

# Bédard proposes new approach

## Machine pretranslation

---

*Montreal Translator and machine translation consultant Claude Bédard has come up with a new idea, designed to combine some of the utility of the machine with the expertise of the human translator. The concept has been named Machine pretranslation or MPT (la prétraduction automatique — PTA in French.)*

The idea first came to Claude Bédard as he reflected on the weaknesses in the interface between machine translation as it has been practised up to the present, and the post-editing stage. The machine helps the human translator by consulting its dictionary, and producing text in the target language, but then proceeds to spoil this by making various errors. The post-editor has to follow the text as generated by the machine, going backwards and forwards in correcting this text, in a manner which becomes inefficient and irritating.

It is this post-editing problem, that the machine actually generates work for the post-editor instead of taking it away, that Claude Bédard set out to solve. He came to realise that very little thought has been given by machine translation developers and users (a notable exception is *Muriel Vasconcellos* of the Pan-American Health Organisation) to the flow of the mental processes of the human translator or post-editor.

Machine translation as conventionally practised, he therefore considers, may be of limited use to monolinguals with little knowledge of the source language, but it is not really a suitable tool for professional translators.

What the professional translator looks for from the machine, especially

in texts where specialised terms play a major role, is for the machine to output the terms accurately in the source language. What the professional translator does not want is the machine trying his hand at the complexities of rendering sentence structure or idiom, and producing, at the worst, gibberish, and at the best target language text which requires a lot of moving around and amendment.

Claude Bédard began by experimenting to see what could be achieved using global search-and-replace routines available on any word processing program. He found that by replacing all source language special terms with their target language equivalents, but leaving untouched any words from general vocabulary or needing thought and care in translation, he obtained a bilingual output text which, though clumsy in appearance, could be a basis for translation. The trick is to read it — by acquiring an attitude of mind — as a source language text with target language terms ready supplied.

This pre-translation is then edited on a word processing program. Because the line of thought of the translator/post-editor is much straighter with such a text, he claims, it is more “translator-friendly” than typical MT output.

Claude Bédard is now working on the development of what he calls his Pre-translation Machine, using the high-level programming language GramR developed by MT researcher *John Chandiooux*.

In his approach only fixed terms (*les mots sûrs*) are translated; anything else, especially words which can have more than one meaning or target-language equivalent (for example, the

English word *number* might be translated into French by *nombre*, *chiffre* or *numéro*), is left alone. No word changes place, no word is added to the text, and no word is deleted.

Expressions which are fixed terms but contain several words (e.g. *charges momentanées*) are tagged with an underline connecting the words (*charges-momentanées*).

When the system is developed it would conduct first a source language processing, with each word being lemmatised and assigned syntactic categories from a large source language dictionary. Simple, local rules would be used to disambiguate homographic words. When these rules are exhausted, some words would still remain homographically ambiguous, and in such circumstances, as indicated above, they would simply not be translated. Agreements (article-adjective-noun or subject-verb) would be made locally only. Verb tenses are limited to third persons singular and plural, and infinitive. Claude Bédard began by covering several tenses, but the effort involved turned out not to be justified. Thus the English *would arrive* will now simply come out in the pre-translation version as *would arriver*. It would be necessary to develop special keyboard commands (jump past next word, etc.) to make post-editing easier. He is also looking at extra modules which would identify repetitive strings, including whole sentences, for processing as such. After two years of experimenting with the PTA approach, Claude Bédard is convinced that it could provide a valuable link between MT and professional human translation. Translators would find that they had a machine which did not “mess up” the language, but which left the real work of translation to them. And because PTA does not intervene as a linguistic tool, it is easier to use it for the many technical texts which do not follow the rules of normal composition (telegraphic style, listings, statements

without verbs, badly written sentences, etc.). It also has the benefits of simplicity.

Finally Claude Bédard sees his approach as a beacon to translators to take a much less passive role in face of Machine Translation. They should take the opportunity to define their own tools, and to demand machines which fit more closely with their own way of working, instead of trying to enslave them to alien thought processes.

It is also a way in which translators could equip themselves to cope with the greater demands, in terms of volume and speed, with which they are going to be faced.