[From: Internat.conf. *Methodology & Techniques of Machine Translation*, British Computer Society, Cranfield, 13-15 February 1984]

NOTES ON INTERACTIVE TRANSLATION

Deryle Lonsdale
French-English Project Manager

Automated Language Processing Systems
190 West 800 North
Provo, Utah
USA 84601

Introduction

Past efforts in machine-assisted translation have been the subject of discussion at conferences like this one for years. Much focus has been put on the evaluation of these systems, with most of the effort in evaluating the quality of output of purely automatic systems. Only a few notable exceptions have attempted an interactive approach to the problem.

Alan Melby has described efforts to establish ITS, a fully interactive computer-assisted translation system (Melby 1980; 1982). Several researchers who participated in this project have developed a spin-off project at ALPS. The principle of interaction has since been fundamental to the system's design and subsequent development.

The first development was a multi-pass system which involved three interactive passes through the source text to address the problems of text segmentation, verb-particle matching, and bilingual lexical transfer. Between these interactive steps certain batch processes were run to accomplish other operations. It was soon apparent that a multi-pass system was not of optimal functionality and desirability. The multi-step system was

integrated into an on-line, interactive system where each segment (usually a sentence) is translated in real time with access to a dictionary; post-editing can be done during the translation or deterred to a later time. The translator is required to pass through the source text only once.

It is not within the scope of this discussion to detail the types of interaction used by the current system and how they complement machine processing. However, it will suffice to say that interactions are used to supply needed supplemental information relating to sentence definition, source language structural analysis on the morphological and syntactic level, bilingual lexical transfer, and target text generation. Any semantic information needed is provided by the translator in the word selection or lexical transfer step. Of particular interest is the resolution of structural ambiguities. A brief discussion on this topic was recently presented by Russell Bateman at the last ASLIB conference.

It would be informative to examine the challenges faced and advantages gained by an interactive approach.

Challenges of Interaction

Many challenges need to be met in the implementation of an interactive approach to machine translation. Unless consideration is given to the following points, interaction may prove not to be a useful activity.

In designing and developing any interactive computerized system, whether or not it is related to natural language processing, one issue of capital importance is that of "user-friendliness". Although a system may be complicated, its user interface must be easily operated and understood. In most cases a translator is computer-naive, so the issue of a system's user-friendliness is not to be overlooked.

There must be little compromise where the "natural" translation environment is concerned. The translator must have at his disposal any or all resources required to accomplish his task. This includes continual access to at least one bilingual dictionary, and access to the source and target language texts for contextual consultations. However, one can also take advantage of the computer's capabilities to perform such useful functions as on-line keyword in context (KWIC) scans through the source text, and word counting programs, which would be difficult to duplicate in a manual translation environment.

An interactive translation system must be designed to minimise the knowledge of computers and their operation. The system commands, keyboard layouts, screen formats, and overall

translation task flow must not cause appreciable operational problems. Extensive help utilities and documentation must be provided, not only with reference to the performance of the system itself, but also concerning such translation-related tasks as dictionary management, word processing, text formatting, and other operations.

Great care must be taken to orient interaction towards the level of a "typical" translator. Although most translators have strong formations in grammatical principles, interactions must not explicitly solicit technical or sophisticated data concerning syntactic, semantic, or pragmatic representations necessary for translation. Instead, such devices as alternate readings or possible parsings could be more profitably implemented. Questions should be unambiguously posed, succinct, and not beyond the linguistic grasp of the translator.

Realistically, one cannot assume that all responses obtained from the user during interaction are correct. Indeed, mistakes may be made by the translator because of insufficient knowledge upon which to judge correctly among given alternatives, because of a misunderstanding of the stated objective of the interaction, or even because of the accidental pressing of a wrong key. In designing an interaction all possible responses should be considered, if possible, and provisions made for the processing of each of the possible answers.

It would seem reasonable to suggest that there is an "interaction efficiency threshold" which cannot be profitably surpassed. In other words if too much time or effort must be expended by the translator in interaction, it may not be worth his/her time to use such a system. It is imperative that any information obtained by interaction actually be used by the system. Indeed, if some such information is not implemented the translator does not feel that his involvement has been useful. This probably agitates his annoyance with interaction. To prevent this reaction there must be a specific purpose for each and every interaction.

An interactive system cannot be considered to be a "black box" in the translation process. For this reason its operation must necessarily be more visible to the translator, whose role is to complement the processing capabilities of the machine. It follows that interactions reflect an inability of the system to resolve certain problems. The translator is called upon to compensate for these inadequacies, and may occasionally become critical of its shortcomings. In addition, any mistakes made by the system are immediately apparent to the translator. In many cases these errors may be of little consequence or easily corrected, but the translator will be bothered nonetheless. Finally, to a user unacquainted with the computational considerations linked with machine translation issues, the questions posed by an interactive system may seem to be "irrelevant" or "unnecessary". Although not all translators will react to interactive systems these ways, it must be expected that

some will feel this way.

As Boitet, Chatelin, and Daun Fraga conclude in their COLING 80 paper, "the human and social aspects should not be neglected". (Boitet et al 1980, 434)

Advantages of Interaction

There are several considerations which would tend to indicate that an interactive approach is a desirable alternative to a fully automatic treatment, despite the aforementioned difficulties in its implementation. Following is a summary of some important advantages which an interactive system affords.

Batch processing is the most widely accepted traditional approach to machine translation systems. In some cases pre-editing of source text is required in order to assure its successful treatment. The actual batch translation is processed automatically, sometimes overnight because of the extensive resources required to translate. A post-editing phase is necessary to correct the translation generated by the computer. Depending upon the quality of the machine-generated output, this may or may not be a difficult task. Translation with an interactive system, on the other hand, produces the target text in real time, and since a human is involved in the translation process, the text may not be as difficult to correct, proofread, and even format during the translation task.

One of the most readily apparent advantages of an interactive approach to translation involves dictionary considerations. For example, dictionary management can be dynamic. At any time during the course of translation, extraneous entries and translations can be suppressed or deleted altogether from a dictionary, missing words can be added, and dictionaries can be consulted. This dictionary flexibility is an extremely important consideration, since a dictionary is one of the translator's most important resources.

Dictionary entries in traditional automatic systems contain complex, syntactic and semantic features in order to assure that certain selection criteria have been met in the bilingual lexical transfer step. However, in an interactive system the repository for semantic and certain syntactic information is the human translator, and not the dictionary. The task of dictionary work is therefore substantially reduced, with respect to both quantity and complexity.

A problematic area of natural language processing by computer is the resolution of ill-formed input such as missing punctuation, orthographic errors, and spurious characters from the input medium. Clearly, though, in an interactive system such anomalies can be handled and treated upon interaction with a translator.

Among the greatest advantages of an interactive system is the fact that the translator's job does not merely consist of proofreading the computer's work, but rather of actively

participating in the production of the translation. The translator is in control.

When a mistake is made by a computer while attempting a translation there is often a "ripple effect" in which errors are compounded throughout a phrase, clause, or sentence. This is particularly true of highly inflected languages where all elements must agree in gender and/or number and/or declension. If the incorrect word were chosen as the translation of a source text term, and if it were not of the same gender as the proper translation, this ripple effect would be readily apparent. An interactive rendering of the translation would provide a "front-end" solution to this problem by assuring that the proper equivalent is assigned (interactively, if necessary). Interaction therefore permits the early correction of errors, thereby reducing subsequent error propagation.

One obvious advantage of interaction is the fact that it integrates the machine capacities for morphological and syntactic processing with the human intelligence necessary to solve any problems encountered in these stages, and to continue processing at the semantic and pragmatic levels. This obviates the need to have recourse to knowledge representation and artificial intelligence processing for the complete treatment of translation-related problems. As technology advances in these fields, processing capability can be associated incrementally into the translation process. For the present, though, proper use of information gleaned from the translator's human intelligence

increases the probability of correctly rendering a translation.

Since interaction reduces system requirements by minimising the processing power necessary to attempt a translation, it is possible to attain increased cost-effectiveness. The system can be implemented on a smaller and less expensive hardware configuration.

A system which is interactive can be very flexible in that it can lend itself to several environments as defined by run-time switches. Such switches permit definition of default values where several possibilities are recognised by the system. For example, in a technical users' manual where the English imperative appears, one could translate it into French using the conjugated imperative, or the more preferred impersonal (infinitival) imperative. For such texts the switch could be set appropriately so that the correct form would be chosen without any interaction. Switches are particularly useful in the bilingual transfer and target language synthesis steps of treatment. This ability to redefine environments assures the system's applicability to the treatment of a wider range of text types, and not just sublanguages.

The proper design and implementation of man-computer interaction can avoid the realisation of Andreyewski's description of more conventional machine translation modes:

> ...(duplicating) the problems
> of a single translator working in a

> vacuum and engaged in little more than vocabulary substitution using a dictionary. (Andreyewski 1981, 60)

Conclusion

Interaction has proven in the past to be an important answer to the difficulties faced in attempting the automated translation of human language. Properly designed and implemented, human intervention can extend the computer's processing capabilities on the morphological, syntactic, semantic, and even pragmatic levels. There are several important considerations, though, which must be taken into account if interaction is to be successfully implemented. When these are heeded, a powerful man-machine team can be created which may serve to bridge the gap between today's technology and tomorrow's innovations in the field of natural language processing.

References

Andreyewski, Alexander. 1981. "Translation: Aids, Robots, and Automation". META, Journal des traducteurs/Translators' Journal, 26 , 57-66.

Boitet, Christian, Philip Chatelin, and P. Daun Fraga. 1980. "Present and Future Paradigms in the Automatized Translation of Natural Languages". COLING 80: Proceedings of the 8th International Conference on Computational Linguistics, 430-36. Tokyo, Japan.

Melby, Alan K. 1980. "ITS: Interactive Translation System". COLING 80*:* Proceedings of the 8th International Conference on Computational Linguistics, 424-429. Tokyo, Japan.

Melby, Alan K. 1982. "Multi-Level Translation Aids in a Distributed System". COLING 82*:* Proceedings of the Ninth International Conference on Computational Linguistics, 215-20. Prague, Czechoslovakia.