

# FREE AND OPEN SOURCE SOFTWARE FOR TRANSLATORS

Corinne McKay

Abstract: Free and open source software (FOSS) is gaining popularity in the translation industry, due to its low cost, flexibility, reliability, security and freedom from licensing hassles. With a range of free and open source translation memory suites, web browsers and office software to choose from, FOSS is an attractive option for freelancers and agencies alike. The past year has seen many new developments in the world of FOSS for translators.

## 1. WHAT IS FREE AND OPEN SOURCE SOFTWARE AND WHY DO PEOPLE USE IT?

Many people think that software that doesn't cost money is automatically Free Software. In fact, FOSS might be free or might cost money, but the real distinction is that its source code is open and publicly available. Free Software Foundation ([www.gnu.org](http://www.gnu.org)) founder Richard Stallman famously paraphrased this as “free as in free speech, not as in free beer.” Free and Open Source Software is often contrasted with proprietary software, whose source code is owned by the company that sells the software, and is not in the public domain. FOSS is also licensed differently from proprietary software. The most common FOSS licenses allow the user to:

- Install the software on as many machines as he/she wants
- Allow any number of people to use the software at once
- Copy the software and give it to anyone
- Modify the software, as long as certain features are kept intact (most commonly the licensing agreement)
- Freely (in the sense of “without restrictions”) and for any purpose distribute or sell the software without paying royalties to the original developer

Probably the best-known example of FOSS is the Linux computer operating system, represented by its mascot, Tux the penguin. For a thorough and readable overview, see Wikipedia's entry on Linux at: <http://en.wikipedia.org/wiki/Linux>. Linux, originally written as a hobby project by Finnish university student Linus Torvalds, was first released to the public in 1991. and today is the market leader server operating system on the public Internet ([http://www.dwheeler.com/oss\\_fs\\_why.html#market\\_share](http://www.dwheeler.com/oss_fs_why.html#market_share)). Wikipedia also notes that “The Linux market is among the fastest growing in the IT industry, and is projected to exceed \$35.7 billion by 2008.” Linux's reputation for security and reliability has made it a favorite OS at installations such as the U.S. National Security Agency, and a 2004 Evans Data survey ([http://www.evansdata.com/n2/pr/releases/Linux04\\_02.shtml](http://www.evansdata.com/n2/pr/releases/Linux04_02.shtml)) found that 92% of Linux installations have never been infected with a virus, and 78% have never been hacked, while only 40% of non-Linux systems had never experienced a security breach. Although still a small portion of the desktop market, Linux is making inroads there too. For example, the user experience consultancy Relevantive ([www.relevantive.de](http://www.relevantive.de)) concluded in a 2004 study that the Linux desktop environment was “equal to Windows XP” in usability for a variety of common desktop tasks.

Many people think that FOSS is a fringe player in the software market; in the past this may have been true, but let's look at some current-day examples, which will show that whether we know it or not, nearly all of us use FOSS every day. The Google search engine and Amazon.com both run Linux almost exclusively (<http://www.wired.com/news/technology/0,1282,54504,00.html>) on their servers, and with over 100,000 Linux servers, Google is probably the world's largest Linux installation. So if you've ever searched on Google, you've used FOSS. In addition, Apache (

[www.apache.com](http://www.apache.com)) has been the market leader web server ever since such data have been collected, and with approximately 70% of the web server market, currently counts three times the market share of the second-ranked server, produced by Microsoft. The Microsoft Office-compatible OpenOffice.org office suite, supported in almost 50 languages, has counted more than 25 million downloads, and the open source web browser Firefox, more than 75 million. Firefox was also voted the PC World “Product of the Year” for 2005. More impressive still is the statistic, revealed in a 2004 article in Information Week (<http://www.informationweek.com/story/showArticle.jhtml?articleID=51201599&tid=5979>) that “Two thirds of companies surveyed use open-source products, and only 17% reported no near-term plans to support open-source products.”

Outside of the United States, FOSS is often even more widely used. A much-publicized example came in 2002, when the German Government inked a contract with IBM and Germany-based SuSE (pronounced “Susie”) Linux to implement Linux on computer systems at the national, state and local levels. While Linux use was not mandated, German Interior Minister Otto Schily commented, “With the contract with IBM we meet three key targets. We raise the level of IT security by avoiding monocultures; we lower the dependency on single software vendors; and we reach cost savings in software and operation costs.” (<http://news.com.com/2100-1001-931027.html>)

## **2. WHY IS FREE AND OPEN SOURCE SOFTWARE IMPORTANT TO TRANSLATORS?**

Translators are heavy computer users, with most of us spending the majority of our work day on tasks that require the computer. In addition, most translators are self-employed and bear the cost of maintaining and upgrading their own computer systems, often at considerable cost. Translators also use, or would like to use, language-specific versions of software that are often difficult to find in proprietary distributions. Still, translators as a group have been reluctant to migrate from proprietary software toward open source alternatives.

This author can't pretend to know everything about why translators are reluctant to switch to FOSS, but here are a few theories gathered from personal experience:

- Translators as a group are somewhat technophobic and would rather concentrate on translation itself rather than on technological tools for translation.
- Translators often don't know much about FOSS, and so aren't often aware of the options available.
- Translators are afraid that using FOSS will cause them to lose clients. They are afraid that by running a different program than the client, or than other translators, they will have a hard time finding work.
- Many CD-ROM dictionaries do not work on a Linux computer.
- The market-leader translation memory software companies do not produce Linux versions of their products. FOSS TM applications are available, but they do not have a great deal of name recognition, and may require some work to make their memories compatible with the market leaders.
- FOSS applications are largely volunteer developed and do not have the financial resources to compete with proprietary applications in terms of marketing, training and support.
- People in general, not just translators, often think that if something is free, it may be less useful or valuable than something that costs money.

As you can see, some of these reasons, such as difficulty of running CD-ROM dictionaries on a Linux machine, are entirely valid, while others, such as lack of dollars to market open source applications, are merely a matter of circumstance and not of the software's usability. Following are some reasons that translators might consider switching to FOSS applications.

- The software is there. FOSS applications exist for every step of the translation work flow; translation memory, word processing, web browsing, accounting, website development, and more.
- Open source is cost-effective. While “it's cheap” is rarely the primary motivation cited by avid FOSS users, “value for money” is often a reason for choosing FOSS. FOSS won't free you from the need to pay for documentation, support and training, just as you do with proprietary software. But it will in many cases free you from the need to pay high prices for software, upgrades, data recovery after a virus attack, tracking licensing paperwork, etc. In addition, having free or low-cost tools available for critical but unexciting tasks like backups makes it more likely that you will actually obtain the tools.
- Open source encourages innovation. When the Mozilla Firefox web browser made headlines in late 2004, Microsoft hadn't released a major redesign of Internet Explorer since 2001, in part because they didn't have to; with a 95% market share, there simply wasn't much of an incentive to integrate tabbed browsing or localize into Albanian. However, a community-developed project like Firefox can take a “don't you wish you had...” feature and integrate it as fast as coders can work.
- Open-source licenses are less restrictive. Most end-users, including translators, never read EULAs (end-user license agreements) before clicking “I Agree” and installing new software. In addition to restrictive conditions such as allowing the software to be installed on only one computer (meaning that a user with a laptop and a desktop is legally required to buy two copies of the software at full price), some EULAs go even further and specify that the user may not publicly criticize or evaluate the software. Network Associates, the makers of McAfee antivirus software, was criticized in 2003 for prohibiting users from disclosing the results of benchmark tests on its software without prior written permission (<http://www.eff.org/wp/eula.php>). Nearly every proprietary software EULA prohibits activities that are generally allowed by law, for example reverse engineering. By contrast, the most popular open source licenses such as the GNU General Public License, specify very basic requirements, such as that any released improvement of the software must also be free software, and that anyone must be allowed to run the software for any purpose.

SDL's recent acquisition of Trados has highlighted some of the reasons why proprietary licensing agreements affect translators. For example, one respondent to the GALA (Globalization and Localization Association) survey “Language Service Provider Reaction to SDL's Purchase of TRADOS” ([http://www.gala-global.org/en/resources/downloads/2005Q3\\_GALA-SDL-TradosReport.pdf](http://www.gala-global.org/en/resources/downloads/2005Q3_GALA-SDL-TradosReport.pdf)) offered the following insight: “Worst case scenario: SDL changes the licensing terms for TRADOS to require information on every client/project where TRADOS is used. Thus SDL would have a list of all the competitive vendors' clients. This is unlikely, but it just gives SDL too much power in our small industry.” As this respondent points out, while this scenario seems far-fetched from a business standpoint, it would in fact be legal, pointing out that proprietary software users are very much at the mercy of the vendors who own their software.

### **3. HOW MUCH LESS DOES FOSS REALLY COST?**

Various studies have compared the total cost of ownership (TCO) of proprietary software and FOSS; the real TCO depends on how you use the software- whether you want documentation on

paper, whether you pay for support and training, etc. Notably, Sun Microsystems and Microsoft have both released “independent reports” stating that their products have the lowest TCO.

However in terms of cost of acquisition and upgrades, it's worth noting that FOSS isn't just a little bit cheaper than proprietary alternatives, it's a lot cheaper, and the price difference is even more apparent in a large installation. Let's look at a computer setup for a typical freelance translator, comparing proprietary software (prices obtained from Amazon.com and Translationzone.com in August, 2005) and FOSS.

### **Proprietary Software**

Microsoft Windows XP Professional: \$262.99  
Microsoft Office XP Professional: \$225.00  
Microsoft Outlook: \$89.99  
Quick Books Basic: \$189.99  
Trados 7 Freelance: \$895.00  
**Total: \$1,662.97**

### **Open Source Software**

Ubuntu Linux: \$0.00 (including shipping)  
OpenOffice.org \$0.00  
Mozilla Mail: \$0.00  
Gnu Cash: \$0.00  
OmegaT: \$0.00  
**Total: \$220.00**  
OR as above, plus:  
CrossOver Office Professional: \$74.95  
Microsoft Office XP Professional: \$225.00  
Wordfast: \$220.00  
**Total: \$ 519.95**

In a larger office such as a translation company, the savings could be even more remarkable. The Australian consultancy Cybersource ([www.cyber.com.au](http://www.cyber.com.au)) produced an in-depth report entitled “Linux vs. Windows: The Bottom Line” which compared the setup costs for a 50 user office needing an operating system, e-mail server, and database server. The study found a total cost of \$69,987 for the Microsoft option, and \$80 for the Linux option, resulting in a savings of \$69,907. Because FOSS licenses generally allow the software to be installed on an unlimited number of computers, a 200-person translation company could set up its computer systems for the same cost, \$80, while the cost of the Microsoft option would balloon to \$282,974.

The cost of upgrading an open source system is typically a fraction of what is required for a proprietary system. For example, when Trados released Trados 7 Freelance at \$895.00 (price obtained from translationzone.com in August, 2005) it offered upgrades for either \$395.00 or \$495.00, depending on the pre-existing version. By contrast, an open source TM application such as OmegaT or Sun Open Language Tools could be upgraded for free. An open source operating system can simply be downloaded again, usually for free, when a new release comes out.

Another important savings related to FOSS is the freedom from license management costs, especially for large offices such as translation companies. In his paper “Why Open Source Software/Free Software? Look at the Numbers!” ([http://www.dwheeler.com/oss\\_fs\\_why.html](http://www.dwheeler.com/oss_fs_why.html)), David Wheeler summarizes “Proprietary vendors make money from the sale of software licenses, and are imposing increasingly complex mechanisms on consumers to manage these licenses. Customers who cannot later prove that they paid for every installed copy of proprietary software (e.g., due to copying by an employee or losing the license paperwork) risk stiff penalties. In short: by using proprietary software, you run the risk of having the vendor sue you. To counter these risks, organizations must keep careful track of license purchases. This means that organizations must impose strict software license tracking processes, purchase costly tracking programs, and pay for people to keep track of these licenses and perform occasional audits.” For this reason alone, FOSS is an attractive option for a translation company that wants to put its resources into language work rather than software record-keeping.

#### 4. WHAT FREE AND OPEN SOURCE SOFTWARE IS AVAILABLE FOR TRANSLATORS?

If you're new to the idea of mixing FOSS and translation, an excellent website to browse is Marc Prior's "Linux for Translators" ([www.linuxfortranslators.org](http://www.linuxfortranslators.org)), which gives an overview of "how to" and "why to" implement FOSS options. The past year has seen major advances in the availability of FOSS for translators.

**OmegaT** ([www.omegat.org](http://www.omegat.org)), a cost-free and open source CAT tool, released version 1.4 on August 11, 2005. OmegaT is written in Java, so will run on Windows, Linux, and Mac OS X. It supports the traditional CAT features such as fuzzy matching, match propagation, simultaneous use of multiple translation memories, and multiple file formats. In addition, unlike CAT tools such as Trados and Wordfast that work from within Microsoft Word and are therefore dependent on it, OmegaT is an independent application, although it is most useful when paired with OpenOffice.org (also free and open source). OmegaT is compatible with other CAT tools at TMX Level 1, and is localized into Catalan, Italian, Afrikaans, and German, with documentation in English, German, French, Italian and Japanese. OmegaT also got some nice press outside of the translation community with an article in the online newspaper NewsForge (<http://software.newsforge.com/article.pl?sid=05/02/11/1831257&mode=nocomment&tid=130>) by translator Dmitri Popov.

**Heartsome** ([www.heartsome.net](http://www.heartsome.net)) , produced in Singapore, is currently the only commercial translation tool that is Linux-compatible. Like OmegaT it is a "single-layer application," meaning that it runs on its own, rather than from within another application such as Microsoft Word. Heartsome runs on Windows, Mac OS X, Linux, Unix and Solaris, and is TMX compliant at Levels 1 and 2. Heartsome also attests that "TMX files generated with Heartsome's tools have successfully been tested with Trados Version 6.5.5, SDLX and Deja Vu X." Like the Sun Open Language Tools, Heartsome is based on an XLIFF translation editor, and it also includes a TMX-based translation memory editor. The personal edition of the XLIFF translation editor alone is \$88, and the full translation suite, consisting of the XLIFF editor plus a TMX editor and dictionary editor, is \$398.

**Transolution** (<http://transolution.python-hosting.com/>) provides a suite of tools similar to the Sun Open Language Tools, but written in the Python programming language. Transolution supports the XLIFF standard and runs on Windows or Linux.

Sun Microsystems released the first installment of its **Open Language Tools** (<https://open-language-tools.dev.java.net/>) project, an XLIFF (XML Localization Interchange File Format) Translation Editor and an XLIFF Filter. These are also written in Java so will run on Windows, Linux, or Mac OS X. At present, the tools can be used with a variety of file formats for translating documentation files in HTML, Docbook SGML, JSP, XML, OpenOffice.org and plain text, and software files in .po, Msg, Java .properties, Java ResourceBundle, and Mozilla .DTD resource file formats. Sun gets a big pat on the back for this one! Following is an excerpt from the "Philosophy" section on the Open Language Tools page, if you're wondering why a for-profit company would release a CAT tool for free: "We believe that computers are tools that can help people - in our case, we want computers to help translators. We strongly believe in the value of shared open standards - everything we do should be based on open standards, and should interoperate via those standards. Lastly writing translation tools is a noble cause : we believe that we should be doing all we can to break down the Digital divide, and language is one aspect of

that. If you can't use a computer because it's not translated to your native language, then we want to provide tools to fix that problem.”