# Machine Translation - a New Dimension and What You Can Expect

## Fred Klein

••••••••••••••••••••••••••••••••••••••••••••••••••••

### A New Dimension

Built-in obsolence? Part of my paper about MT (TC-Forum 4-97, TR 5) must be updated. On December 9, 1997, DEC and Systran A.G. launched AltaVista Translation Service, a free real-time system that can "translate" information from the predominantly English-speaking Web. You could -- and should -- test yourself, though today I faced a disaster.

Instead of "translation", AltaVista offered me unbelievable junk, evidently, an unedited MT version of American promotional material. The text was unreal, the result of a myth: You click a button and the translation is staring at you. You are in the middle of a jungle.

Before you start the test, you should understand the problems, because MT is NOT context-sensitive. For instance:

- The term "fair train stop" was translated as "gerechter Zughalt".

- The noun "fair" (Messe) was rendered as a wrong adjective.

- The term "nipple, stainless steel" was translated as "Brustwarze, rostfreier Stahl" (breast nipple was the only equivalent found).

- The term "miss distance" (the distance by which a missile misses a target) was rendered as "senorita distancia".

- "Tape drive" became "Leukoplasttreibwerk" (Leukoplast is a German trademark for an adhesive plaster).

- And the American term "billion" became "Billione" in German instead of "Milliarde".

Fun? Impossible? You should withhold judgment, and be patient.

The original "classical" MT involves pre- and post-editing (Controlled English) and is INTERACTIVE. Human users must input and code words not found. They may have to create customized specialized bilingual dictionaries and, most important, they must select the proper term from large groups of terms with more than one meaning. Quality is enhanced, but the cost is prohibitive, and post-editing can become a time-consuming nightmare. The investment is very high.

Real-time MT of the AltaVista/Systran type is free or very cheap (1 cent per word). It is unbelievably fast, available 24 hours a day anywhere with a minimum requirement of a 486 PC clone and Internet access. There is no pre- or post-editing, the text is called a "suggestion", and AltaVista offers a needed disclaimer.

The language is uncontrolled and there is no access to databases or online dictionaries. The system offers you a translation for EVERY word--often the wrong one. Words with typos are not translated. You can take the result and edit it conventionally, but that defeats the purpose?

Is AltaVista's real-time MT a ridiculous, flawed experiment? I do not think so. It brings a new dimension to international information transfer, NOT to translation. It is a new niche, something impossible until now. It is application-driven and limited to specific purposes.

If you are a professional translator -- like me -- or a technical communicator, you must radically change your thinking to accept a different approach. No professional translator needs to feel threatened by this type of "translation".

We think in terms of good English, style, and high quality, publication-type translations. Each time I find a mistranslation or typo I jump; it has become a reflex. But we deal with a different product to be judged by other standards. The first time I saw the strange result, I jumped and rejected it. You may do the same, but I suggest that you test and test again at random. Think of it as cheap, instant international (mis)information.

## What Can You Expect from Real-time, Online Translation?

The results are completely unpredictable. I suggest you test 10 different topics. I am not listing Instant Translations here, on purpose. In one case, a medical "translation" was easy to understand; in another test, I, as a professional translator, was unable to make any sense out of the "translation". In a third case I could understand only some of the content, based mainly on a few nouns and some verbs: the total text was jibberish.

Some multilingual Europeans, and most professional translators, will gleefully reject the whole system. Some monolingual Americans will proudly praise the results uncritically, as a breakthrough in MT. Both are wrong.

AltaVista's "Translation Assistant" provides a free service to the Internet community. The company claims that a document is translated into something comprehensible, but not perfect. According to the promotional copy, Systran, a partner of AltaVista, "produces the best and fastest translation software in the world". Fast, yes. But not the best!!!

The sales point of AltaVista and Systran is the word "translation". If we could define it as "international information transfer", I would be more satisfied. Moreover, a disclaimer should be added automatically to each "translation", that "a computer translates everything without any human intervention."

Systran also suggests you may have lots of fun with the mistranslation. Maybe - but you should still be able to use it.

Which type of text is suitable for real-time MT? Web pages, abstracts, headings. Systran suggests newspapers or magazines. My previous experience in "classical" MT tells me that "running text" is not suitable, but then I was thinking of "translation" instead of "information transfer". We are presented with "suggestions"; take it or leave it. If you cannot take the "junk" , real-time MT is not for you.