

## Anapraseus: translators' workhorse

Oleg Tsygany\*

**Abstract:** Article about a new open source Computer Aided Translation tool for the creation, use and management of bi-lingual translation memories in any language, designed as an extension of OpenOffice.org. Features: text segmentation, fuzzy search in the translation memory, terminology recognition, import/export in TMX format, user interface similar to Wordfast's.

**Key words:** computer-assisted translation, CAT, translation memory, terminology recognition, open source, OpenOffice.org.

### Una herramienta excelente para el traductor

**Resumen:** Este artículo trata de una herramienta de traducción asistida por ordenador, de código abierto, que se ha diseñado como una extensión para OpenOffice.org. Se trata de una herramienta para crear, manejar y utilizar memorias de traducción bilingües en cualquier combinación lingüística. Sus principales características son: segmentación del texto, búsqueda aproximada en la memoria de traducción, reconocimiento de la terminología, importación y exportación en el formato TMX y una interfaz de usuario similar a la de Wordfast.

**Palabras clave:** traducción asistida por ordenador, TAO, memoria de traducción, reconocimiento de terminología, código abierto, OpenOffice.org.

Panace@ 2009; 10 (29): 58-60

Modern world is not imaginable without international communication. This is one of the foremost challenges of our times. News, information and culture are increasingly shared across national boundaries. Personal communication has taken on new meanings in the age of the Internet. News media have become global in terms of audiences, courses and ownership. Managers must communicate constantly with customers, suppliers, and shareholders in every part of the world. Students must be equipped to understand and to master this rapidly changing background.

In this environment, ability to quickly translate from one language to another is important. But this goal is hard to reach without special tools: dictionaries, thesauri, even computer translation systems and so forth. In this line, the so-called CAT

tools take a special place. CAT stands for "Computer Aided Translation" and means translation done with specialized software with the capability to increase speed of work while providing quality assurance, terminology management, etc.

The mainstream of these applications includes SDL Trados, OmegaT, Wordfast etc. They are reliable and powerful. Many translators around the world count on them as standard working tools. Let me introduce a small and young but promising application: Anapraseus.

Anapraseus is a computer aided translation tool written in OpenOffice.org Basic (also known as StarBasic). It works with OpenOffice.org Writer, and supports every format Writer can handle. It is designed as an OpenOffice.org extension, and can run in Linux/Unix, Mac OS or Windows operating system.

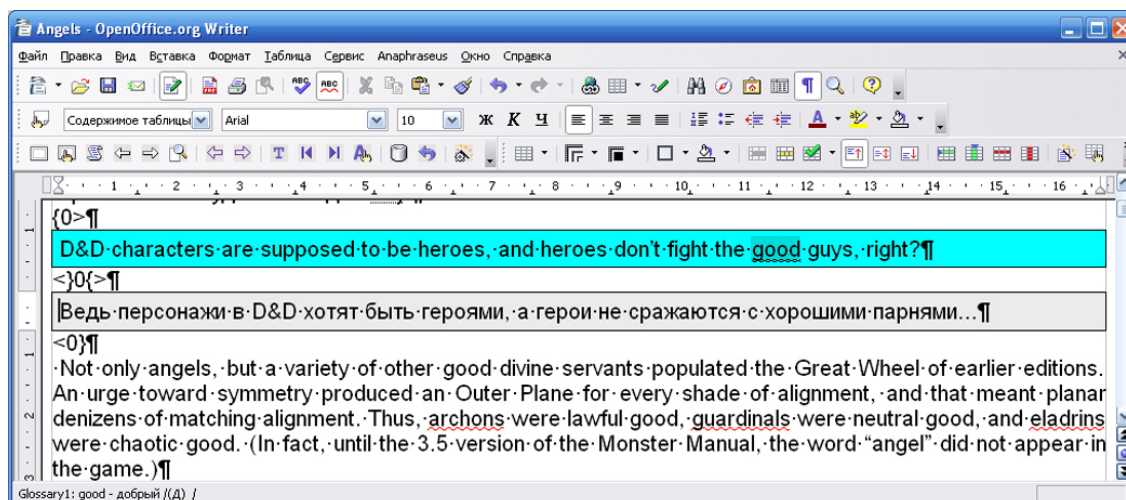


Figure 1: Anapraseus' main window

\* Programmer, hobby translator, Odessa (Ukraine). Address for correspondence: [oleyansen@gmail.com](mailto:oleyansen@gmail.com).

You can ask me... Why choosing Anaphraseus when there are so many other powerful tools? A year ago I worked extensively with Wordfast and liked it much. It is a great tool, really. But it has a serious drawback – it works with Microsoft Word only. There was no similar application for Linux except for OmegaT. So I decided to write my own Wordfast in an OpenOffice.org environment, which I named OpenWordfast. In March 2008, the name was changed to Anaphraseus, because “openwordfast” sounded a bit as a challenge, which was not the case.

At this point, I would like to say thanks to the Wordfast team for an excellent idea and software, “sine qua non” for our project.

Anaphraseus is intended as a replacement for Wordfast on Linux, Mac OS and possibly other systems. It supports main features needed for CAT tools, like text segmentation, translation memory and terminology recognition.

Text segmentation means to divide text to units, such as sentences, paragraphs or topics to work with. Anaphraseus uses segmentation by sentences, but this behaviour could be easily overridden by “Expand translation segment” function. The structure of each text segment is fully compatible with Wordfast. That means you can exchange work between them without problems.

The main idea of tools as OmegaT or Wordfast is storing chunks of text (translation units) you have translated in a special storage, a translation memory, so you don’t need to translate them over and over. Each time you have an exact or similar phrase you’ll have a message about it. It’s very handy. Anaphraseus’ Translation Memory can contain up to 500 000 units. Translation of a medium-sized book will take about 10 000 units, just for comparison.

As for terminology, it includes two glossaries of up to 120 000 units: a main glossary and a user’s glossary. You can add your terms to the user glossary on the fly. This is helpful when working with a new field, where new terms appear virtually in each sentence.

Anaphraseus is fully integrated in OpenOffice.org Writer. A positive side is that you don’t need any other application. A negative side – you need OpenOffice.org to work with. On the other hand, you’ll have incredible flexibility, because you can work on Linux/Unix, Mac OS and Windows using the same interface.

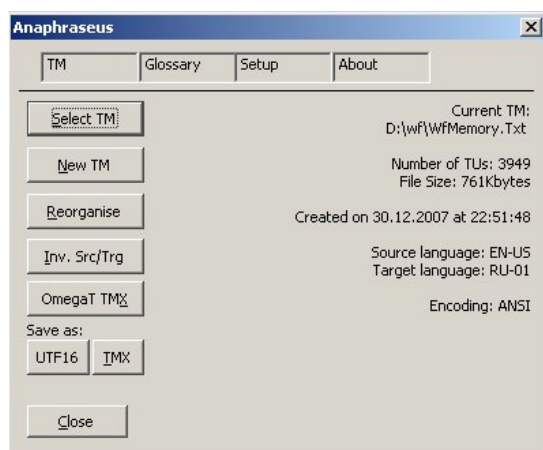


Figure 2: Setup form

Anaphraseus main features:

1. Segmentation of translated text compatible with Wordfast.
2. Translation memory suitable for Wordfast. Encoding of memory either in plain text or Unicode (UTF-16).
3. Terminology recognition: two glossaries of up to 120 000 units. Encoding of glossaries in plain text or Unicode UTF-16. Both compatible with Wordfast.
4. Memory export/import in UTF8, UTF16 TMX formats. Translation memories of OmegaT can be imported.
5. Targeted for students and small business.
6. Ready for localisation in any language.
7. The Anaphraseus extension for OpenOffice.org is covered by the terms of the GNU General Public License version 3. So everybody is encouraged to modify its code according to their own needs.

One of the disadvantages of Anaphraseus relates to the nature of the application. As an OpenOffice.org extension, its performance is slower than Wordfast, especially when handling large files.

The second major disadvantage is that Anaphraseus is still in beta test. It works only with unformatted text, skipping all formulas, pictures and similar objects. There is still a lot to do, and we have great plans for the program.

Let us see what two users think about this application.

#### Dmitri, translator

Anaphraseus is a free and open source application that runs in every environment supported by OpenOffice.org (i.e., Windows, Linux, FreeBSD, Mac OS). The program is absolutely portable, provided that you have a portable version of OpenOffice.org (OOo). Just plug in your USB stick anywhere and go on translating!

*Q. Who can efficiently use Anaphraseus, translators, localizers, etc?*

Anaphraseus basically depends on OOo’s capabilities. But yes, it’s rather for translators than for localizers.

*Q. Anaphraseus or OmegaT, what do you think?*

It depends on your work style. Anaphraseus combines a CAT tool with the power of a word processor. Anaphraseus is also somewhat advanced in respect of glossary handling. Anaphraseus can also re-use Wordfast translation memories directly (with certain limitations). However, I would not put it as Anaphraseus VERSUS OmegaT, it’s rather Anaphraseus WITH OmegaT. Try both and decide, if you choose one of them or combine.

*Q. Could you tell us about your work experience with Anaphraseus?*



I do 85 to 90% of my jobs with OmegaT, 5 to 10% with Wordfast, which leaves 5 to 10% for Anaphraseus. This program is essential when:

- a) my project requires an extensive use of glossary functions,
- b) I work in a portable project, and
- c) I work with Linux, a robust and reliable working environment.

**Sergei, freelancer**

Anaphraseus is a free software. Not freeware but truly free software. There are not so many really free tools for translators out there at present moment. So far, Anaphraseus has its own unique “niche”.

*Q. Who is intended to use Anaphraseus?*

Both translators and localizers who prefer to work in office applications which support the StarOffice Basic programming language (e.g. OpenOffice.org).

*Q. Do you prefer Anaphraseus to OmegaT?*

Anaphraseus allows you to use all the features of the word processor “on the fly”.

*Q. Could you tell us about your work experience with Anaphraseus?*

Due to the fact that Anaphraseus is still in beta, its use is still limited to translating texts with no heavy formatting.

