



A comparative study of OmegaT and Memsource: advantages and disadvantages of their primary features

Student: Laura Moreno Sorolla
Supervisor: Dr. Bartolomé Mesa Lao
Master's degree in Specialised Translation
Universitat Oberta de Catalunya
February 2018

Contents

Abstract	3
1. Introduction	4
1.1 Motivation	4
1.2 Aim	4
2. Literature review	4
3. Difference between CAT tools and cloud-based translation systems	5
4. Materials and methods	7
4.1 Tools	7
4.2 Methods	7
5. Results	8
6. Advantages and/or disadvantages in OmegaT and Memsource	9
7. Conclusions	9
8. Works cited	11
Annex I. Sample text for the comparative study	12

Abstract

This study is aimed at analysing the functions of sentence retrieving from the TM and the features involving translation project management in two free CAT tools, OmegaT and Memsource. Sentence retrieving from the TM is analysed in terms of the matches obtained. Features that concern translation project management are also analysed according to their presence or absence.

Keywords: free CAT tools, cloud-based translation management system, sentence retrieving, features, translation project management, OmegaT, Memsource.

Aquest estudi està centrat a analitzar les funcions de recuperació de correspondència de la MT i les característiques que inclouen la gestió de projectes de traducció en dues eines TAO gratuïtes, OmegaT i Memsource. Les funcions de recuperació s'analitzen segons les coincidències de la MT i les característiques de la gestió de projectes de traducció s'analitzen segons la seva presència o absència.

Paraules clau: eines TAO gratuïtes, sistema de gestió de traducció basat al núvol, recuperació de correspondència de la MT, característiques, gestió de projectes de traducció, OmegaT, Memsource.

1. Introduction

The emerging of CAT tools in the mid-nineties generated an innovation in the field of translation. However, the first CAT tools available in the market were expensive and not all translators could afford the purchase. For those who could, though, they could experience great advantages when translating with the tool, for instance, the improvement of the translation productivity as well as the increasing of the translation consistency and quality. However, the recognition of CAT tools has not been given merely for being a time-saving means, but for all the resources created when a translation is finished.

If CAT tools are an outstanding resource for translators, cloud-based translation management systems will not be less impressive. Since its origins in the early 2000s, this way of translating has been gaining popularity by virtue of its benefits as compared to desktop-based CAT tools. For instance, no license fee in some cases, maintenance free and the updates are made automatically. Many of us are used to use cloud-computing systems, such as online banking services, e-mail (Gmail), or cloud storage (Dropbox), and this technology has also reached the field of translation. It was in 2007 when Lingotek launched the first free cloud-based translation management system. Yet, it was not until 2009 that cloud-based translation systems became widespread: when Google launched its Translator Toolkit.

Free CAT tools, like OmegaT, are a suitable resource for all kind of translators due to the fact that they share many features included in translation software subject to payment. Nowadays, the trend is starting to change gradually and many translators are aware of the benefits that CAT tools offer compared with software subject to payment: free license, possibility of installing it in more than one computer or the compatibility with the most used operating systems, among others.

With cloud-based translation management systems, in this case Memsourc, the trend followed until now resembles the trend experienced by CAT tools. These systems are relatively new, but those who have tested Memsourc declare that the features that are included in the tool are certainly advantageous for translators.

1.1 Motivation

The reason why the topic of this Master's degree dissertation is based on CAT tools is due to the fact that, at present times, these tools offer great support to translators. Not because they are free it means that the features that they provide are less efficient, quite the opposite. In many cases they are supplied with practical resources that commercial CAT tools do not always possess. Therefore, the study of free CAT tools is relevant for the translation community for the reason that they are undergoing a revolution in terms of popularity. I analyse the features that OmegaT and Memsourc include.

1.2 Aim

The aim of this research is to carry out a comparative study of the different features that two free CAT tools offer and to determine the advantages and/or disadvantages that the presence or absence of these features cause. The two CAT tools under study are OmegaT and Memsourc. In the interest of making the study more rigorous, a medical text sample is used to compare both tools. More precisely, we analyse: 1) the functions of sentence retrieving from the TM, as well as 2) the inclusion of other features related to translation project management. These two veins of research comprise the specific aims of the dissertation.

2. Literature review

Throughout the years, many specialists have written articles dealing with CAT tools, but little has been said about cloud-based translation systems yet. In this respect, the literature available is mainly concerned with the history of CAT tools, the real computer needs for translators or the different translation systems. Neither the studies have focused on the sentence retrieving from the TM that these tools offer to translators nor on the inclusion of other features related to translation project management.

Attila (2015) establishes the further features included in CAT tools: «importing/adding one or more source texts, creating/using one or more translation memories, creating/using one or more term bases, [...] the possibility to use various online resources, term banks, or even more than one machine translation aid». (Attila, 2015: 196)

It is in García (2017) that a first comparison between CAT tools and cloud-based translation systems can be found. Specifically, he focuses on the different elements that are created once a translation project is finished: «CAT-enabled translators deliver a target text *plus* reusable language data. [...] Cloud translators deliver a target text *plus* the language data *plus* metadata». (García, 2017: 66) He also states that «cloud computing reshapes the management of labour in ways that unsettle the traditional relations between managers and workers». (García, 2017: 59)

As Gambín (2014) declares: «one of the most important changes that we can appreciate in these ten years is the proliferation of this type of solution, with a clear trend toward cloud-based solutions». (Gambín, 2014: 46)

Aguilar (2013) explains the different advantages that cloud-based translation systems offer, as well as the disadvantages. She also provides a brief, but useful history of the origins of these tools.

Likewise, Muegge (2012b) focuses on the changes that cloud-based translation systems bring to the way translators work. Yet, he attempts to illustrate other functions that these tools offer, but are not present in CAT tools, such as the project management, instant messaging or online payment, among others.

Zydron (2012) is one of the few writers that highlights the reasons why cloud-based translation management systems, as well as SaaS (software as a service) are gaining recognition among translators. He declares that «currently, over 80% of the global translation market is run by small to medium-sized enterprise (SME) organizations». (Zydron, 2012: 20) He also identifies another reason that explains the advantages of using a cloud-based translation management system:

I have never ceased to be amazed at how inefficient the traditional desktop paradigm is for a highly collaborative environment such as translation. E-mailing files backward and forward, shifting work from translators to reviewers and then possibly back again all introduces delays and potential points of failure. (Zydron, 2012: 20)

Muegge (2011) also emphasizes the multiple benefits that translators can take advantage of when working with a cloud-based translation management system, for instance, sharing resources or role-based access, among others.

Yet, despite all the advantages of cloud-based translation management systems above mentioned, there are still some drawbacks that need to be solved if translators want to work with these tools with all its potential: «we can state that this technology has not yet fulfilled its whole potential and the barriers are still the same – internet infrastructure and the inherent hurdles involved in team work with a freelance base». (Gambín, 2014: 47)

3. Difference between CAT tools and cloud-based translation systems

CAT tools are essential for the translation process itself, whereas cloud-based translation systems transcend the translation practice. Memsource, though, is free for freelance translators, but not for enterprises who need to use more features for their projects. However, Gambín (2014) declares that «we now classify CAT tools in two different categories: desktop-based or cloud-based». (Gambín, 2014: 46) As Attila (2015) declares:

The most popular of freeware TEs is probably *OmegaT*, featuring multiple TMs and TBs, multiple file projects, interface to *Google Translate*, glossaries, *MediaWiki*, around thirty file formats, non-Latin alphabets, spelling checker, and most importantly, compatibility with leading proprietary TEs (TMX, TTX, TXML, XLIFF, SDLXLIFF formats). (Attila, 2015: 198)

Among the features that *OmegaT* provides, the translator can use multiple translation memories simultaneously and it is possible to create new translation memories and glossaries from one project. Furthermore, if a translator does not possess a glossary, it can be imported from other websites provided that it is in format UTF8. Still, Gambín (2014) affirms that «*OmegaT*, probably the open source solution with the greatest impact on the CAT tool industry, now offers the

possibility of sharing an online TM». (Gambin, 2014: 47) This means that the translation industry is slowly realising that the online marketplace has multiple advantages and, at the same time, they are starting to introduce some online features to their tools.

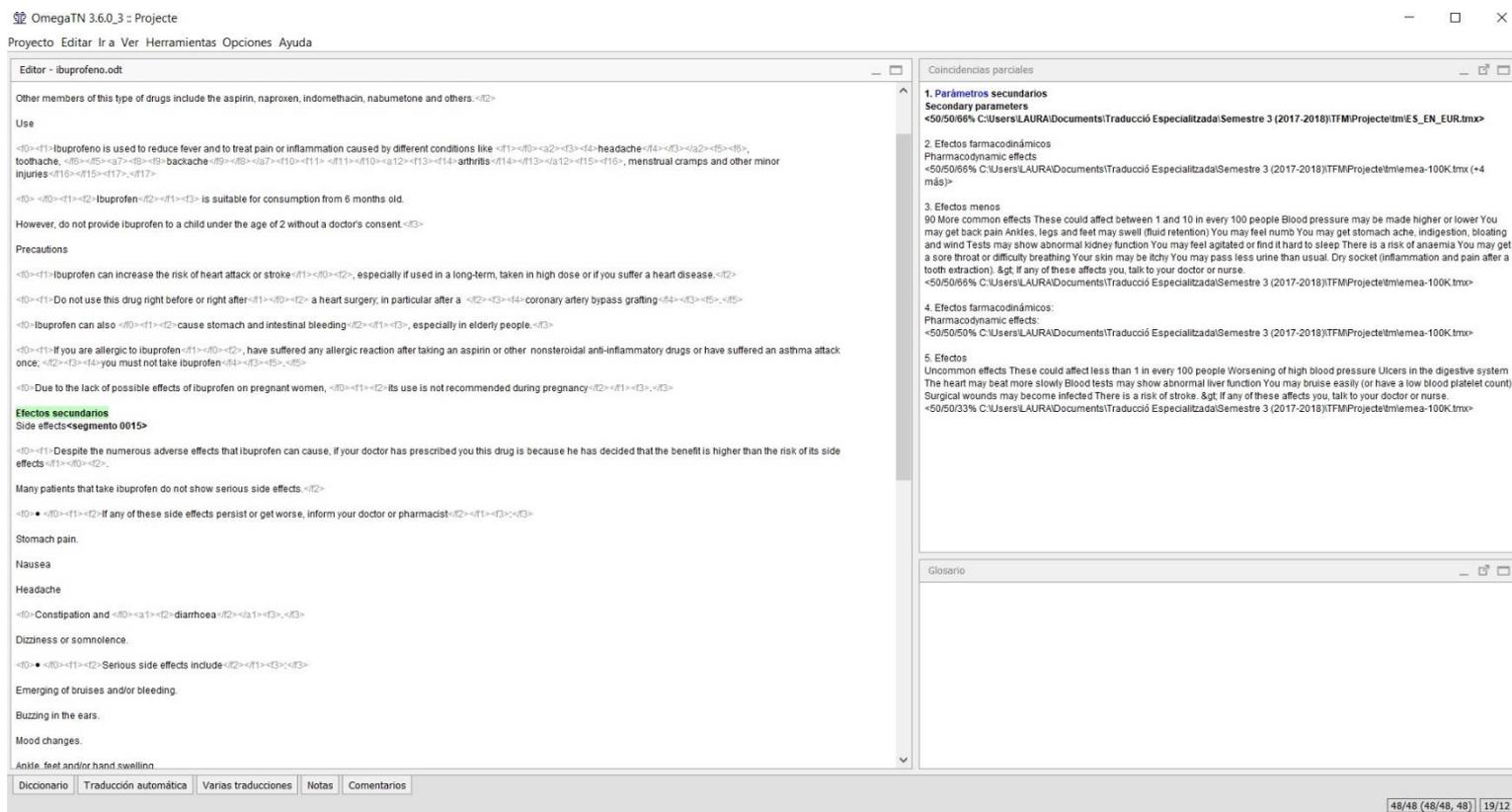


Figure 1. OmegaT's translation editor

Flórez & Alcina (2011) declare that there is another difference between free and payment software that in many cases is not taken into account:

A diferencia del modelo privativo, donde el usuario es visto como un cliente que adquiere un producto terminado, en el software libre los usuarios somos parte importante de la comunidad de desarrollo. [...] La manera más sencilla de contribuir con un proyecto de software libre consiste en utilizar el programa e informar sobre cualquier error que encontremos o función que echemos en falta. (Flórez & Alcina, 2011: 59)

Cloud-based translation management systems are transforming the way in which translators work. As Muegge (2012a) reports:

With a cloud-based TM system where all linguistic assets are stored on a single centralized server, sharing translation memories among multiple linguists is effortless: by granting other users access to your project, either through a simple system setup or via e-mail invitation, teams of almost any size can instantly collaborate. (Muegge, 2012a: 4)

Moreover, there exist other differences regarding CAT tools and cloud-based translation systems. For instance, the workflow is much more effective, as the interface will automatically notify another person (project manager or proofreader) when a specific part of the translation has been concluded. In traditional CAT tools, this step needs to be done by someone.

Zydron (2012) provides a good summary of all the features that translators can benefit from using a cloud-based translation management system: «a proper SaaS TMS/CAT solution will allow you to integrate a customizable web-based ordering and payment system into your own website so that customers can upload, get quotes, pay online and initiate the workflow for translation jobs». (Zydron, 2012: 21) Finally, it is necessary to remember that when using a cloud-based translation management system all the data is securely stored on the server.

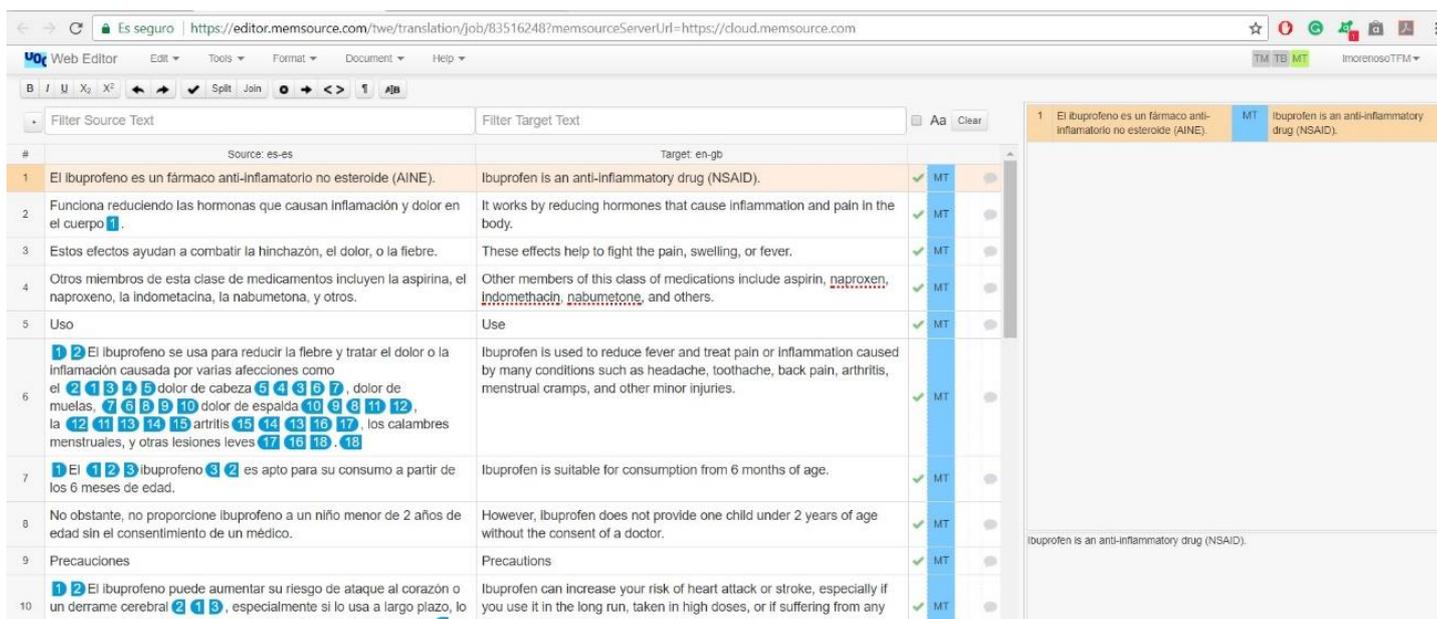


Figure 2. Memsorce's translation web editor

4. Materials and methods

In order to accomplish the comparative study among the two translation tools a sample text in Spanish from the field of medicine has been chosen to be translated into English. The text is an .odt file which has 617 words and 3,756 characters. With this text, as well as a translation memory in TMX format from EMEA (European Medicines Agency) obtained from UOC, the analysis will be focused on the functions for retrieving TM matches and the presence or absence of other features related to project management. In this regard, the text has been translated sentence by sentence.

4.1 Tools

As stated before, OmegaT is a free and open source software. Despite being free, they offer «solutions equivalent to those of proprietary software». (Cánovas & Samson, 2011: 50) Furthermore, open source CAT tools use open standard formats such as TMX and they are compatible with the most used file formats: HTML, XML and XLIFF.

In order to work with Memsorce, it is necessary to obtain a username and a password. Once the translator can access the platform, the steps to create a translation project are rather similar to the ones in other translation tools. The most outstanding feature that translators can find in Memsorce is that the translation project can be managed from the same platform. In this respect, the translator can see the status of the project, the percentage of what has been translated or the people that are involved in that project (project manager, translator and client).

4.2 Methods

On the one hand, to determine the sentence retrieval from the TM, the percentages of matches are explored. The specific matches are the following: repetitions, exact matches, 100%-95%, 94%-85%, 84%-75%, 74%-50% and no match.

On the other hand, the presence or absence of features that allow translation project managing is tested during the process of translating the sample text. While creating the translation project it is going to be taken into account the different steps in which the other members involved in it can do their work in the same tool.

5. Results

With regard to the results of the comparative study of both CAT tools, the function of sentence retrieving from the TM will be analysed first. As the sample text, as well as the translation memory and glossary used have been the same in both OmegaT and Memsourse, the results concerning the sentence retrieving from the TM have been identical. The percentages of matches produced for the two CAT tools are shown in the table below:

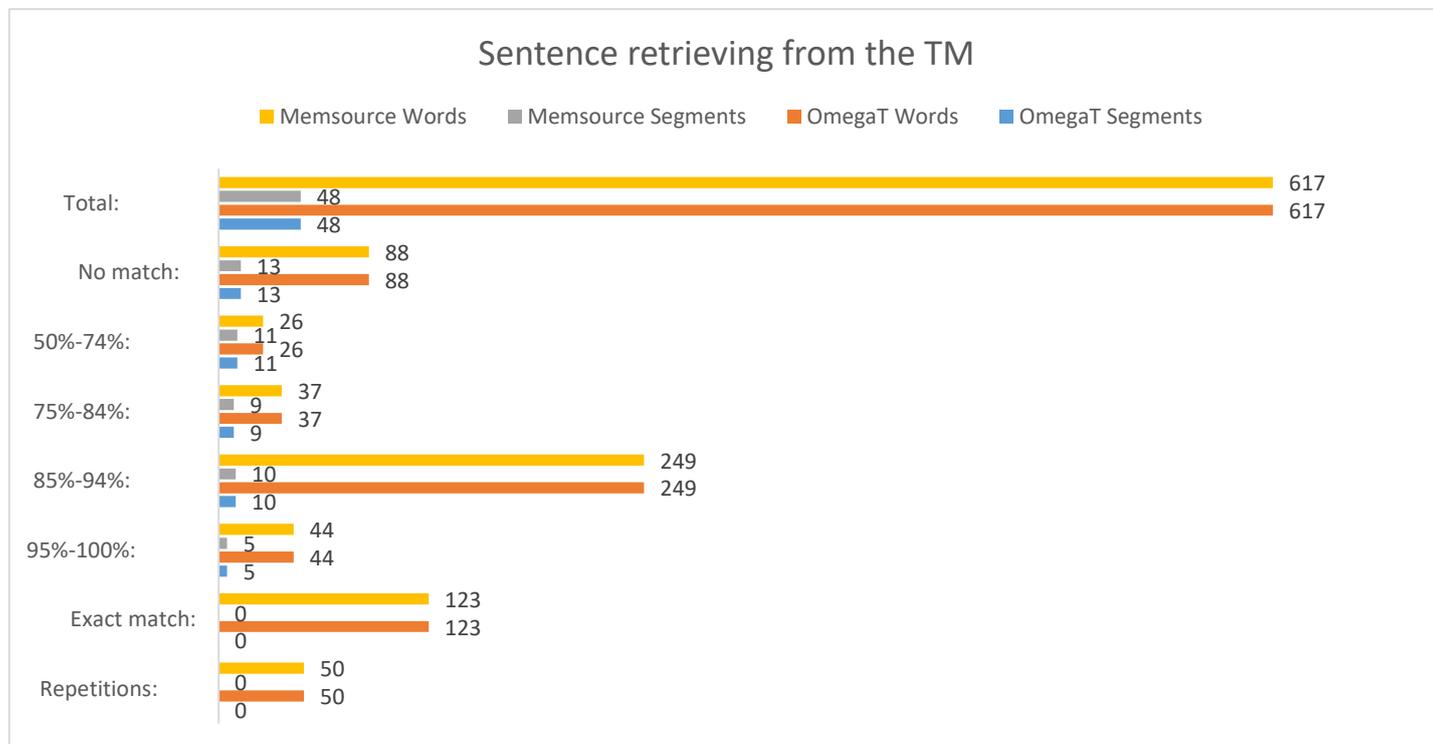


Figure 3. Table with the TM matches obtained with OmegaT and Memsourse

In view of the information that is shown in the table, it can be assured that the sentence retrieving from the TM with a CAT tool and with a cloud-based translation system remains the same. It can be seen in both tables that all the percentages of matches shown are equivalent. This means that the translation work can be done effectively in both tools without major problems. In this respect, both OmegaT and Memsourse provide a favourable, fast and efficient translation experience.

It is in the aspect of the features that allow translation project management when the results differ from OmegaT and Memsourse. Although OmegaT includes many features to conduct an effective translation process, it has been proved that this CAT tool does not include any feature that helps in the management of a translation project. OmegaT has been proved to be outstanding for the complementary features that assist the translator work. In this case, when working with OmegaT, a translator can find extra tools that assist the automatic translation (Belazar, MyMemory, Google Translate, Apertium or Microsoft Translator), as well as the glossary (TaaS). Thanks to these features, the resources for translating are more abundant and thus, time-saving.

Memsourse, though, does include features that allow the management of a translation project. For instance, when creating a translation project, the project manager can also include its due date as well as the status (*new*, *assigned*, *completed* and *cancelled*) and the purchase order. In this first step, it remains clear that Memsourse is provided with features that permit the management of a translation project.

Before finishing creation of the project, Memsourse also includes other settings that automatize the status of the process: set project to *Assigned* once all jobs are *Emailed*, set project to *Assigned* once all jobs are *Accepted*, set project to *Completed* once all jobs are completed, set project to *Completed* once all jobs are *Delivered*, set project to *Cancelled* once all jobs are cancelled and

set project to *Completed* once project is completed by a vendor. To finish, the project manager can assign a price list and the workflow: *translation, revision* and *client review*. In the next step, a linguist can be assigned to the translation, the revision and the client review processes and they are automatically notified.

When the translation process is finished, the status has to be changed to “completed by linguist”. Then, the person assigned for the revision is emailed to let him/her know that the revision process can be started. When the revision is finished, the status has to be changed too. The same steps are followed for the client revision.

Other features that Memsorce offer are the possibility of importing your own translation memories and glossaries, the storing of all the previous translations done and quality assurance checks. Memsorce also offers a desktop Memsorce Editor that can be used when there is no internet connection.

6. Advantages and/or disadvantages in OmegaT and Memsorce

The advantages and disadvantages found in both tools are shown in the following table:

	OmegaT	Memsorce
Advantages	<ul style="list-style-type: none"> -Can be used on different OS. -Acceptance of many file formats. -Updated TMX memory when the translation is finished. -No internet needed. -Better control of the documents 	<ul style="list-style-type: none"> -Project management is possible. -Automatic communication between the participants. -Data stored in the servers. -Import/export TM and glossaries. -Offline editor. -Real-time work. -Can be used on different OS. -Acceptance of many file formats. -Updated TMX memory when the translation is finished.
Disadvantages	<ul style="list-style-type: none"> -No project management is possible. -Data stored in the computer. 	<ul style="list-style-type: none"> -Pay for the full use of its features. -Internet connection is needed. -Fewer control of the documents.

7. Conclusions

Focusing on the translation editor itself, both tools provide an understandable and satisfactory practice. Furthermore, OmegaT and Memsorce share the same features included in the translation editor. This means that if the same TM and glossary is used in a translation, the sentence retrieval from the TM will be the same in both tools.

However, the inclusion of other features related to translation project management is different. On the one hand, OmegaT does not include any of these features. Hence, it works solely as a CAT tool. On the other hand, Memsorce includes features that permit the management of a translation project besides the translation editor, as they have been mentioned in section 5.

However, there are translators and clients who still do not rely on these tools. In this regard it is necessary to re-educate them according to the technological changes that surround the field of translation. Nevertheless, it would be relevant for translators to try the new technology that is developed to examine how helpful it can be. It is claimed that free CAT tools are not sophisticated, but its functionality is usually better.

As the demand for translation is constantly increasing, thanks to the continual development of CAT tools and cloud-based translation management systems, new working environments are being created to fulfil the needs required for translators. Nowadays cloud-based translation management systems offer a complete solution for translators as they have all the necessary elements stored in the cloud. Thanks to their appearance, all the processes involved in a translation project can be centralised in a single tool.

8. Works cited

- Aguilar, A. (2013). "Taking Advantage of the Cloud", *The ATA Chronicle*, XLII, 10-13. doi: http://dx.doi.org/10.1007/978-3-319-61827-2_9
- Attila, I. (2015). "Machine translation and CAT-tools", in Pletl, R. & Nagy, I.K. (coord.), *Nyelvi sokszín ú ség Európában* (pp. 187-203). Cluj-Napoca: Scientia Publishing House.
- Cánovas, M., & Samson, R. (2011). "Open source software in translator training", *Revista.Tradumàtica*, IX,46-56. Doi: <https://doi.org/10.5565/rev/tradumatica.9>
- Flórez, S., & Alcina, A. (2011). "Catálogo de software libre para la traducción", *Revista.Tradumàtica*, IX, 57-73. Doi: <https://doi.org/10.5565/rev/tradumatica.5>
- Gambín, J. (2014). "Evolution of cloud-based translation memory", *MultiLingual*, XXV, 46-49. doi: <http://dig.multilingual.com/2014-04-05/index.html?page=46>
- García, I. (2017). "Translating in the Cloud Age: Online Marketplaces", *Hermes – Journal of Language and Communication in Business*, LVI, 59-70. Doi: <http://dx.doi.org/10.7146/hjlc.v0i56.97202>
- Muegge, U. (2011). *Teaching computer-assisted translation in the cloud: a down-to-earth model*. Lecture presented at International Federation of Translators (FIT) XIX World Congress, San Francisco. Retrieved from https://works.bepress.com/uwe_muegge/57/
- Muegge, U. (2012a). Cloud-based translation memory tools are changing the way translators work and train. *The Big Wave of Language Technology*, I. Retrieved from http://works.bepress.com/uwe_muegge/76/
- Muegge, U. (2012b). "The silent revolution: Cloud-based translation management System", *Tcworld Magazine*, VII, 17-21. Retrieved from http://works.bepress.com/uwe_muegge/67/
- Zydron, A. (2012). "Cloud computing, SaaS and translation tools", *MultiLingual*, XXIII, 20-21. Retrieved from <http://dig.multilingual.com/2012-01-02/index.html?page=20>

Annex I. Sample text for the comparative study

El ibuprofeno es un fármaco anti-inflamatorio no esteroide (AINE). Funciona reduciendo las hormonas que causan inflamación y dolor en el cuerpo. Estos efectos ayudan a combatir la hinchazón, el dolor, o la fiebre. Otros miembros de esta clase de medicamentos incluyen la aspirina, el naproxeno, la indometacina, la nabumetona, y otros.

Uso

El ibuprofeno se usa para reducir la fiebre y tratar el dolor o la inflamación causada por varias afecciones como el [dolor de cabeza](#), dolor de muelas, [dolor de espalda](#), la [artritis](#), los calambres menstruales, y otras lesiones leves.

El **ibuprofeno** es apto para su consumo a partir de los 6 meses de edad. No obstante, no proporcione ibuprofeno a un niño menor de 2 años de edad sin el consentimiento de un médico.

Precauciones

- **El ibuprofeno puede aumentar su riesgo de ataque al corazón o un derrame cerebral**, especialmente si lo usa a largo plazo, lo toma en dosis altas, o si padece alguna enfermedad del corazón.
- **No utilice este medicamento justo antes o justo después de una operación cardiaca; en concreto, después de una derivación aortocoronaria con injerto.**
- El ibuprofeno también **puede causar sangrado del estómago o intestinal**, sobre todo en personas mayores.
- **Si es alérgico al ibuprofeno**, ha sufrido alguna reacción alérgica después de tomar aspirina u otros medicamentos antiinflamatorios no esteroideos, o ha sufrido alguna vez un ataque de asma; **no deberá de tomar ibuprofeno.**
- Debido a que todavía no hay evidencias de posibles efectos del ibuprofeno en mujeres embarazadas, **no se recomienda su uso durante el embarazo.**

Efectos secundarios

A pesar de los múltiples efectos adversos que puede producir el ibuprofeno, si su médico le ha recetado este fármaco es porque ha juzgado que su beneficio es mayor que el riesgo de sus efectos secundarios. Mucho de los pacientes que toman ibuprofeno no presentan efectos secundarios grave.

• **Si cualquiera de estos efectos persiste o empeora, informe a su médico o farmacéutico:**

- Malestar estomacal.
- Náuseas.
- Dolor de cabeza.
- Estreñimiento y [diarrea](#).
- Mareos o somnolencia.

• **Efectos secundarios más graves incluyen:**

- Fácil aparición de moretones y/o sangrado.
- Zumbido en los oídos.
- Cambios anímicos.
- Hinchazón de los tobillos, pies y/o manos.
- Aumento repentino del peso corporal y rigidez en el cuello sin explicación aparente.
- Cambios en la visión.
- Cansancio inusual.
- Cambio en la cantidad de líquido expulsado al orinar.

• **El ibuprofeno también puede aumentar la presión arterial.**

• **En muy raras ocasiones, también puede causar una enfermedad hepática grave.** Indicios de esta complicación incluye: orina oscura, vómitos persistentes, pérdida de apetito, dolor estomacal, náuseas, ...

Los efectos secundarios antes mencionados son sólo una parte de los posible efectos secundarios del ibuprofeno. Lea el prospecto o consulte a su farmacéutico para más información.

Dosis

Siga las instrucciones exactas proporcionadas por su médico, o según se lo a indicado su farmacéutico / receta (según corresponda).

Use la dosis más baja necesaria para tratar su condición. Una sobredosis de ibuprofeno puede dañar el estómago o los intestinos. Tome este medicamento con comida o leche para reducir el malestar estomacal.

- Para dolores leves a moderados, calambres menstruales, o la fiebre, la dosis habitual para adultos suele ser de 200 a 400mg cada 4 / 6 horas.

- La artritis se trata con 300 a 800 mg 3 o 4 veces al día. La artritis juvenil con 20 a 40mg por kg y día en 3-4 dosis.
- Sin seguimiento médico, la dosis MÁXIMA recomendada es de 1,2g al día.
- A excepción de que lo indique un médico, no se debe usar ibuprofeno durante más de 10 días para combatir el dolor, o más de 3 días para combatir la fiebre.
- La dosis MÁXIMA en niños de entre 6 meses y 12 años de edad es de 40 mg / kg al día.