

CORPUS RESOURCES FOR TRANSLATORS: ACADEMIC LUXURY OR PROFESSIONAL NECESSITY?

Lynne Bowker*

ABSTRACT: Since the late 1990s, corpora have become established as popular and useful translation resources within translator training institutes; however, the uptake of corpora in the world of professional translators appears to have been considerably slower. This article explores the different uses of corpora (including translation memories) in these two sectors and attempts to account for these differences. To determine how corpora are used in academics, a literature survey of papers produced by translator trainers is conducted. With regard to the use of corpora in a professional setting, this study focuses on professionals working in Canada. To find out the extent to which Canadian professional translators make use of corpora, two methods are used. Firstly, a literature survey of publications produced by Canadian translators' associations is carried out, and secondly, a database of job advertisements is analyzed to determine how many Canadian employers are seeking candidates who are familiar with corpus-based resources. The resulting data are compared and discussed with a view to uncovering and understanding why corpus use differs in academic and professional circles.

KEYWORDS: corpora, translation memories, hybrid tools, academics, translator training, professional translators.

RESUMO: Desde o final dos anos 90, os corpora se firmaram como uma ferramenta de tradução conhecida e útil nos centros de treinamento de tradutores. No entanto, parece que sua difusão no universo do tradutor profissional tem

* University of Ottawa, Canada.

sido bem mais lenta. O presente artigo explora os diferentes usos de corpora (inclusive as memórias de tradução) nesses dois setores e pretende apontar suas diferenças. Para determinar como os corpora são usados no meio acadêmico, foi realizado um levantamento bibliográfico de artigos produzidos por tradutores aprendizes. Com relação ao uso de corpora no meio profissional, este estudo concentra-se nos tradutores profissionais do Canadá. Para verificar até que ponto eles fazem uso de corpora, foram empregados dois procedimentos. Primeiramente, foi feito um levantamento dos artigos publicados por associações canadenses de tradutores e, numa segunda etapa, avaliou-se um banco de dados de ofertas de emprego para levantar quantos empregadores procuram candidatos familiarizados com o uso de corpora como ferramenta de trabalho. Os dados resultantes são comparados e discutidos com vistas a revelar e compreender por que o uso de corpora é diferente nos meios acadêmico e profissional.

UNTERMOS: corpora; memórias de tradução; ferramentas híbridas; meio acadêmico; ensino de tradução; tradutores profissionais.

0. Introduction

Corpora are resources that can be usefully applied to translation research, training and practice. Essentially, a corpus is a collection of electronic texts that have been gathered according to specific criteria and that can be used to investigate particular linguistic features. Over the past decade, there has been a rising interest in all aspects of the corpus-based approach to translation, as evidenced by the increasing number of conferences,¹ special issue journals² and monographs³ dedicated specifically

¹ For example, the “Corpus Use and Learning to Translate (CULT)” conferences held at the University of Bologna in November 1997 and November 2000, as well as the “Language Resources in Translation Work

to this subject. However, most of the documented activity in this area appears to be taking place within the academic community. This is not too surprising – in many different fields, discoveries are initially made by the research community and once their relevance or usefulness has been demonstrated, they are then adopted by the professional community.

It would seem that the usefulness of translation corpora within academic circles has been firmly established, but it is not yet clear to what extent the corpus-based approach has been adopted by professional translators. The aim of this article is to investigate and compare the use of corpora in translator training institutes, where corpus-based resources have become very popular in recent years, with the use of corpora in a professional translation setting. While the use of corpora for conducting research into the nature of translation will be briefly addressed, the particular focus will be on the use of corpora as a resource for finding solutions to translation problems because this goal is common to both translator trainers/trainees and professional translators.

The article will be divided into three main parts. Section 1 will provide an overview of corpus use in an academic setting, looking briefly at the use of corpora for translation research, before focusing on the use of corpora for translator training. Section 2 will explore the importance of corpora in a professional setting by examining reports in publications produced by professional translators' associations, as well as surveying a corpus of job advertisements for translation-related positions. This discussion of the professional use of corpora will focus primarily on the situation in Canada. Finally, section 3 will compare the use of corpora in academics and the workplace, and will consider possible trends of corpus use for the future.

and Research Workshop" held as part of the 3rd International Conference on Language Resources and Evaluation (LREC) in May 2002.

² The present volume, as well as Laviosa (1998a) and Tagnin (2003).

³ Kenny (2001), Bowker and Pearson (2002), Laviosa (2002), Olohan (forthcoming).

1. Translation corpora in an academic setting

As seen in Laviosa (2002), corpora are typically used in an academic setting in one of two main ways: by researchers as a means of studying the nature of translation, or by translator trainers/trainees as a resource for finding solutions to translation problems. The focus of this article will be on the latter of these two applications; however, for the sake of completeness, we will provide a very brief overview of some research-related applications.

1.1 Corpora and translation research

Researchers in translation studies have devised a number of different types of corpora, as well as corresponding methodologies, for investigating the nature of translation as seen through translated texts. For example, as described by Baker (1995, 1996) one type of corpus that can be used to investigate the nature of translation is a comparable corpus. A comparable corpus has two parts, one that contains original texts written in language A and one that contains texts that have been translated into language A. These two parts can be compared in order to uncover the features of translational text production, such as simplification (Laviosa, 1998b), explicitation (Olohan, 2002) and normalization (Kenny, 2001).

Another type of corpus used by translation researchers is a parallel corpus, which is sometimes referred to as a bitext. This type of corpus typically contains source texts written in language A, which are aligned with their translations into language B. A corpus of this type can be used for identifying translational norms, as well as for contrastive studies, such as investigating similarities and differences in language use (e.g. Ebeling, 1998, Maia, 1998, Kenny, 2001).

A variation of the parallel corpus is one that contains a source text in language A, aligned with several different translations into language B (e.g. produced by different translators). This type of corpus permits the study of differences existing between the various translations of the same original work (e.g. Malmkjaer, 1998).

1.2 Corpora and translator training

The other main application of corpora in an academic setting is as a resource for finding solutions to translation problems. This could include teaching students to identify terminological or phraseological equivalents, as well as to investigate elements of style that are appropriate to a given text type or sublanguage. A literature survey provides evidence that a considerable number of translator trainers have successfully integrated a range of corpus-based techniques into translation classrooms.

1.2.1 Designing and building corpora

Some trainers, such as Pearson (2000a), Maia (2000), Zanettin (2002), Bowker (2002a), Tagnin (2002) and Varantola (forthcoming) have shown that teaching students how to design and build useful corpora is an excellent method for getting these students to reflect on issues such as text typology, register, and domain specificity. As pointed out by Pearson (2000a:237), when they are first learning to translate, students sometimes show poor judgement in sourcing information to use in their translations (e.g. taking information from a source that is outdated, or from a text that is of a different level of technicality than that of the intended target text). Exercises in corpus building provide an excellent opportunity for teaching students the importance of evaluating the quality and appropriateness of materials to be used as translation resources.

1.2.2 Use of monolingual corpora

With regard to applying corpus-based techniques on existing corpora, Bowker (1998, 2000), Lindquist (1999), and Bowker and Pearson (2002) have demonstrated that monolingual corpora created in the target language can be a useful resource for investigating terminology, grammar and style. Bowker (1998), for instance, details an experiment carried out with a group of translation students. Half of the students translated a text using conventional resources (e.g. dictionaries,

TRADTERM, **10**, 2004, p. 213-247

printed parallel texts), while the other half translated the same text using a monolingual corpus and corpus analysis tool. The results of the experiment indicate that the students who used the corpus produced translations that were more idiomatic and more accurate with regard to term choice and the use of idiomatic expressions. The students attributed their success to factors such as the ease and speed with which an electronic corpus could be searched, and the presence of an abundance of contextual material.

Another successful application of monolingual corpora in the translation classroom has been as an evaluation tool. Bowker (2000, 2001) and Pearson (2000b) have found that such corpora can act as a useful benchmark for verifying the correctness of students' linguistic, grammatical and stylistic choices. As reported by Bowker (2000:184), students themselves appreciate this approach to the correction of their translations because the corpus data provides an objective and concrete source of information and feedback.

1.2.3 Use of bilingual corpora

Other trainers have demonstrated that bilingual corpora are also highly useful in a translator training context, where they can be successfully used by students to help with a range of translation-oriented tasks. For instance, both Zanettin (1998, 2001) and Palumbo (2002) have conducted exercises using Italian-English parallel corpora, and they found that their students were able to make more accurate choices with regard to identifying appropriate equivalents and collocations. Tagnin (2002) made similar observations about translation students using parallel corpora in English and Brazilian Portuguese.

Another area where bilingual parallel corpora have proved useful is for helping students come to grips with difficult grammatical points. Schmied (2002), for example, notes that an English-German parallel corpus proved invaluable for providing translators with guidelines and examples about appropriate use of prepositions – a notoriously problematic area for translators working into a foreign language. Hansen and Teich (2002) show

how a parallel corpus can be used to investigate the cross-linguistic differences in the usage of the present and past perfect tenses in German and English. Meanwhile, Frankenberg-Garcia (2002) provided her students with a parallel corpus that allowed them to understand different uses of negative prefixes in Portuguese and English.

Bernardini (2002) reports that parallel bi-directional corpora (Italian and English) have been successfully used in her classroom to help students identify the norms, stylistic preferences and discourse structures associated with different text types.

Pearson (2000c) demonstrates that, in addition to having linguistic applications, bilingual parallel corpora can also be used to reveal important conceptual information. By using search patterns such as “called” in English or “c’est-à-dire” in French, students can retrieve contexts containing definitions or explanations. This is particularly important when working in specialized domains because the students need to understand the specialized concepts in question in order to translate the text effectively.

Finally, another type of bilingual corpus, known as a bilingual comparable corpus, can also be a valuable tool for translators. This type of corpus contains original (i.e., non-translated) texts in two languages. The two sets of texts are comparable in terms of text type, subject matter, and time period, but they are not translations of one another. López Rodríguez (2002) used a bilingual comparable corpus of Spanish and British newspaper articles on primary education to teach her students how to uncover culturally relevant information, and how to develop strategies for translating culture-specific references.

1.2.4 Use of translation memories

Another type of corpus-based resource that deserves special mention is the translation memory. Essentially, a translation memory is a database of aligned source and target texts; it can therefore be seen as a special type of bilingual parallel corpus. L’Homme (1999), Austermühl (2001) and Bowker (2002b), among others, note that translation memory systems (e.g. Trados

TRADTERM, **10**, 2004, p. 213-247

Translator's Workbench, STAR Transit, Déjà Vu, SDLX) are increasing in popularity among translators, and consequently, many translator training institutes⁴ are now teaching students how to use such tools.

The idea behind a translation memory system is that it allows a translator to reuse or 'recycle' previously translated texts. Such systems work by breaking down the source and target texts into segments that typically correspond to sentence-like units (e.g. sentences, titles or subtitles, items in a list, cells in table). The matching source and target language units are linked together and stored in a database. When a translator is faced with a new text to translate, the system also breaks down this new source text into segments, and it then compares each of these new segments against those stored in the translation memory database. If it finds a match (i.e., a segment that has been fully or partially translated on a previous occasion), it presents that match to the translator, who then has a choice of reusing, modifying, or rejecting the previous translation.

One major difference between a conventional bilingual parallel corpus and a translation memory is the degree of automation that can be achieved. Whereas translators who use conventional corpora and corpus analysis tools must formulate their own queries (i.e., they must decide upon and then type in the search pattern), a translation memory system automatically searches for matching segments. Another difference is in the length of the search patterns used. When using a conventional corpus, translators tend to search for terms or short phrases, while a translation memory system will search for longer segments (e.g. complete sentences or even paragraphs). A final difference is the nature of

⁴ Translator training institutes that offer courses in translation technology, which includes the use of translation memory systems, include, among others, University of Ottawa (Canada), Université du Québec en Outaouais (Canada), Université de Montréal (Canada), York University (Canada), Dublin City University (Ireland), Kent State University (United States), Monterey Institute of International Studies (United States), Universität des Saarlandes (Germany), Université de Genève (Switzerland), Imperial College (United Kingdom), University of Swansea (United Kingdom), and Université Rennes 2 (France).

the matches retrieved. Conventional corpus analysis tools typically retrieve only those instances that precisely match the search pattern that has been entered. Therefore, if a translator entered “colour laser printer” as a search term, instances of “laser color printers” would not be retrieved.⁵ In contrast, translation memory systems make use of fuzzy matching techniques, which permit the retrieval of instances that are similar to, but do not precisely match, the search pattern. The degree of similarity that must be achieved in order for a fuzzy match to be established can be set by the translation memory system user (anywhere from 1 – 99%).

As noted above, a number of translator trainers have begun to integrate the use of translation memory systems into their classrooms. For example, DeCesaris (1996) notes how translation memories can be used as a self-learning resource to provide students with immediate access to a variety of possible translation solutions. She goes on to explain that while the idea of providing students with models is not new, the point is that students are often given a single model for specific translations, which perpetuates the view that there is only one correct translation for a given text. A translation memory with fuzzy matching capability can make it easy for trainers to provide students with multiple possible translation solutions. Students can then learn to analyze and to weigh the pros and cons of these different solutions.

Kenny (1999) has identified a number of benefits to be gained from introducing students to translation memory systems. For instance, students can learn more about inter- and intra-textual features by examining source texts and evaluating their characteristics in an effort to determine whether or not

⁵ Note that some degree of flexibility can be achieved by searching with the help of wildcards (e.g. the search pattern “colo?r” would retrieve instances of both “colour” and “color”, and the search pattern “printer*” would retrieve instances of both “printer” and “printers”). Some corpus analysis tools also permit context searches where the user can specify that s/he wants to retrieve all instances of the term “laser” when it appears within a 3-word span of the term “printer”. While wildcards and context searches can indeed be useful, they are still less flexible than the fuzzy matching techniques used in translation memory systems.

they can be usefully translated with the help of a translation memory system (e.g. whether or not the texts are made up of recurrent units). Kenny also points out that translation memory systems could be used by trainers to conduct longitudinal studies of students' progress over the course of their training program. If students begin using translation memories at the start of their degree, and continue using them throughout, then the resulting translation memory database could be used to investigate the development of translator competence.

2. Translation corpora in a professional setting – examining the situation in Canada

Let us now turn our attention to exploring the use of corpora by professional translators. Much of the data and discussion that will be presented in the following sections primarily describe the Canadian situation. This decision was made for pragmatic reasons; because I live and work in Canada, it was easier for me to collect information pertaining to professional practices in this country. Nevertheless, it is hoped that readers from other countries will find this information interesting also.

For this investigation, I used two main sources: the literature produced by some professional translators' associations (e.g. newsletters and magazines aimed at language professionals), and a database of job advertisements for translation-related positions.

2.1 Literature from professional translators' associations

In Canada, there are a number of professional associations to which translators can belong. The two largest are the Association of Translators and Interpreters of Ontario (ATIO) and the *Ordre des traducteurs, terminologues et interprètes agréés du Québec* (OTTIAQ).⁶ These associations regularly publish news-

⁶ Both of these associations belong to the Canadian Translators and Interpreters Council (CTIC), which is in turn a member of the International Federation of Translators (FIT).

letters and magazines aimed at keeping translators informed about the latest developments in the profession. A quick glance through recent issues (January 2000 – December 2002) of these publications shows that translation technology in general is being given a considerable degree of coverage. The following sections will elaborate more specifically on the attention being paid to corpus-based tools in particular, including translation memories, conventional corpus analysis tools (e.g. concordancers), and a new type of hybrid corpus-based tool.

2.1.1 Translation memories in the professional association literature

The type of corpus-based tool discussed most frequently in the professional association literature is the translation memory. According to *InformATIO* (the newsletter of ATIO) and *L'antenne* (the newsletter of OTTIAQ), these associations have provided their members with a number of opportunities to familiarize themselves with translation memory tools. A number of recent events organized by these associations are summarized in Table 1.

ATIO events	OTTIAQ events
<i>October 2000</i> – a professional development day entitled “Technology in the Service of Translation” which included demos of translation memory tools (Cadieux 2000)	<i>October 2000</i> – a workshop on translation memories entitled “À l’aide! Les mémoires de traduction, des outils à retenir” (Quirion 2000)
<i>January 2001</i> – a panel discussion on the “Future of technology in translation” accompanied by demos of translation memory tools (Rosen 2001)	<i>September 2001</i> – a round table on translation memories (Duarte 2001)
<i>April 2001</i> – a professional development day entitled “Are you ready for CAT?” which focused on translation memory tools (Bélisle 2001a)	<i>November 2001</i> – a workshop on translation memories entitled “Le portrait-robot: Les mémoires de traduction” (Dandurand 2002)

Table 1: Translation memory events offered by professional translators’ associations.

In addition to its newsletter, OTTIAQ also publishes *Circuit*, which it describes as a quarterly magazine on language, communication and translation. The magazine focuses on topical subjects concerning the profession, and it includes a number of practical and informative columns and sections. Translation memories are regularly discussed as evidenced by articles from Bédard (2001), Lanctôt (2001), Schwab (2001) and Cohen (2002); however, it is interesting to note that the discussions do not always contain favourable reports. For instance, Bédard (2001) notes that translating sentence by sentence with the help of a translation memory system does not always produce a high-quality result. Schwab (2001) and Cohen (2002) raise the thorny issues of ownership of translation memories and of how translators should be remunerated when they use translation memories. Meanwhile, Lanctôt (2001) discusses difficulties that arise when a group of translators who have different styles and working approaches are forced to contribute to and share the same translation memory.

2.1.2 Conventional corpus analysis tools in the professional association literature

Based on the information available in the publications of ATIO and OTTIAQ, it appears that the professional associations have paid considerably less attention to conventional corpus analysis tools than to translation memories. There was only one event focusing on these tools reported: in April 2002, ATIO held a professional development day entitled “Concordancing tools – a viable option?” in which monolingual and multilingual concordancers (e.g. WordSmith Tools, MultiConcord, ParaConc) were discussed and demonstrated (Evans 2002).

2.1.3 Hybrid corpus-based tools in the professional association literature

Although the use of conventional corpus analysis tools does not seem to be very popular among professional translators, a

new type of tool, which is a sort of hybrid between a translation memory and a corpus analysis tool, does seem to be generating a considerable amount of interest as demonstrated in articles by Arrouart and Bédard (2001), Bélisle (2001b), Lanctôt (2002) and Lavallée (2002). In addition, as part of its Continuing Education Programme, OTTIAQ has advertised an upcoming workshop to be held in April 2003 entitled “À l’aide! Logiterm”, which will focus on one of these hybrid tools.

As previously noted in section 1.2.4, some of the main differences between conventional corpus analysis tools and translation memories include 1) the fact that translation memories perform automated searches on all segments while corpus analysis tools require the user to think up and manually enter the search pattern; 2) the fact that translation memories attempt to match units at the sentence level, while corpus analysis tools typically identify matches at the sub-sentence level (e.g. terms or short phrases); and 3) the fact that translation memories can perform fuzzy matching whereas corpus analysis tools do not have this capability.

A new generation of tools, which includes MultiTrans developed by the company MultiCorpora (Lavallée, 2002) and LogiTerm developed by Terminotix (Lanctôt, 2002), is beginning to combine features of both translation memories and conventional corpus analysis tools. For example, MultiCorpora permits automated searching at the sub-sentence level, while LogiTerm allows fuzzy matching at the sub-sentence level.

In a number of articles, such as Arrouart and Bédard (2001) and Bélisle (2001b), these new hybrid tools are compared directly with translation memory tools, and on the whole, the hybrids seem to fare more favourably in the comparison.

2.2 Database of job advertisements for translation-related positions

Another way of determining the extent to which corpus-based tools are used by professional translators is to examine job advertisements to see if employers are looking to hire candidates who have experience using such tools. To this end, I

TRADTERM, **10**, 2004, p. 213-247

constructed a database of advertisements for translation-related jobs in Canada collected over a three-year period from January 2000 to December 2002. The advertisements were gathered from four main sources:

- 1) Canadian newspapers (e.g. *The Globe and Mail*, *The Ottawa Citizen*, *The Montreal Gazette*, *The Toronto Star*, *Le Devoir*, *La Presse*);
- 2) Canadian online job sites (e.g. www.monster.ca, www.workopolis.com)
- 3) Advertisements distributed by two Canadian professional translators' associations: the Association of Translators and Interpreters of Ontario (ATIO) and the *Ordre des traducteurs, terminologues et interprètes agréés du Québec* (OTTIAQ);
- 4) The job bulletin board at the School of Translation and Interpretation of the University of Ottawa.

In addition to jobs specifically aimed at translators, the database includes jobs for other professionals in the language industries who may be required to perform translation-related tasks.⁷ The different job types, as well as the number of advertisements for each job type, are shown in Table 2.

Job type	Number of advertisements
translator	170
reviser	44
localization specialist	27
translation manager	16
terminologist	14
TOTAL	271

Table 2: Number of jobs advertised for various translation-related positions.

⁷ Note that this database was initially constructed to undertake a much wider study of the Canadian language industries job market so it also includes jobs for other positions, such as interpreters and technical writers. However, in the context of this article, the focus will be solely on the translation-related jobs listed in Table 2.

As illustrated in Figure 1, once the advertisements were entered into the database, they were indexed in order to facilitate future searching.⁸

The screenshot shows a software window titled "Translation Jobs Database - [Advertisement]". It contains a form with various fields for job details. The fields are organized into two columns. The left column includes fields for ID (52), Source (The Ottawa Citizen), Date (13/10/01), Employer (Association of Universities a...), Ad Language (English), Permanent (checked), Full-time (checked), Examination (unchecked), and Professional association certification (unchecked). The right column includes fields for Domain (Education), Job (Translator), Keywords (translation technology), Language (English-French), Location(s) (Ottawa), Qualifications (degree), and Required Skills (computer literate). Each field has an "Add/Modify" button next to it. Below the form is a large text area containing the job description for a Translator position at AUCC. The text describes the role, reporting structure, and requirements. At the bottom of the window, there is a status bar showing "Record: 14" and "49 of 440".

ID	52	Domain	Education	Add/Modify
Source	The Ottawa Citizen	Job	Translator	Add/Modify
Date	13/10/01	Keywords	translation technology	Add/Modify
Employer	Association of Universities a...	Language	English-French	Add/Modify
Ad Language	English	Location(s)	Ottawa	Add/Modify
Permanent	<input checked="" type="checkbox"/>	Qualifications	degree	Add/Modify
Full-time	<input checked="" type="checkbox"/>	Required Skills	computer literate	Add/Modify
Examination	<input type="checkbox"/>			
Professional association certification	<input type="checkbox"/>			

Text

AUCC is a national, non-governmental, non-profit organization that represents Canada's 32 public and private universities and university colleges. We provide a forum for discussion and a framework for action at the federal level, facilitating the development of public policy on higher education. We are presently seeking a Translator.

Translator

Reporting to the Senior Translator/Revisor, you will translate and revise the Association's communications from English to French, including publications, reports, briefs, minutes of meetings. Your prior experience and excellent knowledge of both official languages has prepared you for this position in an organization that values quality communications with its membership. Your familiarity with terminology banks and other computer-assisted translation tools will be a definite asset to the successful operations of the Translation Unit.

If you have a university degree in translation, three years of solid experience, and are interested in joining the AUCC Translation Unit, please send us your curriculum vitae.

You will find additional information on the candidate profiles listed under Current Job Postings on our Web site. Please send your curriculum vitae by October 25, 2001 to AUCC's Human Resources Service.

600-350 Albert,
Ottawa, Ontario K1R 1B1
Tel: (613) 563-1236
Fax: (613) 563-9745
careers@aucc.ca

Record: 14 49 of 440
Form View

Figure 1: A sample entry in the job advertisement database.

This indexing was implemented using a controlled vocabulary drawn up specifically for this project, where each concept was represented by a single term. For example, the idea that might have been stated in the jobs ads using various expressions such as “able to use computer tools effectively” or “famil-

⁸ Because the database was originally constructed to undertake a wider study, the advertisements were indexed for information such as language, domain, location of the job, qualifications, required skills, etc. as shown in figure 1. However, in the context of this article, the focus will be exclusively on whether or not the candidate is required to be computer literate and/or familiar with corpus-based tools.

ilarity with Microsoft Office” was indexed using the controlled vocabulary term “computer literate”. Job advertisements that specifically required candidates to be familiar with specialized translation software (e.g. corpus analysis tools, translation memories) were further indexed using the term “translation technology”. In this way, all advertisements requiring some degree of computer literacy or translation technology knowledge could be easily extracted from the database and further examined. For instance, Figure 2 shows an example of a search for all advertisements for the job of “translator” which indicate that the candidate must be “computer literate”.

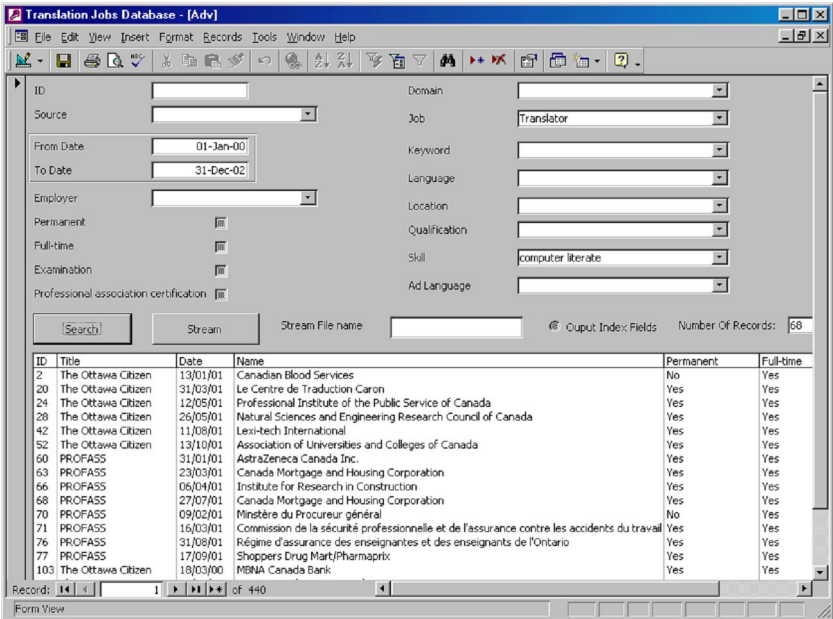


Figure 2: The results of a database search for all jobs for “translator” that require the candidate to be “computer literate”.

2.2.1 Limitations of the database

The database has some limitations that must be kept in mind when interpreting the data. Nevertheless, as pointed out by Atkins *et al.* (1992:6), even an imperfect collection of data can

be a source of useful information, as long as one is aware of its shortcomings.

One limitation of the database is that it can only provide information about jobs that have been advertised. This means that positions that exist but that have not had a turnover in personnel are not represented. In addition, jobs that are not advertised, such as freelance contracts where clients communicate directly with candidates, or jobs filled through competitions, such as those held by the Government of Canada's Translation Bureau, are not represented in the database.

In addition, even though an advertisement may list desired skills (e.g. computer literacy or knowledge of specialized translation software), it is not possible to tell whether the successful candidate actually met these criteria. It may be the case, for instance, that the company in question was willing to settle for a candidate with lesser qualifications, or the company may have been fortunate enough to hire a candidate who surpassed the minimum advertised qualifications. It may even happen that the position was not filled at all.

In spite of these limitations, the database can still provide a reasonable overview of the types of translation-related skills that employers feel are important in today's market. The following sections will focus specifically on computer-related skills.

2.2.2 Computer-related skills desired by employers

Table 3 presents an overview of the number of employers who advertised for candidates with either general computer skills or more specialized skills relating specifically to translation technology (e.g. translation memories or corpus analysis tools).

Job type	Total number of ads for each job type	Number of ads requiring some degree of computer literacy	Number of ads requiring familiarity with translation memories or other corpus-based tools
translator	170	95 (55.8%)	10 (5.8%)
reviser	44	24 (54.5%)	9 (37.5%)
localization specialist	27	27 (100%)	17 (62.9%)
translation manager	16	12 (75%)	2 (12.5%)
terminologist	14	11 (78.5%)	3 (21.4%)
TOTAL	271 (100%)	169 (62.3%)	41 (15.1%)

Table 3: Number of translation-related job ads requiring general or specialized computer skills.

With regard to general computer skills, the type of computer literacy desired by employers fell into several sub-categories:

- employers who simply requested that candidates be comfortable with computer use in general;
- employers who specified that familiarity with word processing software or other general desktop applications (e.g. PowerPoint, Excel, Outlook, QuarkExpress) was a pre-requisite;
- employers who required candidates to be able to use the Internet or the World Wide Web as a research tool;⁹
- employers who specified that familiarity with electronic terminology resources such as TERMIUM or the *Grand dictionnaire terminologique* was required.

⁹ It is possible to view the World Wide Web as a type of corpus and to use a search engine as a sort of concordancer (i.e., to find numerous examples of a term or phrase in context). Therefore, translators could be applying corpus-based techniques without realizing it. Although this is an interesting idea, further exploration of this possibility is beyond the scope of this paper.

In all, 62.3% of the job advertisements for translation-related positions specifically stated that some general degree of computer literacy was required in order to perform the functions of the job. This included more than 50% of the jobs for translators and revisers, more than 75% of the jobs for translation managers and terminologists, and not surprisingly, 100% of the jobs for localization specialists.

For jobs that required a higher degree of specialized computer knowledge, advertisements requested skills such as the following:

- familiarity with “state-of-the-art translation tools”, “computer-assisted translation tools”, “translation software”, “terminology software”, or “terminotics”; or
- familiarity with specific tools, including Trados, MultiTerm, MultiTrans, LogiTerm, and MTX.

Overall, a mere 15.1% of the advertisements were aimed at candidates who had knowledge of specialized translation software. Even in the case of localization specialists, for whom one might assume familiarity with tools such as translation memories to be a pre-requisite, only 62.9% of the ads called for such skills. Perhaps the most shocking revelation is that only 5.8% of translators were requested to have experience using specialized translation software. Translation managers came in slightly higher at 12.5%, terminologists at 21.4% and revisers at 37.5%.

In this day and age, where computers seem to be an integral part of almost every profession, and where translation technology seems to be a popular theme at conferences and in journals and newsletters, it is somewhat surprising that only 62.3% of the advertisements in the database indicated that at least some degree of computer literacy is required, and even more noteworthy that only 15.1% indicated that a more specialized knowledge of translation-related software is desirable.

It is therefore necessary to consider a number of possible explanations for these relatively low percentages. Looking at the remaining advertisements, it must first be noted that 31 (11.4%) of these were extremely minimalist, and did not really list any-

TRADTERM, **10**, 2004, p. 213-247

thing in the way of a job description or desirable skills. Table 4 provides some examples of this type of advertisement.

<p>The Masha Krupp Translation Group Ltd. (MKTG) has job openings for experienced English-into-French translators. MKTG is an Ottawa-based translation firm which was founded in 1992 and is now one of Ottawa-Hull's largest translation firms, with more than 50 full-time staff.</p> <p>Excellent remuneration commensurate with experience and abilities. Fax or e-mail c.v. in confidence to: FAX: (613) 256-4737; E-MAIL: brian@mashakrupp.com</p>
<p>The translation service of Costco Wholesale in Ottawa is looking for an English to French translator for administrative and advertising texts. The person will also act as terminology advisor and will occasionally be asked to translate from French to English. Competitive salary and good work benefits. Requirements: bachelor's degree in translation, or bachelor's degree in another field with a diploma in translation, also a few years of experience. Fax your resume to 613-221-2309.</p>
<p>TRANSLATOR (English/French). Required for National Insurance Organization in Markham. Degree or equivalent Translation designation required. \$40K-\$45K w/benefits. Send CV to: ANNE WHITTEN, 438 University Ave, Suite #308, Toronto, Ontario, M5G 2K8, Fax: 416-598-5127</p>

Table 4: Examples of advertisements giving no real job description or skills requirement.

With regard to the remaining 71 advertisements that did provide a list of desired skills, but did not include any mention of computer literacy or translation technology skills on this list, three other possibilities must be considered. First, it is possible (though increasingly unlikely) that the candidate will not be required to carry out any computer-related tasks as part of the job (e.g. a translator could use a dictaphone and printed dictionaries or other non-electronic resources). Second, it could be the case that the person who composed the advertisement did not fully understand the requirements of the job and did not there-

fore list computer skills as an important component. Third, an employer may regard computer skills to be such an integral part of the job that it is considered to be “understood” that the candidate must be computer literate. For example, many of the advertisements for translation positions did not specifically state that a candidate must have a good command of the source and target languages, though clearly it is not possible to produce a translation in the absence of such knowledge. In some cases, employers may have felt that, like linguistic competence, the ability to manipulate a computer is now such an essential skill for a translator to have that there was no need to mention it.¹⁰ While all of these explanations are viable, there is no way of knowing which, if any, are applicable to any given advertisement in the database.

3. Discussion

This article set out to investigate and compare the extent to which corpus-based translation resources are used in academic versus professional settings. As outlined in section 1, the use of various types of corpora (including conventional electronic corpora and translation memories) in translator training institutes has been well established since the late 1990s. Students are being educated in the design, construction and exploitation of corpus resources. For instance, translator trainers have successfully used corpora to teach students about a wide range of issues, including the identification of appropriate terminological or phraseological equivalents, the exploration of grammatical points, the investigation of stylistic preferences, the evaluation of the suitability of textual resources for a given translation, the

¹⁰ Carliner (2000), in a discussion of the technical writing industry, likens this type of essential skill to “hygiene”. He points out that while a technical writer assumes that good writing is the primary skill needed for the job, “employers view writing as ‘hygiene.’ That is, they assume that we can write, just as they assume we shower before we arrive at work. They only notice the absence of good writing skills, not the presence.”

investigation of cultural references, and the acquisition of conceptual information in specialized fields of knowledge. Based on this evidence, there can be little doubt that corpora are now considered to be valuable resources for both translation research and training in academic circles.

With regard to the use of corpora by professional translators, however, a slightly different picture emerges. Focussing on the situation in Canada, section 2 explored the extent to which professional translators' associations and employers are interested in the use of corpus-based resources.

Looking first at the professional literature, it appears that translators' associations are very aware of the existence of corpus-based tools in general; however, they clearly have more of an interest in translation memories than in conventional corpora. The fact that translators' associations do at least recognize that corpus-based tools exist is heartening as it means that they are keeping pace with the times and have moved beyond the state described by Haynes (1998:viii), who noted that during the late 1990s, many professional translators, and their organizations, were remarkably uninformed with regard to the progress made in translation technology. Haynes went on to observe that at that time, many were also largely unenthusiastic about it – with attitudes lying somewhere between sceptical and scathing. This point was echoed in a report compiled by the Canadian Translation Industry Sectoral Committee (CTSIC), who, in 1999, undertook a survey of the translation industry in Canada. In their final report, the Committee stated that the scepticism of translation professionals toward computer tools could be perceived as a threat to the expansion of the Canadian translation industry (1999:54). As noted in section 2.1, many of the articles published in the professional association literature were critical of some features of translation memory tools, which would seem to indicate that many translators do indeed maintain a healthy scepticism towards technology. However, this is not necessarily always a bad thing as it is important for any competent user to be aware of the limitations as well as the strengths of any tool. What is more, other articles indicate that translators seem to be very enthusiastic about the potential of a new generation of hy-

brid tools that combine features of both conventional corpora and translation memories. Of course, it is important to keep in mind that tools are generally received quite favourably when they are initially developed because they are marketed on the basis of their strengths. It is only through experience using the tools that translators may become aware of potential drawbacks. Therefore, while the hybrid tools are currently enjoying favour because they are the newest tools on the market, some flaws may become apparent once they have been in use for a longer period of time. Nevertheless, the fact that translators and their professional associations are at least willing to discuss translation technology at workshops and round tables, as well as in their newsletters and magazines, can be seen as a positive sign that translators are not simply shunning corpus-based tools without even evaluating their potential.

Turning next to the database of job advertisements, it was observed that 62.3% of the advertisements were seeking candidates who were computer literate in a general sense, but only 15.1% of the advertisements (most of which were for localization specialists) required candidates to be familiar with specialized translation technology such as translation memories or hybrid corpus analysis tools. The fact that only 5.8% of the advertisements for the job of translator were aimed at candidates with translation technology skills was surprising, particularly in light of the supposed need for technologically-skilled translators identified by the Canadian Translation Industry Sectoral Committee (1999:52), which noted in its report that:

The industry expects its professionals to be at ease in using translation-related technologies. This means having more than a theoretical knowledge, as they must have familiarity with various aids for translation. Increased levels of productivity achieved with the efficient use of computerized tools will increasingly be a matter of survival for translation firms.

Based on this statement, one might expect that in the three years since the Committee published the report, there would be

TRADTERM, **10**, 2004, p. 213-247

a considerable number of employers seeking translators with specialized technological skills; however, the data found in the database of job advertisements do not fully support this statement since only 5.8% of the advertisements were seeking translators with such skills.

In trying to answer the question of why corpora and corpus-based tools have received a somewhat less enthusiastic welcome in the professional world than in the academic world, we must consider several possibilities. One factor that has led to a difference in uptake in these two sectors is the ease of access to such tools. This includes both knowledge of the existence of the tools as well as the ability to afford the tools.

Firstly, it should be noted that corpora have long been used in fields such as language teaching or second-language learning (e.g. Johns, 1986; Mindt, 1986), but it is only more recently that their potential as translation aids has been recognized. Academics working in the field of translation are often involved in, or have colleagues who are involved in, language teaching, and as such they may have gained exposure to corpus-based techniques in this way. Many of the existing conventional corpus analysis tools were initially developed by academics who work in language teaching¹¹ simply as a means of helping their own students. This means that while such tools are generally very reasonably priced and may be easily accessible within the academic community, they are not usually widely advertised or distributed to the professional translation community because the people who have created these tools have full-time teaching jobs and are not necessarily interested in becoming big-league software developers or vendors as this would entail providing products that are bug free, have substantial documentation, and come with technical support. In contrast, tools such as translation memories and the new generation of hybrid tools, which have been devel-

¹¹ For example, WordSmith Tools was developed by Dr. Mike Scott, who works in the Department of English Language and Literature at the University of Liverpool in the United Kingdom; MonoConc and ParaConc were both developed by Dr. Michael Barlow, who works in the Linguistics Department at Rice University in the United States.

oped in the private sector by companies that have professional full-time programmers, technical support staff and generous advertising budgets, are actively marketed to working professionals. The fact that conventional corpus analysis tools do not seem to be well advertised in the professional setting may explain, in part, why translators' associations and employers seem to be more aware of the existence of translation memory tools and hybrid tools than they are of conventional corpus analysis tools. This situation may change in the future, however. As established in section 1, the use of conventional corpus-based tools at translator training institutes has become popular since the late 1990s. This means that, at present, most of the translators in the workforce will have received their education during a time when corpus analysis tools were not part of the translator training curriculum. However, over the coming years, the number of technologically-trained graduates will increase and they will bring to the workforce their knowledge of corpus-based techniques. They will be able to share their experience with their senior colleagues and employers and gradually, more and more companies will have translators on staff who have an understanding of such tools.

With regard to affordability of tools such as translation memories and hybrids, which are relatively expensive, educational institutions are often able to negotiate discounted rates (or sometimes even obtain free copies of this type of specialized translation software in exchange for user feedback or beta testing) which may help to explain why these tools are becoming so prevalent in academic circles. In contrast, although professional translators are clearly aware of the existence of translation memories and hybrids, the high price of these tools for private sector users may be one of the reasons why a limited number of job advertisements are calling for translators who can use this technology. Perhaps many employers still feel that the long-term benefits of using such tools have not yet been sufficiently proven and so they are not yet willing to make an initial investment in these expensive products. However, there is a general trend in most areas of computing for prices to come down once products become established. Therefore, once it has been satisfactorily demonstrated that, over the long term, translation memories and

TRADTERM, **10**, 2004, p. 213-247

hybrids can increase productivity without compromising quality, an increasing number of employers who are currently sitting on the fence may be convinced to invest in such tools. In summary, with regard to ease of access of corpus-based tools, we can say that on the one hand, conventional corpus tools are very affordable and are known within academic circles, but are not well advertised to the professional community, while on the other hand, translation memory and hybrid tools are well advertised in both academic and professional circles, but they are more expensive for professional users to obtain. These two sets of conditions give rise to a situation where translator training institutes have relatively good knowledge of and easy access to all kinds of corpus-based tools, whereas professional translators may be less knowledgeable about conventional corpus analysis tools and less able to afford translation memories or hybrids.

Another reason that may help to explain why corpus-based tools are currently more popular in academic institutions than in professional institutions is linked to the different goals of these institutions. Firstly, there is a greater number of relevant applications for corpus-based techniques within academics. For example, research into the nature of translation is the remit of academic institutions. In the professional translation world, the production of translations is the primary goal, and research is typically only of peripheral interest. Another important aim of an academic institution is to provide training for students. This includes imparting knowledge as well as simply teaching technological skills. O'Brien and Kenny (2001:22), who are involved in the translator training program at Dublin City University in Ireland, ask the following questions:

Is it our job to train translators how to use the leading translation memory tool so as to increase their chances of employment on graduation? Or is it our job to impart knowledge of the technology on a more general level and to equip the students with the ability to evaluate and to learn the use of these tools themselves?

They conclude that the latter is more appropriate for an educational setting. Since corpora have proved to be exception-

ally versatile resources that can be used to help teach a wide range of translation-related issues (e.g. evaluation of sources, comparison of different solutions to translation problems, learning to evaluate and use computer tools, learning to identify characteristics of different text types, learning to identify appropriate terms/styles/grammatical usage, etc.), it is worthwhile for students to devote time to tasks such as learning how to design and build corpora, and learning how to use, compare and evaluate a selection of different tools. Furthermore, once students have acquired this type of basic knowledge, they can apply it in other contexts where corpora are not used (e.g. translators can evaluate other types of sources, texts, tools, etc.). Professional translators, meanwhile, are generally expected to come to a job equipped with the knowledge that will allow them to produce translations. Of course, translation is a profession that requires lifelong learning, which means that professional translators are always acquiring new skills, but the responsibility for providing basic translation-related knowledge typically falls to translator training institutes rather than to employers.

One important reason that professional translators are expected to come to a job with this essential knowledge already in place is because they have to work in a high-pressure environment. They are often faced with very tight deadlines that require them to work under serious time constraints. In contrast, in academic circles, deadlines are more generous so it may be feasible for a student to spend a week designing and building a corpus that can be used to help translate a 500-word text because that student is simultaneously learning other valuable skills. Furthermore, students are encouraged to compare and evaluate the strengths and weaknesses of different tools, which means that translator training institutes do not feel the same kind of pressure to identify a single tool to act as a panacea. Ironically, a poorly-designed or problematic tool can often be very informative in an academic context since it can be used to identify or demonstrate translation pitfalls. In a professional setting, however, a translator faced with a 500-word text to translate is more likely to have a deadline of only a few hours, so there is little opportunity for that translator to set about constructing a useful corpus or comparing different tools. Researchers such as

TRADTERM, **10**, 2004, p. 213-247

Bowker (1998:648; 2001:205), Pearson (2000b:550), and Bowker and Pearson (2002:70) have commented that while corpora can certainly be extremely valuable translation resources, the design, compilation and effective exploitation of corpora can sometimes be a very time-consuming process and so it is not necessarily feasible to employ corpus-based techniques in all translation settings. The fact that productivity is the bottom line in a professional setting may also explain why the interest of professional translators seems to lie more with translation memory and hybrid tools rather than with conventional corpus analysis tools. Translators cannot afford to invest time in building up resources or learning how to use tools that do not help them to work more efficiently. Translation memory-type resources can be compiled as a translator is working on a text, whereas conventional corpora typically need to be created as a separate step prior to beginning the translation process. Furthermore, searches are carried out automatically by translation memory and hybrid tools, while conventional corpus analysis tools require translators themselves to perform the extra steps of deciding upon and entering the search patterns. In fact, this need among professionals to be extremely selective with regard to learning and using tools may explain the recent surge of interest in hybrids since these tools have the potential to overcome some of the existing limitations of both conventional corpus analysis tools and translation memories.

Another way in which an academic setting may differ from a professional setting is in the nature and range of text types that are translated. As pointed out by Fraser (1996), in an academic setting, translator trainers often find themselves working with material that covers a variety of different subjects and text types in order to expose students to a wide range of translational issues. When working with such a range of texts, it is likely that at least some of them can be usefully processed with the help of corpus-based techniques (e.g. texts that are very repetitive can be processed with the help of a translation memory). This means that students will have the opportunity to experiment with corpus-based tools on suitable texts. With regard to professional translators, although some do work with a wide range of subjects and text types, they are much more likely to special-

ize and to work with a limited range. This is the situation outlined by Senécal (2002) when describing the case of expert translators working at the Translation Bureau of the Government of Canada. This means that while some translators might regularly work in an environment where material is repetitive or limited to a given subject or text type, and thus could be usefully processed with the help of a translation memory or corpus-based tool, many other translators may only rarely encounter texts that can be usefully processed in this way. For instance, a literary translator will probably not have any need for a translation memory, and a freelance translator or translation agency that deals with a different subject on every job may not find it worthwhile to build up a corpus to help translate a single text. In an environment where corpora or corpus-based tools will not be used on a regular basis, it will probably not be worthwhile for a company to invest money and time in buying such tools and training employees to use them. Of course, repetition between texts is likely to occur more often at the sub-sentential level rather than the sentence level, and this may result in a greater interest in hybrid tools among professional translators since hybrid tools provide the benefits of automation and speed with an increased likelihood of getting term or phrase matches even in texts that do not contain repetition at the sentence level.

4. Concluding remarks

Clearly corpora and corpus-based tools are present in both the academic sector and the professional sector, but not to the same degree. While the use of corpora in professional circles is likely to increase – particularly with the introduction of new hybrid tools such as MultiTrans and LogiTerm – it is unlikely that corpora will ever achieve the same degree of popularity that they enjoy in academic circles. Some of the existing barriers to the use of corpora by professionals are likely to be eradicated. For example, prices may drop, and as more corpus-trained graduates join the workforce, knowledge of how to effectively exploit such tools will be more widespread. Nevertheless, there will always remain some important differences between the academic and

TRADTERM, 10, 2004, p. 213-247

professional sectors, such as their different goals (e.g. research and the transfer of knowledge in academics vs. practice and productivity in the workplace) and different time pressures, which will likely mean that corpora will continue to be more prevalent in academic circles. This is not necessarily a negative thing, however, since much of the knowledge acquired through corpus-based training at an academic institute can be easily transformed into a set of valuable transferable skills that can be usefully applied in other contexts and circumstances in the professional workplace.

Acknowledgements

I would like to thank Peter Bennison, who assisted with the development of the database of job advertisements, and Melissa Ehgoetz and Michèle Healy, who assisted with a portion of the data collection and entry.

References

- ARROUART, C.; BÉDARD, C. (2001) Éloge du bitexte. *Circuit* 73, 30.
- ATKINS, B.T.S.; CLEAR, J.; OSTLER, N. (1992) Corpus design criteria. *Literary and linguistic computing* 7(1), p. 1-16.
- AUSTERMÜHL, F. (2001) *Electronic tools for translators*. Manchester, St. Jerome Publishing.
- BAKER, M. (1995) Corpora in translation studies: an overview and some suggestions for future research. *Target* 7(2), p. 223-43.
- _____. (1996) Corpus-based translation studies: the challenges that lie ahead. In: Somers, H. L. (ed.) *Terminology, LSP and translation: studies in language engineering in honour of Juan C. Sager*. Amsterdam, John Benjamins, p. 175-86.
- BÉDARD, C. (2001) Une nouvelle profession: traducteur de phrases. *Circuit* 70, 29.
- BÉLISLE, G. (2001a) Translating technology. *InformATIO: newsletter of the Association of Translators and Interpreters of Ontario* 30(2), 6.
- _____. (2001b) Text alignment tools and translation memories. *InformATIO: newsletter of the Association of Translators and Interpreters of Ontario* 30(4), 7.

TRADTERM, **10**, 2004, p. 213-247

- BERNARDINI, S. (2002) Educating translators for the challenges of the new millenium: the potential of parallel bi-directional corpora. in Maia, B., Haller, J.; Ulrych, M. (eds.) *Training the language services provider for the new millenium*. Porto, Faculdade de Letras da Universidade do Porto, p. 173-86.
- BOWKER, L. (1998) Using specialized monolingual native-language corpora as a translation resource: a pilot study. *Meta* 43(4), p. 631-51.
- _____. (2000) A corpus-based approach to evaluating student translations. *The Translator* 6(2), p. 183-210.
- _____. (2001) Towards a methodology for a corpus-based approach to translation evaluation. *Meta* 46(2), p. 345-64.
- _____. (2002a) Working together: a collaborative approach to diy corpora. In: Yuste-Rodrigo, E. (ed.). *Proceedings of the Workshop on Language Resources in Translation Work and Research*. Paris, European Language Resources Association (ELRA), p. 29-32.
- _____. (2002b) *Computer-aided translation technology: a practical introduction*. Ottawa, University of Ottawa Press.
- BOWKER, L.; PEARSON, J. (2002) *Working with specialized language: a practical guide to using corpora*. London/New York, Routledge.
- CADIEUX, F. (2000) Toronto: focus on professional development. *InformATIO: newsletter of the Association of Translators and Interpreters of Ontario* 29(4), 3.
- Canadian Translation Industry Sectoral Committee (1999) *Survey of the Canadian translation industry: human resources and export development strategy*. Final Report of the Canadian Translation Industry Sectoral Committee. <http://www.uottawa.ca/associations/csict/rap-e.htm#stre>
- CARLINER, S. (2000) Trends in our business 2000: thriving in the boom years. *Intercom: the magazine of the Society for Technical Communication*, January 2000. <http://saulecarliner.home.att.net/idbusiness/trends2000.htm>
- COHEN, B. (2002) Mémoires et tarification, un débat à finir. *Circuit* 76, p. 16-7.
- DANDURAND, H. (2002) Le portrait-robot: Les mémoires de traduction. *L'antenne* 33(1).
- DECESARIS, J.A. (1996) Computerized translation managers as teaching aids. In: Dollerup, C.; Appel, V. (eds.) *Teaching Translation*

TRADTERM, **10**, 2004, p. 213-247

- and *Interpreting 3: New Horizons*. Amsterdam/Philadelphia, John Benjamins, p. 263-9.
- DUARTE, M. (2001) JMT 2001 à Montreal: Mémoires de traduction. *L'antenne* 32(8).
- EBELING, J. (1998) Contrastive linguistics, translation, and parallel corpora. *Meta* 43(4), p. 602-15.
- EVANS, O. (2002) ATIO offers members professional development on concordancing tools. *InformATIO: newsletter of the Association of Translators and Interpreters of Ontario* 31(2), 7.
- FRANKENBERG-GARCIA, A. (2002) COMPARA – language learning and translation training. In: Maia, B.; Haller, J.; Ulrych, M. (eds.) *Training the language services provider for the new millenium*. Porto, Faculdade de Letras da Universidade do Porto, p. 187-98.
- FRASER, J. (1996) Professional versus student behaviour. In: Dollerup, C.; Appel, V. (eds.) *Teaching translation and interpreting 3: new horizons*. Amsterdam/Philadelphia, John Benjamins, p. 243-50.
- HANSEN, S.; TEICH, E. (2002) The creation and exploitation of a translation reference corpus. In: Yuste-Rodrigo, E. (ed.). *Proceedings of the Workshop on Language Resources in Translation Work and Research*. Paris, European Language Resources Association (ELRA), p. 1-4.
- HAYNES, C. (1998) *Breaking down the language barriers*. London, Aslib.
- JOHNS, T. (1986) Microconcord: a language learner's research tool. *System* 14(2), p. 151-62.
- KENNY, D. (1999) CAT tools in an academic environment: what are they good for? *Target* 11 (1), p. 65-82.
- _____. (2001) *Lexis and creativity in translation: a corpus-based study*. Manchester, St. Jerome Publishing.
- LANCTÔT, F. (2001) Splendeurs et petites misères... des mémoires de traduction. *Circuit* 72, p. 30.
- _____. (2002) LogiTerm: tekno sans amertume. *Circuit* 75, p. 30-1.
- LAVALLÉE, F. (2002) MultiTrans, un outil nouveau genre. *Circuit* 74, 32.
- LAVIOSA, S. (ed.) (1998a) L'approche basée sur le corpus/The corpus-based approach. Special issue of the journal *Meta* 43(4). Montreal, Les Presses de l'Université de Montréal.
- _____. (1998b) The English comparable corpus: a resource and a methodology. In: Bowker, L.; Cronin, M.; Kenny, D.; Pearson, J. (eds.) *Unity in diversity? Current trends in translation studies*. Manchester, St. Jerome Publishing, p. 101-12.

- _____. (2002) *Corpus-based translation studies: theory, findings, applications*. Amsterdam, Rodopi.
- L'HOMME, M.-C. (1999) *Initiation à la traductique*. Brossard, Québec, Linguattech.
- LINDQUIST, H. (1999) Electronic corpora as tools for translation. In: Anderman, G.; Rogers, M. (eds.) *Word, text, translation: liber amicorum for peter newmark*. Clevedon, UK, Multilingual Matters, p. 179-89.
- LÓPEZ RODRÍGUEZ, C.I. (2002) Training translators to learn from news report corpora: the case of anglo-american cultural references. In: Maia, B.; Haller, J.; Ulrych, M. (eds.) *Training the language services provider for the new millenium*. Porto, Faculdade de Letras da Universidade do Porto, p. 213-22.
- MAIA, B. (1998) Word order and the first person singular in portuguese and english. *Meta* 43(4), p. 589-601.
- _____. (2000) Making corpora: a learning process. In: Bernardini, S.; Zanettin, F. (eds.) *I corpora nella didattica della traduzione*. Bologna, CLUEB, p. 47-60.
- MALMKJAER, K. (1998) Love thy neighbour: will parallel corpora endear linguists to translators? *Meta* 43(4), p. 534-41.
- MINDT, D. (1986) Corpus, grammar and teaching english as a foreign language. In: Leitner, G. (ed.) *The English reference grammar: language and linguistics, writers and readers*. Tübingen, Niemeyer, p. 125-39.
- O'BRIEN, S.; KENNY, D. (2001) In Dublin's fair city: teaching translation technology at dublin city university. *Language international* (October 2001), p. 20-3.
- OLOHAN, M. (2002) Comparable corpora in translation research: overview of recent analyses using the translational english corpus. In: Yuste-Rodrigo, E. (ed.) *Proceedings of the Workshop on Language Resources in Translation Work and Research*. Paris, European Language Resources Association (ELRA), p. 5-9.
- OLOHAN, M. (forthcoming) *Corpora and translation studies*. London, Routledge.
- PALUMBO, G. (2002) The use of phraseology for training and research in the translation of LSP texts. In: Maia, B.; Haller, J.; Ulrych, M. (eds.) *Training the language services provider for the new millenium*. Porto, Faculdade de Letras da Universidade do Porto, p. 199-212.
- PEARSON, J. (2000a) Surfing the internet: teaching students to choose their texts wisely. In: Burnard, L.; McEnery, T. (eds.) *Rethinking*

TRADTERM, **10**, 2004, p. 213-247

- language pedagogy from a corpus perspective*. Frankfurt am Main, Peter Lang, p. 235-9.
- PEARSON, J. (2000b) Using specialized comparable corpora to evaluate student translations. In: Lewandowska-Tomaszczyk, B.; Melia, P.J. (eds.) *PALC'99: Practical applications in language corpora: papers from the international conference at the University of Łódź*, p. 541-52.
- PEARSON, J. (2000c) Une tentative d'exploitation bi-directionnelle d'un corpus bilingue. *Cahiers de grammaire* 25 (special edition on "Sémantique et corpus", guest edited by Condamines, A.), p. 53-69.
- QUIRION, J. (2000) Mémoires de traduction. *L'antenne* 31(8).
- ROSEN, N. (2001) Translation technology showcase at the University of Toronto. *InformATIO: newsletter of the Association of Translators and Interpreters of Ontario* 30(1), 3.
- SCHMIED, J. (2002) A translation corpus as a resource for translators: the case of english and german prepositions. In: Maia, B.; Haller, J.; Ulrych, M. (eds.) *Training the language services provider for the new millenium*. Porto, Faculdade de Letras da Universidade do Porto, p. 223-38.
- SCHWAB, W. (2001) Translation memories in transition. *Circuit* 71, 30.
- SENÉCAL, A. (2002) Traducteur expert: le Groupe des Sept. *Circuit* 75, 20.
- TAGNIN, S.E.O. (2002) Corpora and the innocent translator: how can they help him? In: Lewandowska-Tomaszczyk, B.; Thelen, M. (eds.) *Translation and meaning, Part 6*. Maastricht, Hogeschool Zuyd, Maastricht School of Translation and Interpreting, p. 489-96.
- TAGNIN, S.E.O. (org.) (2003) Tradução e corpora. Special issue of the journal *Cadernos de Tradução* no. 9 (2002/1). Florianópolis, Brazil, Núcleo de Tradução, Universidade Federal de Santa Catarina.
- VARANTOLA, K. (forthcoming) Translators and disposable corpora. In: Zanettin, F.; Bernardini, S.; Stewart, D. (eds.) *Corpora in translator training*. Manchester, St. Jerome Publishing.
- ZANETTIN, F. (1998) Bilingual comparable corpora and the training of translators. *Meta* 43(4), p. 616-30.
- ZANETTIN, F. (2001) Swimming in words: corpora, translation, and language learning. In: Aston, G. (ed.) *Learning with corpora*. Bologna/Houston, CLUEB / Athelstan, p. 177-97.
- ZANETTIN, F. (2002) DIY corpora: the www and the translator. In: Maia, B.; Haller, J.; Ulrych, M. (eds.) *Training the language services provider for the new millenium*. Porto, Faculdade de Letras da Universidade do Porto, p. 239-48.

Corpus-based tools

Conventional corpus analysis tools

MonoConc: <http://www.athel.com>

MultiConcord: <http://www.copypatch.freemove.co.uk/multicon.htm>

ParaConc: <http://www.athel.com>

WordSmith Tools: <http://www.lexically.net/wordsmith/>

Translation memory systems

Déjà Vu: <http://www.atril.com>

SDLX: <http://www.sdlintl.com/sdlx>

STAR Transit: <http://www.star-transit.com>

Trados: <http://www.trados.com>

Hybrids

Logiterm: <http://www.terminotix.com>

MultiTrans: <http://www.multicorpora.com>