are particular cases of phrase structure grammars and that pushdown stores are the generative devices for context-free grammars.

Collecting Linguistic Data for the Grammar of a Language

Wayne Tosh

Linguistics Research Center, The University of Texas

Establishing the grammatical description of a language is one of the major tasks facing the technician in machine translation. Another is that of creating the system of programs with which to carry out the translation process. The Linguistics Research Center of The University of Texas recognizes the advantages in maintaining the specialties of linguistic research and computer programming as two separate areas of endeavor.

We regard the linguistic task as a problem in convergence. We do not expect ever to have a final description of a language (except theoretically for a given point in the history of that language). We do expect, however, to begin with almost immediate application of the very first grammatical description. We shall make repeated revisions of the grammar as we learn how to make it approximate better the language text fed into the computer.

The grammatical description of any one language is based primarily on specific text evidence. We are not attempting to describe "the language". We are, however, attempting to make descriptive decisions sufficiently general that new text evidence does not require extensive revision of earlier descriptions.

Corpora selected for description are chosen so as to have similar texts within the same scientific discipline for the several languages. Tree diagrams are drawn for each sentence in detail. The diagrams are inspected for consistency before corresponding phrase-structure rules are compiled in the computer. The grammar is then verified in the computer system and revised as necessary.

Derivational Suffixes in Russian General Vocabulary and in Chemical Nomenclature

John H. Wahlgren

University of California, Berkeley

A grammar based upon a conventional morphemic analysis of Russian will have a rather large inventory of derivational suffixes. A relatively small number of these recur with sufficient generality to acquire lexemic status (i.e., to be what is usually termed "productive"). Names of chemical substances in Russian may likewise be analyzed as combinations of roots or stems with derivational affixes, in particular, suffixes. The number of productive suffixes in the chemical nomenclature is considerably larger than in the general vocabulary. These suffixes derive from adoption into Russian of an international system of chemical nomenclature. A grammar of this system is basically independent of any grammar of Russian. It must, however, be consistently incorporated into the grammar and dictionary which are to serve in a machine translation system for texts in the source language containing chemical names.

Grammatical analysis of chemical suffixes and connected study of general Russian derivational suffixes has raised certain practical problems and theoretical questions concerning the nature of derivation. On the practical side, where a complex and highly productive system is involved, effective means of detecting and dealing with homography have required development. Theoretical consideration has been given to the question of grammaticality in chemical names and to problems of sememic analysis and classification of root and stem lexemes into tactic classes on the basis of co-occurrence with derivational suffixes.

On the Order of Clauses*

Victor H. Yngve

Department of Electrical Engineering and Research Laboratory of Electronics, Massachusetts Institute of Technology

We used to think that the output of a translation machine would be stylistically inelegant, but this would be tolerable if only the message got across. We now find that getting the message across accurately is difficult, but we may be able to have stylistic elegance in the output since much of style reflects depth phenomena and thus is systematic.

As an example, the order of the clauses in many two-clause sentences can be reversed without a change of meaning, but the same is not normally true of sentences with more than two clauses. The meaning usually changes when the clause order is changed. Equivalently, there appear to be severe restrictions on clause order for any given meaning. These restrictions appear to follow from depth considerations.

The idea is being investigated that there is a normal depth-related clause order and any deviations from this order must be signalled by special syntactic or semantic devices. The nature of these devices is being explored.

When translating multi-clause sentences, there may be trouble due to the fact that the clause types of the two languages are not exactly parallel. Therefore the list of allowed and preferred clause orders in the two languages will not be equivalent and the special syntactic and semantic devices available to signal deviations from the normal order will be different. Thus one would predict that multi-clause sentences in language A often have to be split into two or more sentences when translated into language B, while at the same time multi-clause sentences in language B will often have to be broken into two or more sentences when translating into language A.

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