Design of a Generalized Information System

Ronald W. Jonas

Linguistics Research Center, The University of Texas

While mechanical translation research involves the design of a computer system which simulates language processes, there is the associated problem of collecting the language data which are to be used in translation. Because large quantities of information will be needed, the computer may be useful for data accumulation and verification.

A generalized information system should be able to accept the many types of data which a linguist encodes. A suitable means of communication between the linguist and the system has to be established. This may be achieved with a central input, called Linguistic Requests, and a central output, called Information Displays. The requests should be coordinated so that all possible inputs to the system are compatible, and the displays should be composed by the system such that they are clearly understandable.

An information system should be interpretive of the linguist's needs by allowing him to program the data manipulation. The key to such a scheme is that the linguist be permitted to classify his data freely and to retrieve it as he chooses. He should have at his disposal selecting, sorting, and displaying functions with which he can verify data, select data for introduction to a mechanical translation system, and perform other activities necessary in his research.

Such an information system has been designed at the Linguistics Research Center of The University of Texas.

Some Experiments Performed with an Automatic Paraphraser

Sheldon Klein

System Development Corporation

The automatic paraphrasing system used in the experiments described herein consisted of a phrase structure, grammatically correct nonsense generator coupled with a monitoring system that required the dependency relations of the sentence in production to be in harmony with those of a source text. The output sentences also appeared to be logically consistent with the content of that source. Dependency was treated as a binary relation, transitive except across most verbs and prepositions.

Five experiments in paraphrasing were performed with this basic system. The first attempted to paraphrase without the operation of the dependency monitoring system, yielding grammatically correct nonsense. The second experiment included the operation of the monitoring system and yielded logically consistent paraphrases of the source text. The third and fourth experiments demanded that the monitoring system permit the production of only those sentences whose dependency relations were non-existent in the source text. While these latter outputs were seemingly nonsensical, they bore a special logical relationship to the source. The fifth experiment demanded that the monitoring system permit the production of sentences whose dependency relations were the converse of those in the source. This restriction was equivalent to turning the dependency tree of the source text upside down. The output of this experiment consisted only of kernel type sentences which, if read *backwards,* were logically consistent with the source.

The results of these experiments determine some formal properties of dependency and engender some comments about the role of dependency in phrase structure and transformational models of language.

Interlingual Correspondence at the Syntactic Level*

Edward S. Klima

Department of Modern Languages and Research Laboratory of Electronics, M.l.T.

The paper will investigate a few major construction types in several related European languages: relative clauses, attributive phrases, and certain instances of coordinate conjunction involving these constructions. In each of the languages independently, the constructions will be described as resulting from syntactic mechanisms further analyzable into chains of partially ordered operations on more basic structures. Pairs of sentences equivalent in two languages will be examined. Sentences will be considered equivalent if they are acceptable translations of one another. The examples used will, in fact, be drawn primarily from standard translations of scholarly and literary prose. Equivalence between whole sentences can be further analyzed, as will be shown, into general equivalence 1) between the chains of operations describing the constructions and 2) between certain elements (e.g., lexical items) in the more basic underlying structures. It will be seen that superficial differences in the ultimate shape of certain translation pairs can be accounted for as the result of minor differences in the particular operations involved or in the basic underlying structure. We shall examine two languages (e.g., French and German) in which attributive phrase formation and relative clause formation on the whole correspond and in which, in a more or less abstract way, the rules of relative clause formation are included as intermediate links in the chain of operations describing attributive phrases. The fact that in particular cases a relative clause in the one language corresponds to an attributive phrase in the other will be found to result from, e.g., differences in the choice of perfect auxiliary in the two languages.

* 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.