Gouvernement du Canada
Ministère des Communications

Government of Canada
Department of Communications

Le Centre canadien de recherche sur l'informatisation du travail
Canadian Workplace Automation Research Centre

PTT - 3 :

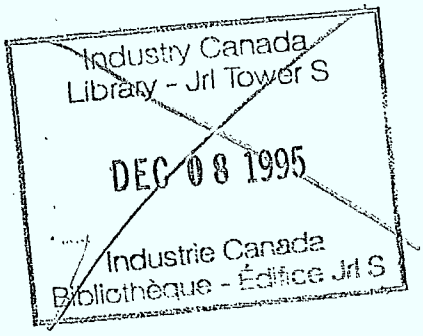
A third version of the CWARC's workstation for translators

Industry Canada

QUEEN
P
309
.M335
1993
c.2

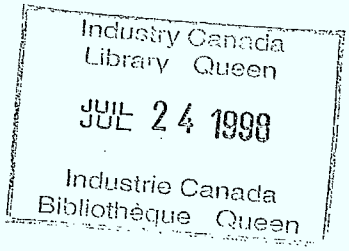
Canada

Queen
JOUR.
P
309
.M335
1993
c.2



PTT - 3 :

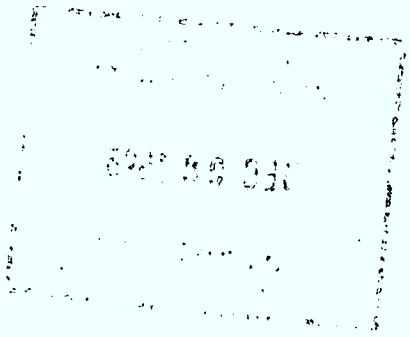
A third version of the CWARC's workstation for translators



Elliott Macklovitch

Canadian Workplace Automation Research Center

March 1993



Cat. No: Co 28 - 1/83 - 1993E

ISBN: 0-662-20504-9

The views expressed in this report are those of the author only.

Ce rapport est aussi disponible en français.

P
309
. M335
1993
C.2

1. Background

The first published proposal for a translator's workstation is generally attributed to Martin Kay of Xerox PARC, in an article entitled "The Proper Place of Men and Machines in Language Translation"; see Kay [4] in the References, Section 5 below. Kay's proposal for a translator's amanuensis, as he called it, is advanced as an alternative to classical machine translation, which is criticized as being overly ambitious, given the limits of our current understanding of how human language works. Rather than attempting to have computers automate the most complex phase of the translation process, it would be more reasonable for us, Kay argues, to take a modest, bottom up approach and offer the human translator an integrated set of computerized aids that would help him perform the routine, peripheral aspects of his job more efficiently. Kay makes some concrete suggestions on how he thinks such a workstation should be assembled; but these details are less important than the overall thrust of his incremental approach to the question of how machines should be used in language translation. "First they will take over functions not essentially related to translation. Then, little by little, they will approach translation itself. The keynote will be modesty. At each stage, we will only do what we know we can do reliably." (Kay 1980, p.11)

1.1 PTT-1

The computer-assisted translation group at the Canadian Workplace Automation Research Centre (CWARC) has been engaged in the development of a translator's workstation, or PTT[1] as it is commonly called here, since 1985. From the outset of the project, the CWARC sought to ensure the cooperation and support of the Translation Bureau of the Department of the Secretary of State, the largest translation service in the country. Early work on the project focussed essentially on defining functional specifications for a translator's workstation: a firm of private sector consultants proposed a first set of specifications; these were then modified and completed with the help of a questionnaire which the CWARC administered to all the translators in the Secretary of State.[2] Based on these functional specifications, the CWARC assembled, between 1988 and 1989, its first operational prototype of a translator's workstation, known as the PTT-1. It is important to note that one of the fundamental goals of the project at this stage was to ascertain to what extent these functional specifications, developed in consultation with the eventual end-users of the workstation, could be satisfied by using inexpensive hardware and commercially available, off-the-shelf software, most of which was not specifically designed for translators. Thus, this first version of the

1. PTT is a French abbreviation that stands for "poste de travail du traducteur".
2. For more details on the early stages of the PTT project, see Macklovitch & Isabelle [10].

workstation was implemented on a PC-AT, equipped with a full-page monitor and a memory extension board, and included software applications for the following tasks:[3]

- bilingual word processing (*WordPerfect*)
- personal or local terminology management (*Mercury/Termex*)
- identifying updates or modifications in two word processing files (*CompareRite*)
- French verb conjugation (*Seconde Mémoire*)
- producing a concordance of a given text, as well as a frequency count of each form in that text, and allowing the user to recall one or more of those forms in context (*TextSearch*)
- converting word processing files from one format to another (*Software Bridge*)
- allowing simultaneous access to all these applications by means of a windowing interface or integrator (*DESQview*)

Between November 1989 and November 1990, this prototype was the object of an operational field trial at two Translation Bureau sites, one in Ottawa and one in Montreal, involving thirteen government translators and revisers. A formal evaluation of this trial was conducted by a small team of specialists, which focussed on the user-friendliness of the PTT's interface, the functionality of its components, and the socio-organizational impact of its introduction into the operational environment at the Translation Bureau.[4]

1.2 PTT-1.1

There were several technical problems with PTT-1, which became apparent even before this field trial ended. The most important of these was an insufficient amount of conventional memory. Despite its memory extension board, the 286-based PC did not allow for enough conventional memory to run the querying software for the CD-ROM version of *Termium*, the Secretary of State's highly reputed terminology bank; or for *Termex*, the glossary management program, to run within the same *DESQview* window as the word processor, thereby forfeiting Termex's attractive cut and paste feature. In July 1990, therefore, Translation Bureau management decided that the participants in the trial would receive an upgraded version of the workstation, called PTT-1.1,

3. Macklovitch [6] provides more details on the constituent components of PTT-1, as well as some justification for the selection of these particular products.
4. The final results of this evaluation are given in Lavoie & Lapointe [5].

which was implemented on a 386-based PC compatible. The adoption of this new hardware platform, in conjunction with *DESQview's* expanded memory manager, was sufficient to resolve both the aforementioned problems. What's more, each new workstation was equipped with its own CD-ROM player, so that translators now had direct access to *Termium* within a *DESQview* window.[5]

1.3 PTT-2

In August 1991, the CWARC delivered to the Secretary of State the technical specifications for a new version of the workstation, called PTT-2. Perhaps the simplest way of defining PTT-2 is to say that it is a networked version of PTT-1.1, augmented with the following new functions: a document management component, plus shared access to lexical and terminological databases over the network. There are many good reasons for wanting to link the individual workstations of a group of translators into a local area network. Generally speaking, a network permits users to share resources in a more rational and cost-effective way than is otherwise possible on stand-alone stations. For example, both PTT-1.1 and PTT-2 provide translators with direct access to *Termium* on CD-ROM; but where the former requires that each workstation be equipped with its own CD player, the latter provides shared access to a smaller number of compact disks by hooking up the players to a central server.[6] As for the document management module, the basic idea is to allow translators to query a bank of previously translated texts, so that they can quickly locate documents that will be helpful for the translation of a new text. For this component of the workstation, we adopted a full-text retrieval system in preference to a standard relational database.[7] The selected product impressed us with its speed and ability to consistently retrieve the most relevant documents among the first two or three it displayed, thereby allaying our fear that such a system would result in an unwieldy number of irrelevant hits. The system also allows the user to introduce a number of fields at the beginning of a text, so that useful information can be added to the database which does not generally appear within the text itself, such as the name of the translator and the reviser, the nature of the text, the date it was delivered, etc.

---

5. Other changes were implemented with PTT-1.1: in particular, *The Genius* full-page monitor was abandoned in favour of a standard VGA colour screen. See Macklovitch [7] for details.
6. PTT-2 uses the same strategy to provide shared access to *Le Robert électronique* on CD-ROM. For those not familiar with it, *Le Robert électronique* is a comprehensive French dictionary, which also offers a verb conjugation program and the ability to search citations by author and/or title.
7. The software we selected is Ardilog's *Naturel*. For more details on this and the other components of the PTT-2, see Macklovitch [8].

The PTT-2 is currently undergoing an operational field trial at the same two Translation Bureau sites at which its predecessors were tested. An evaluation report on the trial is scheduled for the end of March 1993.

2. PTT-3 : What's new?

March 31, 1993 will also mark the end of the current Memorandum of Understanding between the CWARC and the Secretary of State, under which the two departments have cooperated on the PTT project. At that time, the CWARC will deliver to the Secretary of State the technical specifications for a new version of the workstation, called PTT-3. In line with the incremental approach that we have adopted since the beginning of the project, this version of the workstation will include only two significant changes in relation to its predecessors: the *DESQview* integrator will be replaced by *Windows 3.x*, a graphical user interface; and an external e-mail component will be added to the PTT's suite of programs. The following sections focus more closely on each of these additions.

2.1 A graphical user interface

Given the widespread popularity of Microsoft *Windows*, it may appear unnecessary to have to justify the change to a graphical user interface; what may appear questionable to-day is our original decision to employ *DESQview* rather than *Windows* in the earlier versions of the PTT. It should be recalled, however, that the programs that make up the workstation are all text-based applications, and that until recently none were available in *Windows* versions. Within the PTT, the main function of the integrator is to facilitate simultaneous access to the programs, i.e. the switching from one application to another.[8] For this purpose, a simple integrator like *DESQview,* with its pull-down menus, is no less efficient with *DOS*-based programs than a graphical interface like *Windows*. The latter, moreover, requires a more powerful machine, with more random-access memory; and as mentioned above, one of the constraints on earlier versions of the PTT was the limited amount of memory. For all these reasons, *DESQview* was a reasonable choice as the integrator for PTT-1 and 2. Indeed, the evaluation report of Lavoie & Lapointe [5] shows that translators were quite satisfied with *DESQview*, even if they didn't take full advantage of some of its more advanced features (like the program macros).

8. This is not the integrator's only function, however. As we shall see below, it should also facilitate the transfer of blocks of text between applications.

Why then propose a change to *Windows* in PTT-3? Another of the things that the aforementioned evaluation revealed was that translators had some difficulty mastering all the programs included in the PTT, even though they had received specialized training at the beginning of the trial. One of the factors contributing to this problem, we are convinced, is the lack of uniformity among *DOS* applications. One of the drawbacks of our off-the-shelf approach is that it may result in a disparate collection of components, many of which were never intended to cohabitate. As anyone familiar with the *DOS* world knows, different programs often employ different command sequences to effect similar functions, making it difficult for the user to retain these commands. *DESQview* alleviates this problem to some degree, by providing what has been called *access integration*. That is, from the *DESQview* shell, the user can select the applications he or she wants to open by simply clicking on a menu list; he can then switch between application windows, or transfer text from one window to another, all with relative facility. Beyond this, however, *DESQview* has little more to offer. Once an application has been opened, the user is on his own: he must know how to operate each separate program and remember its distinct keyboard or menu commands. The principal advantage of *Windows*, on the other hand, is the coherence it imposes on the applications that are specifically designed for it. In other words, *Windows* does for the PC what the Macintosh's operating system does for that machine: it forces software developers to conform to a standard set of menuing conventions when defining their program's repertoire of commands. And this, we are confident, should definitely lessen the burden of learning and mastering new applications for the user.

At the time of this writing, however, the only component of our PTT that has been released in *Windows* version is *WordPerfect*.[9] A *Windows* version of *Termium* is reportedly being prepared for 1993-94; but as far as we know, the developers of *Le Robert électronique*, *CompareRite*, or *Naturel* have not yet announced a release date for a *Windows* version of their products. Apparently, there is growing pressure on all software developers – even those of small "linguistic" applications with a limited distribution – to bring out a *Windows* version if they want to retain their markets. Thus, we are confident that in the mid-term it will be possible to offer a fully coherent version of the PTT, made up of applications all designed to take full advantage of the graphical user interface. Even if the developers of the above-mentioned programs fail to furnish a *Windows* version, we should be able to find other programs which are functionally equivalent and fully *Windows* compatible.[10]

---

9. Not counting, of course, *Windows*' own *Program Manager*, which subsumes most of the frequent *DOS* commands, or the other minor utility programs that come with *Windows*.
10. For example, we have just begun to examine a *Windows* version of MultiTerm, a glossary management program with many attractive features. We have, in addition, tested two grammar and style checkers that are available in *Windows* version, but neither appeared to us sufficiently reliable for use by proferssional translators.

Still, the PTT-3 that will be delivered to the Secretary of State in March 1993 will necessarily be a hybrid, composed of one or two *Windows* applications, with most other programs running under *DOS* (albeit in a *Windows DOS* window). While this does allow for some of the advantages of the graphical user interface, e.g. reducing a *DOS* application to an icon and restarting it by simply clicking, for certain applications, it may actually result in a slight loss of functionality when compared to the *DESQview* interface. For example, *Termex*, the glossary management tool, is a memory-resident program and one which translators make extensive use of. Within the PTT-2, it runs in the same *DESQview* window as *WordPerfect*. When the translator hits a hot key, *Termex's* pop-up menu appears over the word processor, and the translator can automatically insert a word into the *Termex* search field without having to retype it. Conversely, a relevant record has been retrieved, information can be copied off any *Termex* field and rapidly pasted into the word processing document. Within the PTT-3, on the other hand, TSR programs like *Termex* cannot run within the same window as the word processor. Accessing the glossary management tool therefore involves switching to another *DOS* window; and the process of copying text off the glossary record and pasting it into the word processor requires so many steps that it is really only worthwhile for large blocks of text.

This situation should be contrasted with the full *application integration* that is possible between two *Windows*-based programs. I will illustrate a similar type of consultation using *Word for Windows* and *InContext*, an electronic dictionary module that the CWARC helped develop, but which is not yet being marketed. *InContext* only runs under *Windows* and currently provides access to electronic versions of the following dictionaries: the *Gage Canadian English Dictionary*, *le Multidictionnaire des difficultés de la langue française*, and the *Hayakawa Thesaurus of English*. It also includes a macro command which integrates it as an option within *Word's* Utilities menu. Suppose that the translator has loaded his source text into the word processor, and wants to look up the definition of a given term. To do so, he simply selects the term with the mouse, opens the *Word* Utilities menu and chooses the *InContext* option (see Figure 1 on the next page). A pop-up menu immediately appears which asks him to choose the dictionary or dictionaries in which the term is to be searched; suppose he selects only the *Gage*. A second pop-up screen appears, in which the term highlighted in the word processing window has been automatically inserted into *InContext's* search field, and the dictionary definition displayed (see Figure 2). If the translator wants to copy any portion of this definition into his word processing document, he need only select it with the mouse, click on the word processing window, and pull the text segment off the *Windows* clipboard via the Edit menu or a keyboard shortcut. What makes this kind of consultation so attractive is that *InContext* offers one simple interface to several lexical databases.[11] Moreover, being a standard *Windows* application, *InContext* interacts with other *Windows* applications in a

perfectly standard manner. Once you know one *Windows* application, the others are so easy to learn and use.

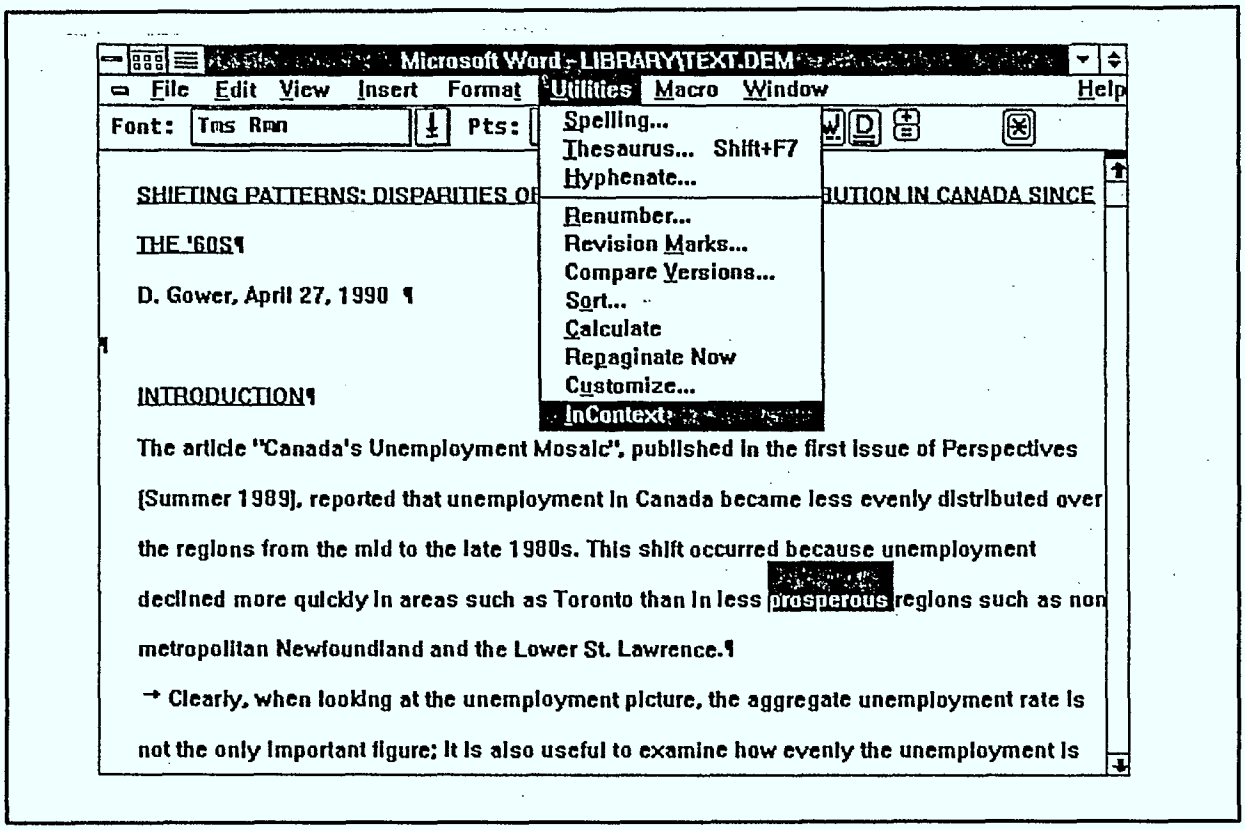


Figure 1

11. Unlike the situation with lexical databases on CD-ROMs, for example, where each CD comes with its own distinct interface.

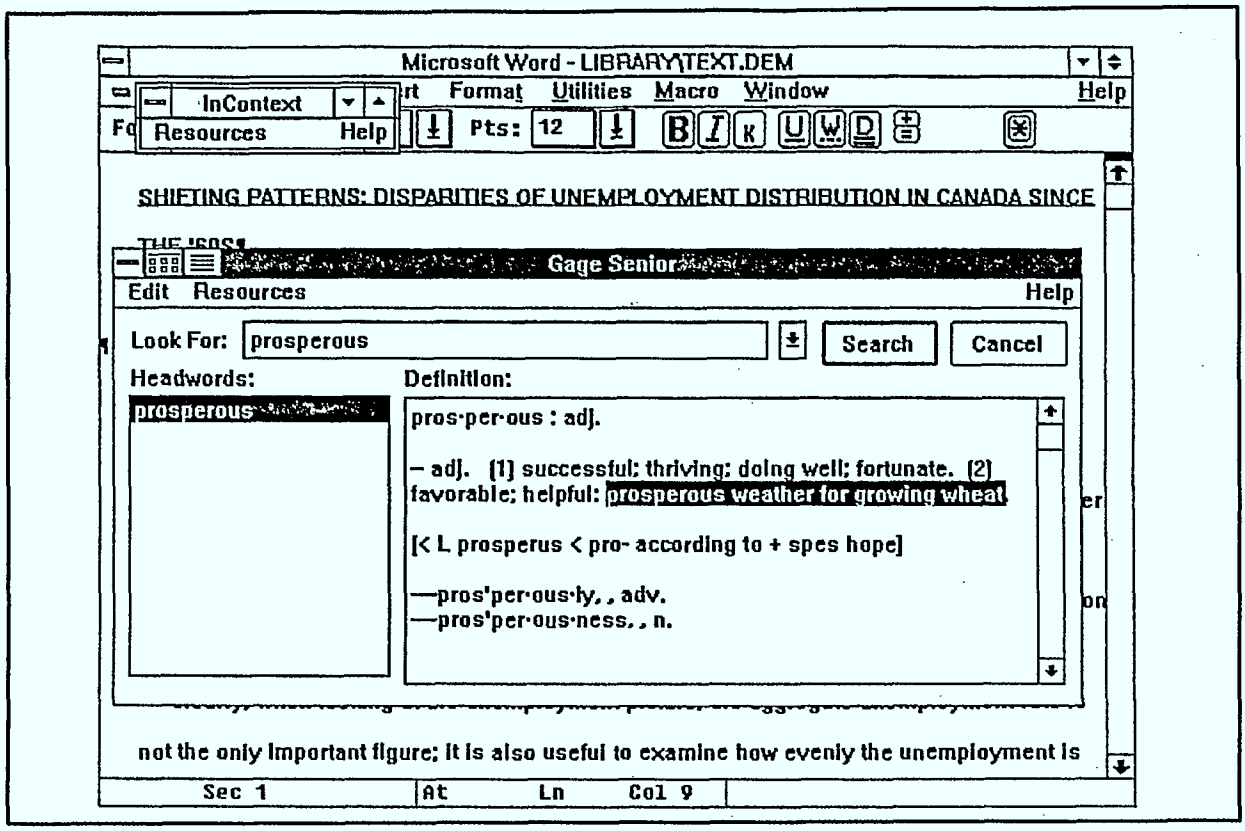


Figure 2

2.2 External E-mail

The Translation Bureau receives the great majority of the texts it handles via regular or internal mail, or by messenger service. Only a small proportion of the texts that make up the Bureau's workload are transmitted electronically, and of these many more arrive by fax than by modem. Most all the texts now drafted in the Public Service are prepared on a word processor, and a growing number of clients are sending the Bureau a diskette, occasionally containing the source text, but more often a blank diskette on which the translation is be returned. While most translation sections are equipped with a PC-linked modem, this is not currently the preferred means of transmitting source and translated texts between the Bureau and its clients.

If electronic transmission were more widely used, there is little doubt that it would reduce the delivery times required for source texts and translations. Since there have been, to our knowledge, no systematic studies of the underutilization of text transmission by modem, we can only speculate as to its causes. People are often heard to complain about how complicated it is to send or receive texts via modem; and how difficult or bothersome it is to get the two machines – and parties – properly synchronized. One of our goals for the PTT-3 is to show that it is possible to introduce a form of electronic text transmission that would not be subject to these kinds of objections, and

would help improve the service the Bureau could offer its clients by reducing delivery times. A second, no less important reason for favouring the introduction of this technology would be to ensure the availability of a greater proportion of source texts in electronic form.[12] For the translator, this would increase the level of support which the workstation could offer; for example, it would allow him to consult *InContext's* electronic dictionaries in the manner described in the previous section. It would also facilitate the systematic archiving of source texts and their translations, which could then be exploited by database querying programs like the one included in the PTT-2, or by the new types of bi-textual tools currently under development at the CWARC.[13]

Technically speaking, the electronic transmission of texts between two computers that may be located arbitrarily far apart is a rather complex question that allows for many possible solutions. The CWARC did undertake a detailed study of these alternatives,[14] and while this is not the place to reproduce its contents, we will mention several of the criteria that helped us make our choice. The selected technology should be simple to learn and use by people who are not computer experts. Furthermore, it should not be too expensive to acquire and maintain, either for the Translation Bureau or for its clients. Finally, the new technology should be perceived as offering sufficientadvantages over the current means of text delivery, so that the Bureau's clients would be willing to introduce it and end users would accept to learn it.

As mentioned above, there are, a priori, several ways in which electronic text transmission could be implemented within a service organization like the Translation Bureau. These include the use of standard telecommunications software, with the possibility of automating several parts of the transmission or reception procedure; using the Canadian government's own electronic messaging and document exchange service, GEMDES;[15] or subscribing to other commercial networks like Compuserve. One particularly natural solution which appealed to us was to extend the internal e-mail facilities already offered translators on the *PowerLan* networks of the two PTT-2 sites at the Translation Bureau, so that they could be used externally, to transmit texts between the Bureau and its various client departments. The e-mail system that comes bundled with *PowerLan* is Lotus' *cc:Mail.* Translators at the PTT trial sites currently use it strictly for internal communication, e.g.

12. By electronic form, we mean one which yields an electronic text file that can be directly processed by a computer. This excludes fax, which generally produces a low quality paper copy.
13. We mention some of these bi-textual aids in Section 3 below; but for a detailed discussion of these and the concept of bi-text on which they are based, see Isabelle [3].
14. See Viau [13].
15. Since GEMDES is available to all government departments, it would appear to be the logical choice for the kind of service we have in mind. However, its interface is far from user friendly, and the support offered with it, far from adequate. By way of example, we had to wait weeks before obtaining incomplete answers to our questions about the cost of setting up a private managment domain (PMD).

to remind members of the section of an upcoming meeting, or to query their colleagues on all sorts of questions; and somewhat to their own surprise, they have come to appreciate even this limited functionality. There now exists a *Windows* version of *cc:Mail*, which can readily be extended to allow for external telecommunications. The use of e-mail to effect the transfer of text files offers several advantages over standard telecommunications software. For one thing, it is far more user-friendly. The user simply prepares and addresses a short note, appends a text file to it, and then clicks on the command that delivers the message to the addressee; the system takes care of everything else. But the main advantage of e-mail is that the user can do this whenever it is most convenient for him. Telecommunications via modem, in contrast, normally requires that both parties and their machines be available at the same time. One user has to be designated the sender, the other has to put his machine in reception mode, and several other technical parameters have to be coordinated between them.[16] With e-mail, on the other hand, each user has his own personal mailbox, in which the system automatically deposits all the messages that are addressed to him, and which he can recover and act on any time it suits him. A zealous translator, for example, could finish a text on the weekend, when no one else was in the office, and send it off to a particular client via e-mail, knowing that it will be waiting in the latter's mailbox first thing Monday morning. And he wouldn't have to be a computer whiz to do it.

For the sake of concreteness, I will illustrate in some detail one possible way in which a hypothetical client could interact with a Bureau translation service located in another site to send off and recover a text to be translated. We assume that the translation service has the PTT-3 in a local area network, and that one of its components is *cc:Mail for Windows*. This network will also include a dedicated PC on which a modem and the program *cc:Mail Gateway* have been installed.[17] For the client, we assume only that he is equipped with a stand-alone PC on which he too has installed a modem, and the program *cc:Mail Remote*. The latter costs $345 CDN, which is not an enormous expense to incur; it serves to ensure the full compatibility of texts transmitted between the client and the translation service. For further details on how the *cc:Mail* software should be configured, see Viau [14].

Suppose that the client wants to transmit a new text to the section he deals with at the Translation Bureau. He begins the session by typing the command "remote" in the *cc:Mail* directory of his PC. He is then prompted to enter his password, after which the program displays the *cc:Mail Remote*

16. To be fair, there exists a mode of telecommunications via modem, called Host mode, which alleviates some of these problems; but it does not offer the user-friendliness of e-mail packages like *cc:Mail*.
17. Note that the program *cc:Mail Dialin*, which is included with the *cc:Mail Platform Pack*, would serve equally well for the scenario to be sketched. The advantage of *cc:Mail Gateway* is that it allows network users to transmit texts to remote users, and to communicate with users on other networks. The dedicated PC on which the *Dialin* or *Gateway* programs run could be an old AT or XT.

main menu (see Figure 3). To send a message, the client would select the option "Prepare new

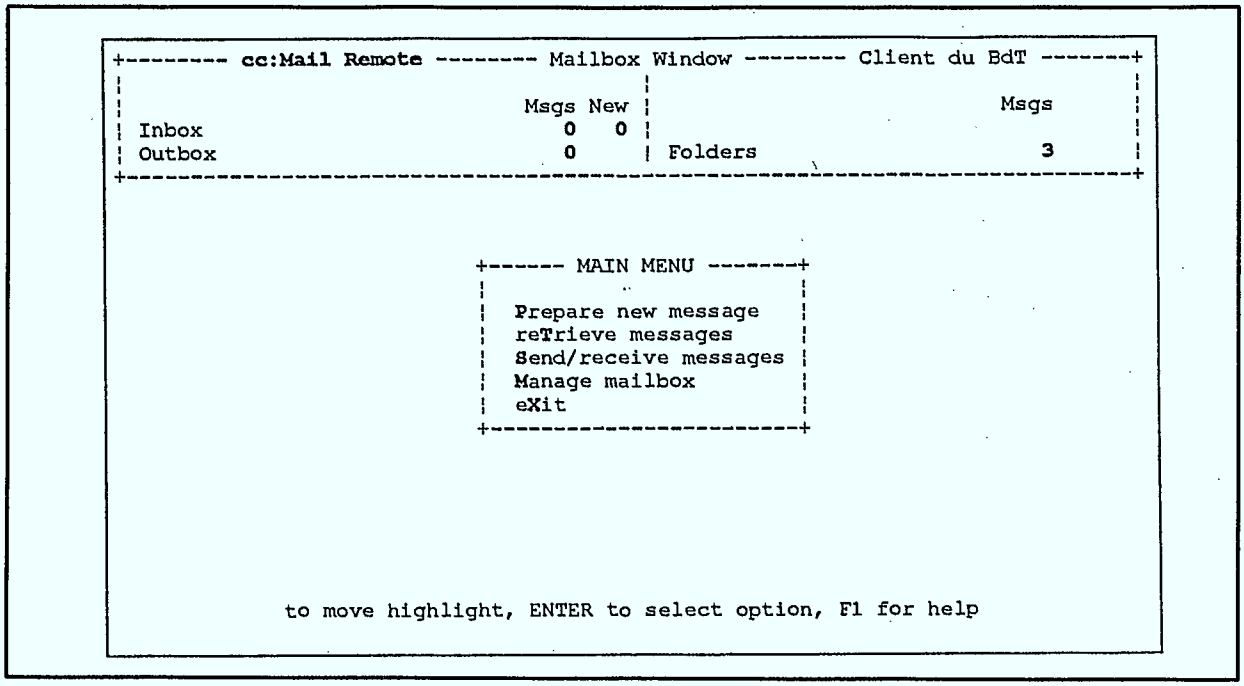


Figure 3

message." This would bring up the addressing menu, where he would find a list of the persons he frequently corresponds with; here, he would simply move the highlight bar to the name of his translator, hit Enter, and then type a short cover note like that given in Figure 4. At this point, the client can choose to attach a copy of a *DOS* file – although this option name may be misleading; it means a file in any *DOS* format, in this case, a word processing file. Once the client selects this option, *cc:Mail* helps him quickly navigate through the directory structure of his disk until he can locate and select the file to be transmitted, once again by simply hitting Enter. The combined message, which now contains two items, the note and the file, is then stored in his outbox, and the client is returned to the main menu, where he can choose to send the message off. All he has to do is verify that his modem is switched on. The system displays messages on the screen to tell him that a call is being placed to the Bureau's *cc:Mail* Post Office (i.e. the network's dedicated PC), and that the connection has been established; it informs him of the number of bytes that remain to be transmitted, and then automatically hangs up the phone, freeing it for voice communication.

The client's message is transmitted over the public phone lines and reaches the Translation Bureau via the modem connected to the dedicated PC that runs the *cc:Mail Gateway* program. This machine is in permanent operation, and so is always ready to forward new messages to the section post office that holds each user's mailbox. The actual elapsed time between the client's sending of the message and the post office's receipt of it will, of course, depend on the size of the file and the

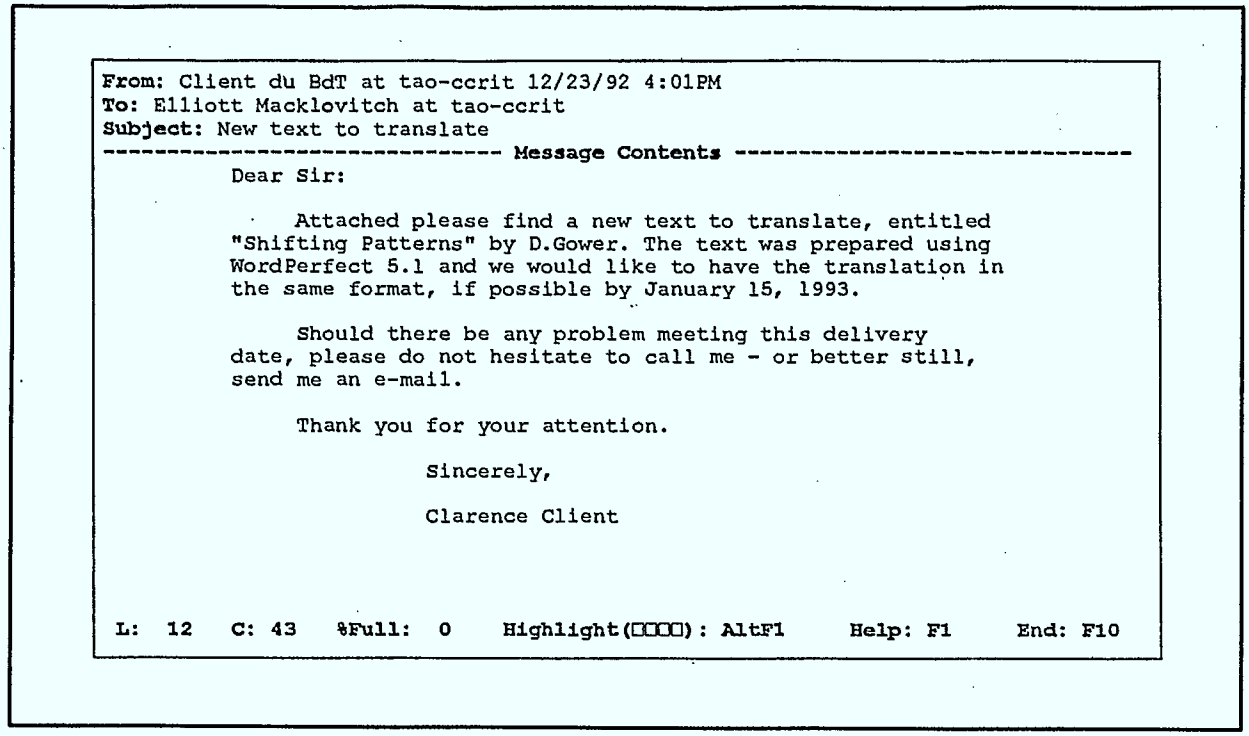


Figure 4

speed of the modem. But some minutes later, the workstation of the translator to whom the message was addressed[18] will emit a beep-beep, and the *cc:Mail* icon flash on his screen, indicating that a new message has just arrived in his mailbox. Whenever he can interrupt what he is doing, the translator double clicks on the icon with his mouse, thereby starting up the *cc:Mail* program. The contents of his inbox is displayed, with a header telling him who the new message is from, the date it was sent, the size of the message and its subject. Again, by double clicking on the message header, the translator can display the message items: in this case, a short covering note and an attached file (see Figure 5). He may then read the covering note, and, by choosing the Save As... option from the File menu, copy the attached text into his *WordPerfect* directory. Starting up that program, he will then load the new file and see it displayed with all its formatting attributes and accented characters intact, exactly as it appeared on the client's PC. All that remains for him to do is to translate it, taking advantage of the full range of resources that the PTT puts at his disposal. When the translation is completed, several options are available for returning it to the client by e-mail; all are as straightforward as it was for the client to send the original source text.

18. Of course, the client's message could be addressed to the Section chief or clerk, and not to the particular translator; these are questions of internal organization that each section will decide for itself. But if the Section chief then wanted to forward the text to an in-house translator, he could use the very same e-mail component. In other words, for the user, the distinction between internal and exteral e-mail is entirely transparent.

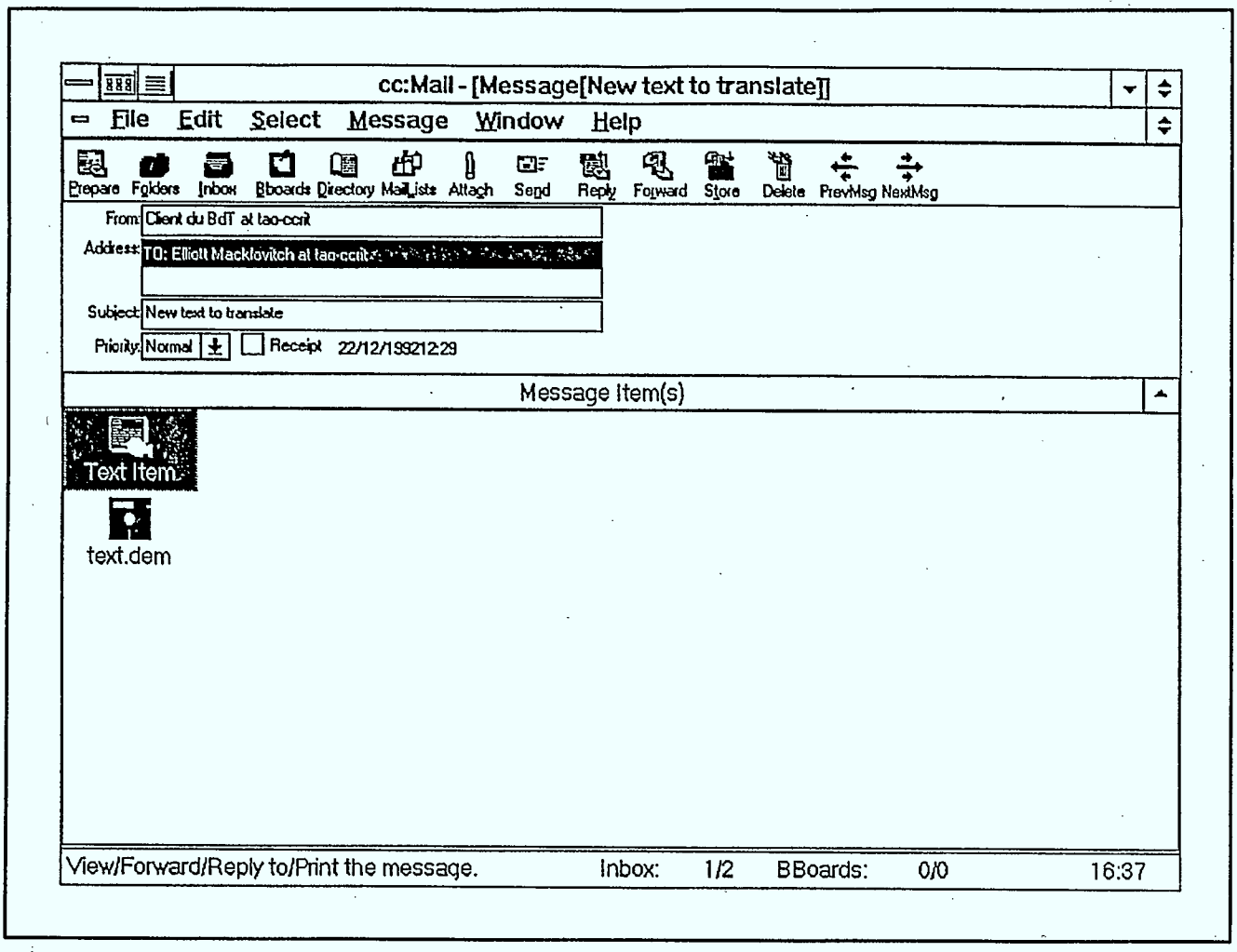


Figure 5

3. Where to from here?

Once the PTT-3 is delivered, we hope that it, like its predecessors, will be the object of a full-scale field trial, so that end-users can provide us with their feedback on the functionality of its new components. Still, we should point out that this version of the PTT[19] marks what we consider to be the end of a development cycle. Although we may yet identify new applications that can help translators perform their routine, peripheral tasks more effectively, with the PTT-3, we seem to be approaching the limits of the off-the-shelf approach that has been the hallmark of our project to

19. Or more precisely, the completed *Windows* version of the PTT, in which most, if not all the applications are designed specifically for the graphical user interface.

date. As mentioned above, many of the components of our PTT are not specifically designed for translators; as general writing aids or office automation tools, they do not directly address the tasks that distinguish translators from other groups of writers. In fact, the only translator-specific components of the PTT are the terminology bank and the glossary managment tool (along with the bilingual dictionary that accompanies it); all the remaining components are regularly used by other types of monolingual English and French writers. Indeed, it could be said that thus far our PTT leaves the job of translation by and large to the human translator. However, as the quotation from Martin Kay cited in Section 1 suggested, the time may now be ripe for the PTT to approach the central and defining task of the translator, which is to reformulate the meaning of a source language text into a grammatically and stylistically correct equivalent in a given target language.

What distinguishes translators as a group from other types of writers is that they operate not on one, but on two texts: the first, a pre-existent source text; and a second text, which they produce in another target language and which maintains a certain equivalence relation with the first. The new orientation of the CWARC's computer-assisted translation group focusses on ways of helping translators establish this equivalence relation. Not, as in classical MT, by having translators reshape or post-edit a "raw" machine translation generated by a system whose understanding of the source text is far inferior to theirs; but by extracting useful information from pre-existent, high quality human translations. Efficient programs have recently been developed that can automatically recreate the correspondences which link source and target texts; see Gale and Church [1] and Simard et al [12]. A set of aligned texts explicitly linked in this way can be referred to as a *bi-text* or a bi-textual database.[20] With the automatic conversion of large masses of previously translated texts into a bi-textual database, it becomes possible to develop a whole range of altogether new types of translator aids. These include bilingual concordancing programs and various types of translation critiquing tools. Such applications are discussed in detail in Isabelle [3] and in Macklovitch [9], and this is not the place to repeat that discussion. Suffice it to say that these new aids, based on the notion of bi-text, appear to us to have enormous potential, both for translators and terminologists. In recycling translators' past production, they render accessible a wealth of solutions to translation problems that are often not addressed in standard reference works. The reaction of those translators who have seen the first prototypes currently under development at the CWARC has certainly been enthusiastic. Our plan is to incorporate these bi-textual tools into future versions of the PTT.

---

20. [Harris 1988] was the first to use the term "bi-text" for this type of parallel corpus, although applications of the same concept are also discussed in [Melby 1981]. The major difference between these proposals and the CWARC's current work is that they do not incorporate a fully automatic alignment algorithm, and hence can only produce large bi-textual databases at great cost. For *OCTA*, our bilingual concordancing prototype, we have aligned about 35 million words of Canadian Parliamentary debates.

4. Acknowledgements

I would like to thank all the translators at the Secretary of State who have participated in the trial of the PTT and who, through their comments and thorough testing, have made it into a better working tool. My special thanks go to Nicolas Viau, who, more than anyone else, has been responsible for assembling the PTT-3. I'm sure that the translators who will have the opportunity of working on it will appreciate what an excellent job he has done.

5. References

[1] Gale, William & Kenneth W. Church (1991), "A Program for Aligning Sentences in Bilingual Corpora," **Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics**, Berkeley, California, pp. 177-184.

[2] Harris, Brian (1988), "Bi-Text: A New Concept in Translation Theory," in **Language Monthly**, no. 54, pp.8-10.

[3] Isabelle, Pierre (1992), "La bi-textualité: vers une nouvelle génération d'aides à la traduction et la terminologie", **Meta**, vol. 37, no.4, pp.721-737.

[4] Kay, Martin (1980), "The Proper Place of Men and Machines in Language Translation," CSL-80-11, Xerox Palo Alto Research Centre, Palo Alto, California, 20 pages.

[5] Lavoie, Richard & François Lapointe (1990), "Poste de travail pour la traduction: Évaluation de la convivialité", CWARC technical report, 110 pages.

[6] Macklovitch, Elliott (1989), "An Off-the-shelf Workstation for Translators," in D. Hammond (ed.) **Proceedings of the 30th Annual Conference of the American Translators Association**, Washington, D.C., pp. 491-498. (Also available as a CWARC technical report.)

[7] Macklovitch, Elliott (1990), "PTT 1.1 : Specifications for an upgraded version of the Translator's Workstation," available from the author.

[8] Macklovitch, Elliott (1991), "A Second Version of the CWARC's workstation for translators," CWARC technical report, 15 pages.

[9] Macklovitch, Elliott (1992), "Corpus-based Tools for Translators," in E. Losa (ed.) **Proceedings of the 33rd Annual Conference of the American Translators Association**, San Diego, California, pp. 317-328.

[10] Macklovitch, Elliott & Pierre Isabelle (1988), "The Translator's Workstation Project," CWARC technical report, 35 pages.

[11] Melby, Alan K. (1981), "A Bilingual Concordance System and its Use in Linguistic Studies," in W. Gutwinski and G. Jolly (eds.), Proceedings of **The Eight LACUS Forum 1981**, Hornbeam Press, Columbia, South Carolina, pp.541-549.

[12] Simard, Michel, George Foster, and Pierre Isabelle, (1992), "Using Cognates to Align Sentences in Bilingual Corpora," in **Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation**, Montreal, Canada, pp. 67-81.

[13] Viau, Nicolas (1992a), "Intégration du courrier électronique dans le poste de travail du traducteur", CWARC technical report, 16 pages.

[14] Viau, Nicolas (1992b), "Implantation du courrier électronique dans le poste de travail du traducteur: Guide d'installation du logiciel *cc:Mail*", CWARC technical report, 14 pages.

CACC / CCAC
33498

QUEEN P 309 .M335 1993 c.2
Macklovitch, Elliott
PTT-3 : a third version of t

JOUR P 309 .M335 1993 c.2
Macklovitch, Elliott
PTT-3 : a third version of t

DATE DUE - DATE DE RETOUR

ISTC 1551 (2/90)

Pour plus de détails, veuillez communiquer avec :

Le Centre canadien de recherche sur l'informatisation du travail
1575, boulevard Chomedey
Laval (Québec)
H7V 2X2
(514) 682-3400

For more information, please contact:

Canadian Workplace Automation Research Centre
1575 Chomedey Blvd.
Laval, Quebec
H7V 2X2
(514) 682-3400