American Journal of Computational Linguistics *Microfiche 12*

THE YERKISH LANGUAGE
FOR NON-HUMAN PRIMATES

ERNST VON GLASERSFELD
Department of Psychology
University of Georgia
and
Yerkes Regional Primate Research Center

Abstract

Yerkish, the language described in this paper, was designed for the purpose of exploring the extent to which non-human organisms (e.g. great apes) could be brought to acquire linguistic skills. First attempts at teaching a spoken language to non-human primates had failed, apparently because of the animals' incapacity vocally to produce the phonemes of a natural language. Subsequent work (Gardner & Gardner, 1971; Premack, 1971) demonstrated that communication could be achieved by means of visual signs or symbols.

Yerkish is a visual language with a lexicon of graphic word symbols (lexigrams), each of which is a combination of discrete recursive design elements. Each lexigram is represented on one of 125 keys of a keyboard. Sentences are formed by pressing keys in successive order. Sentence length, at present, is limited to seven lexigrams. Input from the keyboard is monitored and recorded by a computer that contains in its core the lexicon, a parser, and certain response capabilities.

The grammar is strictly interpretive and was derived from the 'correlational' system implemented in the Multistore parser for English sentences (von Glasersfeld & Pisani, 1970). The parser works on the basis of essentially non-linguistic classifications of items and relational concepts (tables of the operational classes are provided in this paper). It produces a structural analysis in terms of immediate constituents. If an input string yields one comprehensive structure, the string is deemed grammatical. The first lexigram functions as a sentence marker indicating the mood of the utterance: affirmative, negative, interrogative, and imperative. In the case of certain

requests (imperatives), the computer activates mechanical devices that fulfil the request (dispensing food, drink, toys, playing music or a movie, etc.) Given this response capability, a 24-hour learning situation is created in which there is always some potential incentive for the animal to use linguistic communication.

Yerkish lexigrams and sentences are easily translated into English, but semantics and sentence structure are somewhat restricted. Lexigrams always have only one meaning of the corresponding English word, and the rules for their concatenation were designed to reduce ambiguities to a minimum.

The paper explains deviations from English grammar by means of structural diagrams and demonstrates that, in spite of the many simpli fications, Yerkish allows for embedded clauses and is, indeed, capable of expansion in many directions.

Three examples of 'conversations' with the young female chimpanzee that is being taught the use of Yerkish are added as an appendix to the paper.

TABLE OF CONTENTS

TABLES

APPENDIX

The Yerkish Language for Non-Human Primates (1)

Introduction

There are several reasons why it would be cumbersome and even somewhat misleading to organize a description of the language used in our communication study with non-human primates, according to the linguist's traditional pattern, in three more or less independent sections dealing, respectively, with grammar, lexicon, and semantics. Yerkish is an artificial language that was designed for a specific and peculiar purpose to explore to what extent apparently non-linguistic organisms could acquire linguistic skills if they were placed in an environment in which the use of linguistic communication would be to their advantage.

Being an artificial language, the design of Yerkish was subject to constraints that are rather different from those that may or may not have impinged upon the development of natural languages. Computational

---

(1) Research supported by National Institutes of Health grants HD-0616 and RR-00165 and carried out at the Yerkes Regional Primate Research Center, Emory University (Atlanta), by a team of seven specialists from three actively participating institutions: Duane M. Rumbaugh, initiator and Principal Investigator of the project, Timothy V. Gill and Josephine Brown (Dept. of Psychology, Georgia State University) are responsible for behavioral design and experimentation. Harold Warner and Charles Bell (Yerkes Regional Primate Research Center) are responsible for design and engineering of the interface with the computer and the electromechanical devices. Ernst von Glasersfeld and Pier Paolo Pisani (Dept. of Psychology and Computer Center, University of Georgia) are responsible for the design and computerisation of the Yerkish language, its grammar, and the Multistore

linguists will, I am sure, agree that natural languages would have different grammars and different interpretive rules if, from the very beginning of their evolution, they had had to be intelligible to a computer. And that is what Yerkish had to be. For reasons that will become clear when we discuss the research background of the project, the introduction of a computer as monitor of the communication system was one of the salient feature of this research effort (Rumbaugh et al., 1973a).

Other constraints in the development of the Yerkish language will be discussed at those points in the exposition where their explication seemed most appropriate. I have tried to concentrate specific aspects under indicative subtitles. I have no illusions that this has been wholly successful. My main goal, however, was to give the reader as complete as possible a picture, not only of what was done, but also of why it was done. The instrumental aspect of Yerkish as the linguistic vehicle in an experimental communication study must be kept in mind at all times; much of what follows in these pages can make sense only if it is put into that perspective.

Another point that I should like to stress is that the artificial language on which this paper reports is only one of several major efforts that made this communication study possible. Such success as we have had is the result of team work in the fullest sense of that term. The project would never have got off the ground if it had not been for the continuous patient collaboration of seven rather heterogeneous specialists from three different institutions (see footnote 1) and, last but not least, for the perseverance of our female chimpanzee Lana. (Fig. 1)

Figure 1

The chimpanzee Lana working at her keyboard

Background

Yerkish is a visual language of graphic word-symbols, designed for research in communication with non-human primates and, possibly, as a substitute vehicle for humans who, for physiological reasons, could not acquire a spoken language. Lieberman(1968), on the basis of anatomical investigations, came to the conclusion that the vocal apparatus of the great apes precludes the production and modulation of many of the phonemes that make up the repertoire of human languages. Forty years earlier, Robert Yerkes (1925), the founder of the first primatological research institute, in whose honor we have named our language, had already observed this vocal handicap of the non-human primates. In the intervening years, the failure of several long-term efforts to teach a chimpanzee English, Japanese, or Russian, empirically confirmed his observation (for a review, cf. Ploog and Melnechuk, 1971).

The fact, however, that the great apeas are barred from speaking a language does not necessarily mean that they could not understand one, nor that they could not learn to use a linguistic communication system that functions in another sensory modality. There still are, of course, scholars who, defining "language" from a rather anthropocentric point of view, refuse to allow the term for any communication system that does not use the vocal-auditory mode of transmission. Among other things, this would mean that programming languages and other silent communication systems could never be called "language", no matter how adequately they might be described in terms of lexicon, syntax, and semantics. Today, there seems to be a growing consensus that this restriction of "language" to acoustic systems is not a scientifically necessary or useful one (Ploog and Melnechuck, 1971: 640; Lyons, 1972: 64). Interest, thus, has shifted from the question whether or not other organisms can learn to speak a language, to the question whether or not they can learn to handle a communication system that is linguistic in its structure.

Given that there seems to be no compelling evidence that any non-human species on this earth has, in fact, developed a communication system that could legitimately be called "language", one might be inclined to think that attemts to teach a language to a non-human organism are necessarily ~~doomed to~~ fail. This conclusion, however, would be quite unwarranted. Animal trainers in circuses and in the laboratory have shown beyond all doubt that many species have a potential for the acquisition of skills which no one, who observed the species in the wild, would suspect. The fact is that the behaviors an organism manifests in a given environment constitute under all circumstances only a subset of the behaviors which the organism could acquire in different environments (Lorenz, 1974). In the area of cognitive skills, for instance, Kohler's investigations (1925) already indicated that the great apes had been drastically underrated. Since then, and up to the demonstration of "higher mental functions" by Viki (Hayes and Nissen, 1956/1971), especially the chimpanzee's intellectual reputation has continuously grown (Rumbaugh, 1970). Thus it has, indeed, become more and more pertinent to ask just how far a chimpanzee (or other great ape) could be brought in the acquisition of linguistic skills which do not require vocal speech.

The success of the Gardners (1969, 1971) with their chimpanzee Washoe is so well known that there is no need to reiterate the description of their pioneering work. Using American Sign Language (ASL) as a vehicle, they established irrefutably that an infant chimpanzee can be taught to communicate very effectively; and there would seem to be no reason why, given a conducive environment, Washoe's communicatory skill should not continue to grow as she develops towards intellectual maturity.

It has been repeatedly objected, however, that Washoe's successful communications are as yet no proof that she has acquired "language" (e.g. Lenneberg, 1971; Brown, 1971). Most of the skepticism about Washoe's linguistic accomplishments is based on the argument that the strings of signs she produced do not manifest syntactic competence. When Washoe was introduced to ASL, no rigid rules of sign-order were observed and the _relational_ semantics (which, for instance in English, is taken care of to a large extent by word-order) was left implicit in the communicatory event and had to be intuitively gleaned from the situational context by the observer (Gardner and Gardner, 1971). Since a language user's compliance with the syntactic rules of the language is an important criterion in the evaluation of his linguistic performance, the apparent lack of such rules in ASL made it _a priori_ questionable whether Washoe's or, indeed, any other ASL-user's stringing together of signs could be considered syntactic and thus evidence of "language" on the theoretical level. In addition to that, a lack of syntactic rules is the very reason why Washoe's communications could not contain many relational indications. For instance, since the sign system taught to Washoe had no consistent means for designating _actor_ and _patient_ in activity situations (comparable to, say, the subject-verb-object sequence in many natural languages), the assignation of these roles was left to the intuition or the common sense of the receiver.

In retrospect it is easy to see that this relative lack of syntactic rigidity would supply critics with arguments that tend to diminish claims with regard to Washoe's linguistic competence. On the other hand, it is equally clear that the Gardners, when they started on their splendid enterprise, were concerned above all with the formidable task of establishing a viable form of communication with a chimpanzee, and they could not possibly have foreseen all the theoretical reasons why linguists and philosophers of language might doubt that the communication system they chose, and the way Washoe was going to use it, should be called "linguistic". Hence I should like to emphasize that my attempt to clarify the syntactic problem is in no way intended as a criticism of the work accomplished with Washoe, but solely to throw some light on the several ways in which our project, started a few years later, was able to benefit from the Gardners' effort.

The Communication Facility

The basic idea of our project at the Yerkes Primate Research Center was the introduction of a computer as a thoroughly objective monitor of all linguistic transactions. This solved several problems at once. In the first place, it eliminated the problem of subjective or intuitive evaluation of the grammatical correctness of the experimental animal's linguistic products. Incorporating a reduced and suitably adapted Multistore Parser (von Glasersfeld and Pisani, 1970), the computer can "objectively" judge grammaticality. An input string either conforms to legitimate syntactic structure, or it does not. There cannot be any doubt either way. Second, the computer has no difficulty in recording every input and transaction that takes place, be it grammatical or not. Third, thanks to the computer, the communication facility can be kept in operation twenty-four hours a day, without the forbidding cost of several shifts of techinicians and observers.

In order to turn the communication facility into a learning environment that could at least to some extent operate without the presence of a human a system of automatic responses was implemented. By activating one of a set of machine-commanded dispensers, the computer can satisfy a number of requests, provided these requests are correctly formulated by the experimental animal. So far, the automatic responses are limited to the dispensing of various foods and drinks, to opening and shutting a window, activating a movie and a slide projector as well as a tape player. In the future we hope to add something of a question-answering system in order to enable the computer to respond verbally to some questions and, perhaps, also to give some feedback with regard to errors made in the subject's linguistic input.

A full description of the communication facility, as it is in operation at present, has been published elsewhere (Rumbaugh et al., 1973a). Here we shall be mainly concerned with the Yerkish language. A quick survey of the main components of the installation will have to suffice.

Input to the system is effected by means of a keyboard of maximally 125 keys, arranged in vertical panels of 25 each. Four such panels are in use at present, corresponding to a total of 100 keys. Each key represents one lexigram, i.e., a geometrical design which constitutes a word-symbol (lexical item) of the Yerkish language. Depression of a key activates the corresponding item in the computerized lexicon which is permanently incorporated in the Multistore parser. The spatial arrangement of the lexigrams in the keyboard can be easily reshuffled (to prevent the experimental animal from acquiring a fixed motor pattern).

To switch on the system, a horizontal bar, mounted well above the keyboard, has to be pulled down. The bar has to be held down continuously throughout the input of a message. Lana, the female infant chimpanzee with whom we have been working, does this by hanging on to the bar with one hand while using the other to press keys. If the system is switched on and several keys are then pressed in succession, ending with the "period" key (the "end-of-message" signal for the computer), the parser takes this string as a "sentence" and analyzes it in order to establish whether or not is is grammatically correct. If the input string is a grammatically correct request, the machine also determines the object of the request and, if it is within the range of automated responses, satisfies the request by activating the relevant dispenser or mechanism.

Regardless of the outcome of the grammatical analysis, the machine prints out the English word corresponding to each lexigram that has been activated and records, at the end of the string, whether or not it was found to be correct.

Directly above the keyboard in the experimental chamber, there is a row of seven small projectors in which the geometric designs of the lexigrams appear, one by one from left to right, as their keys are being pressed on the keyboard. This provides Lana with feedback as to the part of the message that has already been typed in, and also with a linear representation of the string she is composing. A signal light, on the right of the projectors, lights up when the "period" key has been pressed and terminates the message.

Above this first row of projectors there is a second similar one which serves to display messages that are sent in to Lana from a second keyboard in the technician's station outside Lana's chamber. Messages originating from the technician's keyboard are also recorded by the computer, but they are marked by a code symbol as "operator's messages" and cannot be confounded with Lana's linguistic production.

The Yerkish Lexigrams

The original constraints under which the Yerkish language was to be designed were essentially three.

1) Drawing on the experience of the Gardners (1969, 1971) and Premack (1971), Yerkish had to be a visual language with a lexicon of unitary word-symbols that could be represented singly on the keys of a keyboard.

2) Both lexical items and sentence structure were to be as univocal as possible, because this, on the one hand, would facilitate the automatic parsing of input and, on the other, it was expected to make acquisition of the language easier for our subjects.

3) The structure of Yerkish was to be close enough to English to allow word-by-word translation, in order to make participation in communication events, as well as their evaluation, maximally accessible to technicians and observers.

For a few weeks at the very outset of the enterprise, the author revelled in dreams of an ideal language in which each word was to be composed of semantically significant pleremes (Hockett, 1961). There were to be individual design elements designating the more important recurrent semantic categories, and each concept available in the Yerkish universe of discourse was to be represented by a lexigram (i.e. the visual/graphic counterparts to words in spoken languages) composed of design elements which, in their own right, would designate the major semantic categories to which the concept belonged. Thus, for instance, as the American Indian language Yuchi (Crawford, 1973) has a morpheme that recurs in any word that designates a part of the human body, every Yerkish lexigram designating a part of the primate body would have contained a specific design element. Given that the Yerkish lexicon was, in any case, to contain no more than two or three hundred lexigrams, it seemed feasible to cover at least the major semantic categories with a hundred or so design elements. The reason for doing this was, of course, that such a language would have been an invaluable instrument for testing our subject's classificatory skill and processes of concept formation.

The dream was soon shattered by technical restrictions. The feedback projectors above the keyboard had to be such that each one of them could display every lexigram of the language. Within our budget, this could be achieved only if all lexigrams were designed in such a way that they could be generated by combining design elements of a common set limited to twelve. Under these circumstances it was obviously impossible to maintain the individual design elements semantically constant and a drastic compromise had to be made. By choosing nine graphic elements that could be readily superimposed, one over the other, and three basic colours, a little additional flexibility was gained (see Table 1). By "mixing" the three basic colours we could generate seven discriminable hues. Together with black (absence of colour), this gave us eight background features, and these could be used to colour-code at least some important conceptual categories (see Table 2).

Interpretive versus Descriptive Grammar

The grammar of Yerkish is a direct derivative of the "correlational" grammar that was implemented some years ago in the Multistore parser for English sentences (von Glasersfeld, 1964, 1965, 1970; von Glasersfeld and Pisani, 1968, 1970). It is, therefore, strictly an interpretive grammar and lays no claim to "generative" properties, nor is it "transformational" in the Chomskyan sense of that term.

In the hope that it might dispel some misunderstandings that have haunted the development of correlational grammar since its initial conception by Silvio Ceccato (Ceccato et al., 1960, 1963), I should like to dwell for a moment on a purely theoretical point. While the term "grammar" is predominantly used to indicate the formalized description of a language (e.g. Chomsky, 1965; 9 and 140), "correlational grammar" is, instead, the description of an interpretive system. The main difference between the two, though basically simple, has perhaps not been made sufficiently explicit. An ordinary grammar is expected to account for all grammatical sentences of the language in a more or less axiomatic way, i.e. by demonstrating that every possible grammatical sentence is a case under a formally stated rule or set of rules. An interpretive grammar, on the other hand, is not concerned with demonstrating the grammaticality of any sentence, but with transforming the content of a given piece of language into a canonical form composed of pre-established semantic elements or modules. It is a "grammar" in the sense that it consists of rules that govern this transformation, but these rules describe the language only indirectly, since what they actually describe is a model of the language user in the receiving role. (Note that by "model", in this context, we intend a processor which, given the same input, will yield the same output as the processor to be modelled, regardless of the means it employs to do so.) An interpretive system of this kind, thus, presupposes the grammaticality of its input. But since it is designed to interpret all grammatical pieces of language, it can be used to define operationally as "grammatical" any input that it can interpret, while input that it cannot interpret can be considered "ungrammatical".

When designing a correlational grammar for a natural language, it is a truly enormous problem to bring the grammar's interpretive capability anywhere near the interpretive capability of the native user of the language. In the case of an artificial language, however, this problem is altogether eliminated, because the lexicon, the rules of concatenation, and the interpretive grammar can be designed all at the same time. Since there is no native user, who has a universe of experiential content and well-established semantic connections (by means of which this content is linked to linguistic expressions), the designer is free to tailor the lexicon, as well as the syntax of his language, to the universe of discourse he envisages.

That is, to a large extent, how Yerkish was designed, expecially with regard to the rules of grammar. The result of it is that the user of Yerkish can communicate in grammatically correct lexigram strings no more than the correlational grammar of Yerkish can interpret.

A Restricted Universe of Discourse

Yerkish, as it operates at present, is in fact a compromise in more than one respect. An effort was made to create a potential universe of discourse that would allow a non-human primate to formulate as many communications as possible which, given the particular environment, could be used instrumentally for the attainment of goals (von Glasersfeld, 1974a). Such an attempt is necessarily based on more or less anthropocentric conjecture. There is, however, a certain amount of evidence that non-human primates organize their perceptual world in a way that does not seem incompatible with ours. In actual fact, Lana has already demonstrated that all the items which we assumed would take on the function of goals for her and would, therefore, act as incentive to communicatory activity, were indeed appropriate. Where food and drink were concerned, this could almost be taken for granted. With visual displays such as a movie and slides, with the sounds of music and voices, and with the view through an open window, our anthropocentric hope of analogy was well rewarded. Above all it is gratifying to note that there was never a need to resort to any form of negative reinforcement or punishment. Though there were, especially at the beginning, not very many things that Lana could "say" in Yerkish, she has never tired of saying them.

On the practical side, since the interpretive grammar was to be implemented in a functioning parser, the universe of discourse was strictly limited by the size of the computer that could be obtained within the budget of the project. Because the project is wholly experimental and explorative, it was and is an absolute requirement to leave within the computerized system a certain amount of room for _ad hoc_ modifications and additions that might suddenly prove necessary in view of our subject's actual performance. Thus it was essential that the implemented grammar should never occupy all of the available space within the computer. This is still the case and we hope to be able to maintain this flexibility for some time to come.

Technical Constraints

There are four ways in which the Yerkish universe of discourse is restricted. First, there is the number of lexical items the system can handle. The present version of the Multistore parser can deal with a maximum of 250 lexigrams. The interface that links the computer to the keyboard in the experimental chamber is designed for half that number, i.e. for 125. The keyboard, however, is divided into five panels of 25 keys each and these panels are readily exchangeable. This means that the subject's vocabulary can, in fact, be extended to 250 items, but only a subset of these, namely 125, will be operative during any one session. (Since Lana at present uses a total of 100 lexigrams, there is still much room for vocabulary expansion.).

The second restriction also concerns the vocabulary of lexigrams, but it springs from the grammar of Yerkish and does not limit the number of individual lexigrams but rather the number of conceptual classes to which lexigrams have to be assigned. Because of its interpretive function, correlational grammar requires a classification of lexical items that differs considerably from the word-classification used by traditional descriptive grammars. Lexigrams, in fact, are classified according to certain functional characteristics of the concepts they designate, i.e., according to cognitive characteristics.

The lexicon with which a correlational grammar operates, therefore, is divided, not into a few generic and largely morphologically defined classes such as nouns, verbs, adjectives, etc., but into a much larger number of classes defined in terms of what the designated items can do, i.e., by the role or roles they play in the cognitive representation of experiential situations. In the case of "things" this is, for instance, the kinds of activity which they can perform as actors and the kinds of activity in which they can play the part of patient; and in the case of "activities" it is, for instance, the kinds of change they bring about.

In the implementation of the interpretive system, i.e. the parser, it is the characterization of the lexical classes that occupies considerable space, not the individual lexical items. The total number of classes, therefore, has to be decided a priori. In the present Yerkish parser, the maximum number of lexigram classes is 46. At the time of writing, 35 of these classes have been filled (see Table 3). The remaining 11 are still empty, but they can be made operative at any moment by the simple insertion of new lexigrams and the definition of the functional properties of the items they designate.

The third restriction concerns the number of lexigrams that can be strung together to form one message. The amount of data the parser has to take into account during the processing of a given message, obviously, depends to some extent on the number of lexical items of which the message is composed. This dimension corresponds to sentence length in natural languages. As it was impossible to foresee with any precision just how much work space the parser might require for the analysis of all types of grammatical input strings, we preferred to be on the safe side and limited sentence length to seven lexigrams. On the basis of the experience gathered since then, we can now say that the computer system could, in fact, handle input strings of up to ten lexigrams and, hence, we plan to extend the capacity of the interface hardware in the near future from seven to ten lexigrams.

The fourth restriction involves the number of connectives (see Table 4) by means of which phrases and sentences can be put together. These connectives or correlators are far more numerous in a correlational grammar than are the traditional syntactic functions. This proliferation is again the result of the interpretive purpose of the system. A parser that is intended to extract the conceptual content from pieces of language must be able to identify not only the conceptual items involved, but also the relational concepts by means of which they are connected with one another. Hence, the traditional distinction between syntax and semantics is no longer operative in a correlational grammar, and the few "basic grammatical relations" (e.g. subject-verb, verb-object, etc.) which connect grammatically characterized items, are replaced by a great many "correlators" which are considered the linguistic expression of the relational concepts that link items on the conceptual level.

While our English grammar operated with some five hundred correlators,(2) the grammar of Yerkish in its present implementation is limited to 46. Of these, 34 have so far been specified and are functioning (see Table 4). The remaining 12 will be filled as additions to the grammar become desirable from an experimental point of view.

The Grammar of Yerkish

The interpretive purpose of correlational grammars leads to a shifting of focus from characteristics of words and sentences, qua linguistic items, to the characteristics of concepts and conceptual structures, qua cognitive items. Ideally, a correlational grammar should be a complete mapping of the semantic connections between the elements and structures of a given language, on the one hand, and the elements and structures of conceptual representation, on the other. The bulk of work required to produce such a mapping for a given natural language is so vast as to be almost forbidding. Nevertheless, work in that direction continues under various headings and significant advances have been made (e.g. Schank, 1972, 1973). It will take a good deal more time to map the semantics of an average language user's universe of discourse, but that is hardly a reason for not going on with it, expecially since much of what has been done encourages the hope that the task can, indeed, be completed.

In designing an artificial language with a drastically curtailed universe of discourse, the problem is far more manageable. The semantic connections can be made as univocal as desired and, consequently, the process of interpretation can be thoroughly systematic. In the case of Yerkish, univocality was desirable not only with a view to the size of the automatic parser but also from the point of view of the teaching strategies to be employed with a non-human subject. Hence, Yerkish was

---

(2) cf. Final Scientific Report, Automatic English Sentence Analysis, (December 1969) Grant AFOSR 1319-67, Georgia Institute for Research, Athens, Georgia. (Obtainable through D.O.D.)

made as univocal as possible.

Since both on the linguistic and on the conceptual level we are dealing with _elements_ and their concaternation in _structures_, the interpretive grammar has to specify the connections (a) between linguistic and conceptual elements and (b) between linguistic and conceptual structures. With regard to the elements that are concatenated on the linguistic level, their semantic specification can be given in the lexicon because, here, we are dealing with a fixed set of items, i.e. precisely, lexical items. With regard to structures - phrases and sentences on the linguistic level, and situational representations on the conceptual level - they have to be specified by rules of composition on concatenation, i.e. by a grammar, because language is "open" in that direction and allows of a practically infinite number of individually different word concatenations.

Because Yerkish is based on English and the output of subjects in the experimental environment will be evaluated by English speakers, the lexical semantics of Yerkish lexigrams could be left implicit to a certain extent. Thus, for instance, the Yerkish parser does not have to contain an exhaustive semantic analysis of lexigrams such as BALL or RAISIN, because it can be taken for granted that the reader of the parser's output will be quite familiar with the concepts designated by "ball" or by "raisin" _qua_ experiential items. What the parser _must_ contain, however, is a mapping of those specific characteristics of the concepts which determine these items' potential for entering into structural relations with other items.

In Yerkish, then, the relational characteristics of conceptual items determine the classification of lexigrams. Thus, having decided, for instance, that there should be items that can be eaten and items that can be drunk, the lexigrams designating these items will be divided into _edibles_ (i.e. suitable patient/objects for the activity designated by EAT) and _drinkables_ (i.e. suitable patient/objects for the activity designated by DRINK). Together they constitute the class of _ingestibles_ which, as it happens, is marked by the red hue of the corresponding lexigrams (see Tables 2 and 3).

In short, Yerkish grammar does not require, nor lead to, a complete semantic analysis of lexical items. What it does require is a lexicon in which classes of lexical items are exhaustively characterized as to the specific relations into which their members can enter with members of other classes. This exhaustive characterization is supplied, not by listing all the other classes with whose members connections can be potentially formed, but by a string of indices, each of which specifies a connective relation and the place in it (c.f. below) a member of the class thus characterized can occupy.

Finally, we come to the relational concepts or _correlators_ which are instrumental in the building up of complex structures, both on the conceptual and on the linguistic level. Strictly speaking, a correlator is a connective function that links conceptual items on the cognitive-representational level. Languages indicate these connective functions by a variety of means: prepositions; verbs, nouns, and other types of words that incorporate a preposition; conjunctions and other particles; syntactic "markers" and, very frequently, merely word-order. Since these linguistic elements _indicate_ correlators, we should call them "correlator expressions". However, once it has been made clear that correlators <u>function</u> on the conceptual level and connect concepts with other concepts or combinations thereof, we can in most cases use the term "correlator" for both the relational concepts and the linguistic devices that express them.(3)

In designing an artificial language, the classification of lexical items and the definition or explication of relational concepts must go hand in hand since the first is done in terms of the second. The relational concepts have to be explicitly listed and explicated by some form of paraphrase. In principle, that is what a "case grammar" does. Its cases, basically, are relational concepts (e.g. Fillmore, 1968). However, because correlational grammar attempts to cover as much relational semantics as possible, its list of correlators will be both much longer and more specific than the lists of "cases" which, to my knowledge, have been suggested.

Yerkish; in its present form operates with some thirty correlators and the Yerkish lexicon is classified with reference to these (see Tables 3 and 5).

Given a basic list of correlators and their linguistic expression, the classification of lexical items can be carried out by listing for each item the correlators by means of which it can be <u>potentially</u>

---

(3) One area where the distinction has to be maintained is the semantic analysis of natural languages, because correlator expressions such as prepositions rarely have a one-to-one correspondence to relational concepts; instead, they merely mark the presence of one of a set of relational concepts.

linked to other items. To give an example, there is a relational concept (No. 11) paraphrased as active ingestion of solids involving solid food stuff'; on the linguistic level, this correlator is expressed by the juxtaposition of two lexical items in a certain order. If we now have a lexigram EAT, that designates active ingestion of solids' and another lexigram RAISIN, that designates a subcategory of 'solid food stuff', we can form a compound or <u>correlation</u> with the two lexigrams which can be represented as the structure:

```
(a)        EAT                              RAISIN
           └--------------11--------------┘
```

Because the order of succession of the two items in the linear linguistic expression is obligatory and cannot be reversed, it is not enough for the grammar merely to supply the information that the lexigrams EAT and RAISIN can be linked by correlator No. 11, but the grammar must also specify that, in this correlation, EAT has to be the left-hand piece (LH) and RAISIN the right-hand piece (RH).

This information is part of the permanent lexicon of the system. It is recorded there by means of "correlation indices" (I_c's), which consist of the number of the potential correlator plus the indication whether the items to which this I_c is assigned can function as LH-piece or as RH-piece. In many cases there are, of course, several lexical items that can function in the same place (e.g. NUT, M&M candy, RAISIN, etc., as RH-piece of correlator No. 11). Therefore , I_c's are assigned to lexigram classes, not to single lexical items. Thus, while the lexigram EAT, in the present Yerkish lexicon, is the only member of the class VE ('active ingestion of solids'), the lexigram RAISIN is one of several in the class EU ('solid food stuff'). On the one hand, this indexing of classes, rather than individuals, is obviously more economical with regard to storage space, on the other, it makes it possible to add new lexigrams to the existing classes without in any way disturbing the operative part of the lexicon or the parsing algorithms.

To expand the above example, let us add another correlation. The relational concept that can be paraphrased as 'autonomous animate actor' performing 'stationary activity' is correlator No. 01. The paraphrase 'autonomous animate actor' comprises three lexigram classes in the present lexicon, namely AP ('familiar primates', i.e., the regular technicians TIM, SHELLEY, BEVERLY, and the experimental animal LANA); AV ('visiting primates', i.e., unnamed human or non-human visitors); and AO ('non-primates', i.e., at present ROACH only). The paraphrase 'stationary activity', i.e. activities that do not involve a change of place on the part of the actor, nor a change of hands on the part of a patient, comprises three lexigram classes, namely VE (with the single member EAT), VD (with the single member DRINK), and VA (with several members such as: GROOM, TICKLE, HOLD, etc.).

Given the lexigram sequence LANA EAT, the interpretive grammar finds that LANA, belonging to class AP, bears the I_c: 01, LH, while EAT, belonging to class VE, bears the I_c: 01, RH; and on the strength of this the grammar will allow the correlation:

```
(b)                 LANA                      EAT
                     └----------01----------┘
```

For the parser, "allowing a correlation" means to record it as a possible part-interpretation of the input string. As such it is recorded as a "product" in order to be tested for its potential correlability with other parts of the input.

The information, on the basis of which such first-level correlations (connecting single lexigrams as in a and b) are formed, is contained in the permanent lexicon and the form in which it is stored can be visualised as a kind of matrix (see Fig. 2 and Table 5).

Lexigram Class	Correlators										
	01		02			I1			n	
	LH	RH	LH	RH	.....	LH	RH	....	LH	RH
AP	x									
AO	x									
.										
.										
.										
EU							x			
.										
.										
.										
VE		x				x				
.										
.										
.										
.										
.										
m										

Figure 2

The correlational data required to form examples (a) and (b) is represented by markers (x) indicating the I_c's (at head of column) assigned to the lexigram classes (at beginning of row). In the present implementation of the Yerkish grammar n = 34, m = 35.

Though this information contained in the lexicon covers all correlations involving two single lexigrams, it does not provide for correlations linking phrases or phrases and lexigrams. The system a correlational grammar uses to discover higher-level structures in a given input string is again rather different from that of traditional grammar.

In order to be able to handle phrases, i.e., already correlated lexigrams, or "products", in exactly the same way as single lexical items, each product must be assigned a string of I_c's that represents its particular potential for functioning as component (LH-piece or RH-piece) of a new and larger correlation that links it with other lexical items or phrases.

The procedure that assigns these Ic's to a given product is what might be called the dynamic part of the grammar, because it is governed by a set of <u>operational</u> rules that cannot be stated in a simple formalized way.[4] The reason for this is that the correlability of a given phrase often depends on more than one constituent of the phrase.

An example may help to make this clear. With regard to correlator No. 30 that links the two single lexigrams involved, the phrases

```
THIS            RAISIN          and          THIS             BALL
 └-------30-------┘                           └-------30-------┘
```

are identical. As potential RH-pieces of a correlation formed by correlator No. 11, however, they are not equivalent.

```
EAT            THIS            RAISIN
 |               └-------30-------┘
 └-------11--------------┘
```

would be acceptable and correct, whereas

```
EAT            THIS            BALL
 |               └-------30-------┘
 └-------11
```

would not be acceptable because BALL does not belong to the lexigram class EU defined as 'solid food stuff' and, therefore, is not a potential RH-piece

(4) The operational rules are, of course, always combinations of individually simple rules taken from a relatively small set. This is, indeed the way in which the parsing program compiles them; although this <u>can</u> be called 'formalisation' it certainly is not a simple one.

of correlation No. 11. In fact, if the string EAT THIS BALL occurs as input to the interpretive grammar, it must be rejected as incorrect.

To implement this discrimination, the phrase THIS RAISIN must be assigned the I_c:11, RH, while the phrase THIS BALL must not. And in order to do this, the assignation must be based not only on the particular correlator that links THIS with another item, but also on the condition that this other item is one that belongs to the lexigram class 'solid food stuff'. In other words, there has to be an operational assignation rule that makes sure that a first-level correlation produced by correlator No. 30 is assigned the I_c:11, RH, so that it can be linked in a second-level correlation with the preceding lexigram EAT, which bears the I_c 11, LH; but this assignation must be made contingent upon the condition that the product 30 (P:30) does, in fact, contain a lexical item of class EU as RH-piece; because only if P:30 contains a member of the class 'solid food stuff' can it function as patient of the activity designated by the LH lexigram EAT.

The operational assignation rules, therefore, are of diverse types, some assigning I_c's unconditionally, others assigning I_c's only on condition that the same I_c is present, as the case may be, among those characterising the LH or the RH of the product that is being classified. (See Table 6).

In the implementation of the parser, the assignation of I_c's to products is primarily determined by the particular correlator that is involved in the product to be classified. The assignation rules a particular correlator calls into action, though functionally of three types only, are specific to that correlator and cannot be written in a generalised form: This indeed, is the fundamental reason why a correlational grammar cannot be represented by means of a small number of relatively "powerful" rules. In a correlational grammar there must be as many sets of specific assignation rules as there are correlators; and since the number of correlators in such an interpretive grammar is very much larger than the number of "syntactic functions" in conventional descriptive grammars, correlational grammars connot be written in concise and powerful formulas. As a justification for this lack of formal elegance, however, it can be said that correlational grammar has no need of the otherwise indispensable and somewhat unwieldy adjunct of "selection rules", because it incorporates that very information in its one basic interpretive algorithm.

Peculiarities of the Yerkish Grammar

The grammar of Yerkish had to be kept as simple as possible for the reasons mentioned above. First, given the small size of the computer, it was mandatory to avoid complex constructions and rules of grammar that might require special space- and time-consuming subroutines in the parsing procedure. Second, the rules of the language to which the linguistic behavior of our subject would have to conform, were to be few and consistent from the learner's point of view; nevertheless they were to be such that Yerkish structures could be translated easily and without major structural transformations into comprehensible English. As a result of these objectives, Yerkish grammar may seem somewhat unusual. In the following paragraphs several deviations from English grammar will be discussed.

Yerkish, at present, has only one voice, the active, and three moods, i.e. indicative, interrogative, and imperative. Both the interrogative and the imperative are formed, not by specific verb-forms or word-order (as in many natural languages), but by sentential prefixes or markers, i.e. specific lexigrams that are placed at the beginning of the message. The prefix of the interrogative is the conventional question mark "?", that for imperatives (in Yerkish "requests") is an arrow, translated into English as PLEASE. The keys representing these two lexigrams must be pressed at the beginning of a string and they can appear only in the first feedback projector on the left. The lexigram string following them has the same form as an indicative statement. In fact, if the string is to be interpreted as an indicative statement, i.e.. if it is not preceded by either "?" or PLEASE, the first feedback projector on the left remains blank. Thus:

TIM MOVE INTO ROOM	=	indicative statement;
? TIME MOVE INTO ROOM	=	interrogative;
PLEASE TIM MOVE INTO ROOM	=	request.

A third lexigram that functions as a sentential marker is NO, which corresponds to an over-all negation of the statement. NO TIM MOVE INTO ROOM, therefore, corresponds to the English "it is not the case that Tim moves into the room". However, since Lana has quite spontaneously come to use the lexigram NO to designate what, given the situational context, could be interpreted only as "don't", we may adapt the grammar to her usage and turn this NO into a marker for negative imperatives (see Appendix).

Yerkish, as yet, has no tenses but the present. A simple past and future, however, are foreseen, and they will be designated by particles preceding the activity lexigram.

There are no auxiliaries in Yerkish and the function of the English copula "to be" is taken over by correlator No. 10, which is expressed by juxtaposition of a lexigram belonging to one of the classes of items that are "modifiable" and a lexigram designating a specific state.

e.g. BALL RED. = "the ball is red";

TIM AWAY. = "Tim is away";

The absence of an explicit copula is noticeable also in conjunction with the "naming function", an important instrument in Lana's acquisition of new lexical items. It is used for the ostensive definition of new lexigrams which are placed at the beginning of a string of the form:

XX NAME-OF THIS. = "XX is the name of this".
(where XX is the new lexigram)

Two English constructions that have a specificatory restrictive function, i.e. for instance, "the red ball" and "the ball which is red" are one and the same in Yerkish, and the specificatory relation is designated by a lexigram which can be translated into English as the compound WHICH-IS (correlator No. 31).

e.g. BALL WHICH-IS RED. "the red ball" or "the bal which is red".

For the sake of greater univocality, Yerkish spatial prepositions were strictly divided into locational and directional ones. The first -- e.g. IN, ON, OUTSIDE, etc. -- could designate only the location of items or activities, the second -- e.g. INTO, OUT-OF, FROM, etc. -- could designate only the direction of activities involving a change of place. However, since Lana has spontaneously used a locational preposition to indicate the target of a directional activity, and since this is allowable in many if not all natural languages, we are considering the removal of this restriction with regard to spatial prepositions.

So far, there are no conjunctions in Yerkish, but a somewhat restricted form of "and" and "or" has been worked out and will shortly be introduced into the system.

There are also some minor pecularities that an English-speaker must keep in mind. A Yerkish structure involving correlator No. 17, for instance, implies that the speaker is the receiver of the item that changes hands, unless another receiver is explicitly indicated by a prepositional complement. Thus, if Lana sends the message:

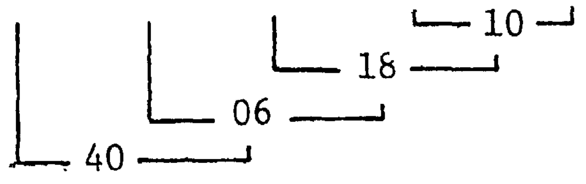
```
PLEASE  TIME  GIVE  MILK.
  |       |     └─17─┘
  |       └─05────┘
  └─40─────┘
```

it must be understood that the milk is to be given to Lana. The receiver, however, can be made explicit by adding a prepositional phrase, which yields the correlational structure:

```
PLEASE  TIM  GIVE  MILK  TO  LANA.
  |      |    └─17─┘     └─22─┘
  |      |       └───21───┘
  |      └─── 05 ───┘
  └── 40 ───┘
```

English resultative verbs, e.g. "to open", "to clean", etc., are broken up in Yerkish. The causative element is rendered by MAKE, the effect by a lexigram designating the resulting state. Also, in Yerkish the agent must be specified. Thus, "Please (Tim) open the window" becomes:

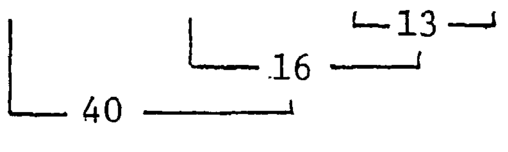
PLEASE TIM MAKE WINDOW OPEN.



Translated literally into English, this should be "Please, Tim, make window be open", since the correlator that links WINDOW and OPEN is No. 10, i.e. the predicative copula equivalent to "to be". But in this case, as indeed in most occurrences of correlator No. 10, the Yerkish string is easily understood without the explicit copula.

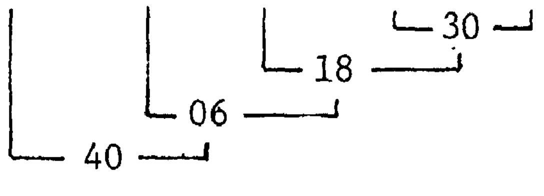
The Yerkish MAKE is not limited to causation of a change of state of specific items, but can be used also to indicate a number of perceptual conditions or events in the environment. Specific sensory events or changes, such a MOVIE, MUSIC, SLIDE, HEAT, COLD, LIGHT, and DARKNESS, are considered the result of activities subsumed by MAKE. In Lana's wholly technological environment, this is not at all unreasonable. It obviously makes sense for her to request:

PLEASE MACHINE MAKE MOVIE.



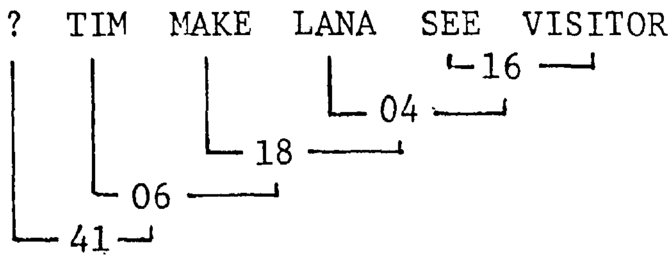
It is, indeed, the machine that causes the projector to start running. Similarly, however, in Yerkish one could correctly say:

PLEASE TIM MAKE LESS HEAT.

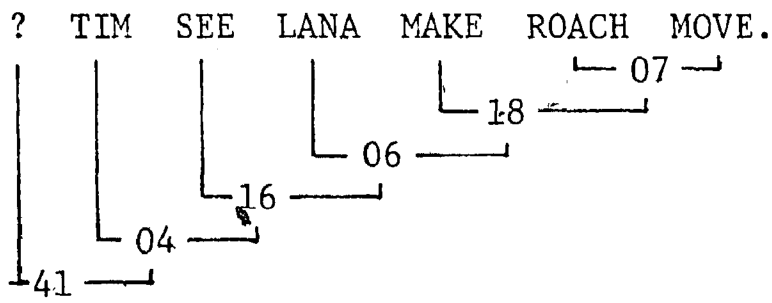


Though in Lana's experience Tim can indeed "cause less heat" by turning down the thermostat, this would hardly be a reasonable request in the "real" world outside the Yerkes Lab.

MAKE also opens the way to embedded constructions, since it can govern a clause. Though Lana has not yet come to this, the grammar foresees strings such as :

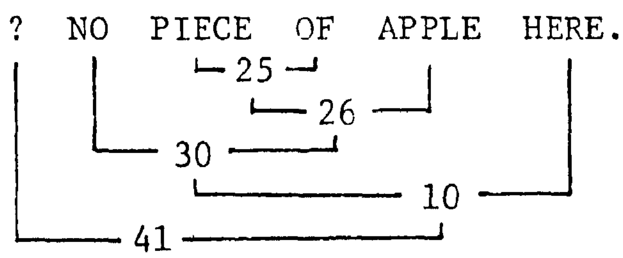


or even a double embedding:

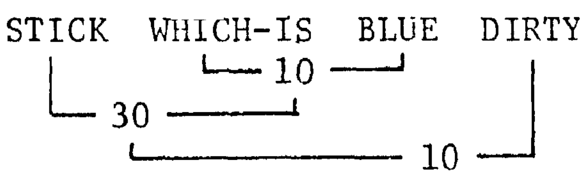


and similar structures are, of course, possible with WANT.

Lest these correlational diagrams create the impression that Yerkish structures are invariably right-branching, here are two examples that contain left-branchings:



which, in English, would read: "Is there no piece of apple here?" And



which in English, would be: "The stick which is blue is dirty", and, as such, roughly equivalent to "The blue stick is dirty."

In one particular the grammar of Yerkish deviates from correlational practice. Prepositions and conjunctions being "explicit" correlator expressions in that they designate relational concepts only, are (in the correlational approach) not items to be linked, but items that do the linking. Thus, in the original Multistore parser they functioned as correlators and not as ordinary lexical items. In the structure diagrams, therefore, they appeared in a node, not at a terminal. Given the very small computer used for the Yerkish parser, as well as the fact that the lexicon was to remain extremely limited (in comparison to English), it was more economical to correlate prepositional phrases in two steps rather than introduce the special routine that had been developed for prepositions and other "explicit" correlator expressions in the English parser. Thus a string such as "move into room" is not constructed as it would be in a proper correlational system, i.e.:

```
MOVE                ROOM.
 └--------into--------┘
```

but rather in two steps:

```
MOVE       INTO      ROOM
 |           └---22---┘
 └----21--------┘
```

where P:21 containing P:22 expresses the conceptual relation designated by INTO.

In all other respects the Yerkish system is similar to the Multistore parser whose characteristic data-compression was, in fact, the feature that made it possible to contain the entire system - lexicon, operational interpretive grammar, and automatic response programs - in less than 5000 machine words of central core.

Conclusion

Though Yerkish is, indeed, an extremely limited linguistic system, the examples of sentence structure I have used above should be sufficient to show that it has a considerable range and flexibility with regard to what can be formulated in it. The reports on Lana's progress that we have published so far (Rumbaugh et al., 1973b, 1974; von Glasersfeld, 1974b; see also Appendix to this paper) leave little doubt that Lana has already acquired a number of skills that certainly belong to what is usually called linguistic competence. The grammar of Yerkish as it is at present allows many structures which are still far out of Lana's reach. She has a long way to go before one might venture to say that she fully exploits the expressive possibilities of Yerkish. That is precisely how it was intended at the outset of the project.

In any case, the range of expression could easily be extended at short notice and without interfering with the existing operational system. Nor would such additions require an inordinate expansion of the lexicon. The introduction of the one lexigram WANT, for instance, has opened the way to a completely new level of expression that may eventually lead to a demonstration of the chimpanzee's capability for conceptual representation. The addition of a Yerkish "if. . . then" would be no more difficult and could perhaps further clarify the cognitive potential of non-human primates.(5) Such additions could also make Yerkish a

---

(5) Premack (1971) reported that his chimpanzee Sara could correctly interpret an "if. . . then" connection between actual activities or states; the greater range and flexibility of Yerkish would make possible the introduction of much more sophisticated hypothetical statements.

valuable communication vehicle for some of those many unfortunate children who, though they are mentally not at all deficient, remain averbal because of some physiological damage. It is towards this end that we are now exploring the possibility of adapting the Yerkish system, its grammar, and the parser, to a form of simple English.

References

Brown, Roger (1971), in Ploog and Melnechuk, 1971.

Ceccato, Silvio, et al. (1970), Linguistic Analysis and Programming for Mechanical Translation, Milan, Italy: Feltrinelli.

Ceccato, Silvio, et al. (1963), Mechanical Translation: The Correlational Solution (Technical Report), Milan, Italy: Center for Cybernetics, University of Milan.

Chomsky, Noam (1965), Aspects of the Theory of Syntax, Cambridge, Massachusetts: M.I.T. Press.

Crawford, James (1973), personal communication.

Fillmore, C. J. (1968), The case for case, in (E. Bach and R. Harms, Eds.) Universals in Linguistic Theory, New York: Holt, Rinehart, & Winston.

Gardner, R. A. and B. T. (1969), Teaching sign language to a chimpanzee, Science, 165, 664-672.

Gardner, B. T. and R. A. (1971), Two-way communication with an infant chimpanzee, in (A.M. Schrier and F. Stollnitz, Eds.) Behavior of Nonhuman Primates, Vol. 4, New York: Academic Press.

Hayes, Keith J. and Nissen, C. H. (1956), in (A.M. Schrier and F. Stollnitz, Eds.) Behavior of Nonhuman Primates, New York: Academic Press, 1971.

Hockett, Charles F. (1961), Linguistic elements and their relations, Language 37, 1, 29-53.

Köhler, Wolfgang (1925), The Mentality of Apes, New York: Harcourt, Brace.

Lenneberg, Eric (1971), in Ploog and Melnechuk, 1971.

Lieberman, P. (1968), Primate vocalization and human linguistic ability, Journal of the Acoustic Society of America, 4, 1574-1584.

Lorenz, Konrad (1974), Analogy as a source of knowledge, Science 185, 229-234.

Lyons, John (1972), Human language, in (R.A. Hinde, Ed.) Non-Verbal Communication, Cambridge, England: Cambridge University Press.

Ploog, Detlev, and Melnechuk, Theodore (1971), Are apes capable of language?, Neurosciences Research Bulletin, 9, 5.

Premack, David (1971), On the assessment of language competence in the chimpanzee, in (A.M. Schrier and F. Stollnitz, Eds.) Behavior of Nonhuman Primates, Vol. 4, New York: Academic Press.

Rumbaugh, Duane M. (1970), Learning skills of anthropoids, (in L. A. Rosenblum, Ed.) Primate Behavior, Vol. 1, New York: Academic Press.

Rumbaugh, D. M., von Glasersfeld, E., Warner, H., Pisani, P., Gill, T., Brown, J., and Bell, C. (1973a), A computer-controlled language training system for investigating the language skills of young apes, Behavioral Research Methods and Instrumentation 5, 5, 385-392.

Rumbaugh, D. M., Gill, T. V., and von Glasersfeld, E., (1973b), Reading and Sentence completion by a chimpanzee, Science 182, 731-733.

Rumbaugh, D. M., von Glasersfeld, E., Warner, H., Pisani, P., and Gill, T. (1974), Lana (chimpanzee) learning language: A progress report, Brain and Language, 1, 205-212.

Schank, Roger (1972), Conceptual dependency: A theory of natural language understanding, Cognitive Psychology, 3, 4, 552-631.

Schank, Roger (1973), Causality and Reasoning, Technical Report No. 1, Castagnola, Switzerland: Fondazione Dalle Molle.

von Glasersfeld, Ernst (1964), A project for automatic sentence analysis, Beiträge zur Sprachkunde und Informationsverarbeitung, 4, 38-46.

von Glasersfeld, Ernst (1964); Multistore - a procedure for correlational analysis, Automazione e Automatismi, 9, 2.

von Glasersfeld, Ernst (1970), The correlational approach to language, Pensiero e Linguaggio, 1, 4, 391-398.

von Glasersfeld, Ernst (1974a), Signs, communication, and language, Journal of Human Evolution, 3, 465-474.

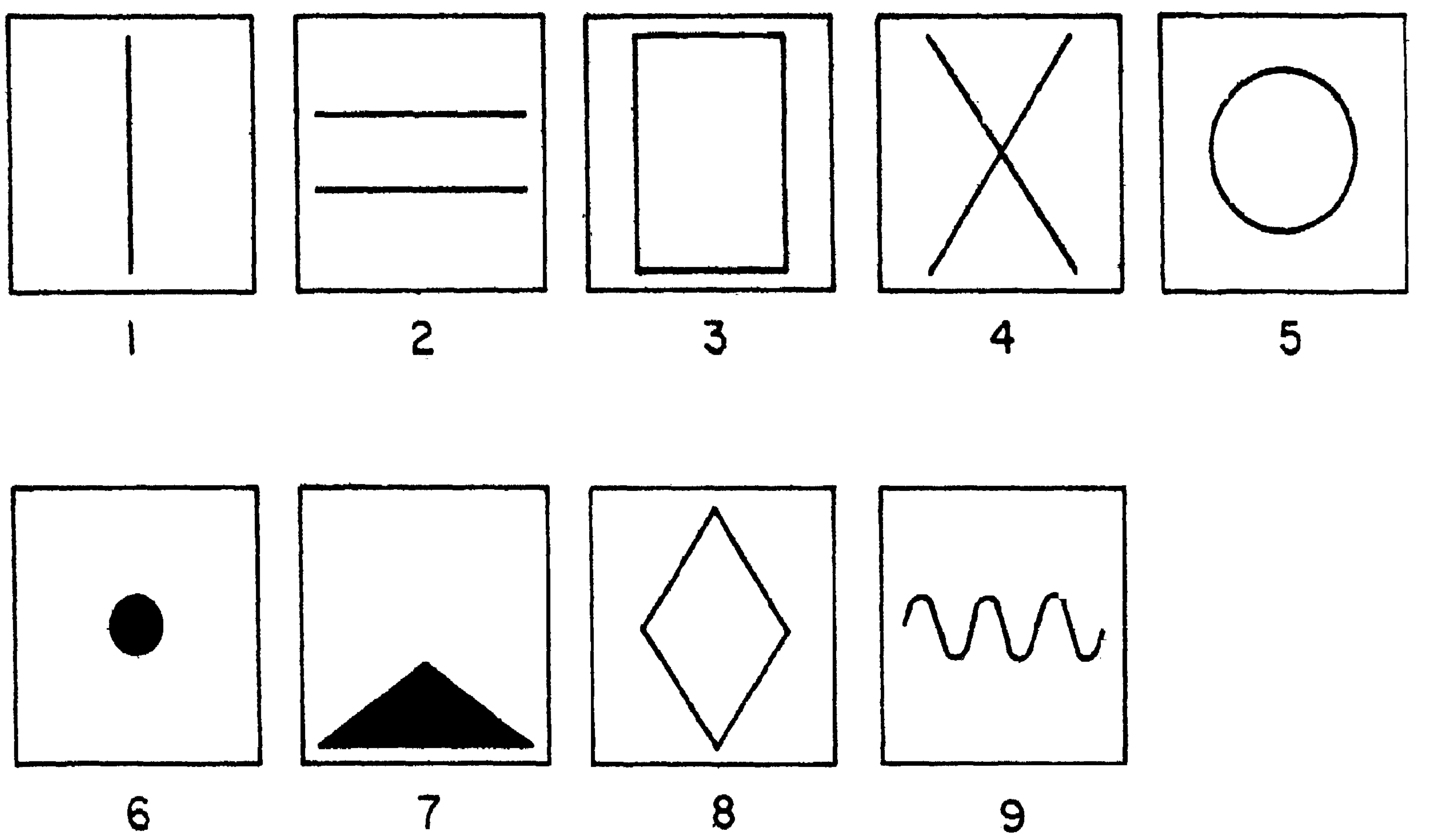
von Glasersfeld, Ernst (1974b), Lana's progress, paper presented at the 12th Meeting of the Association for Computational Linguistics, Amherst Massachusetts, July, 1974.

von Glasersfeld, Ernst, and Pisani, Pier Paolo (1970), The Multistore parser for hierarchical syntactic structures, Communications of the Association for Computing Machinery, 13, 2, 74-82.

von Glasersfeld, Ernst, and Pisani, Pier Paolo (1968), The Multistore System MP-2, Scientific Progress Report, Athens, Georgia: Georgia Institute for Research.

Yerkes, Robert M. (1925), Traits of young chimpanzees, in (R. M. Yerkes and B. W. Learned, Eds.) Chimpanzee Intelligence and its Vocal Expression, Baltimore: Williams & Williams.

DESIGN ELEMENTS



EXAMPLES OF LEXIGRAMS

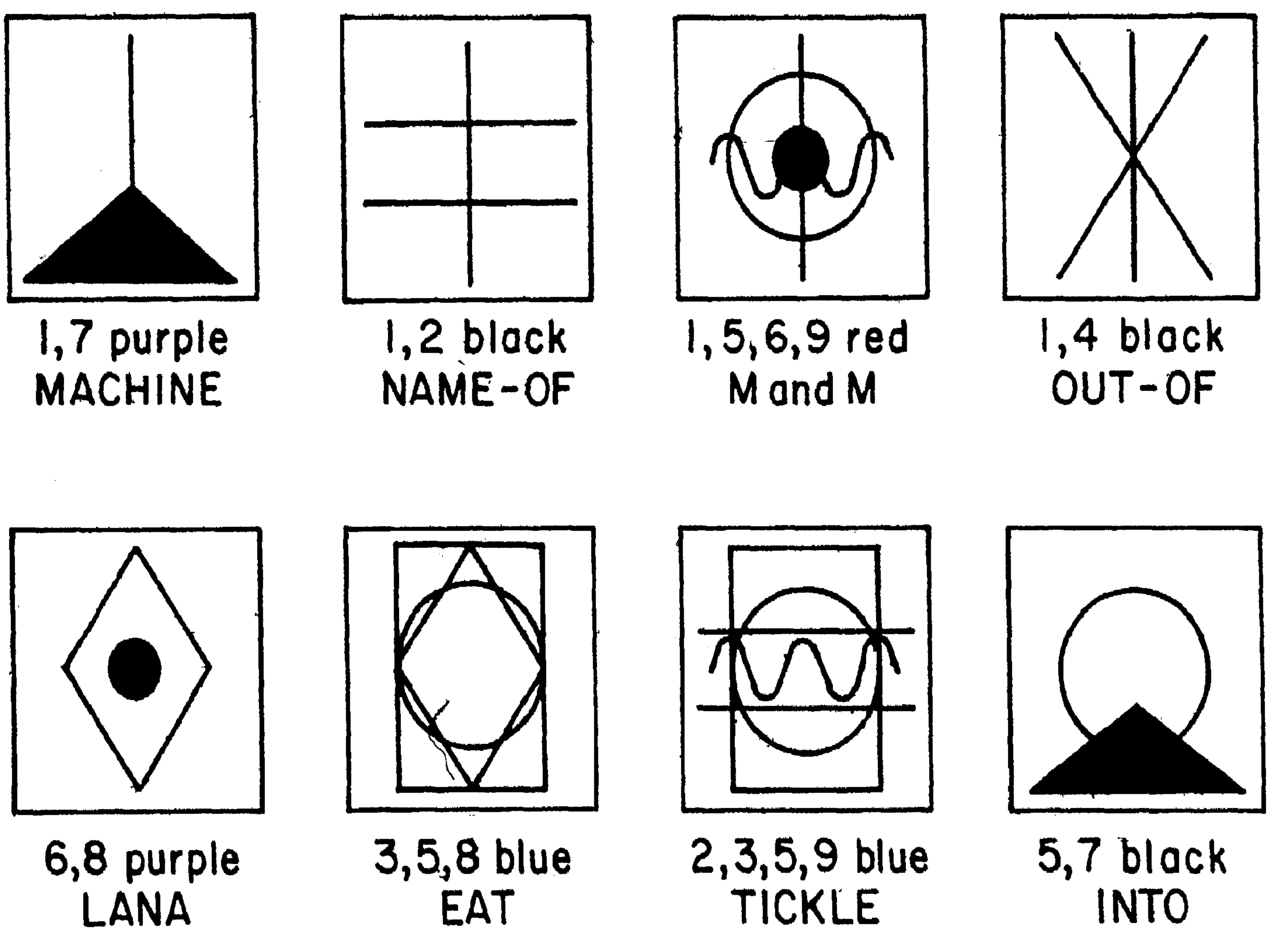


Table 2

Semantic Colour-Coding of Lexigrams

Colour	General Type	Lexigram Classes
violet	Autonomous Actor	AP, AV, AO, AM.
orange	Spatial Objects, Spatial Concepts	FA, FP, TF, CT, WR.
red	Ingestibles	EU, EM, ED.
green	Parts of Body	PB.
blue-grey	States and Conditions	ST, LS, CD.
blue	Activities	VA, VB, VC, VD, VE, VG, VL, VM, VP, VS, VW.
black	Prepositions, Determiners, particles	DC, DD, DQ, DP, LP, ID, NF, PP.
white (+)	Affirmation	"YES"
yellow (+)	Sentential Modifiers	Query, Please, Negation, Period.

(+) White and yellow are available in the first feedback projector (left) only, where the sentential modifiers appear.

Table 3

Operational Lexigram Classes

Autonomous Actors		
AP	Familiar primate: (human)	BEVERLY, SHELLEY, TIM
	(non-human)	LANA
AV	Unfamiliar primate:	VISITOR
AO	Non-primates:	ROACH
AM	Inanimate actor:	MACHINE
Absolute Fixtures (items that cannot move or be moved)		
FA		FLOOR, KEYBOARD, ROOM
Relative Fixtures (items that can move but not change place)		
FR		DOOR, PUSHKEY, WINDOW
Transferables (items that can change place and/or hands)		
TF		BALL, BLANKET, BOWL, BOX, BUCKET, DOLL, PLATE, STICK
Edibles		
EU	Dispensed as unit:	M&M (candy), RAISIN
EM	Dispensed in pieces:	APPLE, BANANA, BREAD, CHOW, COOKIE
Drinkables		
ED		COKE, JUICE, MILK, WATER
Parts of Body (can change place but not hands)		
PB		BACK, EAR, EYE, FINGER, FOOT, HAND, HEAD, MOUTH, TUMMY
States (as attributed to items)		
ST		BLACK, BLUE, CLEAN, COLD, DIRTY, DRY, GREEN, HARD, HOT, OPEN, ORANGE, PURPLE, RED, SHUT, SOFT, WET, WHITE, YELLOW

(Table 3, Lexigram Classes, continued)

Locational States (as attributed to items)	
LS	AWAY, DOWN, HERE, UP
Ambiental Conditions (sights, sounds, smells, etc., are treated as states of the environment that can be caused (MAKE) by an agent)	
CD	COLD n., DARKNESS, HEAT, LIGHT n., MOVIE, MUSIC, SLIDE n., VOICE
Conceptual Categories (as applied to spatio-temporal items)	
CT	BEGINNING, BOTTOM, COLOR, CORNER, END, PIECE, SIDE, TOP
Stationary Activities	
VE Ingestion of solids:	EAT
VD Ingestion of liquids:	DRINK
VA Relational motor act:	BITE, GROOM, HIT, HOLD, TICKLE
Locomotive Activities	
VB Transferring (change of place):	CARRY
VC Requiring contact and application of force:	PULL, PUSH
VG Causing change of hands:	BRING, GIVE
VL Change of place:	MOVE, SWING, TURN
Stative Activities	
VS Maintaining position in place:	LIE, SLEEP, STAND
Conceptual Activities	
VM Causative, creating change:	MAKE
VW Conative activity:	WANT

(Table 3, Lexigram Classes, continued)

<u>Locational Prepositions</u>	
LP	IN, ON, OUTSIDE, UNDER
<u>Directional Prepositions</u>	
DP	BEHIND, FROM, INTO, OUT-OF, THROUGH, TO, TO-UNDER
<u>Partitive Preposition</u>	
PP	OF
<u>Determiners</u>	
DD Demonstrative:	THIS, WHAT
DQ Quantitative:	ALL, MANY, NO (not one), ONE
DC Comparative:	LESS, MORE
<u>Semantic Indicator</u>	
NF Indicating semantic nexus:	NAME-OF
<u>Identity-Difference Markers</u>	
ID	SAME-AS, OTHER-THAN
<u>Attributive Marker</u> (also relative clause marker)	
WR	WHICH-IS
<u>Sentential Markers</u> (at beginning of sentence)	
Request (imperative)	PLEASE
Query (interrogative)	"?"
Negation	NOT
Affirmation	YES
End-of-message marker (at end of sentence)	"." (period)

Table 4.

Operational Correlators

Actor/Activity	LH		RH
01	autonomous actor AP, AV, AO.	performing	stationary activity VA, VD, VE.
02	idem	performing	transferring activity VB.
03	idem	performing	act. requiring contact and force VC.
04	idem	performing	perceptual activity VP.
05	autonomous actor AP, AV, AM.	performing	causing item's change of hands; VG.
06	causative agent AP, AV, AM.	performing	causing change of state VM.
07	actor AP, AV, AO, FP, TF, EU, PB.	performing	change of place VL.
08	item capable of changing spatial location AP, AV, AO, FP, TF, EU, EM, ED, PB.	performing	stative activity VS.
09	conative agent AP, AO, AV.	performing	conative activity VW.
Predicative Copula			
10	item having perceptual characteristics AP, AV, AO, FA, FP, TF, EU, EM, ED, PB, CD, WR.	described by	predicated state ST. LS.
11	ingestion VE.	involving as patient	solid food stuff EU, EM.

(Table 4, correlators, continued)

Activity/Complement

	LH		RH
12	ingestion VD.	involving as patient	liquid ED.
13	stationary motor activity VA	involving as patient	any spatial item AP, AV, AO, FA,FP, TF, EU, ED, PB. (*)
14	transferring VB.	involving as patient	item capable of change of place AP, AV, AO, TF, EU, ED, PB.
15	contact and force VC	involving as patient	any spatial item (same as for 13!)
16	perceptual activity VP.	involving as result	any percept. item AP, AV, AO, FA, FP, TF, EU, EM, ED, PB, CD.
17	change of hands VG.	involving as patient	handable item AO, TF, EU, ED. (*)
18	causing change VM.	involving as result	condition or state CD, CS.
19	conative activity VW.	involving as result	desired item AO, TF, EU, ED, CD, VE, VD, VS.

Activity/Spatial Adjunct

	LH		RH
21	change of place VC, VL, (and P's: 14,15,17)	involving as target	target location LS, (and P's:22)
22	directional prep. DP.	step 1 for corr. 21	specification of target AP, AV, AO, FA, FP, TF, EU, EM, ED, PB.

(*) "Edible materials" (FM), which are dispensed in slices only, are at present excluded from this correlation because we do not want Lana to formulate requests such as: PLEASE LANA HOLD BANANA, since she is normally not given whole bananas, etc.

(Table 4, correlators, continued)

23	stative activity VS, (and P's: 11, 12,13).	localization	specification of location LS, (and P's:24)
24	locat. prep. LP	step 1 for corr. 23	specification of location (same as for 22!)

Relation Whole/Part

	LH		RH
25	item considered 'part' PB, CT.	step 1 for corr. 26	partitive preposition PP
26	item considered 'part' P's:25	part-whole	item considered 'whole' AP, AV, AO, FP, TF, EU EM, ED, PB, DD.

Naming Function	LH		RH
27	semantic indication NF.	step 1 for corr. 28	item to be name AV, DD (*)
28	new lexigram or WHAT.	nexus to be formed	item designated

Conceptual Categorization.

	LH		RH
30	determiner DO, DC, DD.	applied to	any item singled out AP, AV, AO, FA, FP, TF, EU, EM, ED, PB, CD, CT.

Relative Clause	LH		RH
31	item to be qualified AP, AV, AO, FA, FP, TF EU, EM, ED, PB, CD.	attribution	restrictive marker WHICH-IS (WR)

Comparative State	LH		RH
32	quantitat. det. DQ.	qualification	state ST., LS, LP, DP.

(*) DD, i.e. the demonstrative determiner THIS implies ostension of the item to be named.

(Table 4, correlators, continued)

Identity Function	LH		RH
33	identify-difference marker ID	applied to	term of comparison AP, AV, AO, FA, FP, TF, EU, EM, ED, PB, CD, DD.

Sentence Modifiers	LH		RH
40	request marker PLEASE	imperative	expression turned into request.
41	question marker QUERY	interrogat.	expression turned into question.
46	negation marker NOT	sentential negation	expression negated.

TABLE 5

ASSIGNATION OF CORRELATOR INDICIES TO EXISTING LEXIGRAM CLASSES

CORRELATOR INDICIES

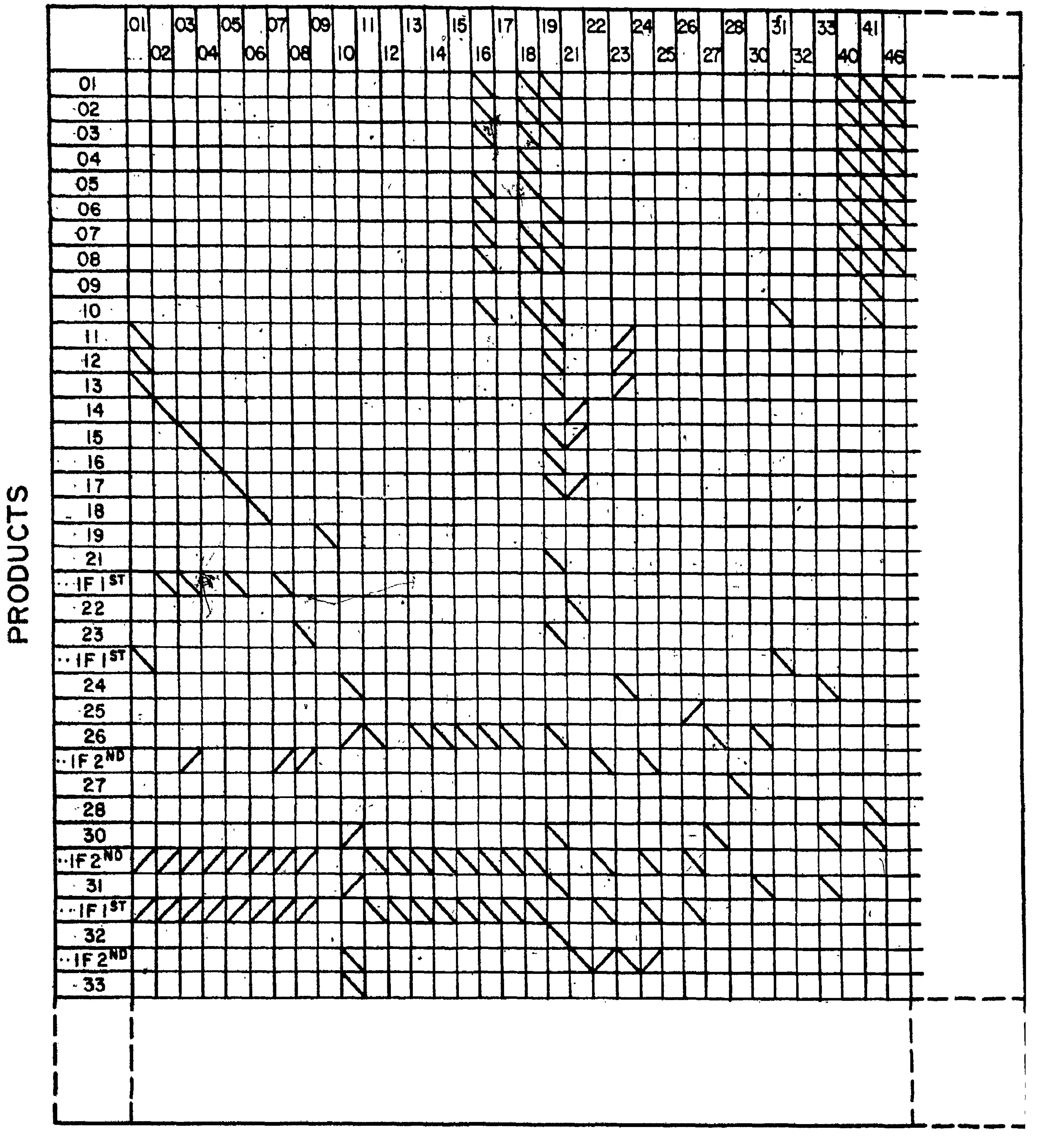
LEXIGRAM CLASSES	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	21	22	23	24	25	26	27	28	30	31	32	33				
AP	/	/	/	/	/	/	/	/	/	/			\	\	\	\					\		\		\		/	\	/		\				
AV	/	/	/	/	/	/	/	/	/	/			\	\	\	\					\		\		\	\	/	\	/		\				
AO	/	/	/	/			/	/	/	/			\	\	\	\	\		\		\		\		\		/	\	/		\				
AM					/	/																													
FA										/			\		\	\					\		\				/	\	/		\				
FP			/				/	/		/			\		\	\					\		\		\		/	\	/		\				
TF			/				/	/		/			\	\	\	\	\		\		\		\		\		/	\	/		\				
EU			/				/	/		/	\		\	\	\	\	\		\		\		\		\		/	\	/		\				
EM								/		/						\			\		\		\		\		/	\	/		\				
ED								/		/		\	\	\	\	\			\		\		\		\		/	\	/		\				
PB			/				/	/		/			\	\	\	\					\		\	/	\		/	\	/		\				
ST										\																	/			\					
LS										\										\		\					/			\					
CD										/						\		\	\								/	\	/		\				
CT																								/			/	\							
VE	\										/								\								/								
VD	\											/							\								/								
VA	\												/														/								
VB														/													/								
VC			\												/					/							/								
VG																	/										/								
VL							\													/							/								
VS								\											\			/					/								
VM																		/									/								
VP																/											/								
VW																			/																
LP																							/				/			\					
DP																					/						/			\					
PP																								\											
DD																									\	\	/	/			\				
DQ																											/	/							
DC																											/	/		/					
NF																										/									
IF																															/				
WR										/																									

/ = LH-piece

\ = RH-piece

TO PRODUCTS (RECLASSIFICATION)

CORRELATOR INDICIES



/ = LH-piece

\ = RH-piece

··IF 1st = Assignation on condition that the P's LH-piece bears the I_C to be assigned to the P;

··IF 2nd = Assignation on condition that the P's RH-piece bears the I to be assigned to the P

Appendix

One of the first of Lana's spontaneous 'generalisations' concerned the lexigram NO. She had learned the use of this lexigram in one specific context, i.e. in answers to questions such as: ? BANANA NAME-OF THIS, when the object ostensively indicated to her was, for instance, her blanket. One morning, Tim Gill, to whose ingenious devising of training situations the project owes a great deal of its success, entered the room with a bowl of banana slices. As he moved around the corner of Lana's cubicle in order to fill the dispenser that responds to the keyboard message PLEASE MACHINE GIVE PIECE OF BANANA, he popped a banana slice into his mouth. Seeing this, Lana adopted a threatening posture and hooted angrily. Then, suddenly, she ran to the keyboard and, three times in succession, vigorously pressed the key bearing the lexigram NO.

Conversation recorded on May 6th, 1974.

On the preceding days Lana had learned the lexigrams for a bowl and a metal can, BOWL and CAN. This had been accomplished by first using objects whose names were already known to her, putting an M&M candy inside them, and asking her: ? WHAT NAME-OF THIS. On May 5th she reliably replied with the correct lexigram when the reward was placed in the bowl or in the can. The next morning Tim came in with the bowl, the can, and a cardboard box. While Lana was watching, he put an M&M candy in the box, and the following exchange took place:

Lana:	? TIM GIVE LANA THIS CAN.
Tim:	YES. (Tim gives her the empty can, which she at once discards)
Lana:	? TIM GIVE LANA THIS CAN.

Tim: NO CAN.

Lana: ? TIM GIVE LANA THIS BOWL.

Tim: YES. (Tim gives her the empty bowl)

Lana: ? SHELLEY - (Sentence unfinished)

Tim: NO SHELLEY. (Shelley, another technician who worked with Lana, is not present)

Lana: ? TIM GIVE LANA THIS BOWL

(Before Tim can answer, Lana goes on).

Lana: ? TIM GIVE LANA NAME-OF THIS.

(A spontaneous generalisation of GIVE, not foreseen by the grammar, since NAME-OF had not been classified as a possible object of GIVE!)

Tim: BOX NAME-OF THIS.

Lana: YES. (Short pause, and then)

? TIM GIVE LANA THIS BOX.

Tim: YES. (Tim gives it to her, she rips it open and eats the M&M)

Conversation recorded on November 22, 1974.

Tim: ? LANA WANT APPLE. (No apple is in sight)

Lana: YES. (Tim leaves the room and, after a moment, returns with an apple)

Lana: YOU GIVE THIS TO LANA.

Tim: ? GIVE WHAT TO LANA.

Lana: YOU GIVE THIS WHICH-IS RED.

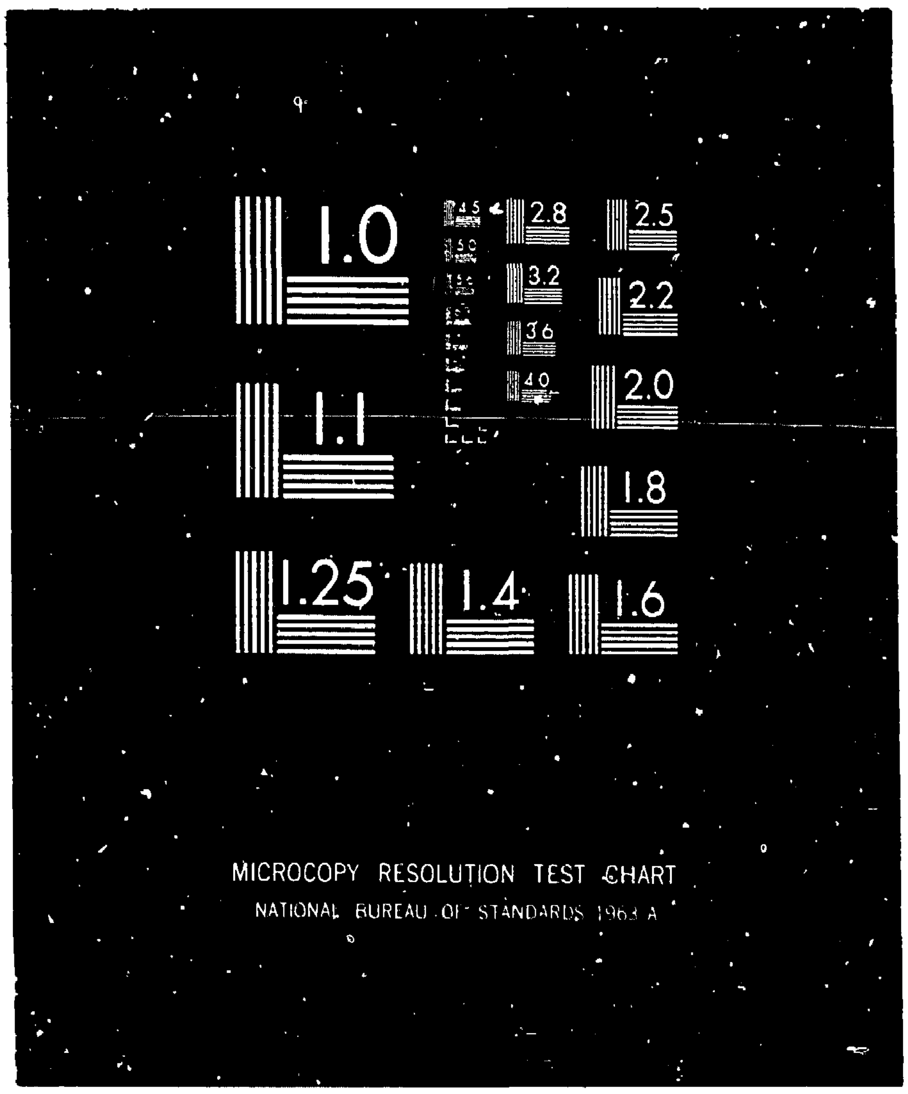
(Since there is no sentence marker, this is an indicative statement and neither a request nor a question)

Tim: ? THIS. (Tim holds up a red piece of plastic)

Lana: ? YOU GIVE THIS APPLE TO LANA.

Tim: YES. (Whereupon she receives the apple.)

END


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A