SOME LOGICAL ASPECTS OF GRAMMATICAL STRUCTURE[1]

BY

HASKELL B. CURRY[2]

1. Introduction. It is a common observation that human progress is most difficult in those fields which do not belong exclusively under one of the accepted major branches of knowledge. With respect to such fields we are frequently in the position of the six blind men and the elephant. According to this story one of the blind men, who got hold of the elephant's leg, asserted that he was like a pillar; a second, who had the animal by the tail, said that he was like a rope; another, who was up against the elephant's side, claimed that he was like a wall; while the remaining three, who touched the elephant's ear, trunk, and tusk, maintained with equal stoutness that he was like a sail, a hose, and a spear respectively.

The field here discussed—general grammar—is preeminently one of that character. It can be approached from the standpoint of linguistics, from that of logic, and from that of psychology; and each of these approaches discloses something not seen from the other. Moreover the logical aspect looks somewhat different from the standpoint of mathematics than from that of philosophy. Even the linguistic aspect is apt to be colored by the particular languages one has specialized in; and there is a suspicion that, in some quarters at least, preoccupation with the Indo-European languages has caused distortion.

In such circumstances progress is probably best made by bold cooperation. By this I mean that we should make as many different approaches as possible without too much fear, at first, of errors due to lack of omniscience; and, that each person should report his observation, in a manner intelligible to the other persons, without worrying about whether the latter have or have not already perceived the same thing.

In this spirit I shall attempt to explain certain impressions of the grammatical elephant as perceived by a blind man who has hold of it in rather an odd place. I do not claim any revolutionary character for these impressions; but, even so, the fact that they arise from the standpoint in question may be of some interest in itself. On that account I have given in §7 some details of the logical motivation.

2. Language and mathematical logic. The standpoint from which this discussion starts is that of a particular aspect of mathematical logic. It is therefore expedient to begin by a discussion of the logical aspects of language in general.

That there is a close connection between mathematics and logic on the one hand and language on the other has been evident for some time, and is brought to a focus even more strongly by the existence of this symposium. The standpoint of mathematical logic has the following advantages. In the first place mathematics is itself a language in some sense—not, to be sure, in the same sense that English or German or Chinese or Eskimo is, but none the less a language with its own grammar and vocabulary.[3] Moreover it is a language which has been especially evolved for purposes where severe logic is essential; and therefore it is natural to suppose that its grammar would reflect the logical requirements more closely that of the natural languages. In the second place, much recent work in logic has been explicitly concerned with artificial languages of an idealized and simplified sort; and it is reasonable that the solution of the problem in these cases would be a useful prelude to the more complicated problems of the natural languages. Of course, the logical standpoint has its limitations; not the least of these is the circumstance that in all the natural languages logic is only one of the factors to be considered. Nevertheless, within its limitations the logical approach may be suggestive and interesting.

To a logician the term 'language' has acquired a sense different from that to which linguists are accustomed. In the most general sense a *language* is a system of objects called *symbols* which can be assembled into combinations called *expressions*. Here the terms 'symbol' and 'expression' are undefined; but for any given language they are defined by the conventions of that language. However, symbols are understood to be objects not too unlike the phonemes of speech or the letters of print, which can be produced in unlimited quantity. Generally the expressions are the finite linear series (or "strings") of such symbols; in that case I shall say the language is a *linear language*. If the language is or can be used for purposes of communication it will be called a *communicative language*; while a language in the sense of linguistics will be called a *natural language*.

3. A simple language. Before entering on a general discussion it will be expedient to consider a simple artificial language, which we shall call Language A. This shall have three symbols, viz.

$$a, b, c\,;^{4}$$

[1] Based in part on a lecture delivered November 23, 1948, to the "comparative literature group", at what was then the Pennsylvania State College, under the title *The logical structure of grammar*. This was mimeographed by a group at the University of Chicago in 1949 and distributed to a number of persons. It was reissued with minor corrections at the Pennsylvania State University, July 27, 1959. The manuscript has been revised, and a new section, §7 has been added; but the elementary character of §§1–6 has been preserved.

[2] Some assistance in connection with preparation of this manuscript has been received from a research grant of the National Science Foundation, U.S.A.

[3] In 1937, and thus just a few years after the publication of his book **[Lng]**, (for explanation of the letters in brackets see the Bibliography at the end of the paper) Bloomfield submitted to the Linguistic Society of America a manuscript of about 300 pp. entitled *The Language of science*. I was one of the persons who had the privilege of seeing and criticizing this document. It contained many interesting observations on the nature of mathematics. It is unfortunate that he was, apparently, unable to finish it.

[4] We shall regard the letters '*a*', '*b*', '*c*' not as themselves the letters of the language A, but as letters being used in the present context as names for the letters of A. For this reason quotation marks about these letters are omitted.

and the language is to be a linear language, so that the expressions are the finite linear series such as

ababcbbca .

We now define a class of expressions called "sams" as the least class such that: 1) *a* is in the class; and 2) if *b* is added after any member of the class then the new expression is in the class. The sams then comprise all and only those expressions of the infinite list

a
ab
abb
abbb
abbbb
. . . . .

Next we define a class of "tettles" as expressions formed by taking two sams and writing *c* between them. Finally we define the class of "tantets" as the least class of tettles such that: 1) *aca* is in the class; and 2) if *A* and *B* are sams such that *AcB* is in the class, then *AbcBb* is in the class. It is easily seen that the tantets are those tettles whose component sams are alike, i.e., a tantet is a tettle of the form *AcA*, where *A* is a sam.

All these definitions have been made without assigning any meaning to any of the expressions of Language A. Everything which we have done relates only to the way in which the expressions are formed from the symbols. Such considerations are called "syntactical" by most modern logicians.

It is now time to explain the meanings to be conveyed by Language A. This meaning will be introduced in three stages.

To begin with, the tettles of Language A are its sentences. This simple explanation tells us what the grammar of Language A is. The sams are evidently something analogous to nouns; *c* is a two-place verb which forms a sentence when flanked by two nouns, like 'strikes', 'pulls', 'loves', 'is' in English; while *b* is a suffix which forms nouns from nouns.

Next, the tantets of Language A are its true sentences. In that case *c* expresses the relation of equality, which we express in ordinary mathematics by '='.

Finally it will be explained that the sams denote numbers: *a* denotes the number 0, *ab* the number 1, *abb* the number 2, etc. In fact the words 'sam', 'tettle', 'tantet' are anglicized corruptions of the Hungarian words 'szám', 'tétel', 'tantét' which mean respectively 'number', 'sentence', and 'theorem'.

The explanation of meaning is now complete. It will be noted that *b*, employed as a suffix, indicates the operation of forming the successor of a number in arithmetic.

4. Semiotics. There has arisen a whole school of philosophers (and mathematicians) who place great emphasis on quasi-linguistic analysis of the sort just illustrated. Some of them call *semiotics*[5] the study of symbolic

[5] See Morris [**FTS**], and also, to some extent, his [**SLB**]; Carnap [**ISm**].

systems in general, whether artificial or natural, and divide the subject into three parts: 1) *syntactics*, i.e., the study of those considerations which depend only on the structure of the expressions as strings of symbols; 2) *semantics*, i.e., the study of those considerations which require a reference to communicative functions; and 3) *pragmatics*, which includes relations between the language and its users (psychological and physiological factors, etc.) Thus the discussion of Language A, in which we talked about sams, tettles, and tantets, was syntactical without admixture of any semantical or pragmatical element; but when we interpreted these terms we were doing semantics. Pragmatics, although it is, in a sense, the basis from which the other dimensions are formed by abstraction, will not concern us here. The semiotical senses of these words must not be confused with the senses which they have in linguistics; the term 'syntactics' is, perhaps, unfortunate, but it has become established.

As the example of Language A shows, we may conceive of semantics as itself divided into three stages. The first stage is concerned with the formation of sentences; let us call this *grammatics*, where the termination suggests that we are dealing with a theoretical science which is a part of semiotics related to grammar. The second stage is concerned with truth; let us call it *aletheutics*. The third stage may be called *onomatics*.[6]

Several remarks will now be made which are pertinent to this discussion.

In the first place the usage of the term 'semantics' may strike you as a little peculiar. There is some diversity of usage in regard to this term. Some authors seem to use it in the sense of 'onomatics'. This appears to have been the intention of Carnap in [**ISm**].[7] He distinguished between rules of formation (the present grammatics), rules of truth (the present aletheutics), and rules of designation (the present onomatics). But the strictly onomatical considerations played only a secondary role, and could be dispensed with altogether; practically all the theorems proved in that book can be considered as purely aletheutical. Likewise the work of Tarski appears to identify semantics with onomatics; the notion of truth is defined in terms of a notion of satisfaction which is onomatical. There is also a widespread feeling that the notion of truth is semantical while the notion of sentence is not. But when I once asked a friend, who was stoutly maintaining this thesis, how he told what sentences were, he replied that in the last analysis it depended on the judgment of native speakers; if that is so, then a communicative (and perhaps also pragmatical) element is involved, and the notion is semantical according to the definition adopted here. Any reference to the use of the language for communicative purposes is semantical from the present standpoint.

The Language A teaches us also that, even though a concept is semantical, there may be a syntactical notion which is equivalent to it. Thus the sentences of A are the tettles, and the true sentences are the tantets; but the notions of

[6] These terms were first proposed in [**LFS**] and [**MSL**]. See also the discussion in [**CLg**] pp. 35–37.
[7] Cf. footnote 5.

tettle and tantet are purely syntactical. In mathematics, and particularly in mathematical logic, we have less trivial examples of the same situation. In the natural languages many persons think that we have something similar in the case of grammar. Although it has not yet been clearly established whether that is so or not, yet it is certainly admissible as an objective for research to seek for such theories. But we should bear in mind that, since there are known to be logical systems which can formulate their own syntax but not their own truth, it may turn out that a syntactical theory of grammar is impossible, or at least not practically attainable.

We shall be concerned here with grammatics—or grammar if you prefer. This may be defined as the considerations which define what the sentences of the language are. However it is not always clear, and I do not know how to make it clear, just what a sentence is. Thus Chomsky maintains that

Sincerity admires John

is not a sentence of English. As a native speaker of English I disagree. To me it seems a perfectly grammatical sentence, although an absurd one. Absurdity, however, is—at least to my linguistic instincts—not a grammatical concept but an aletheutical one. This shows that the line between grammatics and aletheutics is not as sharp as one might suppose. The line is bound to be obscured if one bases his analysis too naively on distribution, since true sentences occur much more frequently in the discourse of sane human beings than false ones. How this difficulty is to be overcome I do not know. Perhaps we shall have to consider still other stages of semantics, intermediate between grammatics and aletheutics,—or other subdivisions of one or the other, corresponding to different levels of grammaticalness. However that may be, I shall suppose for the time being that the category of a sentence is sufficiently well known to form a basis for our present discussion.

After this preliminary discussion let us turn to our main theme, viz. to find concepts and principles for general grammatics.

5. Phrases and their fundamental classification. To begin with, it is evident that (for a linear language) expressions do *not* form a natural class of symbol combinations. Thus in the English sentence

I see both red and blue flowers,

the following expressions occur

see both red,

ed and bl.

On the other hand the words '*both*' and '*and*' together form a semantical unit, and therefore should form a grammatical unit. Let us use the term *phrase* for a symbol combination which forms a grammatical unit, i.e., a unit in the rules for determining what constitutes a sentence. This term we must regard, like 'symbol', 'sentence', and 'expression', as not otherwise defined. A phrase, as we have seen, may consist of detached parts.

Next, what are the main classifications of phrases? In Language A the following classes of phrases occur:

1) Sams. From the grammatical point of view these are essentially *nouns.*
2) Sentences, i.e., tettles.
3) Phrases which combine phrases to form other phrases. Such for instance is the suffix[8] *b* which, when affixed to a noun, forms another noun; and the infix *c* which, when placed between two nouns, forms a sentence. Such phrases were called "functors" by the Polish philosopher Kotarbinski, and the term will be adopted here.

A little thought will show that these are the fundamental categories for any conceivable mathematical language. It may be necessary to consider two or more different categories of nouns (such as individual-nouns, class-nouns, relation-nouns, etc.)[9]; but the nominal character of each of these kinds is clear. *Nouns, sentences,* and *functors* are, then, the basic grammatical categories. The first two categories together we call *closed phrases* in contradistinction to functors.

6. Functors. We now consider the functors more closely. Every functor combines one or more phrases, called its *arguments,* to form a new phrase called its *value.* Functors can evidently be classified as to the number and kind of the arguments and the nature of the value. Thus in Language A the suffix *b* is a functor with one nominal argument, and a nominal value; the infix *c* is a functor with two arguments, both nominal, and a sentential value. But before going further with this it is necessary to introduce some notation.

A functor is by definition a means of combining phrases to form other phrases. The complete specification of a functor must show how the value is to be made up. In other words we cannot represent a functor without blanks or other devices[10] to indicate where the arguments are to be inserted. In some natural languages certain case endings or word order performs this function. We shall use dashes with subscripts; the first argument is to replace the dash or dashes with subscript 1, the second that (or those) with subscript 2, etc. Thus the two functors mentioned for language A are $-_1b$ and $-_1c-_2$ respectively. (These are not the only functors in A; e.g., we have $-_1bb$.)

At this point let us stop to emphasize that the notion of functor is not confined to words or affixes which are placed before or after their arguments, in the case of two arguments, between them. These, with the proper indication of the position of the arguments, are indeed functors; but they are not the most general ones. We have just seen an example, viz. 'both $-_1$ and

[8] The term 'suffix' is used here for a functor which follows the argument or arguments, 'prefix' for one which precedes its arguments, and 'infix' for a functor of two arguments which is placed between them.

[9] The distinction sometimes made between nouns and noun phrases is not taken into account here. From the present point of view these are just different kinds of nouns. Thus the discussion of Hill **[ILS]** p. 175 may be interpreted as meaning there are seven or more different kinds of nouns in English; Harris **[MSL]** Chapter 16 requires four or more.

[10] If a functor is described as a prefix, an infix, or a suffix, this may take care of this point, and in such a case further devices may be unnecessary.

$-_2$' of a functor consisting of detached parts. Even more general forms than this are conceivable. What Harris and Chomsky call transformations are also functors. A functor is *any* kind of linguistic device which operates on one or more phrases (the argument(s)) to form another phrase. A functor may, conceivably, so modify its arguments that even the notations involving blanks are inadequate to describe it.

A notation for describing kinds or "categories" of functors will now be introduced. Let X, Y, Z, U be grammatical categories. Then we can represent by

$$\mathbf{F}XY\,,$$
$$\mathbf{F}_2XYZ\,,$$
$$\mathbf{F}_3XYZU\,,$$
$$\cdots\cdots\cdots$$

respectively, the categories of functors with one argument in X and value in Y; those with two arguments, in X and Y, and value in Z; those with three arguments, viz. in X, Y, Z, and value in U; etc. If N represents the category of nouns, and S that of sentences, then the two primitive functors for Language A, together with the categories to which they belong, are:

$$-_1b, \qquad \mathbf{F}NN\,;$$
$$-_1c-_2, \qquad \mathbf{F}_2NNS\,;$$

while the 'both$-_1$ and $-_2$' of the sentence originally quoted above belongs to the category $\mathbf{F}_2(\mathbf{F}NN)(\mathbf{F}NN)(\mathbf{F}NN)$.[11]

In order to attain complete generality we must admit functors with other functors as arguments. As for the value, we note that

$$\mathbf{F}X(\mathbf{F}YZ), \qquad \mathbf{F}_2XY(\mathbf{F}_3UVWT)$$

are the same categories respectively as

$$\mathbf{F}_2XYZ, \qquad \mathbf{F}_5XYUVWT$$

and so on. In classifying functors we may therefore do either of the following: a) restrict attention to functors of one argument, but allow the value to be a functor, or b) allow functors of any number of arguments but require the value to be closed—in which case we speak of the value as the "closure". The former of these procedures is what is usually followed in linguistics, and is also used in some recent work in logic; the latter is in the spirit of mathematics and the usual symbolic logic. This latter method is more convenient for classification purposes. It should be noted that either method can be unnatural in certain cases. Thus the analysis

$$\textit{both}-_1 \textit{ and } -_2 \qquad \mathbf{F}_2(\mathbf{F}NN)(\mathbf{F}NN)(\mathbf{F}NN)$$

is more natural than either

$$\textit{both}-_1 \textit{ and} \qquad \mathbf{F}(\mathbf{F}NN)(\mathbf{F}(\mathbf{F}NN)(\mathbf{F}NN))$$

or

[11] In the original sentence, the phrase "both red and blue" as well as "red" and "blue" are understood to be adjectives (i.e., to belong to the $\mathbf{F}NN$ category).

$$\textit{both}-_1 \textit{ and } -_2-_3 \qquad \mathbf{F}_3(\mathbf{F}NN)(\mathbf{F}NN)NN\,.$$

The following table gives some examples of phrases belonging to various categories in the above scheme.[12] The first column gives the category; the second column gives examples from ordinary English; the third column gives examples from the technical language of mathematics. It will be noted that functors are classified into two main classes, primary and secondary: primary functors are those all of whose arguments are closed; secondary functors those of which at least one argument is another functor.

EXAMPLES OF PHRASES

CLOSED PHRASES		
N	Proper nouns	$0, 1, 2, \cdots, e$
	Common nouns (tentatively)	
	Pronouns	
S	Sentences in usual sense	$x = y, 1 < 2, 2 = 3$
	Interjections	
	Noun in vocative case	
PRIMARY FUNCTORS		
$\mathbf{F}NN$	Attributive adjectives Noun in genitive case	$\vert -_1 \vert$ (absolute value); $(-_1)^2$ (square)
$\mathbf{F}_2NNN$	Adjective suffixes	$(-_1) + (-_2)$
	Genitive case ending	
$\mathbf{F}NS$	Intransitive verbs	$-_1 > 0$
$\mathbf{F}_2NNS$	Transitive verbs, copula	$-_1 = -_2, -_1 > -_2$
$\mathbf{F}_3NNNS$	Verbs with two objects	$-_1$ is between $-_2$ and $-_3$
$\mathbf{F}SN$	That $-_1$,	
	Verbal noun transformations	
$\mathbf{F}SS$	Negation	$\neg -_1$
$\mathbf{F}_2SSS$	Coordinating conjunctions	$-_1$ & $-_2$,
SECONDARY FUNCTORS		
$\mathbf{F}(\mathbf{F}NN)(\mathbf{F}NN)$	Adverb modifying adjective	Iteration of an operation
	Operation of attribution	Square of a function
$\mathbf{F}(\mathbf{F}NN)(\mathbf{F}(\mathbf{F}NN)(\mathbf{F}NN))$	Suffix '-ly'	Composition of functions
$\mathbf{F}(\mathbf{F}NS)N$	Verbal noun transformations	Least number satisfying a condition
$\mathbf{F}(\mathbf{F}NS)(\mathbf{F}NS)$	Adverbs modifying intransitive verbs	Complement of a condition
$\mathbf{F}N(\mathbf{F}(\mathbf{F}NS)(\mathbf{F}NS))$	Suffixes forming such adverbs from nouns	

[12] For further examples see **[CLg]** pp. 264–265, 274–275.

7. Concluding remarks. The idea of expressing categories by means of **F** was devised to meet the needs of a certain type of logic. It may help us to understand this notion, and some of its limitations, to describe this application briefly.

In the type of logic called combinatory logic[13] the formal metalanguage—i.e., that part of the language being used which serves to name the formal objects and to express the statements derivable within the system—consists solely of the following: a finite number of basic nouns, a two-argument functor, called *application*, forming a noun from two nouns, and a sentence-forming functor of one nominal argument. Yet in this system very complex types of logical system can be represented. In order to explain the paradoxes which have been known to logicians for about half a century, it was realized at an early date that one would have to formalize distinctions of semantical category within the system. In order to do this one needs: (a) nouns representing basic categories; (b) devices for forming derived categories; (c) axioms assigning categories to primitive notions represented by nouns; and (d) means for proving within the system that notions constructed from the primitives by the operations belong to appropriate categories. Since (XY), i.e., the application of X to Y, is interpreted as the value of X (as function) for the argument Y, an appropriate means of formalizing (b) and (d) is to postulate a notion, **F**, called the functionality primitive, and then to say that when X is an $\mathbf{F}\alpha\beta$ and Y is an α, (XY) is a β. On this basis it is, I think, possible to treat from a unified point of view all the devices used by various logicians to avoid the paradoxes.[14]

Similar ideas concerning the semantical categories had previously been entertained. Leśniewski apparently placed emphasis on the idea that avoidance of the paradoxes required study of the semantical categories. From H. Hiż I understand that the idea goes back to Husserl. For an account of this earlier work and its recent developments the most accessible references are Ajdukiewicz **[SKn]** and Suszko **[SSS]**.[15]

So much for the logical origin of **F**. The application of these ideas to linguistics came somewhat as an afterthought. It is natural to expect that the applications for that purpose would require some modification. But I think it is also true that the consideration of them suggests improvements in current linguistic procedures. I shall close with some comments along these lines.

In the first place, combinatory logic was formulated as a system in which the formal objects were rather differently conceived than was the case in standard formalizations of logic. The standard procedure at that time was to demand that the formal objects be the expressions of some "object language"; this means that they be strings formed from the symbols of that object language by concatenation. In combinatory logic these formal objects, called *obs*, were wholly unspecified; it was merely postulated that there was a binary operation of application among them, that the obs be constructed from the primitive objects, called *atoms*, by these operations, and that the construction of an ob be unique. This means that the obs were thought of, not as strings of atoms, but as structures like a genealogical tree. Now of course there are various ways in which such a tree can be associated with a string. Any method of making such a one-to-one association between the obs and a special class of expressions called *wefs* (i.e., well formed expressions) is called a *representation* of the system. In order that a linear language be a representation in this sense it is necessary that each wef indicate a unique construction (i.e., a unique tree): in such a case the language will be called *monotectonic*.[16]

Now this situation suggests that we may think of a language in an analogous fashion. That is, we may think of it, not as a system of expressions, but as a system of phrases, in the sense of §5, which are formed from the primitive or atomic phrases by the functors. We may even push this a little further, since we regard the functors as themselves phrases, and think of a phrase as a construction by the single operation of application of a functor to its first argument. In this way we may conceive of the *grammatical structure* of the language as something independent of the way it is represented in terms of expressions; and this grammatical structure can be studied by means of **F** (and possibly other similar notions). Of course this will need to be supplemented by a study of the way these phrases are represented by expressions. This gives us two levels of grammar, the study of grammatical structure in itself, and a second level which has much the same relation to the first that morphophonemics does to morphology. In order to have terms for immediate use I shall call these two levels *tectogrammatics* and *phenogrammatics* respectively; no doubt someone will propose better terms later.

One or two examples will illustrate what has just been said. Thus the application operator of combinatory logic may be represented by a functor '$(-_1-_2)$' involving parentheses, or by a prefix '$*-_1-_2$' without parentheses. Both these representations can be shown to be monotectonic. We would thus have two languages for representing combinatory logic; these would have the same grammatical structure, but would differ in their phenogrammatics. Another example is the theory of sams, in which the suffix '$-_1b$' can be replaced by the prefix '$s-_1$'.

Harris and Chomsky have suggested that grammar be divided into three stages, "phrase structure" grammar, transformation grammar, and morphophonemics. But if phrase structure grammar means the building up of phrases by concatenation of adjacent phrases, then it has a phenogrammatical aspect. From the standpoint of tectogrammatics I see no reason to put phrase structure and transformation grammar on separate levels, nor to suppose that

---

[13] For further information concerning this system see **[CLg]**, especially Chapters 1 and 8. For briefer treatments see, besides the papers cited in **[CLg]** p. 1, the introductory parts of **[DTC]**.

[14] Thus the theory of types, the "definite" notions of Zermelo, Fraenkel and their successors, the "stratification" of Quine are all specializations of this idea, in the sense that they may be treated, at least in principle, by adjoining to the combinatory theory of functionality special assumptions of the sorts considered under (a) and (c).

[15] For the history of these ideas see **[CLg]** p. 273.

[16] In **[CFS]** I called this property the tectonic property. But I now think the term 'monotectonic' is better, because it leaves the 'polytectonic' for the opposite property.

phrase structure operations necessarily either precede or follow transformation operations.

In the same vein, Lambek[17] has recently proposed a calculus of grammatical structure based on two kinds of functionality notions expressed by slant lines. Thus N/S would mean a functor forming a noun from a sentential argument on its right, while $N\backslash S$ would mean a functor forming a sentence from a nominal argument on the left. This is a classification of expressions in a concatenative grammar rather than a classification of functors in the sense of §6. If we use 'f' as abbreviation for the expression, then Lambek's "'f' is an N/S" would mean the same as "'$f-_1$' is an **F**SN", whereas his "'f' is an $N\backslash S$" would mean the same as my "'$-_1f$' is an **F**NS". Thus Lambek's conception has an admixture of phenogrammatics. Moreover it seems to break down completely with reference to functors which are not either prefixes or suffixes.

It is to be expected that grammatical structure will vary less from language to language than does the phenogrammatics. Different languages use entirely different devices for indicating grammatical phrase composition. In some languages word order is important, as it often is in English; but in Latin the three words in

Puer puellam amat

can be arranged in any of the six possible orders without changing the structure. It is not uncommon in such languages to have functors consisting of detached parts, or of parts widely separated from the arguments.

Finally I shall mention two phenomena which occur in the natural languages but are rare in the artificial languages of logic and mathematics. These may be expected to cause some modification in the concepts of grammatics.

The first of these is the phenomenon of ellipsis. It is a common tendency in the natural languages to omit anything which is not necessary for communication in the particular context. Thus almost any single word in the English language is capable of forming a sentence by itself; and one can imagine a question which could be significantly answered by giving a detached suffix. Most languages have pronouns; and, as Harris has recently pointed out,[18] one may have proverbs and promorphenes of other sorts. Any adequate discussion of grammatical structure must take this into account.

The second phenomenon is that the same phrase may be constructed in different ways. By this I am not referring to homonymity, which is the similarity in form of what most of us would regard as distinct phrases, but to the fact that even phrases which we may wish to regard as semantically identical may still be analyzed in different ways. Thus the sentence

The clams were eaten by the children

is described by Chomsky as necessarily obtained from

The children ate the clams

[17] See Lambek **[MSS]**, also his paper in this symposium.

[18] See Harris **[CTL]**.

by the passive voice transformation, which is an **F**SS. But there is an alternative explanation viz. as formed from the **F**NS.

$-_1$ *were eaten by the children*

by applying it to 'the clams' as argument. This is the only explanation possible in

The clams were eaten by the seashore.

It is true that the 'by' in the second sentence is merely a homonym for the 'by' in the first (and it would be expressed in German by a different word). But I cannot follow the argument which says that the second analysis is correct in the second case but not in the first. It is true that the artificial languages of mathematics and logic have generally the property that every phrase has a unique construction—i.e., these languages are monotectonic. But it is a fallacy to assume a priori that the natural languages have this property. They may be, and I believe most of them are, polytectonic, and the equivalence of different constructions of the same sentence is one of the features which has to be taken account of in an adequate grammar.

These features, as well as the oversimplification in taking N and S as the only basic categories, can, I think, be taken into account. They do not affect the fundamental point, that we can profitably study grammatical structure as such, apart from its representation in terms of concatenation.

Bibliography

Works are cited, with or without the author's name, by abbreviated titles consisting of three letters in brackets, as in the following list. (When the author's name is omitted entirely the author is usually Curry or Curry and Feys.) Journals are abbreviated according to the practice of Mathematical Reviews.

KAZIMIERZ ADJUKIEWICZ

[SKn] *Die syntaktische Konnexität*, Studia Philos. vol. 1 (1935) pp. 1–27.

LEONARD BLOOMFIELD

[Lng] *Language*, New York, 1933.

RUDOLF CARNAP

[ISm] *Introduction to semantics*, Cambridge, Mass., 1942.

NOAM CHOMSKY

[LSL] *Logical structure in language*, Amer. Documentation vol. 8 (1957) pp. 284–291.

[SSt] *Syntactic structure*, The Hague, 1957.

HASKELL B. CURRY

[DTC] *The deduction theorem in the combinatory theory of restricted generality*, Logique et Analyse, 3e Année (1960) pp. 15–39.

[LFS] *Languages and formal systems*, Proceedings of the Tenth International Congress of Philosophy (1949) pp. 770–772.

[MSL] *Mathematics, syntactics and logic*, Mind vol. 62 (1953) pp. 172–183.

HASKELL B. CURRY and ROBERT FEYS

[CLg] *Combinatory logic*, Amsterdam, 1958.

ZELLIG HARRIS

[CTL] *Co-occurrence and transformation in linguistic structure*, Language vol. 33 (1957) pp. 283–340.

[MSL] *Methods in structural linguistics*, Chicago, Ill., 1951.

ARCHIBALD A. HILL
[ILS] *Introduction to linguistic structure*, New York, 1958.
CHARLES F. HOCKETT
[CML] *A course in modern linguistics*, New York, 1958.
OTTO JESPERSEN
[PGr] *Philosophy of grammar*, London, 1924.
JOACHIM LAMBEK
[MSS] *The mathematics of sentence structure*, Amer. Math. Monthly vol. 65 (1958) pp. 154–170.
CHARLES MORRIS
[FTS] *Foundations of the theory of signs*, Chicago, 1938.
[SLB] *Signs, language and behavior*, New York, 1946.
JOSEF SCHÄCHTER
[PKG] *Prolegomena zu einer kritischen Grammatik*, Vienna, 1935 (cf. its review, Weinberg, Phil. Rev. vol. 46 (1937) pp. 334–335).
ROMAN SUSZKO
[SSS] *Syntactic structure and semantical reference I*, Studia Logica vol. 8 (1958) pp. 213–244.

THE PENNSYLVANIA STATE UNIVERSITY,
UNIVERSITY PARK, PENNSYLVANIA

GRAPHIC AND PHONETIC ASPECTS OF LINGUISTIC AND MATHEMATICAL SYMBOLS

BY
YUEN REN CHAO

1. Graphs and logographs. In both linguistics and mathematics ther is a lot of talk and there is a lot of scribbling. The present paper is a pre systematic (and pre-computer) stage of discussion, dealing with the ways i which the talk and scribbling is done in the two fields, with some suggestion as to the directions in which possible future developments may take. Sinc most of the symbolizing in logic has been derived from mathematics, I shall unless otherwise specified, simply say 'mathematics' to include both mathe matics and logic. The plan of the present paper is to look over the usage in linguistics and mathematics in general and summarize them in the form of problems for the future. But before entering upon the subject, I shoul like to point out that I am not planning to discuss linguistics or mathematics but only some relatively peripheral aspects of both, namely, those about the symbolism used in the two fields.

It is well known that the language of any given speech community in the world makes use of from about ten to not over 100 distinctive kinds of sounds called 'phonemes'. The ear can detect several times that number of fine shades of acoustic qualities, but of elements out of which meaningful units of speech communication is made, the total number in any one given language is always fairly small.

Now man spoke for centuries and milleniums before he started to record visually what he spoke. At the same time, he drew pictures and made visual marks to record or represent things and ideas independently of speech. In high antiquity the Chinese are said to have run their government by tying knots and the peoples of the Occidental countries of today are still tying knots around their little fingers to represent errands to do—or at least are represented as doing so in the comics. But within historical times writing has largely been a form of recording speech and only to a minor degree an independent system of symbols which bypasses the act of speech. This has been true of the Egyptian hieroglyphics and even more true of the shell and bone inscriptions of ancient China. The so-called pictographs and ideographs are not symbols of things or ideas, but symbols of words, or logographs.[1]

That this has been so almost follows by definition, since there can be no writing of history without a history of writing. History differs from prehistory in that pre-history is what we say now what men did, as inferred from traces they left—pictures, totems, artefacts—while history is what men

[1] Peter S. DuPonceau, *A Dissertation on the nature and character of the Chinese system of writing*, Philadelphia, 1838, xi and xii; Y. R. Chao, *A note on an early logographic theory of Chinese writing*, Harvard Journal of Asiatic Studies vol. 5. (1940) pp. 189–191.