33 Integrating computer-assisted translation tools into language learning

María Fernández-Parra[1]

Abstract

Although Computer-Assisted Translation (CAT) tools play an important role in the curriculum in many university translator training programmes, they are seldom used in the context of learning a language, as a good command of a language is needed before starting to translate. Since many institutions often have translator-training programmes as well as language-learning programmes within one department or school, this paper explores the possibilities of expanding the usefulness of CAT tools from the Translation curriculum into the Foreign Language Learning curriculum. While it is not expected that CAT tools will replace any other methods of language learning, this paper hopes to show that CAT tools can nevertheless contribute to enhance the language learning experience.

Keywords: computer-assisted translation tools, foreign language learning.

1.   Introduction

In professional translation, CAT tools have gradually become a staple tool and this is increasingly reflected in translator training programmes across universities and schools (e.g. Olohan, 2011, p. 342), often at both undergraduate and postgraduate level. In recent years, we have seen the proliferation of these tools, which can be described as a single integrated system allowing for a more efficient and consistent translation process (cf. Quah, 2006, p. 93).

1. Swansea University, Swansea, UK; m.a.fernandezparra@swansea.ac.uk

Despite their usefulness for translation, CAT tools are seldom used in the context of learning a language, since a good command of a language is usually needed before starting to translate. CAT tools are designed to facilitate the translation process rather than to facilitate language learning. However, since translator training programmes are often delivered in universities or schools where language learning programmes exist alongside translator training programmes, this paper explores the possibilities of expanding the usefulness of CAT tools from the Translation curriculum into the Foreign Language Learning curriculum. After providing an overview of the main features of CAT tools, this paper maps how some of the main components can be used to support and improve a number of skills in language learning.

2.   Features of CAT tools

CAT tools can vary in the functionality provided, but at a basic level CAT tools offer at least *Translation Memory* (including *alignment*) tools or *Terminology Management* tools, or both. At a more advanced level, both the architecture and functionality of the tools are increased (cf. Fernández-Parra, 2014).

2.1.   Translation memory (TM) and alignment tools

A TM consists of a database of texts and their corresponding translation(s), divided into segments, often at sentence level, for future reference or reuse. The main advantage of a TM is that "it allows translators to reuse previous translations" (Bowker, 2002, p. 111) quickly and efficiently. TMs are particularly suited to technical documentation because they allow a fast and easy retrieval of any previously used content (Bowker, 2002, p. 113) and work by comparing the current source text to translate to previously translated documents.

One method of creating a TM is by aligning a source text with its translation. *Alignment* is the process of comparing both texts and matching the corresponding sentences which will become segments, known as *translation units*, in the TM. In many CAT tools, the alignment is carried out automatically by the software.

In this case, it is almost inevitable that some of the segments will be misaligned (e.g. Bowker, 2002, p. 109), but some alignment tools cater for this possibility by allowing manual post-editing of the results of the alignment process.

2.2.   Terminology management

Along with the TM, the terminology database, or *termbase*, is an essential component of CAT tools, as terminology is a crucial task in technical translation (cf. Bowker, 2002, pp. 104-106). A termbase is a database, but it differs from a TM in that it is used to store and retrieve segments at term level, e.g. phrases and single words, whereas the TM is typically used for sentences.

Depending on the level of sophistication of the CAT tool, the termbase can also be used to store and retrieve various kinds of information about the term, such as gender, definition, part of speech, usage, subject field, etc. In addition, the termbases in some CAT tools allow the storage and retrieval of multimedia files, e.g. graphics, sound or video files, etc., much more quickly and efficiently than spreadsheet software such as MS Excel. Further, they can allow for a hierarchical organisation of the information.

3.   Using CAT tools in language learning

There is much literature on the skills needed in language learning, but a number of foundational skills are generally well established in language pedagogy, such as speaking, listening, reading, writing, grammar and vocabulary (e.g. Hinkel, 2011, p. xiii; Widdowson, 2013, p. 632). Similarly, the idea of using computers for language learning is not new. Although Kenny (1999) had already pointed out that the integration of CAT tools into university curricula could open up new areas of research and pedagogy, there has been little research on how CAT tools in particular might be applied to foreign language learning. However, as Rogers (1996) points out, foreign language learners and translators "have a good deal in common when it comes to dealing with words: each must identify new words, record them, learn them, recall them, work out their relationships with

other words and with the real world" (p. 69). It is on the basis of this common ground between the tasks performed by translators and the tasks performed by language learners that this paper aims to 'recycle' the main components of CAT tools, such as the TM and the termbase, in order to support the various stages of the language learning process.

In the following sections, an overview is provided on how the TM and the termbase might support the various foundational language learning skills. However, given that CAT tools mainly deal with written text, focusing on skills such as listening and speaking rather falls outside the scope of this paper. They will nevertheless be hinted at when discussing multimedia files in the termbase. Translation skills have been added to the list of foundational skills, as translation is clearly another skill that CAT tools can contribute to.

It should also be pointed out that the list of suggested activities, which can be incorporated both to classroom learning and private study, remains open-ended in that new skills may be added and, as technology evolves, new CAT tool components may also be added. Further, new ways may be devised whereby a CAT tool feature might be able to support a skill currently not listed. Finally, this paper is not intended to suggest that the language skills should be approached in isolation. Therefore, each activity suggested in the following sections will typically integrate a number of the above skills.

The CAT tools explored in this paper are mainly the SDL Trados Studio 2011 suite, which includes SDL WinAlign and SDL MultiTerm, and Déjà Vu X2, not only because these are two important CAT tools in the translation industry and therefore widely taught in (at least UK) universities, but also because the useful components for language learning in these tools can be accessed as standalone features, without the need to launch the rest of the software.

3.1.   TM, alignment and language learning

Since the TM deals mainly with segments at sentence level, it is particularly suited to supporting language learning skills such as reading, writing and

translation, which are often employed at textual or sentential level. For the same reason, the more advanced language learners would particularly benefit from using TMs as an additional tool in their language learning. An example of TM that can be used for language learning is that of Déjà Vu X2. Figure 1 shows its typical use in translation.

Figure 1.   Example of TM in translation
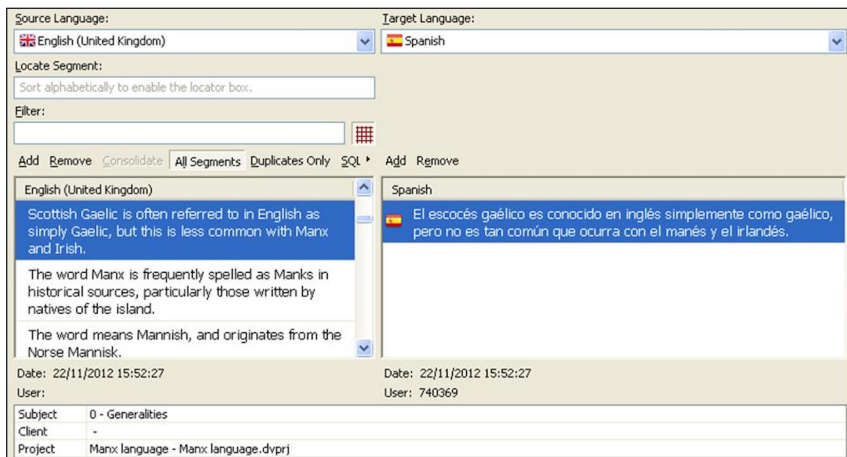
The column on the left corresponds to what translators would use as a source text to translate. This column shows the source text divided into segments.

The column on the right is where translators would type the translation. Language learners could obviously use this as an advanced translation exercise, where the lecturer can provide the text for students to translate either into or out of their first language.

The less advanced language learners can also carry out a variety of exercises, ranging from substitution and gap-filling exercises to all kinds of text manipulation exercises, such as partial or complete text reconstruction, re-ordering words in a sentence, unscrambling, etc., either in the source language or the target language, or both. In short, the students can carry out the type

of computer activities typically associated to CALL or Computer-Assisted Language Learning as accounted for by Blake (2013), for example.

Another kind of activity where TMs can help the more advanced students is writing in the foreign language, for example by helping students to structure their writing and use fluent, natural ways of expressing themselves. One example is shown in Figure 2, where English is the source language and Spanish the target one.

Figure 2.   Writing with a TM
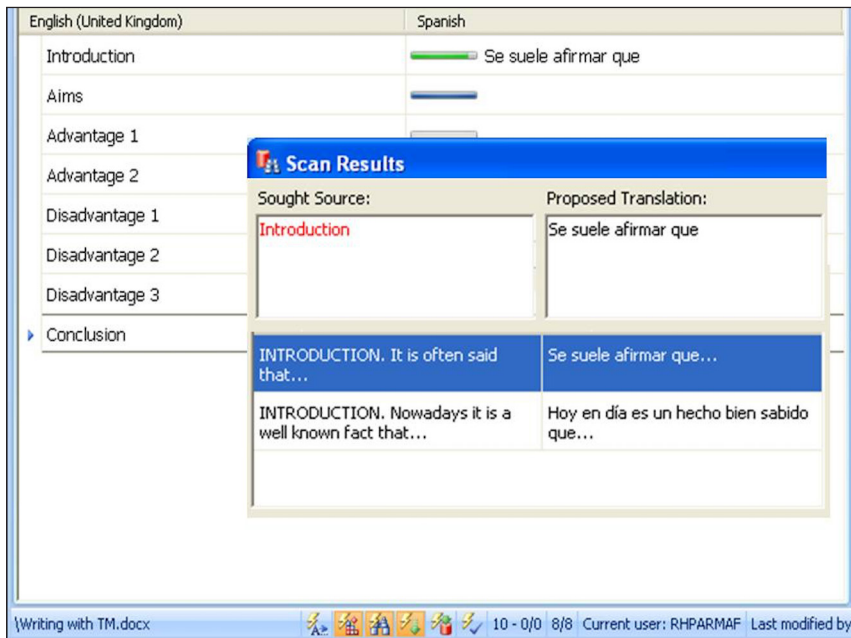
On the left of the screen, instead of a source text, the lecturer can create a file which will be used as a template with headings structuring the essay in a particular way, e.g. *Introduction*, *Disadvantage 1*, etc. There can be different templates for different tasks and students need not adhere to the templates very strictly. The example in Figure 2 shows a possible template for students to write

an essay about the advantages or disadvantages of a chosen topic. Once the template is uploaded into the CAT tool, in this case Déjà Vu X2, the students can search the translation memory for phrases to start and end paragraphs, for example.

In order to do this, students would select the word *Introduction* in the English column and right-click it to search for that word in the TM. This will bring up the bilingual *Scan Results* dialog box also shown in Figure 2. In this case, the dialog box contains a couple of examples of good ways to start a paragraph in Spanish. Students would select one and copy it into the Spanish column and then finish the sentence with their own input.

Of course, this scenario requires a certain amount of preparation of the TM contents for the exercise. Figure 2 shows that every English segment in the TM has been amended to include the label *INTRODUCTION* at the start. This would indicate to students that the phrase can be used as an introductory phrase to start a paragraph. The use of uppercase is deliberate to distinguish the label from the actual phrase. Similarly, other labels could be PRESENTING AIMS, CONCLUSION, INTRODUCING AN OPPOSED VIEW, etc.

The TM in SDL Trados can also be used in this way. An example of the results obtained from searching an SDL Trados TM is shown in Figure 3.

Figure 3.   Writing with SDL Trados

The automatic changes made by the software to the source text, as shown on the left in Figure 3, should not affect the writing task, as the label INTRODUCTION,

placed at the start of the segment, remains unchanged and so do the contents of the target segment, which the students will decide whether they use in their writing task.

The lecturer need not be overwhelmed with the preparation of the TM contents in advance because this can be done by the students themselves using alignment. For example, the lecturer can simply direct the students to a bilingual resource such as the Collins Dictionary (2009), where the central pages contain a wealth of formulaic phrases in both languages. The students would simply need to create a separate file for each language with the required phrases and run both files through an alignment program. One such program to use in this scenario is SDL WinAlign. An example is shown in Figure 4.

Figure 4.   Example of alignment
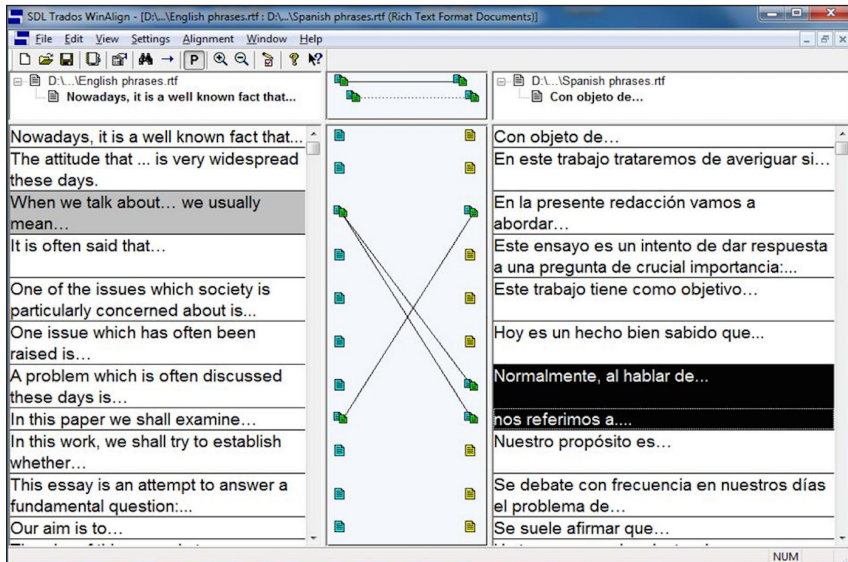
In Figure 4, the English phrases appear on the left, the Spanish on the right. Initially, the program will attempt to match the relevant phrases and it will display them joined with a dotted line. Dotted lines joining segments can easily

be deleted by selecting the relevant option after right-clicking the icon of a particular segment. The suggestions can also be accepted by right-clicking and selecting the relevant option. In Figure 4, all dotted lines have been disconnected and several new re-connections are shown with a continuous line. An alignment exercise such as this could be used at all levels of language learning, as the level will depend on the contents to align, from single words to complete sentences or paragraphs.

Thus far we have assumed a one-to-one correspondence of segments but it may be that one segment matches two, as shown with the segments *Normalmente al hablar de…* and *nos referimos a…* in Figure 4. Both these segments correspond to one segment in English and can easily be joined by right-clicking the relevant option. Following the same procedure, a segment could be split if required.

Once all the segments are matched, the students can export the results as a TM, which they can then reuse in a variety of tasks as explained above. Déjà Vu X2 also has an alignment component, but it does not have the visual, drag-and-drop approach in SDL WinAlign to join the relevant segments which could be particularly helpful in language learning.

3.2.   The termbase and language learning

In translation, the termbase is used to manage terminological tasks, a crucial aspect of the translation process. In language learning, the termbase can also support a crucial aspect: that of vocabulary acquisition, which in turn will support reading and writing tasks. The termbase facilities can vary widely from one CAT tool to another, but one good example of a useful terminological facility for language learning is that of SDL MultiTerm. An example of a terminological entry is shown in Figure 5.

An advantage of the SDL MultiTerm termbase over others, such as Déjà Vu X2, is that it allows the inclusion of multimedia information such as graphics, videos and sound files. The latter two, in particular, can be exploited in tasks such as pronunciation and role plays.

Figure 5.   Example of terminological entry
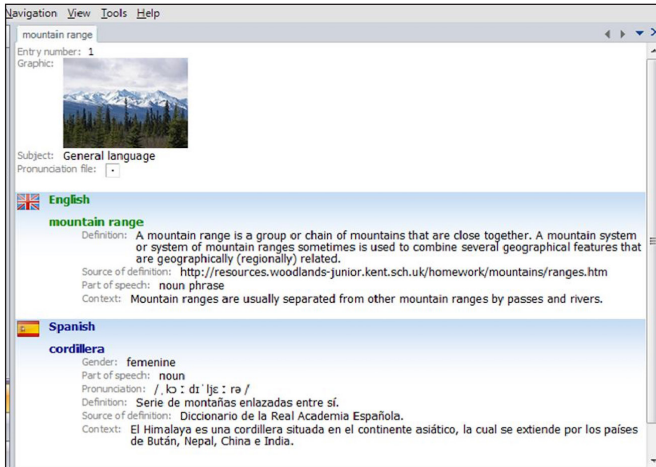
As with the TMs, termbases come empty to start with, and it is up to the user to decide whether a termbase should be bilingual or multilingual and which fields to include in each entry. Many of the fields often used in the termbases for translation, such as part-of-speech and definition can also be used in termbases for language learning. Learners may also benefit from fields such as grammatical information, e.g. regular/irregular verb, collocational restrictions, etc.

Termbases can be created by students themselves. In this case, much of the learning can take place by collecting the information needed and doing the necessary research to include the required information in the termbase. The learning, at any level, is further reinforced by consulting the termbase in subsequent tasks. Alternatively, lecturers can provide students with termbases to use in advanced reading or translation tasks.

4.   Conclusion

This paper explored the possibilities of considering CAT tools as additional language learning tools, especially in universities or schools where CAT tools are

already part of the curriculum, as both students and staff may already be familiar with the software. Some knowledge of the software is therefore assumed, although an effort has been made to present examples from components of CAT tools which can be accessed relatively quickly as standalone components.

The original idea of the suitability of CAT tools for technical translation based on lexical repetition can be recycled as repetition in language learning for the reinforcement of the learning. While it is not expected that CAT tools will ever replace other methods of language learning, this paper hopes to have shown that CAT tools can nevertheless co-exist with such methods and contribute to enhance the language learning experience.

References

Blake, R. J. (2013). *Brave new digital classroom. Technology and foreign language learning*. Georgetown: Georgetown University Press.

Bowker, L. (2002). *Computer-aided translation technology. A practical introduction*. Ottawa: University of Ottawa Press.

Collins Dictionary. (2009). *Spanish-English English-Spanish dictionary* (9th ed.). Glasgow: HarperCollins Publishers.

Fernández-Parra, M. (2014). *Formulaic expressions in computer-assisted translation*. Saarbrücken: Scholars' Press.

Hinkel, E. (Ed.). (2011). *Handbook of research in second language teaching and learning. Volume II*. Abingdon: Routledge.

Kenny, D. (1999). CAT tools in an academic environment: what are they good for? *Target, 11*(1), 65-82. Retrieved from http://dx.doi.org/10.1075/target.11.1.04ken

Olohan, M. (2011). Translators and translation technology: the dance of agency. *Translation studies, 4*(3), 342-357. Retrieved from http://dx.doi.org/10.1080/14781700.2011.589656

Quah, C. K. (2006). *Translation and technology*. Basingstoke: Palgrave Macmillan. Retrieved from http://dx.doi.org/10.1057/9780230287105

Rogers, M. (1996). Beyond the dictionary: the translator, the L2 learner and the computer. In G. Anderman & M. Rogers (Eds.), *Words, words, words. The translator and the language learner* (pp. 69-95). Clevedon: Multilingual Matters.

Widdowson, H. (2013). Skills and knowledge in language learning. In M. Byram & A. Hu (Eds.), *Routledge encyclopedia of language teaching and learning* (p. 631). Abingdon: Routledge.

Resources

Déjà Vu X2. Retrieved from www.atril.com
SDL Trados. Retrieved from http://www.sdl.com/products/sdl-multiterm/desktop.html

New perspectives on teaching and working with languages in the digital era
Edited by Antonio Pareja-Lora, Cristina Calle-Martínez, Pilar Rodríguez-Arancón