

Date: June 30, 2022

Subject: What has been exciting to me in the glyph project
as told by Jon the lead computer assisted translator

I was asked to give you folks an overview of what has been exciting to me in the glyph translation effort so be patient if this is a somewhat technical document. It has been exciting to work with existing tools in this effort as well as new tools designed just for this effort. First I will explain some of the tool we have been using as well as how they help the process.

Topic: machine assisted translation

Computer-aided translation (CAT), also referred to as machine-assisted translation (MAT) or machine-aided human translation (MAHT), is the use of software to assist a human translator in the translation process. The translation is created by a human, and certain aspects of the process are facilitated by software; this is in contrast with machine translation (MT), in which the translation is created by a computer, optionally with some human intervention (e.g. pre-editing and post-editing). CAT tools are typically understood to mean programs that specifically facilitate the actual translation process. Most CAT tools have (a) the ability to translate a variety of source file formats in a single editing environment without needing to use the file format's associated software for most or all of the translation process, (b) translation memory, and (c) integration of various utilities or processes that increase productivity and consistency in translation.

Range of tools

Computer-assisted translation is a broad and imprecise term covering a range of tools. These can include:

Translation memory tools (TM tools), consisting of a database of text segments in a source language and their translations in one or more target languages.

Spell checkers, either built into word processing software, or available as add-on programs.

Grammar checkers, either built into word processing software, or available as add-on programs.

Terminology managers, which allow translators to manage their own terminology bank in an electronic form. This can range from a simple table created in the translator's word processing software or spreadsheet, a database created in a program such as FileMaker Pro or, for more robust (and more expensive) solutions, specialized software packages such as SDL MultiTerm, LogiTerm, Termex, TermWeb, etc.

Electronic dictionaries, either unilingual or bilingual.

Terminology databases, either on the host computer or accessible through the Internet, such as TERMIUM Plus or Grand dictionnaire terminologique from the Office québécois de la langue française.

Full-text search tools (or indexers), which allow the user to query already translated texts or reference documents of various kinds. Some such indexers are ISYS Search Software, dtSearch Desktop and Naturel.

Full-Image search tools (or indexers), which allow the user to query glyphs as well as existing glyph translations.

Concordancers, which are programs that retrieve instances of a word or an expression and their respective context in a monolingual, bilingual or multilingual corpus, such as a bitext or a translation memory.

Bitext aligners: tools that align a source text and its translation which can then be analyzed using a full-text search tool or a concordancer.

Project management software that allows linguists to structure complex translation projects in a form of chain of tasks (often called "workflow"), assign the various tasks to different people, and track the progress of each of these tasks.

Concepts Translation memory software

Translation memory programs store previously translated source texts and their equivalent target texts in a database and retrieve related segments during the translation of new texts. Such programs split the source text into manageable units known as "segments". A source-text sentence or sentence-like unit (headings, titles or elements in a list) may be considered a segment. Texts may also be segmented into larger units such as paragraphs or small ones, such as clauses. As the translator works through a document, the software displays each source segment in turn, and provides a previous translation for re-use if it finds a matching source segment in its database. If it does not, the program allows the translator to enter a translation for the new segment. After the translation for a segment is completed, the program stores the new translation and moves on to the next segment. In the dominant paradigm, the translation memory is, in principle, a simple database of fields containing the source language segment, the translation of the segment, and other information such as segment creation date, last access, translator name, and so on. Another translation memory approach does not involve the creation of a database, relying on aligned reference documents instead. Some translation memory programs function as standalone environments, while others function as an add-on or macro for commercially available word-processing or other business software programs. Add-on programs allow source documents from other formats, such as desktop publishing files, spreadsheets, or HTML code, to be handled using the TM program.

Language search-engine software

New to the translation industry, Language search-engine software is typically an Internet-based system that works similarly to Internet search engines. Rather than searching the Internet, however, a language search engine searches a large repository of Translation Memories to find previously translated sentence fragments, phrases, whole sentences, even complete paragraphs that match source document segments.

Language search engines are designed to leverage modern search technology to conduct searches based on the source words or image of glyphs in context to ensure that the search results match the meaning of the source segments. Like traditional TM tools, the value of a language search engine rests heavily on the Translation Memory repository it searches against.

Terminology management software

Terminology management software provides the translator a means of automatically searching a given terminology database for terms appearing in a document, either by automatically displaying terms in the translation memory software interface window or through the use of hot keys to view the entry in the terminology database. Some programs have other hotkey combinations allowing the translator to add new terminology pairs to the terminology database on the fly during translation. Some

of the more advanced systems enable translators to check, either interactively or in batch mode, if the correct source/target term combination has been used within and across the translation memory segments in a given project. Independent terminology management systems also exist that can provide workflow functionality, visual taxonomy, work as a type of term checker (similar to spell checker, terms that have not been used correctly are flagged) and can support other types of multilingual term facet classifications such as pictures, videos, or sound.

Alignment software

The process of binding a source language segment to its corresponding target language segment. The purpose is to create a translation memory database or to add to an existing one.

Interactive machine translation

Interactive machine translation is a paradigm in which the automatic system attempts to predict the translation the human translator is going to produce by suggesting translation hypotheses. These hypotheses may either be the complete sentence, or the part of the sentence that is yet to be translated.

Augmented translation

Augmented translation is a form of human translation carried out within an integrated technology environment that provides translators access to sub-segment adaptive machine translation (MT) and translation memory (TM), terminology lookup (CAT), and automatic content enrichment (ACE) to aid their work, and that automates project management, file handling, and other ancillary tasks. Based on the concept of augmented reality, augmented translation seeks to make translators more productive by providing them with relevant information on an as-needed basis. This information adapts to the habits and style of individual translators in order to accelerate their work and increase productivity. It differs from classical post-editing of MT, which has linguists revise entire texts translated by machines, in that it provides machine translation and information as suggestions that can be adopted in their entirety, edited, or ignored, as appropriate. Augmented translation extends principles first developed in the 1980s that made their way into CAT tools. However, it integrates several functions that have previously been discrete into one environment. For example, translators historically have had to leave their translation environments to do terminology research, but in an augmented environment, an ACE component would automatically provide links to information about terms and concepts found in the text directly within the environment.

As of May 2017, no full general purpose implementations of an augmented translation environment exist, although we have within the context of glyph translation developed a fully Augmented Translation environment.

How we use the augmented system we developed

When new glyphs or images of glyphs are added to one of the input databases, the first thing that happens is as much information regarding each added glyph. Then at the end of a day adding glyphs to the database, the system scans to see if any of the newly entered glyphs match or are close to some already existing in the database. If yes then these glyphs are noted for people to look at them and decide how to log them in the system. This may also require an in context image search for how the nw glyph is used as well as any existing glyphs like it. From this we can build translations as needed.

God bless all of you who have also helped with your prayers and any funding.

MY WOW MOMENTS IN THIS PROJECT

(from Jon)

The most wow moment for me in this project was to see how the glyphs from different areas and times the glyphs were created contained the same messages for people. Also there is the astonishing fact that there are materials that our modern material science still does not have a grasp of that were used in making some of the artifacts. Additionally it was shocking to see how the fallen angels lead people down the path to that not only is it ok to kill people but human sacrifice to lucifer is a wonderful thing according to the fallen angels. One of the most exciting moments was to discover DNA virus making instructions on an artifact from the middle east and the same DNA virus making instructions in the GENSIX Mexico dig. It was only recently discovered on a glyph scan in the databases and everyone got really excited about it because it begins to prove the validity of the GENSIX effort as well.

WOW MOMENTS FOR ME IN THIS PROJECT

(From SG)

The biggest overall wow thing was how widely and how deeply the fallen angels have invaded every language and culture on the planet, and some of the very subtle tricks used to seduce people into the very depths of depravity without even realizing it. It is also shocking in a way that people who say they are Christians fear so many things in life when in fact God can and does supply all our needs as that is his promise to us and even when we fail God and must ask forgiveness again God is merciful and welcomes us back with open arms. One of the most exciting moments was to discover DNA virus making instructions on an artifact from the middle east and the same DNA virus making instructions in the GENSIX Mexico dig, and also discovered this DNA encoding in artifacts from Europe as well. This in fact proves that the dig in Mexico is valid, and it also proves that our glyph translation efforts are valid as well. Additionally it proves that the large image database for this effort is worth having because of nuggets like this which would not show up in the normal process of a historical dig.