Structure of Language and its Mathematical Aspects

PROCEEDINGS OF SYMPOSIA IN APPLIED MATHEMATICS

VOLUME XII

http://dx.doi.org/10.1090/psapm/012

PROCEEDINGS OF THE TWELFTH SYMPOSIUM IN APPLIED MATHEMATICS

Held in New York City
April 14–15, 1960

SPONSORED BY

AMERICAN MATHEMATICAL SOCIETY
ASSOCIATION FOR SYMBOLIC LOGIC
LINGUISTIC SOCIETY OF AMERICA

ROMAN JAKOBSON
EDITOR

PROCEEDINGS OF
SYMPOSIA IN APPLIED MATHEMATICS
VOLUME XII

STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS

AMERICAN MATHEMATICAL SOCIETY
190 HOPE STREET, PROVIDENCE, RHODE ISLAND
1961

Cosponsored by
THE INSTITUTE FOR DEFENSE ANALYSES
under Contract Nonr 2631(00)
with the Office of Naval Research
and under Subcontract IDA 2631(00)-6

International Standard Serial Number 0160-7634
International Standard Book Number 0-8218-1312-9
Library of Congress Catalog Number 50-1183

CONTENTS

INTRODUCTION

The Symposium on the Structure of Language and its Mathematical Aspects arose through the fortunate initiative of the American Mathematical Society which fully realized that the attention of linguists, logicians and mathematicians has become focused upon problems of mutual interest.

The need for an ever closer contact between linguistics and mathematics was clearly understood at the threashold of our century by the two great anticipators of the modern structural analysis of language—Baudouin de Courtenay, and Ferdinand de Saussure. As early as 1894, Saussure noted: "Les quantités du language et leurs rapports son régulièrement exprimables dans *leur nature fondamentale*, par des formules mathématiques," and somewhat later, discussing the problem of simplest expression for linguistic concepts, he stated: "L'expression simple sera algébrique ou elle ne sera pas." Finally, in 1911, while working on his last course in general linguistics, Saussured pointed out that this science appeared to him as a system of geometry: "On aboutit à des théorèmes qu'il faut démontrer."

Attacking, since the seventies, the crucial questions of the relation between continuity and discreetness in language, Baudouin de Courtenay attempted to utilize in the study of language some of the basic notions of contemporaneous mathematics, and in his historical survey of linguistics, published in 1909, he expressed his conviction that this scholarship would become ever closer to the exact sciences. Upon the model of mathematics it would on the one hand, deploy "ever more quantitative thought" and on the other, develop "new methods of deductive thought." In particular, "just as mathematics converts all the infinities to denumerable sets amenable to analytic thought," Baudouin expected somewhat similar results for linguistics "from improved qualitative analysis."

While Baudouin referred to the mathematical model for the analysis of language, at about the same time, before the 4th International Congress of Mathematicians in 1909, E. Borel discussed the antinomy of the denumerable infinities and appealed to the fundamental role of language in mathematical operations, "car les prétendus systèmes entièrement logiques reposent toujours sur le postulat de l'existence de la langue vulgaire; ce langage commun à des millions d'hommes, et avec lequel ils s'entendent à peu près entre eux, nous est donné comme un fait, qui impliquerait un grand nombre de cercles vicieux, s'il fallait le créer *ex nihilo*." The linguistic inference therefrom became Bloomfield's thesis according to which "mathematics is merely the best that *language* can do" (1933). If "mathematics, the ideal use of language," is a mere superstructure over the common language, the interrelation between this superstructure and its basis must be of primary interest for mathematicians and linguists alike.

Both the theoretician of language and the investigator of languages in their single stages or in evolution, equally as the workers in the rapidly develop-

ing branches of applied linguistics, are attracted by the manifold mathematical disciplines: mathematical logic, in particular, the theory of recursive functions and automata; the topological, algebraic and quantitative facets of mathematics; the theory of communication and probabilistic models. One cannot but agree with the mathematician J. Hadamard, who in 1943 acknowledged the progress of the structural trend in the science of language by declaring linguistics to be a bridge between mathematics and humanities.

ROMAN JAKOBSON

MINUTES OF THE SYMPOSIUM ON THE STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS

NEW YORK CITY, APRIL 14–15th, 1960

SPONSORS: American Mathematical Society
Association for Symbolic Logic
Linguistic Society of America

PROGRAM COMMITTEE: Noam Chomsky, Roman Jakobson (Chairman), H. B. Curry, Hilary Putnam, Henry Hiż, W. V. Quine

SECRETARY OF THE SYMPOSIUM: E. S. Klima, Massachusetts Institute of Technology

FIRST SESSION: Morning, April 14th.

CHAIRMAN: Roman Jakobson, Harvard University and Massachusetts Institute of Technology
W. V. Quine, Harvard University
Noam Chomsky, Massachusetts Institute of Technology
Hilary Putnam, Princeton University
Henry Hiż, University of Pennsylvania
Nelson Goodman, University of Pennsylvania
H. B. Curry, Pennsylvania State University

DISCUSSION LEADER: Max Black, Cornell University

PARTICIPANTS IN THE DISCUSSION:

L. M. Court, Diamond Ordnance Fuze Laboratories, Washington 25, D.C.
Frank Harary, University of Michigan
Henry Hiż, University of Pennsylvania
Roman Jakobson, Harvard University and Massachusetts Institute of Technology
Russell Kirsch, National Bureau of Standards, Washington, D.C.
William P. Livant, University of Michigan
Irina Lynch, Wellesley College and Harvard University
Murray S. Miron, University of Illinois
Leonard Newmark, Ohio State University, Columbus, Ohio
Hilary Putnam, Princeton University
Arthur Sard, Queens College, Flushing, New York
R. J. Solomonoff, Zator Company, Cambridge, Massachusetts

INTRODUCTORY REMARKS

BY

MAX BLACK

The summaries we have heard provide tantalising glimpses of ideas that deserve careful examination and are bound to provoke fruitful discussion. I am sorry to have had no chance to read the papers, and must confine myself

to some general comments.

The topic of the symposium is, 'The structure of language and its mathematical aspects'. Three tasks are plainly involved: the elaboration of well-defined notions of 'linguistic structure', construction of appropriate mathematical systems, and provision of suitable 'co-ordinative definitions' as links between the abstract mathematics and the actual languages studied by linguists. My impression is that sufficient mathematical resources are already available —the theory of recursive functions, or, alternatively, the methods of combinatory logic (for which see Curry's paper) suggest themselves as promising arsenals. Modern algebra, mathematical logic, and possibly topology, may be expected to provide all the necessary mathematical tools.

However, severe conceptual difficulties remain to be resolved in identifying the 'linguistic structure' that is to receive mathematical analysis. The basic distinction between a 'grammatical' and a 'non-grammatical' sentence is still somewhat problematic, as Putnam's comments illustrate. Chomsky's proof of the limitations of the model of 'constituent analysis' that has so long controlled the work of linguists is a striking example of how mathematical analysis can foster the enlargement of the imagination by forcing the empirical scientist to review fundamental presuppositions of method.

Of course, there is the usual risk of excessive distance between the mathematical model and its intended application. Quine has presented a persuasive case for the value to linguists of the distinctions emphasised in mathematical logic. But that subject was invented in the service of the adequate analysis of mathematical discourse and deliberately neglects for its own good purposes much plainly visible linguistic structure (e.g. as shown in the so-called 'token-reflexive words' distinctions of tense and mood, etc.). There is danger here (as possibly also in the results reported by Goodman and Hiż) of the mathematics being pursued for its own sake, with little profit for the linguist. Contrast, for instance, Quine's remarks about the analysis of the phrase, 'The lady I saw you with' with what is implied in such a fragment of familiar speech as the wisecrack, 'That's no lady—that's my wife'. If language has a rigid skeleton, speech remains so flexible—so much a matter of invention and creation—that the communication of the speaker-in-the-street may tax the resources of the most refined grammar to explain. A valuable outcome of the studies that have been reported to us is the healthy tension between linguists and mathematicians they are likely to generate.

COMMENTS

ON QUINE'S PAPER

IRINA LYNCH: If, as Quine states, pronouns may stand for definite participants only, how can one explain the presence of the pronoun "they" in, for example, "They say that the lady I saw you with is somebody else's wife"? Should the 'impersonal' use of pronouns be exempt from Quine's definition, we still could find a number of contexts where a pronoun, be it personal, interrogative, or relative, stands for an indefinite noun. To give just a few

examples:

(1) *Who* is at the door?—*A* man, I saw *him*.—*Whom* did you see?—A man.

(2) This morning I gave a book to a student. *He* thanked me.—*Who* thanked you? A student. *What* did you give *him*? A book.

(3) An apple that falls from a tree is called a "windfall". *What* is a "windfall"?—An apple *that* falls from a tree.

W. V. QUINE: In supposing that her examples conflict with my remarks, Mrs. Lynch is assigning to my words 'stand for' a meaning abruptly at variance with the meaning which I specified and illustrated in the questioned passage. In that passage (the sixth from last paragraph of my paper) I used 'stand for' to mean 'be replaceable by', and contrasted it with the relation of grammatical antecedent. I asserted that pronouns, even when their grammatical antecedents are indefinite singular terms, can be replaced only by definite singular terms. Thus consider her examples. Her first 'him' is replaceable only by 'the man', 'he' by 'the student', and the second 'him' likewise. She is put off by the fact that the grammatical antecedents of these pronouns are the indefinite singular terms 'a man' and 'a student'; but this contrast was my very point.

Mrs. Lynch's last 'that', being a relative pronoun, needs for present purposes to be expanded into 'such that it' (in conformity with my seventh from last paragraph) if we are to isolate its role in cross-reference. This done, we find that 'it' has 'an apple' as antecedent but only 'the apple' as appropriate replacement, true to form.

Mrs. Lynch's 'who', 'whom', and 'what', finally, are interrogative pronouns. These, like the 'they' of her first example, are irrelevant to the contrast I was concerned to draw, for they have no grammatical antecedents at all.

ON CHOMSKY'S PAPER

ARTHUR SARD: What is the motivation of considering many grammars simultaneously?

NOAM CHOMSKY: Consideration of a variety of grammars is motivated by an interest in linguistic universals (i.e., general features of linguistic structure) and in justification of grammars, a task which can be undertaken in a serious way only in terms of a presupposed general theory of linguistic structure (cf. in this connection my *Syntactic structures*, chapter 6; M. Halle, *On the role of simplicity in linguistic descriptions*, this volume).

W. P. LIVANT: I speak as a psychologist interested in trying to validate Chomsky's grammar with human speakers. If one were instructed directly to convert one sentence string into another, would the considerations which lead to putting some strings in the kernel and others in the transform grammar, lead to an inequality in the time taken to expand a kernel, or to collapse the transformed string into the very same kernel? Would the time taken to expand a declarative active into an interrogative passive be longer than the recursive operation?

NOAM CHOMSKY: I am hesitant about commenting on this particular suggestion without further details. In general, I think that the problem to which this question is directed, that of determining the "psychological reality" of linguistic constructs, is as difficult as it is important.

A. G. OETTINGER: If I understood him correctly, the speaker said that the grammars were constructed to apply indifferently to the speaker's or the hearer's point of view. Since the hearer is more concerned with the analysis of sentences than with their synthesis, does this not imply that he must first have a phrase structure analysis of given sentence, then also a means for discovering inverse transformations required to the phrase structure to that of a kernel? Is there in your system any means of obtaining a phrase structure analysis for a given sentence, and then finding the inverse transformation, other than generating all sentences until you obtain one that matches the given one?

NOAM CHOMSKY: A grammar, in the sense in which I have been using the term, specifies the structural description of each grammatical sentence. It is thus neutral as between speaker and hearer, in the sense that it says nothing specific about how either actually operates.

The hearer can be represented (in part) as a device that takes a sentence as input and gives as output the structural description assigned to the sentence by the grammar that the speaker has internalized. One can speculate about various specific mechanisms by which this process could be carried out. For example, the hearer may scan the sentence for hints as to which rules of the grammar are used to generate it, generate a sentence by a route compatible with this initial specification, compare the result with the original, repeating the process, etc., approaching the correct analysis by some sort of successive approximation. And there are many other possibilities, such as that suggested in the question, that should also be seriously considered.

There are no doubt many ways of constructing recognition procedures that can be made to work, to some degree of adequacy. The task of real scientific importance, however, is clearly that of developing a procedure that avoids ad hoc and arbitrary rules, and that makes use of the generalizations about sentence structure provided by a well-constructed grammar to simplify the procedure of analysis.

L. M. COURT: Chomsky arranges his three types of grammars (languages) in a descending sequence, each member of the hierarchy including its successors. Finite sources (even counter devices) are too feeble to generate his Types I and II, whereas automata of this kind suffice for his Type III languages. The extra power of a Type II over a Type III language appears to inhere in the self-embedding character of the former. ("Self-embedding" in the vernacular means that fresh phrases from a selected list can *repeatedly* be introduced to modify existing words.)

One feels intuitively that a self-embedding language is a far richer instrument for conveying ideas. *Infinity*, especially in its mathematical usages, is an advanced concept, and the very finiteness (in the sketched sense) of a Type

III language should militate against the emergence of such notations in civilizations in which this type of language is the medium of discourse. The calculus, which rests firmly on the application of the infinitely small, was crystallized in Renaissance Europe (Chomsky observes that English seems to approximate a Type II language, and probably this is also true of French and German), whereas nothing corresponding to the calculus or the even subtler notion of transfinite numbers can be found in Polynesian cultures.

It may prove profitable, in the light of these remarks, to devise an empirical program to simultaneously compare various languages for the closeness with which they approximate Chomsky's second and third types and the maturity or immaturity of their concepts of infinity. There are other avenues besides the linguistic, e.g., visual and aural experiences, through which these notions enter our consciousness, and their involvement will undoubtedly complicate the relation we are seeking to establish. The physical environment also plays an important role, and the age of a civilization usually qualifies the ripeness of its intellectual pursuits. Despite these obstacles, a program of the sort we are envisioning should throw some light on the influence of language on a people's thought processes.

SECOND SESSION: Afternoon, April 14th.

CHAIRMAN: H. B. Curry, Pennsylvania State University
Y. R. Chao, University of California, Berkeley
Murray Eden, Massachusetts Institute of Technology
Morris Halle, Massachusetts Institute of Technology
Robert Abernathy, Massachusetts Institute of Technology
Hans Herzberger, University of Pennsylvania
A. G. Oettinger, Harvard University
V. H. Yngve, Massachusetts Institute of Technology

DISCUSSION LEADER: Joshua Whatmough, Harvard University

PARTICIPANTS IN THE DISCUSSION:

Andras Balint, Teacher's College, Columbia University
William Fourst, Harvard University
Dennis Fry, University College, London
Saul Gorn, University of Pennsylvania
J. A. Greenwood, Princeton University
V. E. Giuliano, Arthur D. Little, Inc., Cambridge, Massachusetts
C. F. Hockett, Cornell University
Henry Hiż, University of Pennsylvania
Russell Kirsch, National Bureau of Standards, Washington, D.C.
John G. Mackinney, General Kinetics, Arlington, Virginia
Don Mittleman, National Bureau of Standards, Washington, D.C.
Murray S. Miron, University of Illinois
George W. Patterson, University of Pennsylvania

INTRODUCTORY REMARKS

BY

JOSHUA WHATMOUGH

There is, I think, no need for apology for insisting, as I think I was the first to do (fully ten years ago) upon the importance of the Mathematical Theory of Communication for Linguistics, and of Linguistics for Communication Engineering; nor apology for the phrase Mathematical Linguistics, which is precisely what it is called, namely Linguistics studied mathematically, both as theory (what some are now calling type-token mathematics) and practice (which includes automatic translation and various other applications).

In passing I should point out that my theory of Selective Variation in Language, which says

> a linguistic status is produced and maintained by consistent selection which preserves the system from gross inconsistency of variation

was framed even earlier but not published (1951) until I was sure that descriptive and historical (and comparative) linguistics, upon which the theory was first based, had a firm mathematical foundation, which Communication Theory furnished. There is, in fact, a convergence of all these three approaches to the study of language. Philologists had always supposed that language could not be reduced to laws capable of mathematical statement. Attempts to find articulate laws, even on such a trifling matter as tense and aspect a recent writer declares to have occupied him seventeen years, and reaches the conclusion that linguistics must undertake to perform transformations that make it obvious that the theory and practice of descriptive linguistics are not (as at present) incompatible. This can be done only by means of logic and mathematics.

It is perhaps invidious to single out certain papers for comment. I cannot, however, refrain from calling attention especially to those of Chao, with its courageous attempt to obtain some agreement in terminology in our convergent disciplines; of Eden, which is, I think, very promising in the field of bringing writing and speech still closer together in their already close relationship (not always realized to be as close as it is); and of Oettinger, with its fascinating attack that now has been going on for nearly a year, upon problems of syntax, which turn out not to be so terrifying after all once we get away from subjective (descriptive) grammar to the laws of probability.

It would be easy, and I find it tempting, to comment upon these and all the other papers at length. I must await some other occasion. But a general remark may be appended.

Communication theory is not incompatible with linguistics—only with nearly all current structural theory. Structuralism has had its place: for over thirty years I have insisted that its place is neither so important nor so vast as its advocates have urged. But if there can be, as now seems likely, a meeting of minds—the minds of linguists, and of communication engineers, not only will this apparent contradiction be resolved: there will also be discoveries

in store, formulated (I hope and believe) in elegant and economical theorems (instead of in rambling and impenetrable tomes), discoveries of great power and practical application in human communication, and, therefore, in all human affairs, and hence in the development (if not evolution) of man for decades, and perhaps centuries, to come.

COMMENTS

ON HALLE'S PAPER

H. Hiż: The notions of simplicity and of economy may be very misleading and call for much clarifications. One can write any text using only two different symbols. It would be a great economy in the kind of symbols used but not very simple to decode. In practice one has to take into account very many factors and aspects. The notion of simplicity is, thus, close to the concept of tact and good scientific taste.

Morris Halle: It was one of the aims of my paper to give a meaning to the notion "simplicity of phonological descriptions" that would not suffer from the same vagueness and subjectivity as the notions "good taste" and "scientific tact," and that at the same time would also guide us in making the appropriate choice among competing descriptions. This can be done only if one severely restricts the number of descriptions that would be considered adequate. I required, therefore, that admissible phonological descriptions be framed in terms of a fixed set of properties, the distinctive features, which, being linguistic universals like the phoneme or the morpheme, are to be applied in the description of every language and cannot be changed to suit special situations. To choose among the relatively small number of phonological descriptions satisfying this rather severe requirement I suggested the mechanical criterion of counting the number of distinctive features mentioned in the description. Finally, I attempted to show that in specific instances the criterion led us to choose descriptions of greater generality, in the usual sense of that term.

ON HERZBERGER'S PAPER

V. Giuliano: To what extent are the rules for string generation and adjoining independent, in the sense that changing one rule need not result in the necessity of examining and possibly revising several other rules? This question appears to be of particular importance when one is concerned with analysis of unrestricted technical English, when it will be necessary for some time to come to make refinements and revisions in yet imperfect rules.

H. G. Herzberger: The rules would not in general be independent in that sense. The center strings are the generators of the language only with respect to some given adjunction (and replacement) operations. A change at any point in the system requires corresponding changes elsewhere, if the set of sentences is to be held constant. The grammatical rules that are associated

with elementary strings introduce the points of adjunction as well as the adjuncts; and other rules co-ordinate the two. Evidently there will be many interconnections between the two types of rules.

ON OETTINGER'S PAPER

D. MITTLEMAN: What is the relation of the syntactic analysis as presented here to the predictive syntactic of Mrs. Rhodes?

J. G. MACKINNEY: The oral presentation of this paper left the impression that there is only *one* conceptually simple algorithm for converting ordinary algebraic formulae to the Łukasiewicz notation. In a note of limited distribution entitled *An application of sorting to the translation of algebraic coding to machine coding*, A. E. Roberts, Jr. has shown an equally simple algorithm for the assignment of weights to consecutive symbols as a formula in full-parenthetic notation is scanned from left to right. When the symbols are then sorted by weight, the result is in effect a parenthesis-free, prefixed-operator formula. Ill-formedness of the original *may* by detected during scanning or sorting, but a full examination probably can not be made until the parenthesis-free formula is created. Oettinger is of course correct in rejecting methods involving sorting on the grounds of inefficiency as compared with the push-down store.

A. G. OETTINGER: Our experimental approach to the syntactic analysis of Russian and English owes its inspiration to the method of Mrs. Rhodes. The theoretical studies described in my paper were, in turn, inspired by the experimental work.

There is certainly no claim of uniqueness, but Mackinney is correct in contrasting the efficiency of sorting and of pushdown store techniques. Moreover, as I indicate in the paper, the pushdown store technique has the merit of tying together hitherto unconnected experimental observations and theoretical analyses.

THIRD SESSION: Morning, April 15th.

CHAIRMAN: W. V. Quine, Harvard University
Frank Harary and G. E. Peterson, University of Michigan
Joachim Lambek, McGill University and Institute for Advanced Study
H. A. Gleason, Jr., Hartford Seminary Foundation
Benoit Mandelbrot, International Business Machine Corporation
C. F. Hockett, Cornell University
R. S. Wells, Yale University
Roman Jakobson, Harvard University and Massachusetts Institute of Technology

DISCUSSION LEADER: R. M. Fano, Massachusetts Institute of Technology

PARTICIPANTS IN THE DISCUSSION:

Andras Balint, Teacher's College, Columbia University
L. M. Court, Diamond Ordnance Fuze Laboratories, Washington 25, D.C.
Henry Hiż, University of Pennsylvania
Roman Jakobson, Harvard University and Massachusetts Institute of Technology
Russell Kirsch, National Bureau of Standards, Washington, D.C.
R. B. Lees, International Business Machines, Yorktown Heights, New York
Irina Lynch, Wellesley College and Harvard University
A. G. Oettinger, Harvard University
W. V. Quine, Harvard University
J. D. Sable, Radio Corporation of America, Princeton, New Jersey
Mortimer Taube, Documentation, Incorporated

INTRODUCTORY REMARKS

BY

R. M. FANO

It has been pointed out by a number of speakers, and particularly emphasized by Jakobson, that the process of understanding speech or written languages involves sequential decisions made on a probabilistic basis, and that language constraints play a key role in these sequential decisions. I wish to expand further this point by mentioning certain fundamental results of information theory concerning the problem of communicating accurately and efficiently in the presence of random disturbances.

Random disturbances arising from various sources, and of varying degree of severity are always present in the communication process between individuals, regardless of whether spoken or written language is used. The sources of these random disturbances are not necessarily external to the individuals concerned; for instance, the random variations of the physical appearance of the various phonemes in speech and of the various letters in handwriting act as disturbances in the communication process, just as the noise arising from the environment in which individuals communicate.

It is well known that our ability to understand spoken or written language in the presence of random disturbances stems from the redundancy present in the language. Redundancy may be defined, at least for our present purposes, as the fraction of the letters in a written text, or the fraction of the phonemes in a spoken utterance that could be eliminated if no random disturbances were present. For instance, it has been estimated from experiments performed by C. E. Shannon [1] some years ago that at least four out of five letters in printed English are redundant. Of course, I don't mean by this that four out of five letters could be eliminated out-right, but rather that the same text could be recoded, at least in principle, into one-fifth as many letters.

Redundancy is present in linguistic communication in the form of constraints (mostly deterministic) between successive linguistic events. These constraints, known to the listener or the reader, allow him to resolve ambiguities resulting from random disturbances, or as a matter of fact, from homonyms. One of the most important results of information theory is that the span of these constraints in addition to the redundancy resulting from them, is a key factor in overcoming random disturbances. More precisely, it is possible to design artificial codes that insure as accurate a reception as desired by making the span of the constraints sufficiently long while keeping constant the associated redundancy. It is only necessary for this purpose that the redundancy be greater than a certain minimum value which depends on the characteristics of the disturbances that must be overcome.

The net result is that, once the redundancy has been made greater than this minimum amount, the accuracy of reception can be improved by either further increasing the redundancy or by increasing the span of the constraints while keeping the redundancy fixed. A further increase of redundancy implies a further reduction of the rate at which information is transmitted. This is evident when the redundancy is increased, let us say, by repeating each successive element of the message. Increasing the span of the constraint implies increasing the complexity of the encoding and decoding operations, and therefore of the equipment required to perform them. Thus, for a fixed reception accuracy, a greater rate of transmission of information can be obtained at the expense of greater equipment complexity, and vice versa.

Since my knowledge of linguistics is rather limited, I will not explore in any detail the linguistic implications of these theoretical results. However, I will express my belief that the structure of languages, (presumably resulting from evolution) must have been greatly affected by the fact that they must enable individuals to communicate accurately and efficiently in the presence of random disturbances. For this reason, I would expect languages to exhibit structural characteristics similar to those of artificial codes designed to meet the same communication requirements. The point that I want to stress in particular is that, while the encoding operation performed, let us say, by the speaker may well be based entirely on deterministic (as opposed to probabilistic) rules, the rules themselves must reflect the needs of the listener who must make accurate estimates about what the speaker is saying on the basis of probabilistic considerations.

It turns out that the equipment required to generate efficient codes with long constraint spans is only moderately complex. The decoding equipment is inherently much more complex. A decoding procedure that does not require prohibitively complex equipment was suggested three years ago by J. M. Wozencraft [2]. The reason for my mentioning this procedure is that it presupposes a form of sequential encoding with a tree structure somewhat similar to that postulated for languages in some of the papers presented in this session. Roughly speaking, the message transmitted can be regarded as the set of instructions for climbing the tree from node to node. The decoding procedure consists of deciding at each node which branch is most likely to have been followed by the encoder. Each successive decision is made only

after exploring some distance ahead the parts of the tree connected to the different branches involved in the decision. I wish to stress that this procedure is strikingly similar to what I think I do when I am listening to a conversation in a particularly noisy environment, or when I am reading some very poor handwriting. The amount of work that the equipment has actually to do according to the decoding procedures suggested by Wozencraft increases with the severity of the random disturbances. This fact reminds me, of course, of the greater mental effort required in listening or reading under adverse environmental conditions. The moral of my remarks, if there is one, is that linguistics and information theory have still much to offer to each other. Unfortunately, in order to take full advantage of this mutual interaction information theorists have to learn more about languages, and linguists have to learn more about information theory. This is where interdisciplinary cooperation begins to hurt.

References

1. C. E. Shannon, *Prediction and entropy of printed English*, Bell System Technical Journal vol. 30 (1951) p. 50.

2. J. M. Wozencraft, *Sequential decoding for reliable communication*, Technical Report 325, Research Laboratory of Electronics. Massachusetts Institute of Technology, August 9, 1957.

COMMENTS

ON HARARY'S PAPER

W. V. Quine: Harary's quotation from Russell, "Mathematics is that subject in which one does not know what one is talking about nor whether what one says is true," dates from an essay of 1901 (reprinted 1918 in *Mysticism and Logic*). It represents an attitude in which, happily, Russell did not long persist. Already by the time of *Principia Mathematica* his view had changed to one which is well represented by an equally quotable passage in a later book of Russell's, perhaps *Introduction to Mathematical Philosophy*: "The axiomatic method has certain advantages over logical construction, which are precisely the advantages of theft over honest toil."

F. Harary: In this context, the following quotation from the book *Introduction to the Foundations of Mathematics* by R. L. Wilder, 1952, is appropriate:

"Axioms are statements about some concept with which we already have some familiarity. Thus, if we are already familiar with arithmetic, we might begin to set down axioms for arithmetic. Of course the method is not restricted to mathematics. If we are familar with some field such as physics, philosophy, chemistry, zoology, economics, for instance, we might choose to set down some axioms for it, or a portion of it, and see what theorems we might logically deduce from them. We may say, then, that an axiom, as

used in the modern way, is a statement which seems to hold for an underlying concept, an axiom system being a collection of such statements about the concept. Thus, in practice, the concept comes first, the axioms later."

ON LAMBEK'S PAPER

L. M. COURT: At the start Lambek stresses the analogy between the syntactic algebra and dimension theory in physics. But dimensional analysis is largely sterile; it provides a preliminary (superficial) check on the soundness of a physical equation but does not assess the underlying causes. It is only in certain applications, where we possess sufficient insight into a physical situation, but do not comprehend it fully, that it can be applied fruitfully; such is the case when we invoke one of Buckingham's theorems to obtain an expression for the dimensionless Reynold's number from a knowledge of the *complete* list of the physical variables (density, viscosity, etc.) on which the drag (a force) depends, even though we are ignorant of the precise form of the dependence. (See *Exterior ballistics* by McShane, etc., Denver University Press.)

The sentence (s) and the name or "proper" noun (n) are the fundamental entities in the syntactic algebra, corresponding to time, mass and distance in physics; every other word or phrase in the sentence is expressed by this algebra in terms of n and s. Besides the observation that the *same* word in the *same* sentence can have two (or more) distinct expressions in this algebra, which Lambek proceeds to identify, there is little ground for believing that two words in separate sentences, having *identical* formal expressions (in n and s), necessarily exercise similar grammatical functions. Syntactic algebra is therefore unlikely to be an apt instrument for comparing one sentence with another and even less so for shedding light on the logical or semantic validity of a sentence.

J. LAMBEK: The sterility of dimensional analysis in classical physics does not weaken my comparison. After all, grammatical analysis is also sterile, in the sense that the grammatical correctness of a sentence does not guarantee its validity. A similar comparison could have been drawn between grammatical sentences in English and covariant equations in relativity theory. The second part of Court's comment seems to be based on a misunderstanding. It has nowhere been claimed that the words or phrases of a language can be expressed in terms of n and s (and other primitive types), only that their grammatical types (adjective, adverb, conjunction, etc.) may be so expressed. Surely, if the word "sound" occurs as an adjective in two different sentences then it exercises a similar grammatical function in those two sentences, even though it may occur as a noun or verb elsewhere.

H. HIŻ: The paper is a substantial improvement in the line of thought that originated with Husserl (informally) and Leśniewski and was later popularized by Ajdukiewicz. One essential extention done by Lambek consists in considering also functors that have their arguments to the left rather than to the right. One, of course, could extend Lambek's techniques to the cases

when there are several arguments to the left and several arguments to the right.

In many cases in order to decide what are the arguments of a functor one has to examine more of the text than just these arguments; one has to see how the expressions are situated in the entire sentence.

It is often assumed that a modifier of an expression occurs adjacent to the expression it modifies. The analysis presented in Lambek's work is based on this kind of assumption. An expression acts on another expression from the left, or from the right, but always in juxtaposed role. This is a proper grammatical analysis for languages in which concatenation plays a fundamental grammatical role like it does in English. But there are non-concatenative languages, e.g., Latin. In a Latin sentence a masculine noun can be modified by an adjective in the masculine form nearly in any place of the sentence, provided that other nouns in the sentence are feminine or neuter. Incidentally, for non-concatenative languages the tree analysis, like that presented by Yngve, will not be suitable either.

J. LAMBEK: As to the first point of Hiż's comments, one can ascribe types of functors of several arguments without having to extend the techniques already described. Thus "likes" in "John likes Jane" has type $(n\backslash s)/n$. Similarly $((x\backslash s)/y)/z$ would be the type of a functor having one argument of type x on the left and two arguments of types z and y on the right.

His remaining comments do bring out an essential limitation of phrase structure and type theory grammars. Deviations from some idealized word order may be easily described by transformation rules, but only with great difficulty or not at all by type assignments.

ON GLEASON'S PAPER

R. JAKOBSON: The question of genetic kinship between roots or affixes of different languages cannot be handled without an elucidation of their semantic relationship. Problems of lexical and grammatical meanings, to the same extent as the examination of external form, require a thoroughly critical approach and methodological skill from any comparativist.

ON MANDELBROT'S PAPER

L. M. COURT: That there is a common core to all languages, and that it, at least initially, should be the focal point of linguistic studies, is the burden of Mandelbrot's paper. Since man is a biological species and since human psychology in its fundamentals (when the frills are stripped away) is everywhere alike, the thesis is more than warrantable. Every language has terms corresponding to nouns or the names of things and to verbs or words implying action. Modifiers of these two basic types (adjectives and adverbs) probably also occur in all languages, although the forms which they assume may vary. Without exception every culture must somehow take account of number, time and personal and social relationships, and in one way or another these

concepts or interconnections are embodied in all languages. The paramount question is exactly how any one of these physically or biologically real entities and relationships is communicated in a particular language, i.e., what device or grammatical structure is employed. We may find, for example, that two real-world distinct relationships receive identical grammatical treatment in one group of languages, whereas in another they are managed by unmistakably different devices. The writer believes that such a correlation between real entities and their grammatical instrumentations should be the heart of language structure study, at least in its embryonic phases.

B. MANDELBROT: I do not assert that there is a common core to all languages from the grammatical viewpoint. There is likely to be one, but I am not concerned with it. I think that there are ways in which the structure of natural language or rather discourse differs from that of artificial sign systems, but I have concentrated on *non*-grammatical features of similarity between different discourses and of dissimilarity between natural and artificial discourse. My study of the species-genera relationship is perhaps closest to the problem raised by the speaker.

ON HOCKETT'S PAPER

R. B. LEES: During the previous session Yngve advanced the hypothesis that if sentences are produced by speakers in the manner of a certain simple finite-state device, and if these sentences are assumed to have the usual immediate-constituent phrase-structure (parenthesized phrases of words), then the length of left-branched constituent-tree paths would have to be limited to some small number of nodes lest the amount of temporary storage, or memory, exceed reasonable limits, and that the principle syntactic devices in the grammar serve primarity to ensure this asymmetric limitation on "depth". Professor Hockett announced his approval of Yngve's hypothesis. But today he has described a reasonable model of behavior of the *hearer* in which constituent structure is assigned to successive element of an incoming sentence in such a way that for an increasing number of nodes on a *right*-hand branch of the constituent tree the amount of temporary storage required by the hearer would increase. Thus if Hockett's model for the hearer is correct, and it seems quite plausible, then Yngve's hypothesis, which he recommended so positively, could *not* be, despite *its* superficial plausibility. There are also certain other arguments against Yngve's hypothesis, such as the following: (1) There are predominantly *left*-hand branching languages, such as Turkish or Japanese, in which the favorite constituent structure pattern is a *re*gressive construction. (2) In ordinary speech, in both right- and left-branched languages, one never observes node-chains longer than some small number, whether to the left *or* to the right. (3) Those left-branched constructions which *are* widely used in English such as strings of attributive adverbials (as in: Any of those not too overwhelmingly grandiose structures $\cdots$), are not any more difficult to understand or produce than are the more frequently used right-branched

constructions; and exactly the same is true in the other order for left-branching languages such as Turkish.

V. H. YNGVE: Lees' comment appears to be directed more to my paper than to that of Hockett. It is my understanding that Hockett's approval extended only to the fact that I had proposed an hypothesis that could be tested, and not necessarily to the predictions of the hypothesis itself, since he has not put the hypothesis to a careful test. It is not necessarily true that a model from the point of view of the hearer must be the reverse of a model from the point of view of the speaker, and that a past-to-future asymmetry in the one model must become a future-to-past asymmetry in the other. In point of fact, I am not aware that Hockett's model contradicts mine in any way. As to Dr. Lees' other comments:

1. Even if the "favorite constituent structure patterns" of Turkish and Japanese are regressive constructions, this need not necessarily contradict the depth hypothesis. I have a few preliminary indications that depth phenomena consistent with the hypothesis may be found in Arabic, Chinese, Hidatsa, Japanese, Shilha, Toba-Batak and Turkish. We will not be sure, however, until careful and thorough investigations have been carried out by experts in these languages.

2. Ordinary speech in English frequently involves node chains that extend many more than seven steps in the right-hand or progressive direction, but not in the regressive direction, as I pointed out in my paper.

3. The quoted English construction has a depth of four. The fact that it is easy to understand and to produce is in perfect harmony with the hypothesis.

A. G. OETTINGER: There is a strong kinship among the hypotheses advanced by Yngve, the theory and practice of predictive analysis as described in my paper and the proposals of Hockett. The behavior of the prediction pool (approximately a pushdown store) corroborates, at least qualitatively, the nesting and depth analysis of Yngve. Most of Hockett's proposals seem to be already embodied in the predictive analysis system, where the use of the so-called hindsight provides a systematic means of detecting the presence of more than one alternative structure for a sentence.

ON WELLS' PAPER

J. D. SABLE: I take the phrase "necessarily true propositions" to mean analytic statements. Can the "Informativeness" measure be extended to synthetic statements? If so, will it then be equivalent to the "Content" measure of Carnap and Bar-Hillel on synthetic statement and hence be a generalization of "Content"? What relation does "Informativeness" have with the notion of "quantitative credibility and confirmation" of Hemphill and others?

R. WELLS: The informativeness of necessarily true sentences may as well be kept entirely separate from that of contingent truths. For if their respective degrees are placed in the same one-dimensional series, we would still

want to keep them separate by picking an arbitrary number such that all degrees of contingent informativeness are less, and all degrees of necessary informativeness more, than this number—or else vice versa. We would then also want to say that only contingent informativeness depends on serial degree or metrical amount of confirmation as explicated by Hempel, Carnap, etc.

INDEX

CDEFGHIJ–CM–89876543210