

A photograph of a person with dark hair and glasses, seen from the side, working at a computer. The person is wearing a dark jacket and is looking at a laptop screen. The background is a plain, light-colored wall.

LogiTerm Part II

By Naomi J. Sutcliffe de Moraes

Please note that the reviewer is not affiliated in any way with the products being discussed. This review does not represent an endorsement of the product by ATA. All opinions are solely those of the author.

In the November/December issue I described what LogiTerm does, its search interface, and how it stores and searches terminology files. For those who did not read Part I of my review of this tool, I will just summarize these points here.

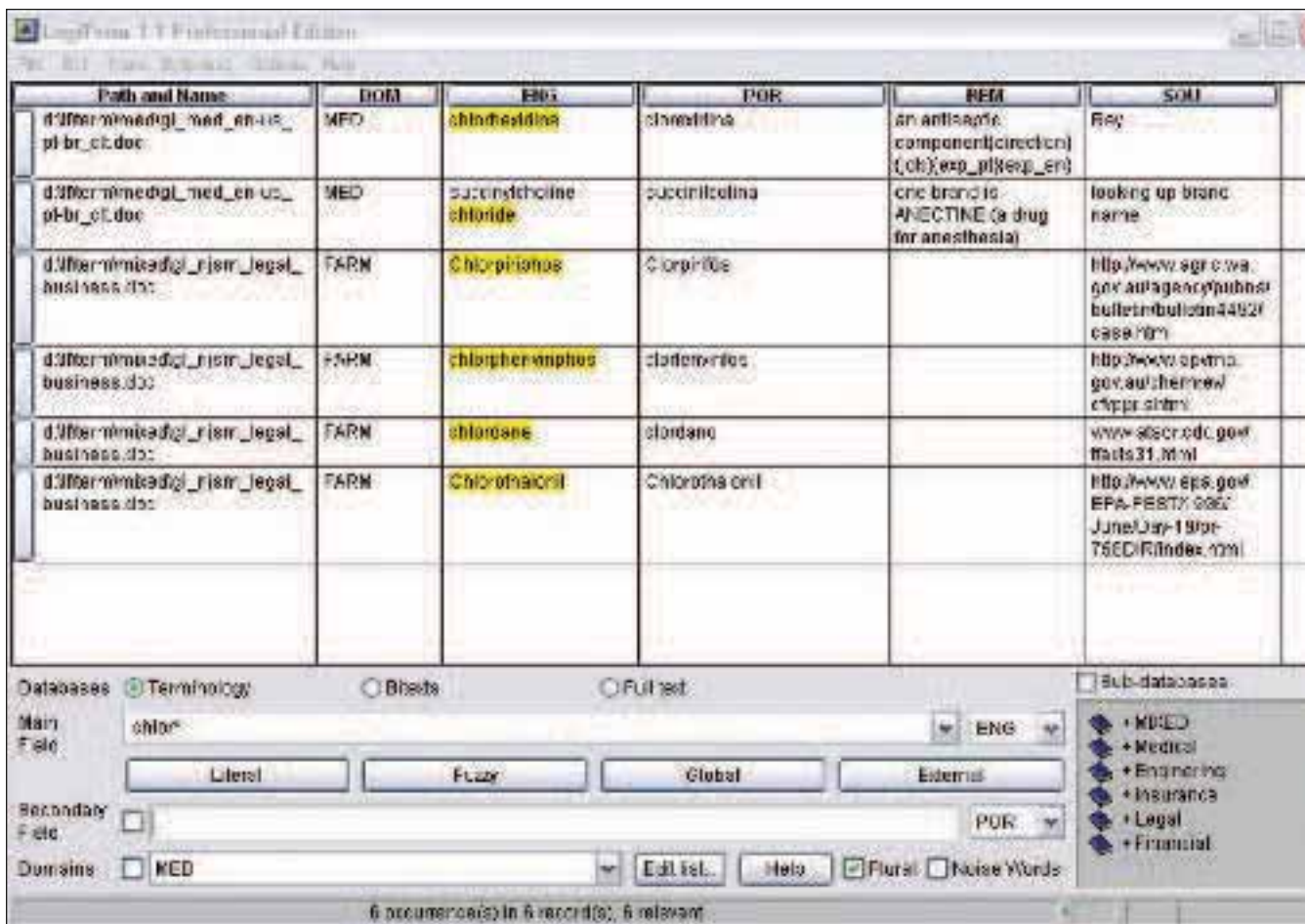
LogiTerm indexes terminology files, aligns translations and reference files, and allows you to search these files from a single interface, as shown in Figure 1 on page 23. The results shown here are for a fuzzy search on “chlor*” in the terminology database. Double-clicking on a cell will copy its contents to the clipboard. (See the November/December issue for more information on how LogiTerm stores data and searches for information.)

Bitext Files

Creating and searching terminology files was addressed in Part 1 of this review. The second type of file that can be indexed and searched is what LogiTerm calls a bitext, which is a file containing an original text aligned with its translation. LogiTerm stores bitexts in HTML (HyperText Markup Language), in a two-column format. Its alignment tool is extremely accurate, and I rarely have a problem with the resulting bitext alignment. The advantage of searching bitexts rather than a translation memory is that you get more than just a random segment—you can click on the segment and see the entire paragraph, or even the entire document. LogiTerm is not the only corpus-based tool. You can find a general discussion of this type of tool and others in the article “Taking the Plunge” in the June and July 2007 issues.

Bitext Generation: LogiTerm’s automatic alignment feature is excellent, as it uses an algorithm that takes into

Figure 1: LogiTerm search interface showing fuzzy search results



account more than just punctuation. It can align up to 25 pairs of files at a time. LogiTerm does not currently include a manual alignment tool, though I have never needed one. Bitext files are stored in HTML format so they can be edited with MS Word or an HTML editor. The only time it did not align two files correctly, I realized I had left an entire page out of one document when converting it from paper. If you are interested specifically in automatic alignment, see Terminotix's AlignFactory tool. It has a manual editor and has no restriction on the number of files. Note that LogiTerm can index—and align—MS Word, WordPerfect, HTML, Excel, PowerPoint, and PDF files (when not created from an image).

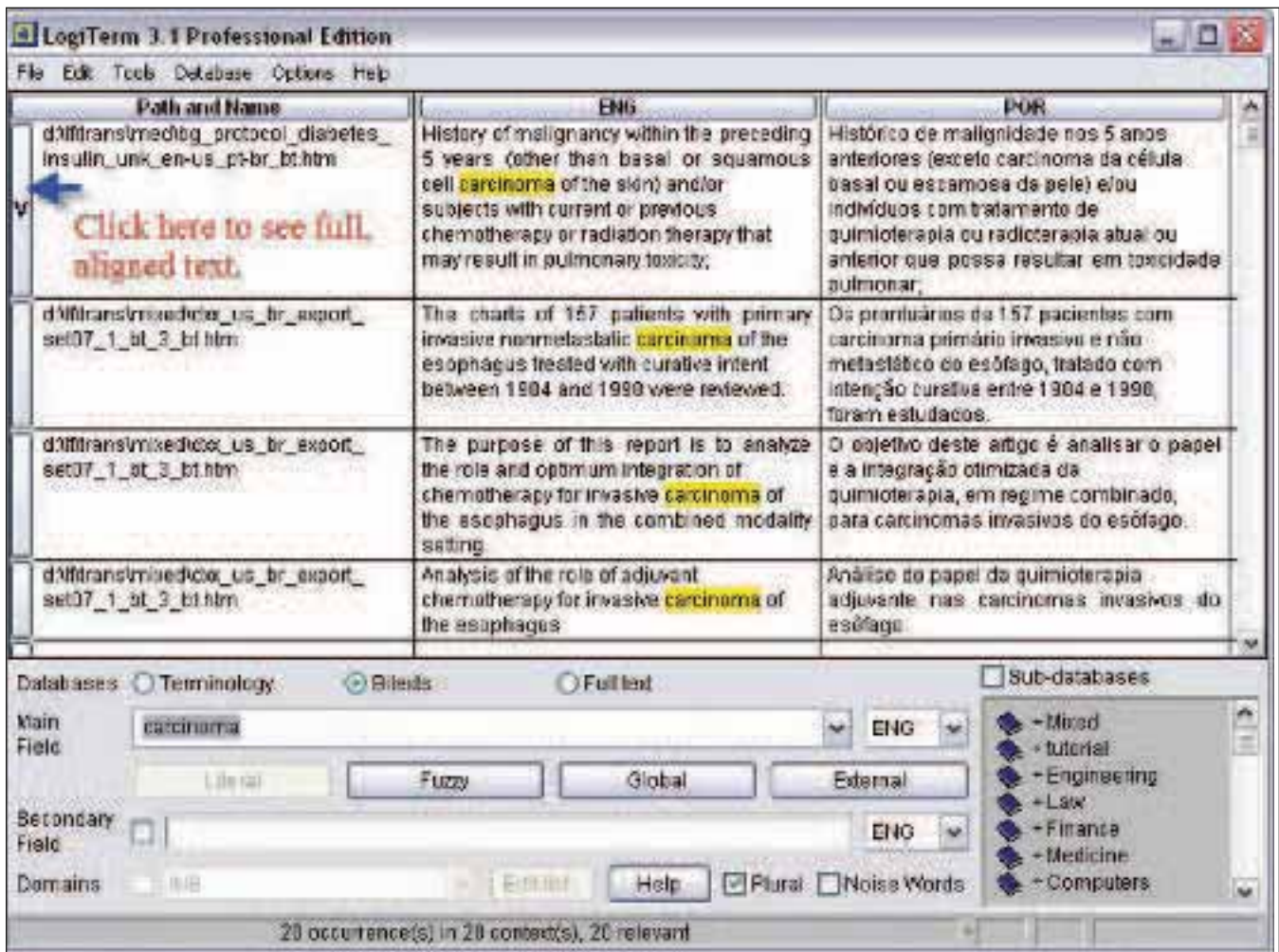
These files do not need to be converted before indexing. Having such a reliable bitext generator led me to search on the Brazilian and Portuguese government websites for official translations of laws, and I was able to align them and add them to my bitext collection.

Translation Memory Conversion: LogiTerm can import data created in translation environment tools such as SDLX, SDL Trados, and Déjà Vu X (DVX) if it is first exported to the Trados or TM/2 formats by these tools. This feature was very convenient for me because I have years of data stored in translation environment tools, and I was able to import it in two steps. LogiTerm can also export

aligned files into the Trados or TM/2 format for subsequent import by other tools, and its automatic alignment tool is much faster than the manual alignment tool provided with most translation environment tools.

Editing Visualization Tool: If you are sent a translation to edit, with the source text in one file and the translation in another, you might consider using the bitext generation feature to create an aligned file, since reviewing an aligned file makes it easier to spot errors. This two-column format is one of the reasons so many translators like translation environment tools like across, DVX, and SDLX. (See the article "Taking the Plunge" in the

Figure 2: Bibtex search results



June and July issues for a more complete description of translation environment tools.)

Search Results Table: Figure 2 shows the result of a bibtex search for *carcinoma* in English. The order of the result grid columns can be configured, and does not indicate the order in which the file was translated. I use a code in the filename to tell me which language is the source language. The search results table shows only the original text and its translation, but clicking on the bar on the left opens an unformatted copy of the entire aligned file, providing more context for the segment. You can also search for content in a second field.

LogiTerm also provides a way for you to create terminology records from the bibtex search results table or from a bibtex file open in MS Word,

LogiTerm's automatic alignment feature is excellent, as it uses an algorithm that takes into account more than just punctuation.

which is useful when reviewing translations done by others that are aligned as reference files. The name of the bibtex file from which the term was extracted is automatically inserted in the terminology record.

As with terminology files, aligned translations do not need to be in LogiTerm's native format to be indexed. You can throw a file in any format into a folder containing bibtexs and tell LogiTerm to index it. When search results are shown, the two lan-

guage columns will be blank, but the path and filename will be displayed. Clicking on the bar to the left will open the original file at the point where the search word appears, showing the original and the translation.

Comparison with Translation Memory Tools: Searching the translation memories of translation memory tools is somewhat complicated if you are not translating in the tool (either because the source text is not in electronic

Figure 3: Déjà Vu X TM search results

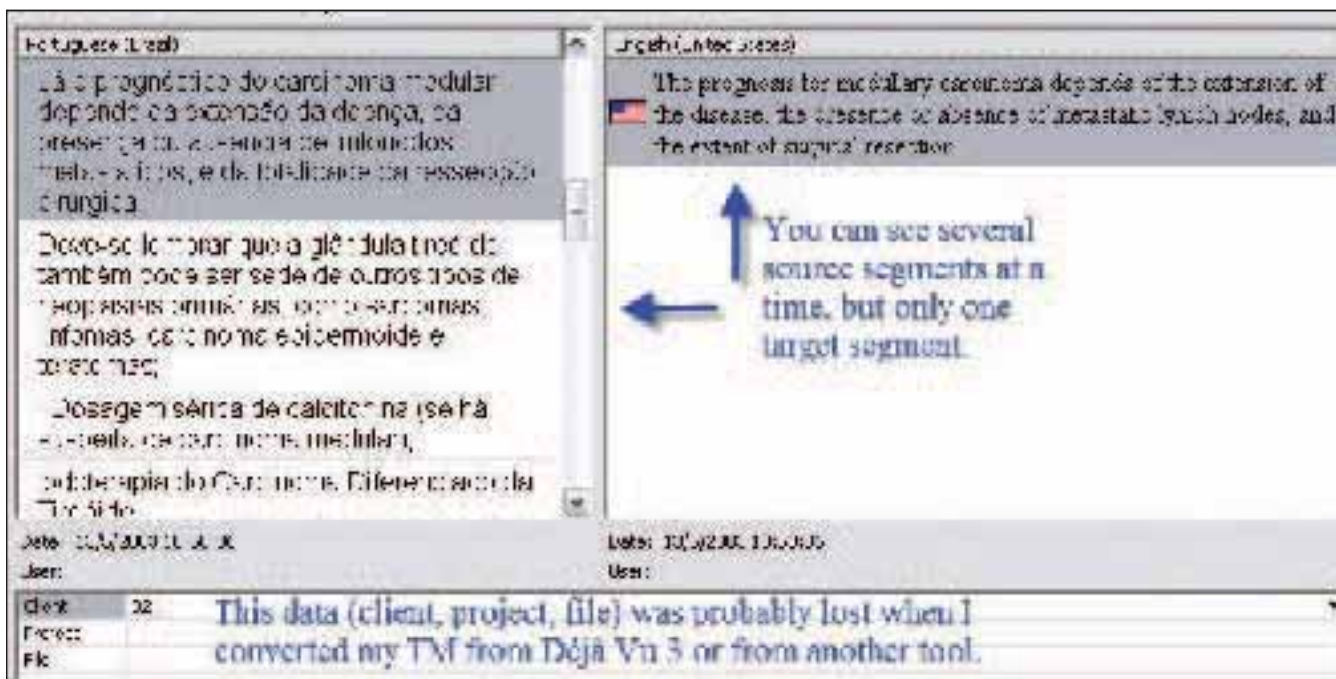
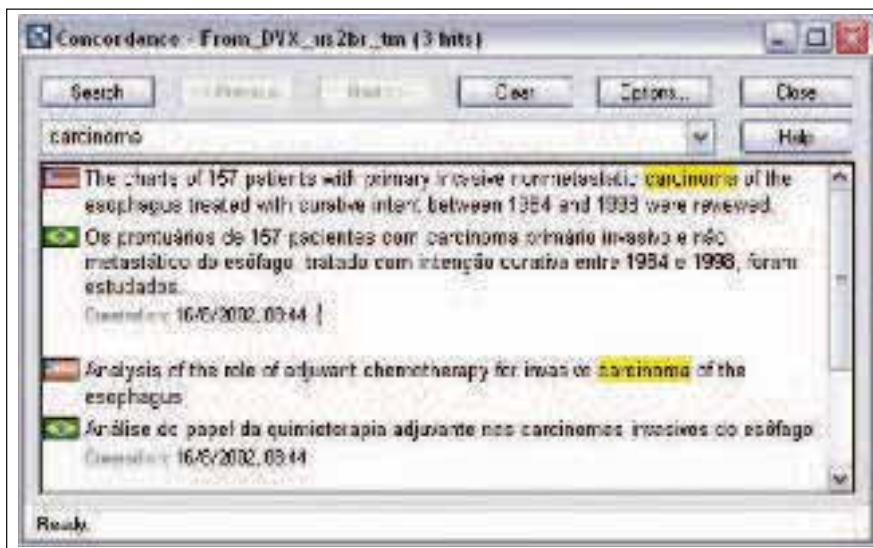


Figure 4: SDL Trados TM search results



format, or because the client requested that a specific tool be used and you normally use a different tool). DVX's TM lookup interface is shown in Figure 3. I rarely used it even before I purchased LogiTerm, although I have had DVX for years. Its search mechanism allows you to use SQL (a structured query language designed for the retrieval and management of data in relational database management systems) commands to perform complicated searches. However, despite having studied SQL a few years ago, I never remember the

commands and only perform simple searches. First, in DVX, the left column shows the segments found based on my search on *carcinoma*, but the word (*carcinoma*) is not highlighted in any way. Second, the right column shows the translation for only the selected cell on the left. So, you must click-click-click to see the different translations one by one, rather than just scrolling as in LogiTerm. The data in the bottom window would provide some context information (client, project, filename) if it had not been lost somewhere along

the way. Another advantage of LogiTerm is that you can put as much or as little information as you want in the filename to indicate external information.

SDL Trados' TM lookup interface is shown in Figure 4. It allows only simple searches, and shows only date/time information and the name of the user who added the segment to the TM if the segment was created from within SDL Trados (rather than imported, as these two entries were.). The name of the file or project is not shown. The interface does allow you to scroll—rather than click—through the segments.

In summary, the main differences between LogiTerm's bitext search and the two translation environment tools SDL Trados and DVX are:

1. LogiTerm's automatic alignment tool is extremely efficient. I have used it to align hundreds of files with no errors and no manual interaction.
2. LogiTerm allows you to use special search functions (described in the November/December 2007 issue), such as searching for quotes, wild cards, or for one word in the source text and another word in the target text (only segments matching ➡

Figure 5: LogiTerm toolbar in MS Word



both criteria are shown, and both words are highlighted).

3. LogiTerm displays the results from most appropriate to least appropriate, which is similar to Google. Highest priority is given to the exact sequence of words in the search field, then for the search words, but in any order, and then for any of the words in any order.
4. LogiTerm shows the results in a two-column format that you can scroll through quickly.

Reference Files

LogiTerm can index reference files in many formats, including PDF files (see the section on bitext generation earlier in this review). There is no standard format for reference files, so LogiTerm has no way of knowing which language is in a given file. I

usually put the language and/or country in the name of the file, or create separate folders/sub-databases for different languages/countries.

When I studied translation, one of my teachers taught me the importance of background reading in the target language to absorb terminology, register, and word collocations. Searching reference texts using LogiTerm allows me to speed up this process, rather than having to read through entire reference documents. One example of how I use reference files is when translating patient information leaflets for drugs. Sometimes I can find a patient information leaflet in the target language online or in print for a similar drug, or even for the same drug manufactured by another company. In these cases, the unknown original text will not be exactly the same as my original text, but the translation will be similar to what I

must produce, so the terminology, register, and collocations will be helpful. I search my target-language reference texts on one or two key words and I am taken to the file(s) with similar text. If I were to use Window's search tool, it would only tell me that a certain file in the searched folder contained the search word. I would then have to open the file and do another search within the file for each occurrence.

Another example is when I translate diagnostic laboratory results. The physician is writing about a specific patient, but I can usually find information online about the exam performed and how to understand the results. I stash this away in my medical reference folder and can call it up by searching on the name of the exam. (See "Teaching Medical Translation into English" in the January 2004 issue for more information on how to

Figure 6: Sample output (with standard color coding) after pretranslation of terminology

Example 1: Translated terms added in brackets

The Parties [parte-] agree [concordar] that any dispute [controvérsia+], claim [reivindicação] or controversy arising under [nos termos de], out of, or in relation to this Agreement [contrato] shall [deverá] be submitted for adjudication and/or settlement by arbitration proceedings [autos-+] in accordance with [de acordo com] the Rules [norma-] of the American Arbitration Association, and any determination thereon shall [deverá] be binding [vinculante] upon the Parties [parte-] hereto with the same force and effect as if rendered by a court of competent jurisdiction [juízo competente], and judgment [decisão+] thereon may [poderá] be entered by any Party [parte].

Example 2: Source text replaced by translated terms

The [parte- concordar] that any controvérsia+, reivindicação, or controversy arising nos termos de, out of, or in relation to this contrato deverá be submitted for...

Example 3: Terms found in terminology database underlined

The Parties agree that any dispute, claim, or controversy arising under, out of, or in relation to this Agreement shall be submitted for adjudication and/or settlement by...

Example 4: Terms found in reference bitexts or monolingual reference documents

No judgments, liens, or security interests will be outstanding at the time of the closing against the Seller or against its business or any assets thereof, except those to be paid and discharged out of the purchase price at closing and approved by the Purchaser's attorney.

use reference texts in medical translation.)

MS Word Toolbar

The MS Word toolbar (see Figure 5 on page 26) is not actually needed to search the LogiTerm databases or to create terminology files in the native format. However, it does provide functions that speed some things up. For example, you can select a word in MS Word and click on a button, which

terms marked, as shown in Figure 6 on page 26.

In the first example in Figure 6, a sentence in English has been pretranslated and the terms found in the specified terminology files (in Portuguese in this example) have been added in brackets next to the English. The first term (Parties) exists in the database in the singular, which the program indicates by adding a tilde to the translated term: [*parte~*]. Note that the

ture allows you to use terminology files specific to a client or type of document for terminology pretranslation, and can be used by an agency to indicate the terminology that should be used when sending out a job, or in-house by the lead translator when working with interns.

In the second example in Figure 6, LogiTerm replaced the source term with the translation. This kind of substitution could be useful when the source and target languages have similar grammatical structures, and/or when the translator translates many similar documents and has well-defined terminology.

In the third example in Figure 6, LogiTerm simply indicated that a match was found in the database. For the first two examples, I told LogiTerm to look only at legal terminology, but in this example I told it to look at all terms, and it underlined a few extra words, such as “settlement” and “by.”

In the fourth example, LogiTerm looked for terms not in the terminology records, but rather in the bitexts. You can configure LogiTerm to look in both the terminology records and bitexts at the same time, in effect combining Examples 3 and 4, with blue underlining for terms and green underlining for bitext matches. In most translation environment tools, you must select terms and click a button to ask the tool to perform this type of search (called Scan in DVX, Concordance in SDL Trados, and Concordance or Fuzzy in *across*).

Export of Glossaries: LogiTerm’s Create Glossary function can export your terminology records into other formats, such as an XML file, tab-delimited text file, or a two-column format (just source and target terms) in RTF (Rich Text Format), with fil- ➡

LogiTerm provides a way for you to create terminology records from the bitext search results table or from a bitext file open in MS Word, which is useful when reviewing translations done by others that are aligned as reference files.

sends the word to the LogiTerm search screen shown in Figure 1 on page 23. It also has shortcuts to aid in the creation of terminology records and for opening favorite terminology files.

Pretranslation and Other Linguistic Functions

LogiTerm performs many of the same functions translation environment tools perform, but by preprocessing the source text rather than doing so interactively, segment by segment, within the tool.

Pretranslation of Terminology: This function searches the terminology files you specify (only those in LogiTerm format) for all terms in a given source text. The result of the search is a copy of the source file with the identified

same term appears in the singular at the end of the sentence and no tilde is added. When the terminology database contains more than one translation for the term, as is the case for “proceedings,” the program provides one possible translation and adds a plus sign: [*autos+*]. In this case, the translation *autos* is not a good choice, but the + tells me I can look the term up with the search interface to see what other options I have stored. In this example, 14 of the 17 suggested translations are useful. There is a way to tell LogiTerm about plural inflections on a case-by-case basis, but it is beyond the scope of this article.

LogiTerm allows you to indicate a priority terminology database, made up of one or more terminology files, for pretranslation, in addition to the core set of terminology files. This fea-

tering by domain. This feature allows me to export all my terms into one file that is compatible with my translation environment tools, so I can update my terminology files in only one place.

Batch Search for Entire Segments in Bitexts: A function called LogiTrans can search your bitexts for matches or partial matches of your source text and prior translations stored as bitexts. This feature is similar to what translation environment tools do, but LogiTerm preprocesses the source file and marks the text that has already been translated. As with the pretranslation of terminology, you can configure the program to mark the text with highlighting, or you can have it insert the translation for you. You can also run the output (marked-up) file from this function through the terminology pretranslation function, and the highlighted areas will be ignored during processing. It can also indicate which of the bitext files in your database are most similar to the new source text.

Term Extraction: *This function only works if French or English is the source language.* It scans the selected text or texts for “terms” and outputs a file showing the terms and their frequency, including collocations, if requested. It can also search for the terms in specified terminology files and underline them if they are found. It only works for French and English

because it is programmed with specific language information—such as how these two languages form plurals—that aid the identification of terms. LogiTerm is a Canadian tool and French/English is presently its principal market.

External Search Function: *This function only works for French, English, and Spanish.* The External Search Function currently connects only to the Termium database (online or on CD), which only contains terms in French, English, and Spanish. Termium is the Canadian government’s linguistic database, and is available through a subscription. LogiTerm simply provides a convenient interface to it. According to Terminotix, the next upgrade will allow the user to search other terminology sites as well. Unfortunately, this will not be user-configurable.

Miscellaneous

I have been impressed by Terminotix’s pre- and post-sales support. I have never had any problems with the program, but sometimes have found it hard to figure out how to search in a certain way, or how to convert files from a specific format. They have held my hand through such periods (mainly before and shortly after I purchased the program), even though they had every right to tell me simply to read the manual.

I think that this tool is probably best for specialist translators who must maintain a good deal of terminological data. It is also ideal for translators who deal principally with printed or scanned source texts that cannot easily be fed into a translation environment tool. My most important tip is to make sure you set up a system for naming your files so you will be able to tell at a glance how reliable a source is, which languages a file contains, what country the text is from, and similar information.

The Professional version of LogiTerm has other features that I have not mentioned here due to space limitations. LogiTermWebPlus also has features not included in LogiTerm Professional. The company (Terminotix) says that LogiTerm is compatible with Windows Vista and MS Word 2007, but I have not tested it. Please see the company website for details.

If I could have only one translation tool, it would be LogiTerm.

References:

LogiTerm Professional Edition v. 3.1

www.terminotix.com
Price: 450 Canadian dollars
(approximately US\$ 450)

Termium

www.termiumplus.gc.ca



ATA Spanish Language Division 5th Mid-Year Conference

Stay tuned to www.ata-spd.org for more information! To learn about advertising, exhibit, and sponsorship opportunities, please contact Milly Suazo-Martinez at millysuazo@gmail.com.

March 28–30, 2008
Doubletree Hotel
Philadelphia, Pennsylvania