1 1

LOGLAN AND THE
OPTION OF CLARITY:
A GENUINELY USER-FRIENDLY LANGUAGE
REED D. RINER   ## FOR HUMANS AND THEIR MACHINES

B ETWEEN 1920 and 1950 Edward Sapir and Benjamin Whorf raised
some disturbing questions about the role of language.(1,2) Suppose,
they suggested, that while we must use a language to communicate, human
thought is limited and directed by whatever particular language one may
choose to think and speak. Suppose further that the structure of thought,
and the way we represent reality to ourselves and others, is pulled in some
directions, pushed in others, and generally warped by the language in ways
beyond control of the thinker. This notion, the Sapir-Whorf Hypothesis,
suggests that our language more than our actual experience may determine
our conception of "reality." It is the same fundamental notion that was par-
tially resurrected in McLuhan's caveat "the medium is the message."(3)

The Sapir-Whorf Hypothesis further proposes that each language cre-
ates its own distinctively structured reality. So, as languages evolve and the
distance between them increases, then the ease of translation, even the pos-
sibility of mutual comprehension, diminishes.

As the world becomes smaller, and as margins of tolerable error or mis-
understanding diminish, as international and cross-cultural communica-
tions become a daily affair, the questions raised by the Sapir-Whorf
Hypothesis become increasingly important. Do we live in arbitrarily differ-
ent "realities" with none of them doing justice to the real world? If our lan-
guages indeed are such jumbles of contradictions, does this bode well for
our images of the real world? Is sustained, coherent, understandable

* Reed D. Riner is Associate Professor of Anthropology at Northern Arizona University.
A former editor of *Cultural Futures Research,* he teaches courses on applied anthropol-
ogy and futures research, and consults on problems of strategic planning, organization,
and community development. He is Past President of the High Plains Society for Applied
Anthropology and a director of the annual CONTACT Conferences.

communication between cultures and their governments really possible, as technology daily pares away the margins of tolerable error?

Sapir and Whorf were not alone in their suspicion that language itself is perhaps the greatest impediment to clear thinking. A contemporary of theirs, the immigrant Pole Alfred Korzybski, developed his general semantics as a way of identifying and reducing the ambiguities, vagaries, and insanities embedded in the English language.(4,5) Perhaps Korzybski's general program was seriously flawed, despite his contribution to a heightened sensitivity to muddled thinking, because he attempted to "fix" a natural language.(6,7)

Natural Languages, Natural Limits

Natural languages, primary among the products of human culture, may be compared to layer cakes. A natural language rests on the material foundation of its system of sounds: it must have neither too few sounds nor too many, and they must be easily producible by the speaker and as easily recognizable by the listener. This is the phonological layer of the language cake.

On this phonemic foundation rests the next, the morphological, layer of the cake. This concerns what the parts of the language are: what constitutes a word, and how are words built, recognized, connected, arranged, and rearranged? Resting on the morphology is the lexicon, the layer that determines the "parts of speech," the pigeonholes in which the vocabulary of the language is to be sorted.

The next layer bears the forbidding label "grammar." To be "grammatical" means only that a string of words is understandable – school teachers notwithstanding, propriety ain't pertinent here; so long as the intent is understandable, the grammar's acceptable. And crowning this layer cake is the most nebulous layer of a language, its semantics – the knowledge about the language that tells us how the senses of two grammatically identical words (like "this" and "that") differ from one another, and finally what the acceptably formed string of words refers to, stands for, or says about the world.

There are no very tight or necessary connections among the several layers of this linguistic layer cake. The sounds change independently of the meanings, and the meanings independently of the sounds that carry them, while the organizational features of morphology, lexicon, and grammar slip-slide around in the middle. Natural languages are very ramshackle affairs; they tend to grow randomly in all directions, by accretion, as the product of contradictory processes. On one hand there is pressure for ever more precise ways of communicating ever more fine-grained messages through modification, qualification, and elaboration, while on the other hand there are persistent efforts to avoid direct expression of truth by poets, politicians, and children. Languages tend both to accumulate new words and construc-

tions, some of which will be lost, and to elaborate their structures, some of which will disintegrate; but always languages tend to become more complex and more chaotic. Hence, there is the potential for language to prejudice the structure and content of thought.

As a result, each of the approximately eight thousand natural languages, regardless of the number of its speakers, is a very large affair. Like whales and elephants, natural languages make very poor experimental animals, ungainly to get into a laboratory. And why would one *want* to get a natural language into a laboratory? To test the Sapir-Whorf Hypothesis perhaps, and to answer the question: In just what ways is human thought limited and directed by the language in which one thinks?

To Test the Hypothesis

The world in 1955 was already small enough, and the margins of tolerable error and misunderstanding were narrow enough, that social-psychologist James Cooke Brown began to ponder, over that Christmas holiday, how the Sapir-Whorf Hypothesis could possibly be tested.(8) Testing it would require the proverbial "something else on which to stand," an alternative to the natural languages. The task would require a language capable of separating the presumed linguistic causes from the predicted cultural effects. To release the Whorfian effects – the signs, if they existed, of increased potential for clear thinking – the language would somehow have to guarantee a kind of metaphysical and culture neutrality. It would have to minimize the kind and number of built-in premises about the world; it would have to minimize ambiguity and maximize clarity.

To do this, the constructed language would have to be structurally as different as possible from English, say, as well as from the other natural languages with which it was going to be experimentally compared. And of course it would have to be a speakable language, as useful in everyday discourse as a natural language.

My reconstruction of Brown's reflections on how to build an instrument for testing the Sapir-Whorf Hypothesis encapsulates a substantial chapter in the birth and development of experimental linguistics over the last thirty years. The ensuing complement would be the development of artificial intelligence research. And, in fact, the story of Loglan, which starts out in the domain of experimental linguistics, ends in the realm of computers and artificial intelligence.

Some readers will recall Brown's public preview of the design of Loglan in an article by that name in the June 1960 issue of *Scientific American.* Others will recall that Mycroft Holmes, the hero computer of the lunar penal colony of Robert Heinlein's *The Moon Is a Harsh Mistress* (1965), spoke Loglan. After these brief appearances, Loglan slipped from public view again until its second official appearance in 1975 in *Loglan 1: A Logical Lan-*

guage, third edition.(9) The first- and second-edition predecessors were microfilmed publications distributed only to libraries and the project-tracking aficionados of the language. This first bound volume reached an estimated three thousand readers, some of whom persevered to become the first generation of speakers of what is now affectionately referred to as Old Middle High Loglan.

Loglan

Loglan has distinctive phonological, morphological, lexical, and grammatical features. Each aspect of the language has been designed by applying our contemporary understanding of the structure of natural languages and the nature of human logic, and engineered wherever possible by testing the designed structures in the laboratory.(10)

Phonology: Loglan has a phonemic alphabet, in that 26 letters of the western alphabet – the most commonly used alphabet in the world – are assigned to its 26 phonemes on a one-to-one basis. Thus Loglan has almost perfect audio-visual isomorphism, each word being pronounced the way it is spelled, and spelled the way it is pronounced. Twenty-three of these sounds, the 6 vowels and 17 consonants – a, e, i, o, u, y and b, c, d, f, g, h, j, k, l, m, n, p, r, s, t, v, z – are "regular" sounds in the sense that they (and only they) occur in the ordinary, unborrowed words of Loglan. The letters representing these 23 sounds are pronounced as they most commonly are in English, with 4 exceptions: "i" is pronounced as in "machine," "c" is [sh], "j" is [zh] as in the French "Jean," and "y" is the [u] in "huh," the linguist's "schwa." These 23 sounds are all easy to pronounce and probably for that reason are widely distributed in human languages.

Three irregular sounds are rather less widely distributed and more difficult to pronounce. They are spelled with the remaining letters q, w, x of the 26-letter alphabet; "q" has the sound of [th] in English "thin," "w" the sound of the French [u] in "plus" and the German [ü] in "Mühler," and "x" the sound of the Scottish and German [ch] in "loch" and "Bach." These latter two sounds don't occur in English while the former, the [th], does not occur in either French or German. They are indeed less widely distributed. In fact these "weird" sounds occur in Loglan only to make it possible to pronounce certain borrowed scientific words phonemically. For example, "wolfram" and "xenon": the international chemical symbols for these two elements are "W" and "Xe," so Loglanists will have to be able to pronounce them in some way. The three irregular sounds provide for this – and allow a more literal spelling of certain borrowed names, e.g., **Qi'ydor Bax** (Theodore Bach).

Despite its great flexibility, the Loglan phonological system, with its 26 phonemes, is rather small by world standards. It compares, for example, to a count of 45 phonemes estimated for American English. This feature

of flexibility as a result of regularity and simplicity is characteristic of each layer of Loglan's linguistic cake.

Morphology: In Loglan there are only three kinds of words. There is not the endless proliferation of nouns, adjectives, articles, prepositions, adverbs, verbs; rather, there are only the classes of (a) little operator words, (b) proper names, and (c) predicate words. Each class of word has a distinctive logical function, and an equally distinctive consonant-vowel (C-V) pattern is applied in constructing them.

The "little words" provide all of the logical, grammatical, numerical, and punctuational operators for Loglan. They are of two kinds: the simple little words and the compound little words. Each simple little word is a single or double vowel and may have a preceding consonant. Thus all simple operator words look like V, VV, CV, or CVV, for example **a, ai, da,** or **tai.** Compound little words are simply strings of these forms, for example **pana** or **anoi.** All little words end with vowels and have no adjacent consonants.

Proper names, because they are for the most part borrowed, show a lot of irregularity, but all members of this class are "marked," distinguished by a final consonant, and in writing also indicated by an initial capital. These two features distinguish proper names from both other classes of Loglan words.

The predicate words in Loglan do what the nouns, verbs, adjectives, and most of the adverbs and prepositions do in English. All predicate words are vowel-final, contain at least one pair of adjacent consonants, and have at least two syllables. These features distinguish them absolutely from names and little words. Predicates are of three sorts: primitives, complexes, and borrowings. A primitive is composed in either [CCV'CV] (**mrenu** = "is-a-man") or [CV'CCV] (**fumna** = "is-a-woman") form. Complexes are either the expansions of these by the addition of three- or four-letter affixes, for example **tarsensi** = "star science," or strings of such affixes, for example **senmao** = "is-a-scientist." Borrowings are the "residual" category: they are all the predicate words that are neither primitive nor complex, for example **iglu** and **protoni.**

Predicate words do most of the work in Loglan. In fact, from a logician's point of view, Loglan is a speakable predicate calculus. Each predicate word may therefore have from one to five (or more) place referents, called "arguments" in logic. For example: **Da mrenu,** where "Da" is a generalized third person pronoun (she/he/it/they) but most conveniently translated "X," reads "X is-a-man." This is a logically complete statement; there is nothing more to be said. So we can say that the predicate **mrenu** has only one argument. But **vedma,** to sell, is a four-argument predicate that begs, although in Loglan does not require, the speaker to specify who (**da**) sold what (**de**) to whom (**di**) at what price (**do**) – the complete linguistic and economic equation being **Da vedma de di do!**

Vocabulary: Because predicates do so much of the work in Loglan, it is

important that they be as immediately recognizable and as culturally neutral as possible. Thus the eight hundred "primitive predicates," the simplest, undefined terms, were composed with specific consonant-vowel sound patterns closest to their counterparts in eight of the world's most widely-spoken natural languages: Chinese, English, Spanish, Hindi, Russian, French, Japanese, and German, representatives of six language families. Sound combinations were tested for recognition on listener populations. Loglan **mrenu** "is-a-man" recalls some of "ren," "man," "ombre," "manushya," "homme," and "mensch" with an average "recognition score" of about 50 percent for native speakers of each of these eight languages. The trial inclusion of Arabic, representing the Semitic language family, did not measurably improve, or detract from, recognizability.

The primitive predicates refer to eight hundred of the most basic and universal of all human experiences and concepts, words that appear as simple words in all of the donor languages. All of the more complex notions in Loglan are based on metaphors, as they are in some of the natural languages (in German and Chinese, for example). Cultural and worldview bias is minimized in the primitives, and is made maximally explicit, open to display and question, in the transparency of the metaphors behind the complex predicates. Thus **sesmao** is immediately decipherable as **sensi madzo** ("science maker") by any Loglanist; and it is the humble poetry behind their word for "scientist." Similarly, **telbie** decodes as **terla bidje** = "earth edge," and is the astronaut's image behind their word for "horizon."

One consequence of Loglan's planned phonological and morphological regularity is that a word's function is immediately apparent, even if the meaning of the particular word is not. Another consequence is that word boundaries are easily (more importantly, unambiguously) resolved in the soundstream. The "I scream/ice cream" kind of confusion can't occur in Loglan. The lexemes are exclusive; no word of the language belongs to any more than one of the fundamental categories; there are no homographs or homophones, such as "bow" and "bow" and "bough." Nor are there the "They are flying planes" or "He made three banks" kinds of ambiguity. And there is no occurrence of the "Time flies like an arrow/Fruit flies like a banana" kind of syntactic ambiguity.

Lexicon: Loglan's three morphological classes of words resolve into just 54 mutually exclusive subclasses of grammatically interchangeable words, categories that Loglan grammarians call "lexemes." But the 54 lexemes are distributed in a strange and powerful new way. Two of the morphological categories, the names and the predicates, are each assigned one lexeme. That is, these most numerous of words are entirely contained in just two lexemes of the language. The other 52 lexemes in Loglan are occupied by its little words, and many of these by exactly 1. For example, the little word **no** (=not/non-) has a unique distribution in the language, and is the sole member of its lexeme.

At 54, the number of Loglan lexemes compares favorably, and very efficiently, with the one hundred fifty or so such classes for English. Thus, if the learner masters the uses of two great categories of words, names, and predicates, and 52 typical little words, **da** has mastered all of the essentials of the lexicon – and the rest is vocabulary expansion.

Grammar: A grammar of a language is the set of rules – often imaginary – that define the permissible utterances of that language. A good grammar describes how the many kinds of words can be combined in meaningful ways; it discriminates all comprehensible utterances of the language from the incomprehensible ones. Efforts to write systematic grammars for English and German have produced lists of 3, 6, 8, and 10,000 rules. These kinds of efforts are still incomplete, and are bound to remain so because of the random growth-and-disintegration and chaotic constitution of all natural languages, and because all the grammars we have for natural languages are by nature descriptive, after-the-fact creations.

Loglan's grammar, by pleasing contrast, is *a priori* complete, and composed of just 220 rules. That is, 220 phrase-structure rules suffice to generate and to parse all permissible utterances in the language. This small set of rules is "all about clarity." It is all about predicates and what can be done to them, replicating in this respect the essential features of the predicate calculus of modern logic. As far as we can yet know, Loglan can accommodate precisely and unambiguously the native ways of saying things in any natural language. In fact, because it is logically rigorous, Loglan forces the speaker to make the metaphysical (cultural, worldview) premises in and of the natural language explicit in rendering the thought into (disambiguated) Loglan. Those assumptions, made explicit, become propositions that are open for critical review and amendment – so not only can the Sapir-Whorf Hypothesis be tested, but its details can be investigated with Loglan.

As a research instrument, Loglan is manageably small. Individuals learn it quickly; 54 typical words and 220 grammar rules suffice to introduce all grammatical possibilities. The rest is vocabulary, which seems to fall into place – the early portions, anyway – at a rate of about four to five hundred new words a week. Thus, a couple of months' efforts learning Loglan seems to provide equal or greater fluency than a two-semester college program in any natural language. Loglan is so readily learnable because the language minimizes – indeed, effectively eliminates – arbitrariness, and maximizes structural efficiency. But as a result of this economic regularity Loglan also maximizes optionality and flexibility; and logic and clarity are among the options of how and what one can think and say.

The "Log" in Loglan

The name "Loglan," readers will have inferred, is a contraction from "logical language." Contracted or not, the word "logical" says at once too little

and too much. Loglan includes speakable provision for the propositional calculus, for quantification theory, for an unambiguously clear distinction between designation and predication, and for a word-classification scheme that treats all predicates as members of a single part of speech and so allows all claims to be expressed in the predicate calculus. Loglan does not, however, make provision for a theory of types or for removal of the ambiguities that result from the absence of hierarchical notation. Thus, Loglan is less than "speaking symbolic logic"– but not by very much.

Loglan is "logical" in the specific, narrow sense of facilitating logical transformations – that is, the transformation of sentences into other sentences in such a way that if the first sentences are true then so also are the second. As an aid to performing the logical transformations correctly, Loglan has been made free from syntactic ambiguity; that is, one and only one "deconstruction" of a sentence can be made in the mind of the hearer. Spoken instructions are unequivocal; the auditor/reader can always reproduce the one – and only one – way that the speaker/writer could have produced/intended the utterance. Every grammatical utterance has one and only one parse tree; the inadvertent packaging of 18 different possible intentions in "It's a pretty little girls school" is precluded in Loglan.

But Loglan's logic is not compulsory. Shorthand expressions are readily available. **Da blanu:** (it (or s/he, or they) is blue. But clarity and completeness are just as readily available and, by implication, asked for. **Blanu** is, logically, a comparative relationship and therefore a two-place predicate: **Da blanu hu?** (X is bluer than what?). The soul of Loglan is not its rationality but its optionality. It can be as optionally illogical as it can be logical. It is, in short, whatever its speakers choose it to be. The normal Trobriand, or Hopi, or Japanese way of saying things can be reproduced in Loglan just as clearly as the intentions of the western logician.

Early Workshops

Following the publication of *Loglan 1: 3rd Edition* in 1975, a journal of Loglan studies called *The Loglanist* was initiated, and potential participants for Loglan workshops began to be identified. The first workshops were conducted in 1972 and again in 1977–78. They produced the first small community of Loglan speakers, the first spontaneous conversational events, and the first informal tests of the Sapir-Whorf Hypothesis. These tests were necessarily informal and did not focus on identifying limiting factors in the participants' first language, English. Rather the workshops provided the first opportunity to look for signs of the expected releasing and enabling effects of Loglan on thinking. Observing these effects, Brown identified a half-dozen traits that seemed to characterize the English speech of the participants in these earliest encounters with Loglan. They were: (1) the richness and oddity of metaphor; (2) the unusually frequent designation of

previously "unheard of," or unthought about, individuals and phenomena; (3) an increased awareness of ambiguity as evidenced by jokes or other usages that call attention to it; (4) a taste for neologisms or for bizarre or overliteral usages; (5) the invention of inflected (or de-inflected) forms that do not exist in ordinary usage in the speaker's native language but are in principle possible (e.g., "couth," "idiosyncrat," "ert," "qualitiedly," "therapped grouply," "encomiast," to list a recent sample); and (6) a heightened sense of fun with one's own and other people's English – that is, with the often comic contrast between what people actually say and what they think they are saying.

The Machine Grammar Project

As implied by the whimsical phrase "Old Middle High Loglan," a major transformation of the original Loglan was already afoot by 1978. Since the inception of Loglan in 1955, another, and more familiar, piece of later-twentieth-century technology had achieved public currency – the personal computer. Loglan, because of its disambiguity, was perfectly preadapted to computer use.

So just as the 1978 workshops were ending, Project MacGram began. This was the effort to write a machine grammar for Loglan. In 1975 the mathematicians Alfred V. Aho, Stephen C. Johnson, and Jeffrey D. Ullman, working at Bell Laboratories, had demonstrated the validity of a constructive proof algorithm for testing ambiguity in LALR[1] (Look Ahead Left and Right Parsing) type languages.(11) Loglan was not such a language, but it was possible that a "machine dialect" of the language of the LALR[1] type could be created. If so, this powerful new algorithm would permit the development on computers of a grammar of Loglan that was demonstrably unambiguous, and therefore "intelligible" to machines. Work began. A team assault was mounted. Resolutions of the existing ambiguities were soon found, necessitating the introduction of some new punctuation words and linkage patterns. Fortunately, the basic grammatical design of the language did not have to be changed. In 1982, the MacGram Project was completed. A conflict-free grammar had been developed that successfully parsed the whole of the known language, and the grammar of Loglan has remained in conflict-free condition ever since.

The Present Prospects

As early as the 1978 workshops, Loglan had begun to demonstrate its suitability as an instrument to test the Sapir-Whorf Hypothesis. Its potential as a mind-expanding alternative to natural language, with a theoretical potential to enhance rates of invention and discovery, was also demonstrated. Further, Loglan's disambiguity recommended it as the best candidate lan-

guage for all kinds of "low tolerance for error" situations. This is especially valuable in the widely-sought-after "world language," the planetary, universal, culture-free auxiliary or second language, intended to be the protector of the endangered languages and of the ethnicities and the cultural diversity that they maintain. A properly chosen world auxiliary second language could well protect the many endangered linguistic species, conserving diversity in that domain, as environmental conservation efforts are seeking to preserve genetic diversity in the gene pool. The front-running candidate to be that auxiliary second language is, of course, Esperanto. But Esperanto is not as good a candidate for any of these tasks, as it is a hybrid of natural languages, and so carries the weight of natural-language bias.

But since completion of the MacGram Project in 1982, a new role for Loglan has become increasingly apparent. Loglan is a natural interface between humans and their machines. It is computer listenable, computer readable, and computer speakable. Its unequivocal grammar matches the machine's incorrigible literality – without sacrificing at the bottom level its compact code. Being a fully described and demonstrably speakable human language, Loglan can also provide artificial-intelligence investigators and hobbyists with at least the beginning of a model of how human speech generation and understanding actually works. Loglan grammar not only is known but is already written in a machine-parsible code (parsers are available from the Loglan Institute in both IBM-compatible and Apple Macintosh software).

Loglan is in itself the beginning of an AI program. This recommends Loglan as an international translation medium. Using Loglan as intermediary would ensure against any possible syntactic ambiguity between any two natural languages. Once the sense of any document has been satisfactorily rendered in Loglan (a more painstaking process than translating back from Loglan), then the resulting document will be unambiguous in every sense of the word. And it will be ever so much more readily accessible to all other natural languages. While programming computers to translate natural text into Loglan would seem to be beyond our capacities now (too much "philosophical" analysis of what the text "really means" would seem to be involved), programming them to translate Loglan texts into almost any natural language would seem to be quite possible. Because all information stored in Loglan is in a predicate calculus form, Loglan would be an ideal information storage and retrieval medium. It devolves upon us to develop AI programs for the "very human art" of scholarly reading, the kind that will create "books that talk to each other."(12)

NOTES

1. Edward Sapir, *Language: An Introduction to the Study of Speech* (1921; reprint, New York: Harcourt, Brace, 1949.
2. John B. Carroll, ed., *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf* (Cambridge: MIT Press, 1956).
3. Marshall McLuhan, *Understanding Media: The Extensions of Man* (New York: McGraw-Hill, 1964).
4. Alfred Korzybski, *Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics*, 4th ed. (Lakeville, Conn.: International Non-Aristotelian Library, 1958).
5. Alfred Korzybski, "The Role of Language in the Perceptual Process" in *Perception: An Approach to Personality*, ed. by Robert R. Blake and Glenn V. Ramsey (New York: Ronald Press, 1951).
6. Ross Evans Paulson, *Language, Science and Action: Korzybski's General Semantics: A Study in Comparative Intellectual History* (Westport, Conn.: Greenwood Press, 1983).
7. Anatol Rappaport, *Semantics* (New York: Crowell, 1975).
8. James Cooke Brown, "Loglan," *Scientific American*, 202(40):53–63 (June 1960).
9. James Cooke Brown, *Loglan 1: A Logical Language*, 3d ed. (Gainesville, Fla.: Loglan Institute, 1975).
10. James Cooke Brown, *Loglan 1: A Logical Language*, 4th ed. (Gainesville, Fla.: Loglan Institute, 1989).
11. Alfred V. Aho, Stephen C. Johnson, and Jeffrey D. Ullman, "Deterministic Parsing of Ambiguous Grammars," *Communications of Association of Computing Machinery* 18(8):441–52 (1975).
12. E. A. Feigenbaum, "Toward the Library of the Future," *Long Range Planning* 22(1):118–23 (1989).

SOURCE

Loglan 1: A Logical Language, 4th edition, by James Cooke Brown, is available from The Loglan Institute Inc., 1701 Northeast 75th St., Gainesville, Fla. 32601.

A general introduction and description of the Loglan language, this is the long-awaited revision of the Institute's basic book. The fourth edition is twice the size of the third, a fact that reflects the reengineering of the language since 1975 as well as the inclusion this time of pronunciation guides, historical notes, specimen translations, 65-page word-lists, and a much higher degree of descriptive completeness. The book has a new chapter on testing the Sapir-Whorf Hypothesis. Loglan software is also available from the Institute in both Apple Macintosh and IBM-PC versions.