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.