) ica (ke [no (mi <hijra va>)] ki [tu kecdri])  
You will have come, and/or both I'm not there and you are sorry.

HH13 Ice mi sui fa hijra va  
ice ([mi sui] [fa (hijra va)])  
And I, too, will be there.

### II. Causally-Connected Clauses (7) (i- ki-)

II01 Ti pa rodja, ikou tu cuidru da  
(ti [pa rodja]) ikou (tu [cuidru da])  
This grew because you watered it.

II02 Nukouki ti pa rodja, ki tu cuidru da  
nukouki (ti [pa rodja]) ki (tu [cuidru da])  
Therefore this grew, you watered it.

II03 Tu pa cuidru ti, inukou da rodja  
(tu [pa (cuidru ti)]) inukou (da rodja)  
You watered this; therefore it grew.

II04 Mi pa cuidru da, imoi da pafa rodja  
(mi [pa (cuidru da)]) imoi (da [pafa rodja])  
I watered it so that it would grow.

II05 Kouki tu pa cuidru ti, ki da rodja  
kouki (tu [pa (cuidru ti)]) ki (da rodja)  
Because you watered this, it grew.

II06 Mi pa danza ti, imoi mi pafa ponsu ta  
(mi [pa (danza ti)]) imoi (mi [pafa (ponsu ta)])  
I wanted this so that I could have that.

II07 Mi ditca, imoi mi bremao lo junti  
(mi ditca) imoi (mi [bremao (lo junti)])  
I teach so that I may prepare the young.

### JJ. Indirect Designation; Foreign Quotation (14) (lae sae lie)

JJ01 Eo kambei mi laeli, Loglan Nen lu  
eo (kambei [mi (lae <li loglan nen lu>)])  
Please bring me "Loglan 1". (Please bring me the thing of which 'Loglan Nen' is a sign.)

JJ02 Eo kambei mi laelie gai, War and Peace, gai  
eo (kambei [mi (lae <lie (gai , war and peace , gai)>)])  
Please bring me "War and Peace". (Please bring me the thing of which 'War and Peace' is a sign.)

JJ03 La Djan, melaelie gai, stingy, gai  
(la djan) (me [lae (lie <gai , stingy , gai>)])  
John is whatever the word 'stingy' means in English.

JJ04 Ei tu pa ridle laeli, Loglan Nen, lu  
ei (tu [pa (ridle <lae (li loglan nen lu)>)])  
Have you read "Loglan 1"?

JJ05 Ei tu fundi lae laeli, Loglan Nen, lu  
ei (tu [fundi (lae <lae (li loglan nen lu)>)])  
Do you like the referent of the referent of 'Loglan Nen' (i.e., Loglan)?

JJ06 Da pa cutse sae lepo la Djan, pa prano  
da (pa [cutse (sae <lepo ([la djan] [pa prano])>)])  
He said some signs meaning that John ran.

JJ07 Ei tu vizka sae la Djan  
ei (tu [vizka (sae <la djan>)])  
Do you see signs of John?

## Notebook 3 – Teaching Corpus

- |   |   |
|---|---|
| <p>JJ08 Ei tu vizka laeli. La Djan, lu<br/>ei (tu [vizka (lae &lt;li la djan lu&gt;)])<br/>Do you see (any) referents of 'La Djan'<br/>(i.e., John)?</p>  | <p>Certainly not; you are the "mother" of Loglan.</p>   |
| <p>JJ09 Ei tu vizka lae la Djan<br/>ei (tu [vizka (lae &lt;la djan&gt;)])<br/>Do you see meanings of John?</p>  | <p>KK03 I la Uorf, pa farfu jo de<br/>i ([la uorf] [pa «farfu jo» de]))<br/>And Whorf was its "father", (...the<br/>"father" of it.)</p>  |
| <p>JJ10 Ei tu vizka lo sanpa je la Djan<br/>ei (tu [vizka (lo &lt;sanpa (je [la djan])&gt;)])<br/>Do you see signs of John?</p>   | <p>KK04 Da smina huigro tojo<br/>da (smina [huigro tojo])<br/>He's a "mental giant".</p>  |
| <p>JJ11 Ei tu vizka lo nu sanpa je la Djan<br/>ei (tu [vizka (lo &lt;(nu sanpa) (je [la<br/>djan])&gt;)])<br/>Do you see referents of John?</p>   | <p>KK05 De smina huicma jo<br/>de (smina [huicma jo])<br/>He's a mental "dwarf".</p>  |
| <p>JJ12 Ei tu vizka sae sae sae la Djan<br/>ei (tu [vizka (sae &lt;sae (sae [la djan])&gt;)])<br/>Do you see signs of signs of signs of<br/>John?</p>   | <p>KK06 La Djan, pa brahea rajo na lepo la Loglan,<br/>brana<br/>(la djan) (pa [(brahea rajo) (na &lt;lepo ([la<br/>loglan] brana)&gt;)])<br/>John midwifed, in a manner of speaking,<br/>at the birth of Loglan.</p>               |
| <p>JJ13 Ei tu vizka lo sanpa je lo sanpa je lo sanpa<br/>je la Djan<br/>ei (tu [vizka (lo &lt;sanpa (je [lo (sanpa &lt;je<br/>(lo [sanpa (je &lt;la djan&gt;)]&gt;)]&gt;)]&gt;)])<br/>Do you see signs of signs of signs of<br/>John?</p> | <p>KK07 Ia no. I la Djan, pa bremao ckela ditca<br/>tejo la Loglan<br/>ia no . i ([la djan] [pa &lt;(bremao ckela)<br/>(ditca tejo)&gt; &lt;la loglan&gt;)])<br/>Certainly not; John was a "prep-school<br/>teacher" of Loglan.</p> |
| <p>JJ14 Ei tu vizka lae sae lae la Djan<br/>ei (tu [vizka (lae &lt;sae (lae [la djan])&gt;)])<br/>Do you see referents of signs of referents<br/>of John?</p>   | <p>KK08 La Djan, pa bremao ckela ditca la Loglan,<br/>rajo<br/>(la djan) (pa [&lt;bremao ckela&gt; ditca] (la<br/>&lt;loglan rajo&gt;)])<br/>John was a prep-school teacher of Loglan,<br/>in a manner of speaking.</p>             |
| <p><b>KK. Metaphor-Marking or "Figurative<br/>Quotation" (9)</b><br/><b>(ja)</b></p>  |   |
| <p>KK01 Mi farfu jo la loglan<br/>mi ([farfu jo] [la loglan])<br/>I am the "father" of Loglan.</p>  | <p>KK09 La Djan, pa rajo bremao ckela ditca la<br/>Loglan<br/>(la djan) ([pa rajo] [&lt;bremao ckela&gt; ditca]<br/>(la loglan)])<br/>John was, speaking figuratively, a prep-<br/>school teacher of Loglan.</p>                    |
| <p>KK02 Ia no. I tu matma jo la Loglan<br/>ia no . i (tu [(matma jo) (la loglan)])</p>  | <p><b>LL. Letter-Variables and Acronyms (17)</b><br/><b>(-z-)</b></p>   |

## Notebook 3 – Teaching Corpus

- LL01 Le mrenu pa vedma le kangu le cmalo nirli  
(le mrenu) (pa [vedma (<le kangu> <le (cmalo nirli)>)])  
The man sold the dog to the little girl.
- LL02 Nei fa cluva kei  
nei (fa [cluva kei])  
n (the girl) will love k (the dog).
- LL03 La Ned, farfu nei  
(la ned) (farfu nei)  
Ned is the father of n.
- LL04 Nai cluva nei, e fundi lo kangu, enumoi fa cluva kei  
nai ([[cluva nei) e (fundi <lo kangu>)] enumoi [fa (cluva kei)])  
N (Ned) loves n, likes dogs, and therefore will love k.
- LL05 Nai merji la Meris, ice Mai no fundi lo kangu  
(nai [merji (la meris)]) ice (mai [(no fundi) (lo kangu)])  
N is married to Mary; and M doesn't like dogs.
- LL06 Bai merji la Alis, ice Ama fundi lo kangu  
(bai [merji (la alis)]) ice (ama [fundi (lo kangu)])  
B is married to Alice; and A likes dogs.
- LL07 Inusoa ba nu dakli lepo Mai no fa cluva kei  
inusoa (ba [(nu dakli) (lepo <mai (no [fa (cluva kei)])>)])  
Therefore it is likely that M will not love k.
- LL08 Nai nu perpli laelie gai, Central Intelligence Committee, gai  
nai ([nu perpli] [lae (lie <gai , central intelligence committee , gai>)])  
N is employed by (works for) the "Central Intelligence Committee".
- LL09 Nai nu perpli la Cailza  
nai ([nu perpli] [la caiiza])
- N works for the CIA.
- LL10 Mai nu perpli, la UNai mai ([nu perpli] [la unai])  
M works for the UN.
- LL11 Nei danza lepo fa nu perpli, la UNaiESaiCaiO  
nei (danza [lepo (fa <(nu perpli) (la unaiesaicaiO)>)])  
n (the little girl) wants eventually to work for (the) UNESCO.
- LL12 Ipou nei danza lepo na nu perpli la SaiPaiCaiA  
ipou (nei [danza (lepo <na ([nu perpli] [la saipaicaia])>)])  
However, n wants to work now for the SPCA.
- LL13 Mao ponsu ba jio kasni  
mao (ponsu [ba (jio kasni)])  
M' (Marvin) owns a cow.
- LL14 Lo HaitoO bi lo cutri  
(lo haitoo) (bi [lo cutri])  
(The) H2O is water.
- LL15 Lo HaitoSaiOfo bi lo melaelie gai, sulphuric acid, gai (/haitosaiOfo/)  
(lo haitosaiOfo) (bi [lo (me <lae (lie [gai , sulphuric acid , gai ]>)])  
(The) H2SO4 is whatever 'sulphuric acid' means in English.
- LL16 Ta HaitoSaiOfo  
ta haitosaiOfo  
That's sulphuric acid.
- LL17 asi groda bei cei  
asi (groda [bei cei])  
a is bigger than b by c.

### MM. Predicates as Names & Vocatives (7) (No new LWs.)

## Notebook 3 – Teaching Corpus

- MM01 Hoi Mrenu  
hoi mrenu  
Hi, Man!  
mi (pa prano) . ifa (mi sucmi) . ifa (mi fleti) . ifa (mi dzoru)  
I ran. Then I swam. Then I flew. Then I walked.
- MM02 Hoi Ganfua go Redro nu Herfa  
hoi ganfua go redro nu herfa  
Hail, Lady with the Red Hair!  
NN02 Mi pa prano, e sucmi, e fleti, e dzoru  
mi ([<pa prano> e sucmi) e fleti] e dzoru)  
I ran and swam, and flew, and walked.
- MM03 Ei tu vizka Hoi Ganbra, lepo mi pana nu  
ganble turka letu fothaa  
ei (tu [(vizka <hoi ganbra>) (lepo <mi (pana [(<nu ganble > turka) (letu fothaa)]>)]))  
Do you see, O Noble One, that I have worked admirably on your castle?  
NN03 Mi pa takna da, e de, e di, efa do  
mi (pa [takna (<(da e de) e di> efa do)])  
I talked to X and Y, and to W, and then to Q.
- MM04 Ei tu pa vizka la Mrenu  
ei (tu [pa (vizka <la mrenu>)])  
Did you see (the) Man?  
NN04 Muvdo, icanoi da redro, icaci de nigro  
muvdo icanoi ([da redro] icaci [de nigro])  
Move if (either) X is red or Y is black.
- MM05 La Garfua go Redro nu Herfa ga matma mi  
(la [garfua go (redro <nu herfa>)]) (ga [matma mi])  
(The) Lady with Red Hair is my mother.  
NN05 Muvdo, icanoi da redro, ica de nigro  
(muvdo icanoi [da redro]) ica (de nigro)  
(Either) Move if X is red, or Y is black.
- MM06 La Blabi Garfua pa korji la Redro Garfua lepo godzi la Hatro Sitfa  
(la [blabi garfua]) (pa [korji <la (redro garfua)> <lepo (godzi [la (hatro sitfa)]>)]))  
(The) White Queen told (ordered the) Red Queen to go to (the) Hot Place.  
NN06 Muvdo, icanoige da redro, ica de nigro, ica di vegri  
muvdo icanoige ([da redro] ica (de nigro)] ica [di vegri])  
Move if (either either) X is red or Y is black, or W is green.
- MM07 La Musmu pa djadou la Ratcu lepo de fa damgoo la Vrici  
(la musmu) (pa [djadou (<la ratcu > <lepo (de [fa (damgoo <la vrici>)]>)]))  
Mouse told (informed) Rat that he (Mouse) was going to go down River.  
NN07 Da muvdo, anoi redro, aci nigro  
da (muvdo anoi [redro aci nigro])  
X moves if (either) red or black.
- NN. Grouped & Ungrouped Afterthought Connections (12)**  
**(i- -ge -ci)**
- NN01 Mi pa prano. Ifa mi sucmi. Ifa mi fleti. Ifa mi dzoru  
mi (pa prano) . ifa (mi sucmi) . ifa (mi fleti) . ifa (mi dzoru)  
I ran. Then I swam. Then I flew. Then I walked.  
NN08 Da muvdo, anoi redro, a nigro  
da ([muvdo anoi redro] a nigro)  
X (either) moves if (it is) red, or (it is) black.
- NN09 Da muvdo, anoige redro, a nigro, epa vegri  
da (muvdo anoige [(redro a nigro) epa vegri])  
X moves if (it is both, either) red or black, and before that, green.
- NN10 Mi godzi da, anoi de, e di  
mi (godzi [(da anoi de) e di])  
I go to X if Y, and (to) W.

## Notebook 3 – Teaching Corpus

- NN11 Mi godzi da, anoi de, eci di  
 mi (godzi [da anoi (de eci di)])  
 I go to X if to (both) Y and W.
- NN12 Mi godzi da, anoifage de, a di, a do  
 mi (godzi [da anoifage <de a di> a do])  
 I go to X if earlier to (either, either) Y or  
 W, or Q.

### OO. Spelling (8) (No new LWs)

- OO01 Liu artomi purkaa liu atom, e nu leasri li,  
 a. rei. tei. o. mei. i, lu  
 (liu artomi) ([purkaa (liu atom)] e [(nu  
 leasri) (li a . rei . tei . o . mei . i lu)])  
 The word 'artomi' is derived from the  
 word 'atom', and is spelled 'eigh. are.  
 tee. oh. em. eye'.
- OO02 Liu atmo sui purkaa liu atom, e nu leasri  
 li, a. tei. mei. o, lu  
 ([liu atmo] sui) ([purkaa (liu atom)] e [(nu  
 leasri) (li a . tei . mei . o lu)])  
 The word 'atmo' also is derived from the  
 word 'atom', and is spelled 'eigh. tee.  
 em. oh'.
- OO03 Ie le to po purkaa ga nu fundi tu  
 (ie [le to (po purkaa)]) (ga [(nu fundi) tu])  
 Which of the two derivations do you like  
 better?
- OO04 Liu protoni purkaa liu proton, e nu leasri  
 li, pei. rei. o. tei. o. nei. i, lu  
 (liu protoni) ([purkaa (liu proton)] e [(nu  
 leasri) (li pei . rei . o . tei . o . nei . i lu)])  
 The word 'protoni' is derived from the  
 word 'proton', and is spelled 'pea. are.  
 oh. tee. oh. en. eye'.
- OO05 Liu purkaa purkaa li purda kamla lu, e nu  
 leasri li, pei. u. rei. kei. a. a, lu  
 (liu purkaa) ([purkaa (li purda kamla lu)]  
 e [(nu leasri) (li pei . u . rei . kei . a . a  
 lu)])

The word 'purkaa' is derived from 'purda kamla', and is spelled 'pea. you. are. keigh. eigh. eigh'.

- OO06 Artomi. a. rei. tei. o. mei. i  
 artomi . a . rei . tei . o . mei . i  
 Artomi. eigh. are. tee. oh. em. ee.
- OO07 Atmo. a. tei. mei. o  
 atmo . a . tei . mei . o  
 Atmo. eigh. tee. em. oh.
- OO08 Protoni. Pei. Rei. O. Tei. O. Nei. I  
 protoni . pei . rei . o . tei . o . nei . i  
 Protoni. pea. are. oh. tee. oh. en. eye.

### PP. Sentences in VOS Order (4) (goa)

- PP01 Goa seidjo lue no nu trecymro bekti ji vi  
 lo rardza goa lea surpernova  
 goa (seidjo [(lue <(no [nu trecymro])  
 bekti>) (ji <vi (lo rardza)>)]) goa (lea  
 surpernova)  
 Among the most interesting objects in the  
 universe (all-existence) are the  
 supernovae.
- PP02 Goa gudbi lo cutci lopo pueblo lo naldi  
 goa lo mroza  
 goa (gudbi [(lo cutci) (lopo <pucblo (lo  
 naldi)>)]) goa (lo mroza)  
 Better than shoes for driving (push-  
 hitting) nails are hammers.
- PP03 Goa gudbi lo cutci lopo pueblo lo naldi  
 goa lo mroza  
 goa (gudbi [(lo cutci) (lopo <pucblo (lo  
 naldi)>)]) goa (lo mroza)  
 Better than shoes for driving (push-  
 hitting) nails are hammers.
- PP04 Na levi delnai goa bilti lo blabi flora goa  
 rui re lea floryclu lo redro  
 (na [levi delnai]) (goa [bilti (lo <blabi  
 flora>)]) goa [(rui <re (lea floryclu)>)] (lo  
 redro)])

## Notebook 3 – Teaching Corpus

These days, more beautiful than white  
flowers, are, according to most flower-  
lovers, red ones.