

New York, October 22, 1987

An open letter about glossaries

Dear Colleagues:

This letter is intended as a query on a topic of interest to many translators. No claims or statements of any sort are being advanced — rather it is an attempt to articulate my own sense of puzzlement in public. May I therefore request your patient compassion as well as any corrections you may care to make.

We sometimes read of learned mathematicians presenting the final proof amidst welcoming applause that such and such an equation can never be solved. In this spirit, discovering the limits of one's ignorance can be as valuable as learning something new. Language does not at least yet lend itself to such exactitude as mathematics, but I cannot help wondering about a few things as I read of tomorrow's computer systems, impending global exchanges of electronic glossaries on all subjects, and remarkable new computer insights into the translating process. All of these are developments that will influence our lives and work in many ways. But are all of these events truly due to occur on as total a scale as projected, or could there just be some key piece of the equation that doesn't mesh, leaving it unsolved in some final sense on both the philosophical and practical levels?

Like others in the field, I follow publications on computers, AI, linguistics, and related areas. I am more than aware of the promise held out by CD ROM, and in general it looks like the hardware part of the equation is fully in order. Software should prove even easier to deal with — indeed several products are already on the market — so where can the mistake in the equation lie, if there is one?

I wonder if it might lie — and this is my query to colleagues — on the human level: in creating and maintaining the necessary reference material, the so-called glossaries in all

subjects, into and out of all major languages. The existence of these glossaries is just beginning to come to the attention of translators in this country. What is already beginning to happen is a bit reminiscent of the views of Yehoshua Bar-Hillel, who at one time believed that Fully Automatic High Quality Translation (FAHQT) was possible. Later, after he studied the problems more thoroughly, he changed his mind and is now regarded as a proponent of the opposite point of view.

But even if we leave to one side the possibility of Fully Automatic High Quality Translation, I wonder if there might not still be some serious problems involved in creating and arranging these glossaries. Let's take a simple example involving only two languages: a complete German-English two-way glossary for all the sciences and social studies. German and English have been chosen for presenting minimal linguistic problems: they are closely related languages and have shared a key role in the evolution of modern knowledge and science. And let's also leave out the science-fiction idea that the system can be so perfect that even a non-expert non-translator should be able to handle it. In other words, let's assume that someone moderately knowledgeable, either about the subject matter or the languages involved, is trying to translate a text using such a system.

Even, if we place such extreme limitations on our basic demands, I still wonder if an adequate universal glossary of the sciences could be constructed and kept updated with the necessary degree of consistency to eliminate such errors as even a qualified translator or specialist might make, to say nothing of the linguistic lay person. The reasons for this are not hard to explain. How many translators, on looking up an expression in a technical dictionary, have had the experience of finding closely related words and phrases but not the exact one being sought? In such cases the correct translation

could only be achieved by considerable ingenuity or consultation with a specialist in the field. Sometimes the dictionaries could contain outright errors. And sometimes a translator could be working in a field so new that there are no dictionaries. Does anyone suppose that the advent of electronic glossaries will suddenly and definitively change all of this?

And what of terms that have one meaning in, say biology, but have a quite divergent meaning in an adjoining science, say bio-chemistry or neuro-biology? What of texts on botanical subjects written by chemists or those on climatology written by astrophysicists, to name only two examples? Will our all-embracing German-English glossary be able to keep all this straight and flag the user appropriately, or will it ail still be up to the translator/expert to solve?

And this is what we confront in just dealing with German and English. What happens when we open the gates of omni-directional translation into and out of a great number of the world's languages, not all of which necessarily share western epistemological and ontological underpinnings? Is it just possible — and here I am hoping my colleagues can help me with their own insights — that even with the German-English example we may already be dealing with the linguistic equivalent of painting the Brooklyn Bridge? As soon as the workers finish painting one end of the bridge, they have to go back to the other end and start painting all over again. Except that where it might help hiring ten times as many workers to paint the bridge, this will not work with computer glossaries compilers, as in our field the extra workers can actually get in each other's way or even destroy each other's contributions. And if our relatively simple bilingual glossary already presents such problems, where does this leave various speculations about universal grammar and deep structure—while perhaps technically correct, could they possibly turn out to be irrelevant to larger linguistic realities?

All of which might seem very philosophical and abstract, except for the fact that such glossaries are already struggling to come into being. One reads appeals from software manufacturers for translators to exchange glossaries or set up a glossary bank. At the United Nations INFOTERM is putting together its own

glossary dedicated to a magnetic version of the Magna Mater dubbed MATER, with its PC stepsister MICROMATER trailing close behind. Surely such glossaries can be helpful to translators even if an ultimately perfect system is not feasible, but a number of questions need to be answered early on. Jean Datta touched on some of these problems in the September 1986 issue of *Language Monthly*.

Precisely how will such glossaries be put together, and under what economic arrangements?

Will they be sold to translators or given out with specific programs or rented on-line?

Will they be the work of scholars or graduate students or bureaucrats or clerks?

Will different translation software systems have interchangeable glossaries (if so it would make a first for compatible standards in the computer field)?

Who will see to it that glossaries are arranged in a standard way or that idiosyncratic macros and abbreviations or outright errors are edited out?

And perhaps most important, will translators willingly donate their own laboriously constructed glossaries to large companies (or even to other translators), or will they expect to be paid?

If the latter, how will payment be determined, by outright purchase or by royalties? (After all, publishers of dictionaries expect to go on collecting royalties — why should a translator not do the same?)

Or should such glossaries even be undertaken by a laissez-faire system — might they perhaps develop more harmoniously under the aegis of national, international, and professional organisations?

And, finally, what of access to these glossaries — is it to be limited to translators and specialists, or to anyone able to pay a fee, or should it more properly, as the heritage of all human culture, be freely granted to all interested parties?

Thus, there are two sets of questions that may require resolution — the larger philosophical one and its detailed practical consequences that are already being felt in our field. Does anyone have any answers?

Sincerely,
Alex Gross, Chairperson,
Machine Translation Committee
New York Circle of Translators