Computer-Assisted Interpretation

Everyone talks about how computers can help with printed translation. Here's a fresh look at implications for the spoken word.

by Daniel Glon

When I recently shared this title with my colleagues, I heard a lot of "you mean computer-assisted *translation,* right?"

Wrong. I mean computer-assisted conference interpretation. Alas, I am indeed talking about our livelihood. However, I am by no means pessimistic by nature and on this subject will refrain from merely constructing theories of which the only real merit would be imaginative richness.

The computer processing of a spoken message and its reformulation in a different target language involves the following phases:

1. Recognition of a spoken word, regardless of who the speaker is and whatever his accent may be.
2. Analysis of the message, whatever imperfections it may contain (mistakes being more frequent in oral than in written expression).
3. Translation of a comprehensible source-language message into a comprehensible target-language message at the speed of a normal speaking voice.
4. Vocal reproduction of the result in the target language.

Let us consider these four points going from the simplest to the most complex, hence, in the opposite order of the above list.

Vocal Reproduction

Synthesized speech is already a reality. Even if the sound is sometimes metallic, it is quite easy to understand. The problem of the speed of speech does not seem to present any difficulty.

Translation

One of the most advanced machine-translation (MT) systems among those already on the market is the Systran system used by the European Union. The various users of the software confirm that for five years the quality of rough translations has progressed enough to allow for a 95 percent comprehension rate. Total comprehension is hindered by obstacles inherent to the source message and obstacles inherent to the software itself.

In the past five years, the software's performance has improved considerably and everything points to MT's progress continuing, especially now that more and more competent translation professionals are taking an active part in its development. Systran allows for a translation speed of 300,000 words per hour. Here, too, speed is apparently no obstacle.

Message Analysis

The main problem lies with the source message. We have often taken a backward approach to the problems involved in MT, settling for post-editing by a human being instead of opting for preventive measures. But by pre-editing material before it is processed through MT, we can eliminate potential "garbage" at its source and avoid extensive post-editing. In fact, the nature of computer-assisted interpretation makes it impossible to carry out post-editing, since the result must immediately be clearly expressed by the synthesized voice

system. So, disregarding errors made by the system (which are likely to become extremely rare), how can errors attributable to the message source be eliminated? By artificial intelligence (AI)?

While it is already superior to human stupidity, AI is not yet up to the task at hand. The only remaining solution is natural or human intelligence. We could quite easily foresee a human interface between the message source and the translation software. This "re-expressor" or "pre-editor" would be able to understand perfectly the meaning of the original and express it simultaneously in the *same* language, keeping in mind both the rules of the language being processed and the structure required by the software. This person would be a language professional (like all of us) whose role would be to simplify the form of the message for the machine, which incidentally is not a very complicated task. Like us, he should be able to understand a subject, whatever its level of complexity, so as to be able to reformulate the message, but he would have no need whatsoever to know a language other than his native tongue.

In this way, a French-speaking "re-expressor" in the "French booth" would merely reformulate the interventions of the French-speaking delegates and would follow the delegates speaking other languages, if necessary, by using the computerized interpretation system. The level of skill involved would be quite inferior to that required of a simultaneous interpreter, the task being of lesser difficulty. We would see an increase in the number of booths manned by only one person and a decrease in the number of jobs, while the number of candidates for those jobs—monolingual speakers or interpreters forced to change over—would rise.

With the problem of message-source imperfection resolved, MT would be able to take place under the best possible conditions. The quality of the final result would be proportional to the error-prevention measures taken.

Word Recognition

Word recognition is the prerequisite to the subsequent phases, and it is true that it presents serious difficulties. At present, the most high-performance systems only recognize a limited number of words. The majority of these "mono-speaker" systems can recognize only one voice, and that only after a preparation phase in which the speaker pronounces each of the words

A potential avenue in automating simultaneous interpretation: a human "re-expressor" or "pre-editor" who understands perfectly the meaning of the original and expresses it simultaneously in the same language, prior to computer processing.

once or twice, allowing them to be analyzed and then stored in memory.

Many language professionals are relieved once they read this. Taking into account the multitude of accents that exist within any given language, plus the accents of non-native speakers, our work is made much easier with the mono-speaker system. Accent problems and the presence of multiple speakers are taken into account by our "re-expressor." A mono-speaker system is all you need.

So the hard part is over. There is so much research being done in the field of word recognition and so much progress that has already been made in word-recognition applications that enormous progress should be made in this field in the years—indeed the months—to come. Of course, a great deal of work will be necessary to arrive at spontaneous language recognition, but experts agree that the problem is not one of feasibility but rather of the amount of time allowed.

When we finally solve the problem of word recognition in five, 10, or 15 years (probably sooner at the rate things are moving), the conditions necessary for the success of all four phases will be met and automatic interpretation will be a reality.

The grace period allowed for by the time it will take to overcome the difficulties of phase one will be used by the experts to improve the quality of the software. The new generation of software will no longer be limited to translating from one language to another, but will translate into and out of a meta-language. Efforts in this direction—including invented languages like Esperanto, natural languages like Aymara (an Indian language of Bolivia), and a French-government project based on algorithms—have already been tried.

There is one obvious objection to be made here: these systems work from words and not from ideas. While a word-based system might increase terminological consistency, it relies on meaning which has already been digested, clarified, and put within the computer's reach by the pre-editor. In another respect, we would not be off-track to assume that the impoverishment of the English and French languages will continue, indeed accelerate. For these generations are used to reading poorly translated computer manuals and communicating in a technical and poorly formulated *Franglais.* They will not have the least problem understanding a message transmitted by a computer—they may even find it too literary for their liking.

A Final Thought

The computer's arrival may actually make it possible to impose discipline on the participants where we fail to do so. The cooperation demanded by the unforgivingly rigid structure of computers will force them to work in an orderly fashion and to submit in advance a list of the technical terms they plan to use. The computer either has the words in its memory, or it does not; it cannot just come up with something in a pinch.

This will also give these delegates a chance, if they are linguistically inclined, to consult a list of equivalent terms to make sure they are using the terms appropriately. The machine, through its very existence and through the delegates' attitudes towards it, will create the ideal working conditions that we human interpreters have had such difficulty creating.

All that is left to do is to await the advent of a reliable voice-recognition system. And even if it is only a mono-speaker system, it might be all we need.

Daniel Glon is a currently a visiting professor at the Monterey Institute of International Studies. A seasoned translator and interpreter, Glon served as senior terminologist and translator for the Atlanta Committee for the Olympic Games in 1996. email him at dglon@miis.edu