Journal des traducteurs
Translators' Journal

Toward Computational Linguistics?

Reinhard R. Hartmann

Aller au sommaire du numéro

Découvrir la revue

Citer cette note

TOWARD COMPUTATIONAL LINGUISTICS ?

Reinhard R. HARTMANN, Vienne *

Many engaged in linguistic pursuits readily jump to conclusions whenever the term *computer* is mentioned. Such comments are, e.g., "No machine can do the work as well as I", "Computers add to unemployment", "Calculators can perform mathematical operations, but language work ? No !", "The human mind cannot be replaced". On the other hand, we do not want to advocate the extreme opinion often raised, "The electronic brain is the solution to all problems".

Certainly a more objective look at the mechanims and possibilities of that fabulous gadget would be in order. There is already considerable literature available. But the translator, as many of his colleagues in the so-called arts and humanities, tends to be suspicious of any technical device, particularly when it threatens to undermine his conventional thinking patterns. In view of the speed and extent with which our world is subjected to scientific scrutiny, this attitude is quite outdated.

Strong emphasis must be placed on the fact that machines, in this case electronic digital computers, can only do what they are told to do. In other words, a program must be formulated, which consists of a logical series of instructions causing in the machine certain operations to be carried out.

In practice, this process is much more difficult than it sounds. The present writer hastens to admit that as an economist and linguist he is still puzzled by many aspects of the transition from (1) formulating the problems to (2) designing the programs to (3) controlling the processes to (4) receiving the results.

Since its inception, many branches of knowledge have been aided by electronic *data-processing* machines. The reader may be familiar with their application in business and documentation, in mathematics and technology. Needless to say, this development was only possible and fruitful with the cooperation of such experts as mathematicians, communication engineers, business machine producers, and others.

(*) M. Reinhard R. Hartmann, de Vienne, était le délégué officiel de *l'Osterreichischer Dolmetscher-Verband Universitas,* au congrès de la FIT, à Dubrovnik, en septembre dernier. En échange d'un article que M. Marcel Paré lui a fait parvenir pour le bulletin de l'association des traducteurs d'Autriche, M. Hartmann a bien voulu nous envoyer quelques notes relatives aux études qu'il poursuit en ce moment en Angleterre sur la traduction mécanique.

One of the most recent adaptations of the electronic computer is in the area of linguistic research and mediation. It must be stressed that the advantages of a computing machine, such as its tremendous speed, storage capacity and reliability, may only be fully utilized, if the object under investigation can be expressed in a set of rules and supported by statistical data.

We need not go into the intimate workings of the computer here [1]. Suffice it to say that the information must be coded in some form (usually the binary system) on cards or tape. This is called *input*. The data (numbers, instructions, words) are stored in the memory unit and manipulated in the arithmetic unit. The control unit directs all operations and the *output* part reads the results out.

The main branches of linguistic computation or *computational linguistics* can be categorized as follows : (1) mechanical operations such as counting, arranging, comparing and exchanging, (2) research goals as textual analysis, index compilation, authorship determination, deciphering and composition, frequency listing, (3) formal language elements such as letters, words, phrase, clauses, sentences, structure patterns, (4) multilingual projects such as translation and lexicography.

Some preliminary work has been done in most of these fields at centres in London; Cambridge; Washington, D.C.; Cambridge, Mass.; Ann Arbor and Detroit, Mich.; Philadelphia, Pa.; Seattle, Wash.; Santa Monica, Calif.; Jerusalem; Milan; Paris; Grenoble; Leningrad; Moscow, etc. There is international contact and overlap by methods, aims, languages, disciplines, and institutions.

We want to mention but a few achievements : concordances, i.e. alphabetical and frequency word lists with line references, to works of great men-of-letters (Plato, Thomas Aquinas, St. Paul) have been compiled; glossaries and indices as well as dictionaries of various works, sizes and languages have been made; structural and stylistic features of various texts have been measured and compared; letters and wordgroups have been classified according to certain criteria; doubtful manuscripts have been decoded and reconstructed from fragments; systematic vocabularies have been collected; etc.

Machine Translation is only one of these applications of the computer to language problems. Here too, the process must be analyzed to its basic routine operations and transformed into a sequence of logical binary instructions. The computer memory must be supplied with the specialized terminology plus a set of syntactical and morphological rules. The text is read in on special media, analyzed according to semantic and structural characteristics and corresponding patterns are looked up. This procedure is carried out by sentences and the equivalents are printed out.

The translations produced in this way are still rather crude. But they can be understood by experts in their respective fields. Automatic

---

1 — Voir pp. 103-112 : *La Calculatrice électronique, etc.*

language translation is a very young art and usually limited to specialized vocabularies. Programming linguistic phenomena is a complicated affair. Computer time is very expensive. Input of texts is tedious and trouble-shooting difficult. Moreover, there is not sufficient linguistic material available to make reliable programs taking into account all possible shades of meaning and grammatical structures.

In spite of these shortcomings, it seems highly probable that many more linguistic questions will be resolved with the help of automatic de-vices. Analysis, composition, translation and instruction will profit from the high speed and accuracy of the machines. If carefully handled, com-puters can make an important contribution to empirical linguistics.

∗