[*The Linguist*, October 2000]

What do translators do?
And what machines can not.

By Paul F. Wood M.A.

"Table lamp high a ca. 42 cm with transparent glass screen and flowers in the youth style engraved once should broken go. The screen, should be also that no problem, for you can amount to see is worthwhile it individually with me nachbestellen. The delivery costs for this DM 11.- within the BRD favor you also my further offers at. It yourself again and again".
—State-of-the-art EBay e-commerce German—>English translation software / 18 August 2000 (Note 1)

"For almost fifty years, the promise—even the certainty—of machine translation taking over from humans was a recurrent part of the grand computer dream, one component of an all-enveloping 'artificial intelligence' destined to organise our menial tasks, our language problems and our daily driving. But during the past ten years or so [&ldots;] this dream has slowly receded as even MT and AI experts have come to grasp the true scope of the problems they have undertaken." (Note 2)

Ask any translator, though, what they think of machine translation (MT) and the answer is likely to be less sanguine.

But that, in a nutshell, is the whole point. Computational linguists, artificial intelligence researchers, language engineers and (in their latest incarnation) cognitive neuroscientists have never really bothered to ask translators what it is they do. The reason? Most language engineers are monolingual, trapped in the microcosm of their own native tongue with a smattering of schoolboy Spanish or French. They are not going to voluntarily broadcast this deficiency by collaborating with professional translators. Secondly, at least in my experience, the art of translation is widely perceived to be a feminine activity, the "science" of language engineering to be masculine.

In a work I first saw advertised in these pages - *The Translator's Handbook 3rd Edition* [Aslib: 1996] -there is an interesting article entitled "Machine Translation" by Harold Somers and Clare Rutzler of the Department of Language Engineering at UMIST, Harold Somers being the Professor of Computational Linguistics there. What troubled me most was the fact that the bibliography attached to the article - fifty or so references - did not include any mention of a human translator. *(Note 3)* It is, I'm afraid, part of the Orwellian aspect of the whole MT discussion that MT enthusiasts do not seem to be uncomfortable with terms such a "human translator", "language engineering", "knowledge engineering" and (my favourite) "shallow linguistic techniques". I mean: we don't talk about "human violinists", do we?

So I wrote to Prof. Somers who replied, albeit non-specifically, that "about a dozen of the works cited were written by translators or translation managers". *(Note 4)* Well, quite. He also added - rather tellingly, I thought - that "in all [his] years of talking to translators, [he] had never found one who can explain what they do in a manner remotely useful for a computational solution". In the event, he did not respond to my suggestion that perhaps he wasn't asking the right questions and that possibly he didn't wish to hear what translators really have to say. Either way, I'm indebted to Prof. Somers for the title of, and inspiration behind, this article.

So: before I analyse what I think I do when I translate German into English, let's start at the very beginning. *(Note 5)*

Though the notion of Machine Translation (MT) had been around quite a few years, it is generally held to have come into being about 1949 — at the instigation of Warren Weaver, a US mathematician attempting to answer the famous communication model question set by the US political scientist Harold Lasswell: *Who says what to whom with what effect?* One can only wonder from the literary point of view whether Lasswell was guided in any way by George Orwell's definitive six questions of clarity - posed only three years earlier - because, if he was, it is a pity he ignored the sixth: *Have I said anything that is avoidably ugly? (Note 6)* Weaver worked closely with a telephone engineer called Claude Shannon, better known to Garry Kasparov as the father of the chess playing machine, the progenitor of Deep Blue. On the other hand, the Cold War climate was clearly having its own kind of influence as well. *(Note 7)*

And there you have it in another nutshell: a male configuration of mathematicians, engineers, political scientists and the US military state complex. In retrospect, it is hardly surprising that Chomsky - a mathematician and politicologist funded by the military - should go on to proclaim, erroneously, the existence of deep structure and universal grammar. This is what the engineering community of the 1950s wanted to hear. They wanted to *master* language without the effort of learning foreign languages in the search for the Atlantis of MT. It is a poignant counterfactual question which will never be resolved: just how much time, effort and money could have been saved if the advice of translators had been sought out in the first place? Ironically enough, the IBM Deep Blue software developers who went on to beat Kasparov were not so stand-offish. They consulted a whole range of, er, human chess grandmasters and integrated the findings.

The name of "machine translation" reflects its birth at a time when computers were mainframes about the size of the White House. However, what the MT obstetricians in the 1950s failed to predict was the rise of word processing software and translation memory tools. Now, since Microsoft partly owns Trados, the leading memory tool providers here in Germany, one can only speculate that Bill Gates & Co. must have looked into the feasibility of MT and concluded that it is unprogrammable, unstorable or unprofitable. Or all three of those things. Considering the amount of jittery software which is prematurely thrown on to the market - ViaVoice, Naturally Speaking, AOL 5.0, Windows 98 and Word 2000 etc. - considering all the jittery software already on the market, I bet Microsoft would have produced something if it had as much as half-worked.

That they haven't can only be seen as further proof of severe obstacles standing in the way of MT progress, obstacles which are clearly insurmountable to people in the know: translators. The combined power of the world's computer industry has thus had five decades to come up with an acceptable solution ... and it hasn't. The Alta Vista Babelfish Translator on the Internet, for example, like the EBay translation software, is simply dreadful. *(Note 8)* And it's awful testimony to the dimness of the people who put it there that they are not aware of how weak it is. But, as an American colleague of mine always says, *they do not know that they do not know. (Note 9)*

Cognitive aspects of translation

In the review footnoted above, Alex Gross lists seven "cognitive" aspects of translation as identified by Dr. Isabelle Schrade as late as 1998:

- memory
- general knowledge

- linguistic knowledge
- understanding & analysis
- recipient-oriented reformulation
- human intuition and
- creativity.

Mm. Not especially incisive, is it? Apart from the fact that the seven categories could easily be brought to bear on the art of opera singing, one can only wonder, again, what MT apologists have been doing for the last fifty years.

Permit me then to describe how I set about translating most of my German into English assignments. I am talking about the large pool of work available in the technical, administrative and commercial fields which are not so felicitously composed: the result perhaps of their authors never having been trained to write. This type of prose, what might be defined as technocratic, is deliberately austere. It works in German but, for an English reader, it is heavily substantivated, passivised, nonidiosyncratic, nonnavigated, depersonalised, not always a model of clarity. German also operates on an effect-to-cause syntagm more or less the opposite of that in English. And underneath, further problems, lie the intricacies of German lexis, word order, tense and grammar.

Tackling a German to English translation, the translator might:

• desubstantiate	verbalise the *-ung* forms, providing more verbs for an English flow
• depassivize	generally switch the mood to active
• colourize	use collocations, idioms, a lively turn of phrase
• resyntacize/reorder	proceed from cause to effect, from a to b to c to d, from input to output
• remap	insert pointers like "however", "subsequently", "therefore" etc.
• repersonalize	put in a few "yous" or even "ones", create a dialogue with the reader
• clarify	talk to the author.

There is no way any software can handle any of that. But that's not all:

I also find myself consciously, musically, recasting the rhythm, especially in the more advertising oriented assignments since German in-sentence lists or structuremes form different numbers of segments than English which seems to prefer clustering in triads (I reckon). *(Note 10)* Information in an English sentence is tacked on at the beginning and the end, as well as in the middle, whereas the German sentence tends to focus on the centre. At the same time, English copywriters do more with layout and formatting whilst their German counterparts exploit grammatical and syntactical controls to keep the text pressing forward, a subject I touched on in my article for *The Linguist* earlier this year. This rhythm and order level
(= style? avoidance of ugliness?) is difficult - which is why you need a translator - and all the MT botched jobs I have seen are remarkable for that very cruel fact: they stutter so badly that they never take off.

We'll have to leave out the translation of tense here except to say by way of an example

that, particularly in oral narration, German seems to operate on a present/perfect level of interaction whereas English might be said to work more at a perfect/preterite interface. Either way, a structural 1:1 translation of tense is just not possible. But this a truly an enormous topic in itself, one which will baffle monolingual language engineers who probably aren't aware that, according to Michael Halliday, English has 36 tenses. *(Note 11)*

Tolerance of low standards

To be fair, it is not as though the language engineers are completely unaware of some of these difficulties. *(Note 12)* Nevertheless, their strategies for coping require a brief examination for they involve the acceptance of a translation quality - or lack of quality - which no human client would tolerate in the work of a human translator for as much as one second.

The basic strategy is to pre-edit the source language text and, after it is has been through their translation machine, to post-edit the output text. This is the reasoning behind all those demands to “simplify” or “control” or “reduce” the input language — i.e. pare it down so that the machine can handle it, so that there is no room for ambiguity. Big brother is editing you. In other words, for a machine translation to work, the source language has to be rewritten by a SL expert and the output text has to be revised by the computer engineer. Presumably, the services of a human translator have to be invoked anyway to check the post-edit with the pre-edit as well.

Can anybody tell me the point of this lengthy, awkward and expensive exercise? Has it been costed? Remember: one aim of MT was to replace HT and cut costs. It is a profoundly odd aspect of the human condition that human engineers will go to enormous lengths to tolerate mistakes in their machines. Mistakes they would not put up with elsewhere — and certainly not in a human translator. Another related strategy is to do away with the post-edit and *accept* a lowering of quality, sometimes known as the “cheap but cheerful approach”. I have two responses to this. Firstly, I recommend that, the next time such strategists go to the dentist’s for a filling, they ask for two price quotes: one for professional treatment and one for the “cheep and cheerful” approach, the latter made up of a drilling without anaesthetic and sterilised instruments. Secondly, I suggest that, the next time they fly, they should ask for a plane which has been “cheaply but cheerfully” serviced. The choice is theirs. But after the crash I would like to interview them about their attitude to deprofessionalisation.

The ultimate tack is a wait-and-see one: the whole issue is put on the back burner until the “AI complete solution” has been found by which the MT enthusiasts mean that software translation will become a possibility once the human mind has been replicated by a computer. Well, pigs may speak.

And, regrettably, another waste of resources. If only all these highly talented computer engineers and computational linguists would devote their energies to improving current systems. A straightforward e-mail system which functions, for instance, without unMIMEs, STUFFiTs and all the other unspeakable multicasing neologisms. Turn their efforts to producing computers that never go wonky. Or, in word processing, how about making the dictionaries, spell checkers, grammar assistants and thesauri much more powerful than they currently are? I could also do with a number of other tools such as a collocation bank, an idiom lexicon and a halfway efficient filefinder. Plus, whilst Microsoft are at it, a PowerPoint, Access and Outlook which really are *user-friendly*, the most inaccurate word of the late 20th century.

Leave translation to the translation experts. “It yourself again and again”.

Acknowledgements

Thanks to Amy S. Bryant Dipl.-Üb. for the CNET NEWS tip-off and her stern proof-reading skills.

Thanks to Daniel Loehr of Georgetown University Washington who kindly let me have a private copy of his pro-MT paper, a useful round-up of latest developments.

Thanks also to the Essen Translators' Round Table, convened by Shawn Christoph M.Sc. and Abdel Ismail Ph.D., to whom I presented the outline of this article in September 1999. About twenty highly qualified, widely experienced translators attended, representing several language families and various subject areas — including computer assisted translation and translation memory tools. Not one of them disagreed.

NOTES:

Note 1. Quoted in an article for CNET News by Rachel Konrad at http://news.cnet.com/news.

Note 2. Alex Gross in his review entitled *Two German Books About Machine Translation* [Translation Journal: Vol. 3 N° 3 July 1999] downloadable free from the on-line site at http://accurapid.com/journal/09review.htm. Even Alex, a teacher of translation history at New York University, who has written extensively about MT with an open mind, is leaning towards long-term scepticism. His website at http://language.home.sprynet.com has more details of his career and an article humorously attacking the whole concept of transformational-generative grammar: *Forty- Four Reasons Why The Chomskians Are Mistaken*. Here too, the all-important question is asked: Why don't linguists such as Chomsky and Pinker consult translators? As Alex points out, "linguist" used to mean someone who could speak one or more foreign languages (as in the title of this journal); it now basically refers to someone monolingual who theorises about languages.

Note 3. Randomly opening the book at Page 277, I find the title "Pre-editing and the use of simplified writing for MT: an engineer's experience of operating an MT system" by P.J. Pym.

Note 4. Private communication [June 1998].

Note 5. A very good place to start. See J. Andrews in *The Sound of Music* [Rogers & Hammerstein: 1964].

Note 6. See his seminal essay *Politics and the English Language* in: *Inside the Whale and Other Essays* [Penguin 1957: p152]. Both Laswell's wording and final six-part model are curiously similar to Orwell's.

Note 7. This extra information gleaned from the *Encyclopaedia Britannica* Vol. 16 pp 623-4.

Note 8. See Alex Gross's website for a demonstration of Babelfish weakness with translations into Spanish and French plus a hyperlink if you wish to experiment.

Note 9. A similar line is taken by Norman Thomas di Giovanni, one of Jorge Luis Borges's preferred Spanish > English translators, in his review of a new translation of Borges's fictions. See *A Bad Translation* in The Literary Review [February 1999: p19a].

Note 10. See James Wood's (no relation) review of *The Spell by Alan Hollinghurst* in The London Review of Books [16 July 1998: p18 a/b] for a brief mention of another "lyrically English triad on the model of Thomas Gray's 'mute inglorious Milton'" i.e. adjective-adjective-noun "this smooth cluster".

Note 11. See pp153-54 in Halliday: *System and Function in Language* edited by Gunther Kress [OUP: 1976].

Note 12. Can Simultaneous Interpretation Help Machine Translation? pp 213-224 in: David Farwell, Laurie Gerber, Eduard Hovy (editors), *Machine Translation and the Information Soup* [&ldots;] Lecture Notes in Artificial Intelligence Vol. 1529 / Springer-Verlag Berlin Heidelberg 1998.