the majority of the rules are dictionary rules, this problem is closely related both to the problem of constructing microglossaries and to the subsequent problem of choosing a particular microglossary suitable to a given text.

Our current approach to this problem entails the construction of key word lists in the first stage of analysis which guide the computer in its choice of a previously constructed microglossary. Work to date indicates adaptations of this technique may not only contribute to the solution of storage and access problems but also facilitate analysis and simplify problems of semantic resolution.

Word-Meaning and Sentence-Meaning*

Elinor K. Charney

Research Laboratory of Electronics, Massachusetts Institute of Technology

A theory of semantics is presented which (1) defines the meanings of the most frequently occurring semantic morphemes ('all', 'unless', 'only', 'if', 'not', etc.), (2) explains their role, as semantically interdependent structural-constants, in giving rise to sentence-meanings, (3) suggests a possible approach to a sentence-by-sentence recognition program, and (4) offers a feasible method of coordinating among different language systems synonymous sentences whose grammatical features and structural-constants do not bear a one-to-one correspondence to one another. The theory applies only to morphemes that function as structural-constants and their interlocking relationships, denotative terms being treated as variables whose ranges alone have structural significance in sentence-meaning. The basic views underlying the theory are: In any given sentence, it is the particular configuration of structural-constants in combination with specific grammatical features which produces the sentence-meaning; the defined meaning of each individual structural-constant remains constant. The word-meanings of this type of morpheme, thus, must be carefully distinguished from the sentence-meanings that configuration of these morphemes produce. Sentence-synonymy is not based upon word-synonymy alone. Contrary to the popular view that the meanings of all of the individual words must be known before the sentence-meaning can be known, it is shown that one must comprehend the total configuration of structural-constants and syntactical features in a sentence in order to comprehend the correct sentence-meaning and that this understanding of the sentence as a whole must precede the determination of the correct semantic interpretation of these critical morphemes. In fact, the structural features that produce the sentence-meanings may restrict the possible meanings of even the denotative terms since a structural feature may demand, for example, a verbal rather than a noun phrase as an indispensable feature of the configuration. Two or more synonymous sentences whose denotative terms are everywhere the same but whose structural configurations are not isomorphic express the same fundamental sentence-meaning. The fundamental sentence-meanings can be explicitly formulated, and serve as the mapping functions to co-ordinate morphemically-unlike synonymous sentences within a language system or from one system to another. The research goal of the author is to establish empirically these translation rules that state formally the structural characteristics of the sentence configurations whose sentence-meanings, as wholes, are related as synonymous.

Translating Ordinary Language into Symbolic Logic*

Jared L. Darlington

Research Laboratory of Electronics, Massachusetts Institute of Technology

The paper describes a computer program, written in COMIT, for translating ordinary English into the notation of propositional logic and first-order functional logic. The program is designed to provide an ordinary language input to a COMIT program for the Davis-Putnam proof-procedure algorithm. The entire set of operations which are performed on an input sentence or argument are divided into three stages. In Stage I, an input sentence 'S', such as "The composer who wrote 'Alcina' wrote some operas in English," is rewritten in a quasi-logical notation, "The X/A such that X/A is a composer and X/A wrote Alcina wrote some X/B such that X/B is an opera and X/B is in English." The quasi-logical notation serves as an intermediate language between logic and ordinary English. In Stage II, S is translated into the logical notation of propositional functions and quantifiers, or of propositional logic, whichever is appropriate. In Stage III, S is run through the proof-procedure program and evaluated. (The sample sentence quoted is of course 'invalid', i.e. non-tautological.) The COMIT program for Stage III is complete, that for Stage II is almost complete, and that for Stage I is incomplete. The paper describes the work done to date on the programs for Stages I and II.

The Graphic Structure of Word-Breaking

J. L. Dolby and H. L. Resnikoff

*Lockheed Missiles and Space Company***

In a recent paper[1] the authors have shown that it is possible to determine the possible parts of speech of

* This work was supported in part by the National Science Foundation, and in part by the U.S. Army Signal Corps, the Air Force Office of Scientific Research, and the Office of Naval Research.

** This work was supported by the Lockheed Independent Research Program.

[1] "Prolegomena To a Study of Written English," J. L. Dolby and H. L. Resnikoff.

English words from an analysis of the written form. This determination depends upon the ability to determine the number of graphic syllables in the word. It is natural, then, to speculate as to the nature of graphic syllabification and the relation of this phenomenon to the practice of word-breaking in dictionaries and style manuals.

It is not at all clear at the start that dictionary word-breaking is subject to any fixed structure. In fact, certain forms cannot be broken uniquely in isolation since the dictionary provides different forms depending upon whether the word is used as a noun or a verb. However, it is shown in this paper that letter strings can be decomposed into 3 sets of roughly the same size in the following manner: in the first, strings are never broken in English words; in the second, the strings are always broken in English words; and in the third, both situations occur. Rules for breaking vowel strings are obtained by a study of the CVC forms. Breaks involving consonants can be determined by noting whether or not the consonant string occurs in penultimate position with the final *c.* The final *e* in compounds also serves to identify the forms that are generally split off from the rest of the word.

A thorough analysis is made of the accuracy of the rules given when applied to the 12,000 words of the Government Printing Office Style Manual Supplement on word-breaking. Comparisons are also drawn between this source and several American dictionaries on the basis of a random sample of 500 words.

Writing of Chinese Recognition Grammar for Machine Translation

Ching-yi Dougherty

University of California, Berkeley

Our approach to this problem is based on the stratificational grammar outlined and the procedures proposed by Dr. Sydney Lamb. How the theory and the procedures can be applied to written Chinese is briefly discussed. For the time being our research is limited to the particular kind of written Chinese found in chemical and biochemical journals. First the Chinese lexes are classified by detailed syntactical analysis, then binary grammar rules are constructed for joining two primary or constitute classes. How a more and more refined classification can eliminate one by one the ambiguity resulting from all possible constructions arising from juxtaposition of two distributional classes is discussed in detail.

The Behavior of English Articles

H. P. Edmundson

Thompson Ramo Wooldridge Inc.

Machine translation has often been conceived as consisting of three steps: analysis of source-language sentence, transformation of analyzed pieces, and synthesis of target-language sentence. This paper is concerned with one aspect of the last step, namely, the rules of behavior of English articles. Since the classical definitions of definite and indefinite articles are operationally imprecise, proper mechanistic rules must be formulated in order to permit the automatic insertion or non-insertion of English articles. The rules discussed are of syntactic origin; however, note is also taken of their semantic aspects. This paper describes the methods used to derive these rules and offers ideas for further research.

On Representing Syntactic Structure

E. R. Gammon

Lockheed Missiles and Space Company

The idea of sentence depth of Yngve (A Model and an Hypothesis for Language Structure, *Proc. Am. Phil. Soc*., Vol. 104, No. 5, Oct. 1960) is extended to the notion of "distance" between constituents of a construction. The distance between constituents is defined as a weighted sum of the number of IC cuts separating them. Yngve's depth is then a maximum distance from a sentence to any of its words.

Various systems of weighting cuts are investigated. For example, in endocentric structures we may require that the distance from an attribute to the structure exceeds the distance from the head to the structure, and in exocentric structures that the distances from each constituent to the structure are equal.

Representations of constructions are considered which preserve the distance between constituents. It is shown that it is impossible to represent some sentences in Euclidean space with exact distances, but a representation may be found if only relative order is preserved. If more general spaces are used then exact distances may be represented. It follows that for a wide class of sentence types, there is a weighting, and a space, in which the distance preserving representations are identical with the diagrams of traditional grammar.

La Traduction Automatique et l'Enseignement du Russe

Yves Gentilhomme

Centre National de la Recherche Scientifique, Paris

Les recherches effectuées depuis quelques années en vue de la Traduction Automatique ont conduit à des méthodes de travail et à des résultats intéressant la pédagogie des langues.

Une expérience d'enseignement du russe a l'usage des scientifiques fondée sur ces données a été poursuivie pendant deux ans à Paris (Centre National de la Recherche Scientifique et Faculté des Sciences), et a abouti à la publication d'un manuel.