Regional Survey: M(A)T in North America[1]

Elliott Macklovitch
RALI - DIRO
Université de Montréal
Canada

Abstract

We examine two North American case studies, each of which illustrates a different strat-egy for coming to terms with high-volume, high-quality translation. The first eschews MT in favour of translation memory technology; the second employs a controlled language to simplify the input to an MT system. Both strategies betray a certain dissatisfaction with the current state of machine translation, although neither alternative, it turns out, fully lives up to its expectations.

1 Caveat Emptor

Allow me to begin with a few disclaimers. The first are not my own, but are taken from the Web sites of two well-kown commercial MT developers whose identities need not be revealed here.

> **"A word about language translation software:**
>
> Because human language is complex, results will vary depending on the source text for each translation. Therefore consider your translations as drafts which enable you to quickly and inexpensively handle day-to-day business communications in foreign languages."

I think we can all agree that human language is complex and that this is indeed one of the principal reasons why machine translation is so very difficult. This being said, we surely *do* want our MT results to vary with the different source texts we submit. Just imagine if the contrary were the case! Ah, but perhaps I am being too literal here? Perhaps this caveat is actually an illustration of another of the difficulties that plagues machine translation: the fact that people often don't say or write what they really mean. Here, for instance, what the writer probably meant was that the quality of the MT results will vary with the quality of the source text. In fact, a linked page on this same site states that "the best way to help the quality of your translations is to make sure your source (original) text is well written." But as our citation illustrates, this is easier said than done; and. as we shall see below, a text that is well-written for a machine is not necessarily the same as a well-written text for humans. Human translators understand that it is the author's *intended* meaning which they must translate, and when they complain about a text being poorly written, it is often because they have difficulty reconstituting that meaning from the source text. MT systems, on the other hand, never complain; but that is only because they have as yet no way of realizing that discrepancies may exist between the intended meaning and the literal source text. The latter is all they have to go on.

The following caveat is found among the Frequently Asked Questions on the Web site of another major MT vendor:

> "MT cannot replace a human translator, nor is it intended to.... The two most important things to remember in machine translation is that no automated translation is perfect nor is it intended to replace human translators."

This immediately raises a number of interesting questions, including what MT is intended for, if not to replace human translators. I will skirt this question for the moment and instead step back in order to ask just what is going on here. The very fact that major MT developers feel obliged to include these kinds of messages on their Web sites is, I would maintain, very significant. Quite clearly, these disclaimers represent an attempt to educate the public as to what can and cannot be reasonably expected from current MT technology. That such messages are posted on the developers' Web sites is also significant, since it is the Web (or more generally the Internet) which is responsible for the fact that more people are using MT today than ever before. AltaVista's online translation service, in particular, has greatly increased the general public's exposure to MT. Barely five months after its introduction, according to Yang and Lange (1998), the Babelfish Systran service was already processing over half a million translation requests per day. The great majority of these turn out to be short texts of under five words submitted by Web surfers who only occasionally have recourse to MT. Be that as it may, the Internet is unquestionably revolutionizing our field.

[1] The abbreviation M(A)T stands for both machine translation and machine-aided translation, taken together. The latter is used to refer to various types of translator support tools, including so-called translation memories.

By bringing vast numbers of people into contact with languages they do not fully master, it has created an enormous new demand for translation which cannot possibly be met by traditional translation services. Computer users naturally turn to computer translation for a solution, or "quick fix". The Internet has thus democratized MT. What used to be an arcane science has suddenly become one of the hottest selling items at the corner computer store. MT service providers are now faced with the unenviable job of explaining the current state of the art in MT and NLP to millions of people who have little or no understanding of just how complex natural language is. These disclaimers are valiant attempts at expectation management on a planetary scale.

Which brings me to my own disclaimer. The conference organizers have kindly invited me to present a survey report on the current state of MT and MT-related products in North America. Five years ago, this would have been a much easier task. The MT community was a like a tight-knit family then. One had only to attend one of a small number of conferences – AMTA or TMI, or even the exhibit at the American Translators Association – in order to meet all the major North American MT vendors; and the same was largely true for most of the large-scale industrial or institutional users of machine translation. Alas, this stable, almost cozy situation no longer obtains. Nowadays, the major developers and vendors of commercial MT systems are literally all over the map. It's not so much that there are more of them; on the contrary, there has been a tendency towards mergers and acquisitions in recent years, with a few large players acquiring smaller, specialized MT firms. What seems to be disappearing, however, is the notion of *local* MT supplier, i.e. companies that are situated, say, in North America and limit their market largely to that continent. Like so many other sectors, the translation industry is undergoing globalization. Increasingly, there is only a single worldwide market and on that global stage, all the major players are attempting to provide MT systems that cover, at the very least, the world's principal commercial languages. MT salesmen rarely come knocking on our doors any more. Instead, they post demo versions of their software on the Web, which can be downloaded regardless of where the supplier is located. Online translation of short texts is now available in a matter of seconds and longer texts can be electronically routed to translation services almost anywhere in the world. The upshot of all this is that it is now more difficult to clearly define who is to be included among North American suppliers of MT and MT-related products. (It's a little like asking whether a Toyota assembled in Nashville with parts manufactured just about everywhere counts as a North American car.)

The situation is slightly different, however, when we consider the major *clients* of translation, those companies and agencies that generate tens of millions of words of translation a year. Occasional users of MT may be able to make do with less than perfect translation quality; and the expansion of the Web promises to open up exciting commercial applications for current MT technology. But all this effervescence over MT on the Internet should not lead us to lose sight of the fact that the demand for publication-quality translation is not diminishing; very much to the contrary. To what extent is current MT technology meeting the needs of large-scale clients of translation, those whose volume is such that conventional human translation is simply not a viable option? Thus put, the question is perhaps too vast and vague to allow for a coherent answer. What I propose to do, therefore, is to consider in some detail two North American cases studies, one in the public sector and one in the private sector, in order to examine some of the strategies that are currently being adopted by such organizations in their efforts to reconcile high-quality, high-volume translation with existing MT technology.

2 Two Case Studies

2.1 Translation memory

Several years ago, the Government of Canada decided to entirely revamp its outdated and excessively complex job classification standard. This was an enormous undertaking and a highly sensitive one as well: all the nearly two hundred thousand positions within the various departments and agencies of Canada's Public Service must be defined within this classification system, and employee pay scales are negotiated in terms of its job types and levels. Revamping the general classification standard meant that every civil servant's individual work description, which particularizes a generic job type to the specific duties of a given position within some department or agency. would also have to be rewritten. Once rewritten, moreover, many of these work descriptions would have to be translated into either English or French, in order to comply with Canada's Official Languages Act. The Universal Classification Standard (UCS) project, as it was called, thus created an enormous new demand for translation. Exact figures are difficult to obtain, but according to informed estimates the total number of words to be translated exceeded 45 million. In December 1998, less than six months before the scheduled implementation date of the UCS, Treasury Boars sent out a memorandum to all government departments in which it recognized that the volume of translation generated by the project exceeded the capacity of the government's own internal Translation Bureau and, in fact, that of the entire translation industry in Canada. The memorandum went on to suggest a range of remedial measures that would help prioritize translation requests and encourage the sharing of translated work descriptions between departments. What is interesting for our purposes, however, is that machine translation was never seriously considered as offering a potential solution to the chronic bottlenecks created by the UCS project, neither by Treasury Board, the agency coordinating the project, nor by the federal Translation Bureau, which remains the largest translation service in Canada and which inherited the bulk of the workload associated with the UCS project. Why is this?

Over the years, the Translation Bureau has had considerable experience with MT, both in evaluating various commercial systems and exploiting specialized MT systems, such as Météo and Lexium.[2] The source texts on the UCS project, however, tend to be far more complex than those translated by the latter two systems: sentences are much longer and often include run-on enumerations; there is a much greater variety of syntactic constructions; and the semantic domain is almost completely open-ended. Moreover, the required level of quality for these translations is extremely high, seeing that the text of a work description may be cited in various grievance procedures. Given these constraints and the project's extremely tight deadlines, the Bureau's Director of Technology concluded that there was simply not enough time to adapt a commercial MT system so it would produce translations of the required quality with minimal post-editing. On the other hand. a significant degree of repetition among work descriptions within the same occupational groups was anticipated. For this reason, the Bureau decided to employ a commercial translation memory (TM) system as the two centres it designated to handle the English-to-French and French-to-English translation of all the work descriptions in the UCS project. Novel sentences, i.e. those for which an identical or near-identical match could not be found, would be given to humans to translate, with their translations later being added to the project databases.[3] Now obviously, the cost-effectiveness of this strategy depends on the size of the database being searched. On the UCS project, the Translation Bureau expected that major savings would accrue once the databases reached substantial proportions. In this, they have been slightly disappointed for after nearly a year on the project, the averaage sentence repetition rate is somewhere between 35-40% according to the Director of Technology.

On this particular project, translation memory technology is being used as an *alternative* to machine translation; but this needn't always be the case. In particular, the two technologies can be configured to complement each other, in ways that can be mutually advantageous:

> "Using translation memory in conjunction with machine translation plays to the strengths of both these tools - good quality translation from human translators and the speed of machine translation. Using MT together with translation memory minimizes the impact of the variable quality of MT." (Westfall 1998, pp. 502-503)

If new texts are first filtered through a translation memory, the latter in a sense simplifies the task of machine translation by reducing the number of sentences the MT system has to translate and post-editors subsequently have to revise. There are, of course, other ways of easing the burden on machine translation, including the use of a controlled language to simplify the source texts to be translated. This is the focus of our second case study.

2.2 Controlled language

When C.K. Ogden first proposed the idea of Basic English in the 1930's, the aim was not to facilitate translation, but rather to promote a small subset of English as the one international medium for science and commerce. And when large multinational corporations later took up the idea of controlled language (CL), the goal remained essentially the same, i.e. to publish technical documentation in a form of English so simplified that it would be easily comprehensible to native speakers of other languages whose command of English was often quite limited. With time, however, this idea has become politically less acceptable; now clients in the non-English speaking world increasingly expect product documentation in their own language. But if controlled languages can no longer be used as a substitute for translation, perhaps the techniques employed to maximize their comprehensibility – the limitation of vocabulary and the restrictions on syntax and style – can still be employed to simplify the source texts that have to be translated by MT systems. After all, MT systems, like non-native speakers, also have a limited understanding of the source text; and the sublanguage approach to MT has shown that simplified input can improve the quality of MT output, thereby making translation more cost-effective. CL promises to achieve the same effect; only instead of being based on a naturally occurring sublanguage (like weather bulletins), the elimination of ambiguity is achieved by artificially imposing a limited vocabulary and range of syntactic constructions. In a well-defined CL, according to Heald & Zajac (1998), there should ideally be a one-to-one correspondence between each term of the vocabulary and a single concept; and there should be no constructions that allow for multiple parses, or ambiguous syntactic analyses.

The rub, however, is that it is not at all obvious that such ideal CLs *can* be defined, at least not for today's complex technical domains. In Caterpillar's well-publicized experience with CL, for example, it has not proven possible to eliminate or paraphrase all polysemous terms, so that interactive disambiguation routines have had to be incorporated into their conformity checker; cf. Kamprath et al. (1998), p. 7. The possibility of eliminating, or even automatically detecting all syntactic ambiguities would seem to be even more difficult.[4] What is more, there may be a fundamental tension, or opposition, between the two stated goals of CL: improving the comprehensibility of technical texts for human readers and reducing the degree of ambiguity for machine translation. In particular, certain recommendations which aim to eliminate constructions that machines cannot reliably parse – e.g. avoid the use of gerunds in sentence initial position; or make sure all pronouns have a clear reference – can result in

[2] For more on Météo, see Grimaila (1992); on Lexium. see Chandioux & Grimaila (1996).

[3] Although they were provided with the output of the batch memory search, most of the translators did not actually use the TM system interactively to produce their translations.

[4] Although Bernth (1998) reports on some very impressive work in just this direction.

texts that are highly unnatural and otherwise difficult for humans to read.

However, these quasi-theoretical objections have not discouraged dozens of major corporations from turning to CL as a means of improving the cost-effectiveness of their MT operations. One of these is a Canadian-based multinational which, for the purposes of this presentation, would prefer to remain anonymous. At its Montreal office alone, company X, as we shall call it, translates between 100 and 140 thousand pages a year, approximately 75% of which is technical documentation. The company began to study the introduction of a CL in 1995. It has defined its own variant of controlled English in which all its technical source documents are now drafted before being translated into multiple target languages. There has been a major effort to standardize terminology and to make this terminology available to all the company's offices around the world, as well as to the MT suppliers that contract with the company. Unlike Caterpillar, however, company X has not developed a specialized MT system in parallel with its variant of controlled English, but sends the texts that are drafted in CL to an outside translation service which uses a commercial MT system.[5] And whereas Caterpillar also had its CL checker developed by the same group that developed its AMT system, company X has contracted with a third party to develop the interactive software designed to ensure the texts drafted by its technical writers formally comply with its CL.

How has the company fared thus far in its experiment with CL? It is difficult to say – and not just because the tentative conclusions are politically explosive. One of the constants that seems to emerge from the literature on MT that is being used in conjunction with CL is how difficult it is to evaluate the cost-effectiveness of combining these technologies. In her slide presentation at the last MT Summit, Kamprath (1997) notes that that the "AMT output cannot be evaluated independently of [the] quality of CTE input."[6] This is certainly true, but it is also true of any mode of translation, whether or not CL is being used to simplify the input; cf. the discussion of the first disclaimer in section 1 above. By the same token, one could ask if the cost-effectiveness of AMT at Caterpillar can be calculated independently of the cost of CTE development. And just what are the costs of the combined CL + MT chain to be compared to: human translators working only with word processors on texts that haven't been simplified? The answers to these questions aren't at all obvious. At company X, in any case, the results obtained thus far have been somewhat disappointing. Not only has the combination of MT used in conjunction with CL and an automated compliance checker failed to deliver the anticipated cost savings; but the quality of the translations produced by the outside MT provider on the first source texts drafted in CL was actually poorer than that produced by the MT system on similar texts in standard language. (The system's analyser was apparently not prepared for the "simpler" CL input.) More recently there have been some modest gains in productivity, on the order of 20-25% in comparison with the cost of human translation. But the principal advantage in sending texts out to MT, according to the manager of the company's Canadian translation service, is the reduced turnaround times. The quality may not be fully comparable to that of human translation, but for certain types of documentation, the volumes and deadlines are such that there really is no alternative. For commercial documents and even users manuals, on the other hand, the company has found it best to draft them in standard English and translate them internally. Another recurrent theme in the CL literature is how difficult it can be to successfully introduce CL into a large corporation and to bring technical writers to comply with it. Kamprath et al. (1998) insist on the considerable investment that Caterpillar has made in promoting its CTE internally and in providing ongoing training to its technical writers. But even this is not enough. As Wojcik & Hoard (1996) point out: "Because the style, grammar and vocabulary restrictions of a CL standard are complex, it is nearly impossible to produce good, consistent documents that comply with any CL by manual writing and editing methods.... Automated checkers [must be developed] that help writers conform to the standard easily and effectively." This is certainly a challenging problem, especially if we expect the automated checkers not only to flag non-compliance, but also to suggest compliant paraphrases. To do this properly requires nothing less than a complete computational grammar of the source language. In the absence of an "omniscient" checker, ambiguities will persist in the CL source text which may well manifest themselves as errors in the MT output. Hence, one shouldn't image that recourse to a CL will automatically eliminate the need for MT post-editing.

3 M(A)T Providers

A survey of this sort would be incomplete if it did not attempt to sound the major providers of M(A)T technology on *their* views as to the major factors currently affecting the field. To this end, I sent out a short questionnaire to the better-known North American-based M(A)T companies, asking them what they felt had been the most significant developments since the last MT Summit. Among those who replied, there was near unanimity: once again, it's the Internet. For Systran, the Internet has not only fundamentally changed the way in which MT is delivered; it has revolutionized the whole field. The Babelfish site has obviously provided Systran with enormous exposure, significantly increasing direct online sales of software and providing interesting leads for new applications of MT, according to Reba Rosenbluth. For Scott Bennett

[5] This is true for the English-to-French translation which is done in Canada. Translation to other targets is done abroad and I'm not sure whether the MT systems involved are exploited in-house or not.

[6] AMT is the abbreviation for Caterpillar's MT system and CTE stands for Caterpillar Technical English, the company's own CL.

of Logos Corporation, the Internet has helped resolve the thorny question of multiple platforms. Like other MT providers Logos now offers its system in a client-server configuration that is TCP/IP compatible; translators no longer need to be in proximity to a particular machine, since the Internet (or an intranet) is used as the preferred platform of delivery.

Another interesting development that emerged in the answers to my mini-questionnaire is that the majors now tend to be offering not just translation software, but something more akin to a one-stop, full service approach to translation. Who would have predicted five years ago that companies like Logos and Lernout and Hauspie would now be providing services that include customized consulting, glossary building and standard human translation, in addition to MT and MT post-editing. Personally, I have long been of the opinion that if commercial MT systems performed as well as their vendors claimed they did, then MT companies should be able to wade into the translation service business and underbid the unautomated competition. The proof of the pudding, in other words, should be in the pricing.[7] This is in effect what these companies are now doing by exploiting their own systems internally for certain customers. As Scott Bennett puts it, "The MT helps sell the service and vice versa."

Turning now to machine-*aided* translation products, there seems to be a growing consensus on the usefulness of combining MT and translation memory, as we have seen above. Some MT companies have even developed their own translation memory software, while others have collaborated with the best-known TM providers to ensure easy access to their MT system. It may be altogether fortuitous, but none of the leading developers of TM software are based in North America, although Trados now has two sales offices in the United States, one on the east coast and one on the west. In fact, this company's rapid expansion has been extremely impressive and very encouraging for someone like myself, who believes that the most promising avenue for high-quality translation lies with translator support tools. Founded fifteen years ago by two IBM programmers, Trados has sold, as of last year, over ten thousand copies of its Translator's Workbench.

On the other hand, there are some indications suggesting that the manner in which repetitions processing is handled within current TM systems may be approaching its limits. To begin with, existing TM systems are all essentially sentence-based, i.e. they seek in a bi-textual database an identical, or near identical string match for each new sentence to be translated. Now those of us who were raised in the generative school of linguistics will certainly recall Chomsky's arguments on the creativity of the language faculty, which go all the way back to *Syntactic Structures*[8]. A priori we should be skeptical of the likelihood that numerous sentences in some new text will reappear identically in a database of previous sentences, no matter how large that database. In fact, it is principally in the very particular context of updates that such verbatim repetitions tend to occur with significant frequency. Even on such projects as the Universal Classification Standard, where the same limited occupational grid was being applied to different departments, the rate of repetitions, as we saw above, was disappointingly low. How productive can such an approach hope to be, then, in the great majority of translation situations that do not involve updates to pre-existing documents?

4 Prospects for the future

The limitations of current TM products have led some people to suggest that the next generation of translation memories will have go beyond or, more precisely, below the level of the sentence:

> "There is a clear need for matching phrases (as well as sentences) in translation memories and for composing retrieved fragments into coherent sentences. Current commercial translation memories are sentence-based and are restricted essentially to the presentation of potential examples of translations, which translators must themselves adapt for incorporation. Searching for sentence fragments (phrases) is clearly beyond current statistical methods; it would require some linguistic analysis... Likewise, the automatic restructuring of selected fragments into well-formed sentences demands a level of linguistic knowledge not yet available in commercial workstations." (Hutchins, to appear)

I agree with John Hutchins' assertion that more sophisticated analysis techniques will be required to identify linguistically coherent phrases below the level of the sentence. But even if the next generation of TM tools manages to do this, and to automatically align as well the corresponding target phrases for subsequent storage in the database[9], it is not a priori obvious that the results will be significantly more useful than those of current TMs. Technical terms are by definition stable units of translation; and full sentences are also relatively stable translation units, as current TM systems have shown. The phrase, on the other hand, is less so. Its translation, like that of the word, tends to vary according to the larger context, making the strategy of one unit : one translation more difficult to retain. As for the possibility of automatically composing correctly identified fragments into a well-formed target sentence, this, as I understand it, is the goal of the Example-based MT research paradigm. As far as I know, it remains a challenging and largely unresolved research problem.

I have expressed certain reservations about the possibility of extending the impact of repetitions processing technology. I would not want this to be inter-

[7] In fairness. it should be mentioned that ALPS has long used its own MAT software within the company's world-wide network of service bureaus.

[8] Cf. Chomsky (1957), pp. 15-17.

[9] Sentence-based TM systems can unobtrusively get the translator to pair up corresponding source and target units. To do so for sub-sentence units, however, will be much more difficult.

preted, however, as a general pessimism regarding the potential of all types translation support tools. On the contrary, the encouraging conclusion is that *other* types of machine-aided solutions will soon be required in order to assist human translators in the production of high-quality non-repetitive translations. There is exciting research being conducted in several North American universities, including my own, which has precisely this objective. Here, I can do no more than cite a few references: Russell (to appear); Macklovitch and Hannan (1998); Foster et al. (1997); Zajac (1996).

Returning now to the North American providers of machine translation software, how do they see the future? Those who replied to my questionnaire seemed quite optimistic. Their business has been growing over the last few years and they see promising potential in the extension of their systems to new language pairs, based on market demand. At the same time, commercial MT developers deplore the fact that too little money is available for research. In the view of Systran's Laurie Gerber, MT providers find themselves caught up in something of a vicious circle. With little money for research, they cannot make significant improvements to their product; and until they can offer a better product, they're stuck with a small number of successful applications that are not hugely profitable.

Commercial MT developers are not the only ones, of course, to complain about the lack of funding for MT research. In my own presentation at the last MT Summit, I lamented the fact that in Canada, the government seemed to be abandoning its long-standing commitment to fund both research and development in machine-aided translation. And in the United States, the MT research situation has been only slightly better: research groups at the five or six American universities with a well-established tradition in MT have been suffering due to a lack of funding.[10] Now, however, there are some indications that the tide may be turning, at least in the U.S., where DARPA has recently announced a new program called TIDES (for Translingual Information Detection, Extraction and Summarization) that features machine translation as one its three major topic areas. (If only the same were true in Canada!) This DARPA initiative is particularly promising, especially when we recall the tremendous boost to MT research that resulted from the Agency's previous involvement in MT. Hence, there seems to be grounds for a certain amount of prudent optimism regarding the future of MT research. Why, even Microsoft is mentioning machine translation in the job offers it has recently been advertising within its Natural Language Group.

5 Conclusion

In this paper, I have examined two case studies, each of which illustrates a different strategy for coping with large volumes of publication-quality translation. The first employs translation memory as an alternative to MT, and the second, controlled language as a way of simplifying the input (and the task) of machine translation. In a sense, both strategies betray a certain dissatisfaction with the current state of machine translation for the purposes of high-quality translation. If MT were entirely up to scratch, in other words, the two clients involved would probably abandon these strategies: the first might at least consider MT as an alternative to human translation, and the second might not go to the trouble of introducing a controlled language. I would hope at future MT Summits to be able to present a more optimistic picture of the current state of the art, in North America and elsewhere. This will only happen, however, if there is a significant increase in funding for MT research. For the first time in many years, there is a real possibility that this may happen. Hence, prospects for the future look relatively good. The market is crying out for new solutions, researchers are champing at the bit, and the field – particularly in the case of machine-aided translation – is benignly. open.

[10] That these groups have still managed to conduct top-rate research is entirely to their credit. Cf. among others: Hogan and Frederking (1998); Knight and Al-Onaizon (1998); Palmer et al. (1998).

6 Acknowledgments

I wish to thank the following people who kindly responded to my questionnaire or otherwise provided information and input that helped me prepare this paper: Scott Bennett, Patrick Drouin, Laurie Gerber, Michael McCord, Roger Racine, Reba Rosenbluth, Graham Russell, Edith Westfall and Remi Zajac. None, of course, are responsible for any errors of fact or interpretation the paper may contain.

7 References

Bernth, A. (1998) "Easy English: Addressing Structural Ambiguity", in Proceedings of AMTA-98, Langhome, PA, pp. 164-173.

Chandioux, J. and Grimaila, A. (1996) "Lexium", in Proceedings of AMTA-96, Montreal, Canada, pp.261-263.

Chomsky, N. (1957) Syntactic Structures. Mouton, The Hague.

Foster G., Isabelle P., Plamondon P. (1997) "Target-Text Mediated Interactive Machine Translation", Machine Translation, 12:1-2, 175-194.

Grimaila, A. (1992) "Made to measure solutions", in J. Newton (ed.) Computers in Translation: A Practical Appraisal, Routledge, London, pp.33-45.

Heald, I. and Zajac, R. (1998) "Controlled Languages", tutorial notes, COLING-ACL'98, Université de Montréal, Montreal, Canada.

Hogan, C. and Frederking. R. (1998) "An Evaluation of the Multi-engine MT architecture" in Proceedings of AMTA-98, Langhorne, PA, pp. 113-123.

Hutchins, J. (to appear) "The origins of the translator's workstation". in Machine Translation.

Kamprath, C. (1997) "Using MT in a Corporate Setting", slide presentation at MT Summit VI, San Diego, CA.

Kamprath, C. et al. (1998) "Controlled Language for Multilingual Document Production: Experience with Caterpillar Technical English", in Proceedings of the Second International Workshop on Controlled Language Applications (CLAW '98) Pittsburgh, PA.

Knight, K. and Al-Onaizan, Y. (1998) "Translation with Finite-State Devices", in Proceedings of AMTA-98, Langhorne, PA, pp. 421-437.

Macklovitch, E. (1997) "MT R&D in Canada" in Proceedings of MT Summit VI, San Diego, CA, pp.200-210

Macklovitch, E. and Hannan, M-L. (1998) "Line'Em Up: Advances in Alignment Technology and their Impact on Translation Support Tools", in Machine Translation, vol. 13, no. 1, pp. 41-57.

Palmer, M. et al. (1998) "Rapid Prototyping of Domain-Specific Machine Translation Systems", in Proceedings of AMTA-98, Langhorne, PA, pp. 94-102.

Russell, G. (to appear) "Errors of Omission in Translation", to appear in the Proceedings of TMI-99, Chester, U.K. August 23-25, 1999.

Westfall, Edith (1998) "Integrating Tools with the Translation Process", in Proceedings of AMTA-98, Langhorne, PA, pp. 501-505.

Wojcik, R. and Hoard, J. (1996) "Controlled Languages in Industry", in R. Cole, (ed.) Survey of the State of the Art in Human Language Technology, http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html .

Yang, Jin and Lange, Elke D. (1998) "SYSTRAN on AltVista: A User Study on Real-Time Machine Translation on the Internet", in Proceedings of AMTA-98, Langhorne, PA, pp. 273-285.

Zajac, R. (1996) "Towards a Multilingual Analyst's Workstation: Temple", in Proceedings of AMTA-96, Montreal, Canada, pp. 280-284.