

Comparing the Effectiveness of Google Translate and MateCat Tools in the Translation of Scientific Texts from English into Arabic

مقارنة بين فاعلية أدوات جوجل للترجمة ومايت كات في ترجمة النصوص العلمية من الإنجليزية إلى العربية

Rami BOUOUDEN¹, Saida KOHIL²

¹TRADIL Laboratory, Translation Department, Badji Mokhtar Annaba University (Algeria),
rami.bououden@univ-annaba.org

²TRADIL Laboratory, Translation Department, Badji Mokhtar Annaba University (Algeria),
kohsaida@yahoo.fr

Received: 15/08/2022

Accepted: 05/09/2022

Published: 24/12/2022

Abstract:

Translation technological tools are receiving major updates continuously. Even though machine translation (MT) is one of the most used technologies, computer-assisted translation (CAT) tools proved to be more reliable in aiding translators with their challenging tasks. However, combining these two into what is called Machine Translation Enhanced Computer Assisted Translation is an interesting trend. This paper primarily focuses on introducing the MateCat tool and explaining how it works. Then, it employs a descriptive approach to compare two translations of a science article from English into Arabic using Google Translate and MateCat, respectively. This process aims to see how well these tools carry out this task in certain areas. The results show that the translation quality of the MT-enhanced MateCate was better than that of the stand-alone Google Translate tool.

Keywords: Translation; Translation Technology; Machine Translation; MateCat; Google Translate.

ملخص:

تتلقى الأدوات التكنولوجية الخاصة بمجال الترجمة تحديات مهمة باستمرار، ورغم أنَّ الترجمة الآلية تعدُّ من بين التكنولوجيات أكثر استعمالاً، فإنَّ أدوات الترجمة بمساعدة الحاسوب أثبتت بأنه يمكن الاعتماد عليها أكثر حينما يتعلق الأمر بمساعدة المترجمين في مهامهم الصعبة، ولكن دمج هاتين الأداةين فيما يُسمَّى بالترجمة بمساعدة الحاسوب المعززة آلياً يعد توجُّهاً مثيراً للاهتمام. يُركز هذا البحث بصورة أساس على التعريف بأداة "مايت كات" وكيفية عملها، مع اعتماد نهج وصفي بُغية المقارنة بين ترجمتين من الإنجليزية إلى العربية لمقال علمي، وذلك باستعمال جوجل للترجمة ومايت كات تبعاً، والهدف من هذه المقارنة يتجلى في معرفة مدى جودة أداء كلتا الأداةين في نواحٍ مُعينة، وقد أظهرت نتائج الدراسة أن جودة ترجمة أداة مايت كات المزودة بالترجمة الآلية كانت أفضل من ترجمة أداة جوجل المستقلة.

كلمات مفتاحية: الترجمة، تكنولوجيا الترجمة، الترجمة الآلية، مايت كات، جوجل للترجمة.

1. Introduction

With various technological tools in the field of translation, translators can now perform their translation tasks in a way that makes the workflow more flexible and productive. In this respect, computer-assisted translation (CAT) and machine translation (MT) are two of the most prominent tools in the industry. The former consists of a variety of convenient features that, most of the time, lead to a satisfying result when combined with proper human intervention. Despite its effectiveness in some remarkable aspects, the latter still does not perform effectively in other areas compared to CAT tools, in which the core element is the translation memory (TM). This paper focuses on the *Machine Translation Enhanced Computer Assisted Translation* by introducing the MateCat tool and its main features. The practical section uses a descriptive approach to compare two translations of an article from English into Arabic, which were done using Google translate and MateCat, respectively. The main goal of this comparison is to show how the two systems perform in some specific aspects, which will be explained in the following sections of this paper.

Previous studies discussed different means of translation, including various systems and tools. Some of these studies compared these tools, while others focused on tool vs. human translator comparisons. Many papers compared human translation and machine/automated translation (Ahrenberg, 2017; Precup-Stiegelbauer, 2013; Xiu & Xeauyin, 2018), while others compared MT and TMs (Milad, 2021) or MT versus CAT (Wang, 2014). Furthermore, Some publications discussed the differences between paid and free tools (Apriliana, Kurniawan, Ferianda, & Kastuhandani, 2016). Unlike those studies, the current paper tries to tackle a different set of tools by comparing the performance of MT with Machine Translation Enhanced Computer-Assisted Translation in translating a particular type of document. Thus, conducting this research will help understand other aspects of these two tools, and bridge the gap in the presented literature.

2. Translation Technologies

2.1 Machine Translation Enhanced Computer-Assisted Translation

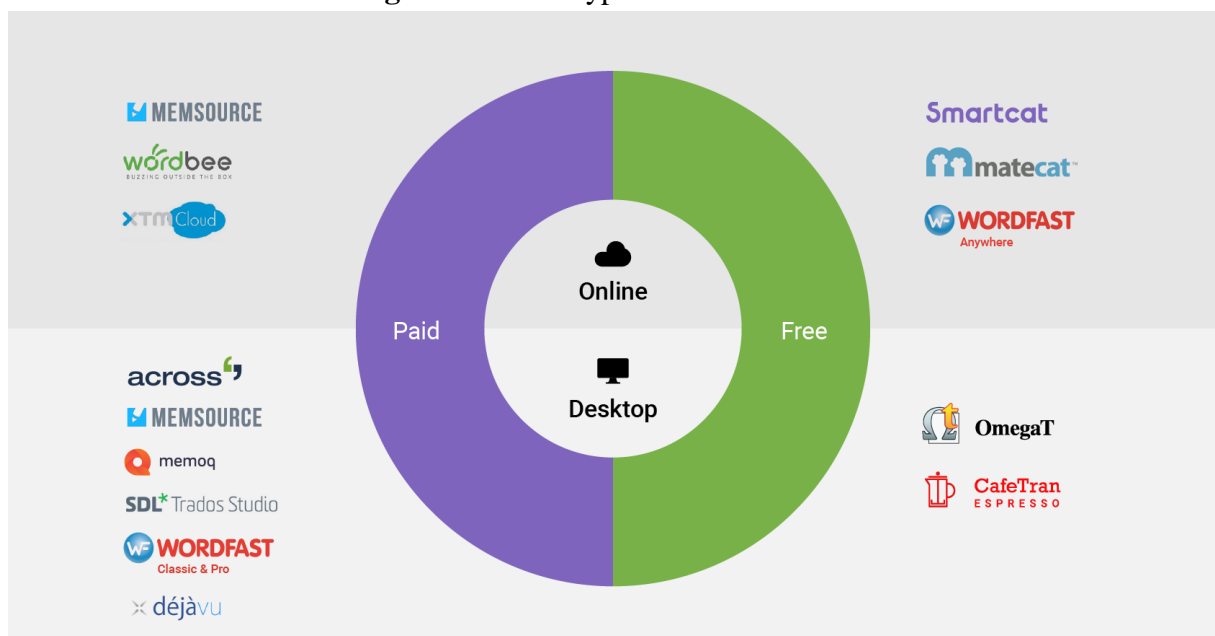
Computer-assisted/aided translation has become a trendy and valuable resource for translators worldwide. With its diverse functions, whether free or paid, software or web-based, this technology made the translation industry livelier than ever. It consists of utilizing the available technological tools to perform translation tasks, and the process mainly falls under the purview of the human element (Bowker & Fisher, 2010). This means that the user is the one in charge of rectifying linguistic inaccuracies occurring during the process (Christensen & Schjoldager, 2017).

CAT tools contain several add-ons or accessories, such as TMs, terminological databases (glossaries), spellcheckers, and MT engines. Unquestionably, TMs are the essential integral part of any Cat tool, as they continuously save what is being

translated to be exploited in future works (Doherty, 2016). It is worth mentioning that translation memories, during first-time utilization, are vacant, and they gradually store segments after each translation job, or they can be enlarged via importing data from external sources, such as parallel corpora (Bowker & Corpas Pastor, 2015). Consequently, when a CAT tool has a considerable amount of source and target data, similar translations will be suggested by the system. There is a myriad of available CAT tools. Figure 1 shows some of the available CAT tools in the market. They are divided into four interwoven categories: paid, free, online, and desktop.

According to figure 1, popular paid CAT tools include the desktop programs, such as SDL Trados Studio, memoQ, Across, Wordfast Pro, and Déjà Vu. There are also other paid programs in the form of an online platform, like Wordbee, XTM Cloud, and Memsorce, which can also be downloaded for desktop use. Despite their different features and accessories, These paid tools are categorized by their fast and reliable performance and extended licenses (Apriliana et al., 2016). In addition, users can directly install the desktop versions of these tools on their devices and use them without being connected to the internet. On the other hand, online tools require an internet connection and can provide an excellent cooperative space for group projects. Besides, Free CAT tools, although not as complex as their paid counterparts, offer a decent alternative with basic free-of-charge features that are very useful to a certain degree. Free well-known online CAT tools include Smartcat, MateCat, and Wordfast Anywhere, while OmegaT and CafeTran Espresso fall in the desktop category.

Fig.1. Different Types of CAT Tools



Source: (Pearse, 2019)

On the other hand, Machine Translation Enhanced Computer-Assisted Translation is a very interesting topic. MateCat, which will be discussed in section (3), is a good example of this technology. This specific tool works conveniently; “When the

translator opens a segment, the CAT tool tries to propose possible translation suggestions, originating from the translation memory and/or from a machine translation engine” (Bertoldi, Cettolo, & Federico, 2013, p. 36). Therefore, this confirms that integrating an MT engine with a CAT tool can be extremely helpful because the translator is free to select the suggested MT option or ignore it based on its quality. Bertoldi et al. (2013) further add that it is entirely up to the translator to either choose one of the results or redo his work.

2.2 Machine Translation

Regardless, MT, as a stand-alone tool, does not give the same advantages compared with being integrated with a CAT tool. Without human involvement, MT generates translations entirely via technology (Precup-Stiegelbauer, 2013). Unlike CAT tools, automatic translation is done directly through specific MT engines. These engines use different approaches to analyze and translate the source text. Among the well-known engines, there is Google Translate¹, Bing Microsoft Translator, DeepL, Reverso Translation, and Amazon Translate. Translating a sizeable number of data is one of MT's advantages, yet the literal translations that are separated from context are its notable inconvenience (Peng, 2018). Moreover, despite having some analogous features like CAT tools, MT still cannot match the former in terms of TM available options, terminological matches, and miscellaneous sentence structures (Wang, 2014).

3. The MateCat Tool

3.1 Introducing MateCat

MateCat² is the acronym for "Machine Translation Enhanced Computer Assisted Translation" (Federico et al., 2014). It is an open-source online (web-based) CAT tool that provides several free features and resources for a variety of users, including students, professional translators, freelancers, and project managers. Before its release, this tool started as a research project supported by the European Union to enhance the interaction between human translation and MT (Federico, Koehn, Schwenk, & Trombetti, 2013). It is also suitable for working in groups and teams on one or several projects. MateCat also consists of a plethora of assets that make it stand out from other CAT tools. For instance, some MT engines are available to users, such as Google Translate, MyMemory, Yandex Translate, Intento, Apartium, Moses, and others. This CAT tool allows translators to use a public collaborative TM provided by MyMemory³. It also gives them the freedom to import their TM TMX files and glossaries that fit their subject matter and project criteria. TMX is the abbreviation of Translation Memory eXchange. TMX files are used for sharing TM data (Bussey, 2020). Besides, MateCat supports up to 79 file formats and Google Drive

¹ <https://translate.google.com/>

² <https://www.matecat.com/>

³ <https://mymemory.translated.net/>

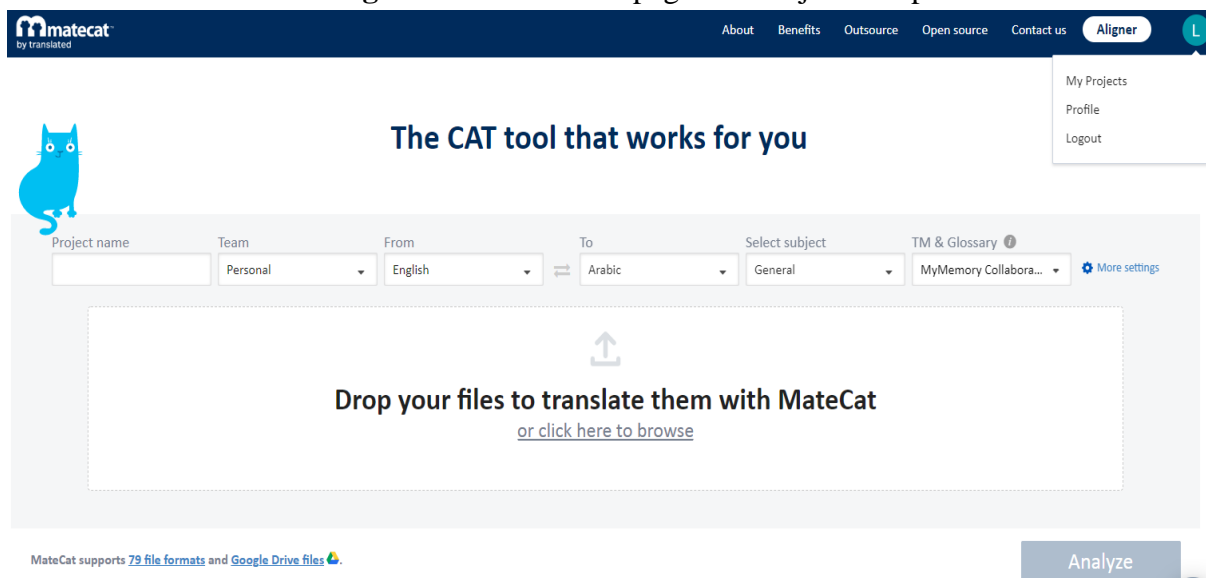
files. Like many other CAT tools, MateCat is not only a good translation tool but also a suitable post-editing space for the segmentation of the document. This facilitates choosing the suggested translations from the previously mentioned sources, correcting errors, and proofreading. MateCat can also process documents that contain Mark-up Tags and accepts Unicode (UTF-8), which can read non-Latin alphabets and right-to-left scripts (Federico et al., 2014).

3.2 How to Use MateCat

MateCat has a user-friendly interface. It is simple and provides many options to choose from. Consequently, students and beginners, in general, do not need specialized training or coaching to learn about this tool. The purpose of this section is to give a brief tutorial about the key steps of translating a document on this CAT tool.

When accessing the website (matecat.com), the main page directly pops up for the user. It is possible to start the preparation of a project without signing in, but it is advisable to do so for a more convenient experience (figure 2).

Fig.2. MateCat's Homepage and Project Setup



After signing in, users can start preparing their projects by dragging and dropping their files into the upload box or by browsing and choosing the concerned files. Subsequently, users have six fields that can be altered, including **Project name**, **Team**, **SL** (source language), **TL** (target language), **Subject Field**, and **TM & Glossary**. Naming a project is always up to the translator, who should specify it to find it easily. There are more than 200 languages to choose from for the SL and TL fields. Next, the user can choose the area that suits his document in the *subject matter* field. The TM and Glossary field include several options. By clicking on *Create Resource* or *More settings*, several options appear to help organize the project's resources (Figure 3).

Fig.3. MateCat's TM, Glossary, and MT settings

The screenshot shows the 'Settings' page in MateCat, with the 'Translation Memory and Glossary' tab selected. At the top, there is a checkbox for 'Pre-translate 100% matches from TM' which is checked. Below this, there is a section for 'Active Resources' with a '+ New resource' button. Under 'Active Resources', there are two entries: 'MyMemory: Collaborative translation memory shared with all MateCat users.' and 'Private TM and Glossary'. The 'Private TM and Glossary' entry has a text input field containing 'Private TM and Glossary', a unique ID '2656f1c658e60a68e996', and an 'Import TMX' button. Below the active resources, there is a section for 'Inactive Resources' with a 'Filter resources' button. Under 'Inactive Resources', there are two entries: 'English - Arabic Personal TM' and 'Personal English - Arabic TM'. Each entry has a checkbox, a text input field, a unique ID, and an 'Import TMX' button. At the bottom of the active resources section, there is a 'Select TMX file to import' section with a 'Choose File' button and the text 'No file chosen'.

As it is shown in figure 2, users can choose a public TM or integrate their TM by importing a TMX file from their computer. It is also possible to import a TMX file containing a glossary that can assist in finding the specialized terms according to the subject matter. The next step is choosing an MT engine from the list displayed in figure 4 underneath.

Fig.4. MT Engines in MateCat

The screenshot shows the 'Settings' page in MateCat, with the 'Machine Translation' tab selected. Below the tab, there is a section for 'Add MT engine'. A dropdown menu is open, showing a list of MT engines. The list includes: 'Choose provider...', 'ModernMT', 'AltLang', 'Apertium', 'Google Translate', 'Intento', 'IPTranslator from Iconic', 'Microsoft Translator Hub', 'Moses', 'MT-HUB', 'SmartMATE', 'Tauyou', 'Tilde MT', and 'Yandex.Translate'.

In the advanced options section (figure 4), users can click on the *Dictation* option to activate the speech-to-text feature, which is very useful depending on the person

Comparing the Effectiveness of Google Translate and MateCat Tools in the Translation of Scientific Texts from English into Arabic

working with it. *Guess tag position* and *QA by lexiqa* help organize the translation by placing tags, punctuation, symbols, etc. Another functional option is *Cross-language matches*. It allows users to get help from familiar tongues as a suggestion. Finally, *Segmentation Rules* are how parts of the document are translated by choosing General, Patent, or Paragraph.

Fig.5. Advanced Options Section

The screenshot shows the 'Advanced Options' section of the MateCat interface. It features a navigation bar with 'Translation Memory and Glossary', 'Machine Translation', and 'Advanced Options'. Below the navigation bar, the 'Advanced Options' title is displayed. The settings are organized into five rows, each with a title, a description, and a control element:

- Dictation:** Improved accessibility thanks to a speech-to-text component to dictate your translations instead of typing them. Inactive
- Guess tag position:** Enable this functionality to let MateCat automatically place the tags where they belong. [Supported languages](#) Active
- QA by lexiqa:** Linguistic QA with automated checks for punctuation, numerals, links, symbols, etc. [Supported languages](#) Active
- Cross-language Matches:** Get translation suggestions in other target languages you know as reference. Primary language suggestion: Secondary language suggestion:
- Segmentation Rules:** Select how sentences are split according to specific types of content. General

After finishing this preparation step, clicking on **Analyse** will immediately take the user to the analysis report of the document (figure 6). This page shows the document's analytical information, including word count, weighted words, job URL, and the split option. These details are handy when performing a paid project. Then, the user needs to press **Translate** to start executing his work (figure 5).

Fig.6. Analysis Report Page

The screenshot shows the 'Analysis Report Page' in the MateCat interface. It features a 'Volume Analysis' section with the following details:

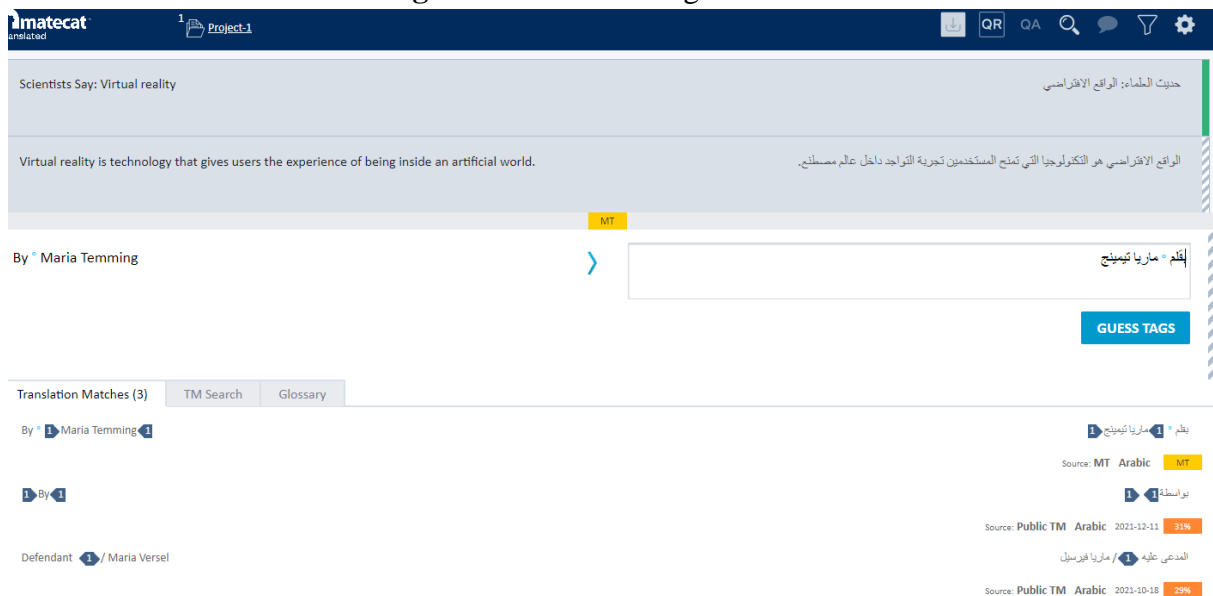
- Translation_1**
- Analysis: complete** | [Download Analysis Report](#)
- 12% Saving on word count** (37 work minutes at 3,000 w/day)
- MateCat gives you more matches than any other tool thanks to a better integration of machine translation and translation memories.

	Total word count	Industry weighted	MateCat weighted	
ID: 5824976 English > Arabic	263	256	231	<input type="button" value="Split"/> <input type="button" value="Translate"/> Buy Translation by translated.

URL: <https://www.matecat.com/translate/translation-1/e...>

The last phase in this process is editing the translation. As shown in figure 7 below, matches with percentages are presented by both MT and public TM. This gives the user more options, allowing him to get the most suitable translation. TM search can also be performed to look for more translations in the previously added TMs. Additionally, glossaries can be created and updated throughout the process. Getting tags managed by pressing on **GUESS TAGS** and then on **Translated** is also possible to finish up the segment. When the translation and editing process is completed, the target document can be obtained via pressing the **download icon** on the top right of the window (figure 7).

Fig.7. MateCat's Editing Section



Last but not least, MateCat's project manager helps users keep track of their current and past projects (figure 8). They can assign the job to other translators, set a password, export TMX files, archive, or cancel the project.

Fig.8. MateCat's Project Management Section



4. Methodology

The analysis in this research paper followed a descriptive approach. It covered two translations of a science article obtained from "Science News Explores" magazine⁴. The translations were done by Google Translate (MT) and MateCat (TM system), from English to Arabic, to compare the performance of these two tools in terms of format, punctuation and ideas structure, vocabulary, and errors. This comparison aims to see how well MateCat, a free TM system, executes the translation process from English to Arabic compared to the MT engine Google Translate. The analysis also tried to highlight both advantages and shortcomings of the two systems.

The original text is an article entitled "Scientists Say: Virtual Reality" by Maria Temming and published in "Science News Explores" magazine in English on June 27, 2022, at 6:30 am (Temming, 2022). It is a science article that revolves around a technological topic (VR). This piece of writing contains some specialized terms and primarily uses simple language understandable to a wide range of readers.

In this process, it is relevant to note that the text was copied directly into Google Translate online tool. It was not translated into the form of a document since the final translation contained many flaws, such as word misplacements and errors in the target language (Arabic). On the other hand, MateCat was provided with some resources including, two translation memories, a personal and a public collaborative TM provided by the website, and a machine translation engine (MyMemory). No glossary fitted the type of article, so it was not integrated.

5. Results

Table 1. Translation Results by MateCat and Google Translate

Source Text (ST)	Target Text1: Google Translate (TT1)	Target Text2: MateCat (TT2)
<p>Scientists Say: Virtual reality</p> <p>Virtual reality is technology that gives users the experience of being inside an artificial world.</p> <p>By Maria Temming</p> <p>June 27, 2022 at 6:30 am</p> <p>Virtual reality</p> <p>Virtual reality, or VR, is an artificial world created by a computer. Users can enter such artificial worlds by strapping on</p>	<p>يقول العلماء: الواقع الافتراضي</p> <p>الواقع الافتراضي هو تقنية تمنح المستخدمين تجربة التواجد داخل عالم اصطناعي.</p> <p>بقلم ماريا تميمينغ</p> <p>27 يونيو 2022 الساعة 6:30 صباحًا</p> <p>الواقع الافتراضي</p> <p>الواقع الافتراضي، أو VR، هو عالم اصطناعي تم إنشاؤه بواسطة الكمبيوتر. يمكن للمستخدمين الدخول إلى مثل هذه العوالم الاصطناعية عن طريق ربط سماعة رأس VR. تحجب السماعة مشاهد</p>	<p>حديث العلماء: الواقع الافتراضي</p> <p>الواقع الافتراضي تكنولوجيا تتيح للمستخدمين تجربة الدخول إلى عالم اصطناعي.</p> <p>بقلم ماريا تميمينغ</p> <p>27 جوان 2022 الساعة 6:30 صباحًا</p> <p>الواقع الافتراضي</p> <p>يعد الواقع الافتراضي، أو ما يُعرف بـ VR، عالمًا اصطناعيًا أنشئَ بواسطة الحاسوب. بإمكان المستخدمين دخول هذه العوالم الاصطناعية من خلال ارتداء خوذة</p>

⁴ <https://www.sciencenews.org/>

a VR headset. The headset blocks out sights of the real world. It also displays views of the simulated environment. Those views seem to have depth. This gives a user the sense that they are inside a real, 3-D space. (VR headsets create that illusion with a trick called stereoscopy.) Motion tracking allows a VR system to change a user's view of a virtual world as they turn their head. And users can interact with objects in VR using handheld controllers. Adding sounds can make the experience even more immersive.

Let's learn about virtual reality

VR hasn't just made video games more realistic. Immersing people in calm VR settings can help soothe pain. And facing fears in VR can help people conquer those fears in real life. Plus, creating 3-D models of historic sites in VR could preserve artifacts lost to sea level rise. And rendering extinct animals in VR could help bring history to life.

So far, VR has mostly been limited to sights and sounds. But some scientists are working to add touch, or haptic sensations to VR. That could make computer-generated worlds even more immersive.

In a sentence

With new technology, people who can't hold hand controllers could navigate virtual reality with facial expressions

من العالم الحقيقي. يعرض أيضًا طرق عرض البيئة المحاكاة. يبدو أن تلك الآراء لها عمق. يمنح هذا المستخدم إحساسًا بأنه داخل مساحة ثلاثية الأبعاد حقيقية. (تخلق سماعات الرأس VR هذا الوهم بخدعة تسمى التنظير المجسم). يسمح تتبع الحركة لنظام الواقع الافتراضي بتغيير نظرة المستخدم للعالم الافتراضي أثناء إدارة رأسه. ويمكن للمستخدمين التفاعل مع الكائنات في الواقع الافتراضي باستخدام وحدات التحكم المحمولة باليد. يمكن أن تجعل إضافة الأصوات التجربة أكثر غامرة.

دعونا نتعرف على الواقع الافتراضي

لم يجعل الواقع الافتراضي ألعاب الفيديو أكثر واقعية. يمكن أن يساعد غمر الأشخاص في إعدادات الواقع الافتراضي الهادئة في تهدئة الألم. ومواجهة المخاوف في الواقع الافتراضي يمكن أن تساعد الناس في التغلب على تلك المخاوف في الحياة الواقعية. بالإضافة إلى ذلك، فإن إنشاء نماذج ثلاثية الأبعاد للمواقع التاريخية في الواقع الافتراضي يمكن أن يحافظ على القطع الأثرية المفقودة بسبب ارتفاع مستوى سطح البحر. ويمكن أن يساعد تحويل الحيوانات المنقرضة في الواقع الافتراضي إلى إحياء التاريخ.

حتى الآن، اقتصر الواقع الافتراضي في الغالب على المشاهد والأصوات. لكن بعض العلماء يعملون لإضافة لمسة أو أحاسيس لمسية إلى الواقع الافتراضي. يمكن أن يجعل ذلك العوالم التي تم إنشاؤها بواسطة الكمبيوتر أكثر غامرة.

في جملة

باستخدام التكنولوجيا الجديدة، يمكن للأشخاص الذين لا يستطيعون حمل أدوات التحكم اليدوية التنقل في الواقع الافتراضي باستخدام تعابير الوجه.

الواقع الافتراضي، بحيث تحجب هذه الخوذة العالم الحقيقي وتعرض مشاهد من البيئة المحاكاة، فيبدو أن لتلك المشاهد عمقا، مما يجعل المستخدم يُحسُّ بأنه داخل مساحة حقيقية ثلاثية الأبعاد. (تُنشئُ خوذة الواقع الافتراضي هذا الوهم بخدعة تسمى التصوير المجسم). كما تسمح تقنية تتبع الحركة لنظام الواقع الافتراضي بتغيير ما يراه المستخدم داخل عالم افتراضي حينما يدير رأسه، ويستطيع المستخدمون التفاعل مع الأشياء داخل الواقع الافتراضي عن طريق أجهزة التحكم المحمولة، فيما ستزيد إضافة الأصوات من إمكانية الإنغماس أكثر في التجربة.

لتتعرف على الواقع الافتراضي

لم يجعل الواقع الافتراضي ألعاب الفيديو أكثر واقعية فحسب، بل إن الانغماس في بيئة افتراضية هادئة يمكن أن يساعد الأشخاص على تخفيف الألم، ومواجهة المخاوف في الواقع الافتراضي يمكن أن تساعدهم على التغلب على تلك المخاوف في الحياة الواقعية. إضافة إلى ذلك فإن إنشاء نماذج ثلاثية الأبعاد للمواقع التاريخية داخل الواقع الافتراضي قد يساهم في الحفاظ على القطع الأثرية المفقودة بسبب ارتفاع منسوب مياه البحر، ثم إن إعادة تجسيد الحيوانات المنقرضة داخل الواقع الافتراضي قد يساهم في إعادة إحياء التاريخ.

وإلى الآن فإن الواقع الافتراضي عموما قد اقتصر على المشاهد والأصوات، غير أن بعض العلماء يعملون على إضافة لللمس أو الإحساس اللمسي إلى الواقع الافتراضي، مما قد يزيد من مستوى الإنغماس داخل العوالم المنشأة بالحاسوب.

باختصار

وبفضل التكنولوجيا الجديدة، فإنه بإمكان الأشخاص الذين لا يستطيعون حمل أجهزة التحكم اليدوية التنقل داخل الواقع الافتراضي باستخدام تعابير الوجه.

Table 1 displays the results of the two translations of the article (Scientists Say: Virtual Reality) using Google Translate (TT1) and MateCat (TT2) from English to Arabic, respectively. As the practical part of this paper, the analysis of these results highlights some aspects which fall under the umbrella term of "Translation Quality." These points are further supported by examples taken directly from the table.

Instead of focusing on the translation or linguistic facet per se, the researcher sees that the final product format is also to be analyzed. This is because the current translation technologies are more than capable of preserving the format of the original work.

The ST format is peculiar, containing specific colors, font sizes, Italics, and Tags. On the one hand, MateCat kept the same format as the original, which is immensely important. Some clients often recommend that the translation should keep the same format as the original, yet others may not do so. On the other hand, Arabic is a language where Italics are not widely found or used, so that they can be easily eliminated. The same goes for Tags since they usually contain online links that directly lead to another article or website in the original language. Nevertheless, they can be a good reference for further reading. One last point in the format department is spacing between words. The spacing in the translation done by MateCat was flawless and performed adequately. On the contrary, Google Translate has some format issues. The translation was performed on the website directly due to the previously mentioned inconvenience in the methodology section. The format of the ST was not preserved in the TT, and four (4) unnecessary spaces occurred throughout the process. This may cause many difficulties, mainly when translating a specific type of document. Examples of these unnecessary spaces include (باستخدام التكنولوجيا) , (بالإضافة إلى ذلك، فإن) (حتى الآن ، اقتصروا) ، (الجديدة ، يمكن

The second aspect that should be tackled is punctuation and ideas structure. Since English and Arabic are two different languages, punctuation usually takes a different direction when translating. Google's machine translation kept the same punctuation in the Arabic version of the text. Commas and full stops separated the ideas based on the structure of the English language rather than Arabic, affecting the coherence of ideas in the target language. For example, in the first and second paragraphs, the punctuation marks in the TT were entirely kept as full stops, taking into consideration that they almost expanded the same idea. In this case, the commas are a better alternative, bearing in mind that Arabic is a language that supports using them instead of continuously relying on full stops. Usually, full stops mean that the idea is finished when used in Arabic. It is worth mentioning that MateCat solved this problem by segmenting the translation units, adjusting the punctuation marks, and adapting those that suit the essence of the target language. This helped the target text to be more organized in meaning and form. For instance, Translation Memories provided some

useful conjunctions to use instead of the repetitive punctuation marks or phrases, e.g., prepositions (و، ف، أو) and coordinators (ثمَّ إنَّ، كما، حيثُ)، as displayed in table 1.

The next and pivotal aspect of this analysis is translation. It is of paramount importance to mention that the quality of translation in Google Translate is heavily reliant on the language pair, which has affected the current translation. In addition, it uses statistical means to accumulate data obtained from translations performed by other sources. This means that it is hard to produce good-quality translations based on the various contexts it provides. Thus, the translation was utterly literal and contained a considerable amount of errors, both in meaning and spelling. This tool has no spellcheck or way to indicate Arabic spelling mistakes. It only suggests some alternatives in the dialogue box. There are also other types of grammatical and semantic errors. The translation results of MateCat, on the other hand, were satisfactory. Thanks to integrating two Translation memories, a personal and a public one provided by MyMemory to the website users, and a machine translation engine provided by the same source, the translation into Arabic was satisfactory. The dependence on the previously mentioned resources resulted in a good translation that respected the ST's meanings and context. Although the translation was devoid of errors, there is a minor inconvenience in MateCat, which is that it does not provide a spellchecker. This caused one small spelling mistake to appear throughout the process.

Fig.9. The Percentage of Errors

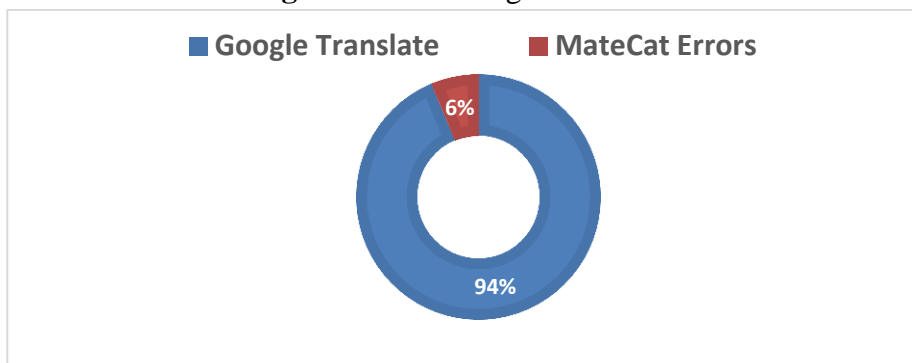


Figure 9 displays the percentage of errors recorded between the translations done by Google Translate and MateCat. The overall number of errors reported is 16, with the former having "15" errors and the latter with only "1" spelling mistake.

The following table shows some errors in google translate along with the analysis. This is performed with reference to the correct translations provided by MateCat.

Comparing the Effectiveness of Google Translate and MateCat Tools in the Translation of Scientific Texts from English into Arabic

Table 2. Translation Results by MateCat and Google Translate

Source Text	Google Translate Error	Analysis
1. By <u>Maria Temming</u>	بقلم ماريا تممينغ	A spelling mistake occurs in the Proper name. MateCat provided a correct spelling for the name " <u>ماريا تيمينغ</u> ".
2. <u>Created</u> by a computer	تم إنشاؤه بواسطة الكمبيوتر.	The Arabic passive voice from "تمّ إنشاؤه" is a common mistake. The correct option is "أنشئ", as it was translated with MateCat.
3-4. <u>Strapping</u> on a VR <u>headset</u>	ربط سماعة رأس VR	This is both a literal translation and a mistranslation (two errors). The word "headset" in this context does not only provide sound but also image. In addition, the abbreviation was not translated correctly, unlike in MateCat, where it was transferred correctly as "خوذة الواقع الافتراضي" with reliance on the meaning.
5. <u>The</u> <u>simulated</u> <u>environment</u>	البيئة المحاكية	This is another grammatical error in which the word "المحاكية" is supposed to be translated as "المحاكاة" as it was transferred correctly in MateCat.
6. Those views <u>seem</u> to <u>have</u> <u>depth</u> .	يبدو أن تلك الآراء لها عمق.	The error found here is related to meaning, which was entirely changed. The word "views" in the ST means "scenes", yet it was translated as "آراء". The correct translation is "مشاهد" as occurred in MateCat.
7. Adding sounds can make <u>the</u> <u>experience</u> <u>even</u> <u>more</u> <u>immersive</u> .	يمكن أن تجعل إضافة الأصوات التجربة أكثر غامرة.	The phrase was translated literally, which resulted in a weak construction. MateCat's translation memory provided a better alternative "إمكانية الإنغماس" "أكثر في التجربة".
8. <u>Soothe</u> Pain	تهدئة الألم	This is yet another Literal translation. A more suitable alternative could be "تخفيف الألم" as suggested by MateCat.

In table 2, eight of fifteen errors were discussed. These errors included grammatical, semantic, and spelling mistakes found in the translation performed on Google Translate. The only spelling mistake found in the second translation by MateCat was in the word "انغماس". This simple mistake is in the Hamzah letter "!", which represents the "Glottal Stop", and the correct form should have been the Connecting Hamzah "!". As mentioned, this is due to the absence of an Arabic language spellchecker tool in MateCat. Nevertheless, such mistakes can be corrected by running spellcheck after the translation, but an integrated spellchecker could have been more practical and time-saving.

The final section of this analysis investigates the vocabulary aspect of the two translations. There are noticeable vocabulary differences between the final products. Most of them were reasonable and acceptable to a certain degree, while others were somehow inadequate. Google Translate is wholly dependent on previously stored translations from different sources. Contrastingly, two translation memories and a machine translation engine supported the MateCat translation. In this vein, the first inadequate translation is "technology". The term was translated as "تقنية" by Google Translate and "تكنولوجيا" by MateCat. The term "تكنولوجيا" suits the original word as they have the same meaning, rather than "تقنية", which is usually the translation of "technique". The second inadequate translation by Google translate is of the word "objects". It was translated as "كائنات", which means "creatures", and this does not suit the current context. MateCat provided a more acceptable alternative by translating it "أشياء". That word holds the same contextual meaning as "objects". The third term is "stereoscopy", which was translated as "التنظير المجسم" by Google Translate. Conversely, in TT2, a more accurate equivalent was provided by MateCat is "التصوير المجسمي".

Despite these inconveniences, there are other different yet acceptable translations. The two translations provided a different cultural equivalent for the month "June". Google Translate gave "يونيو" (used in places like Iraq and the Levant) as a translation, while MateCat provided "جوان" (used in Algeria and Tunisia), which are both correct and represent two different varieties of the Gregorian calendar. Furthermore, the words "computer" and "controllers" were respectively translated as "كمبيوتر" and "أدوات التحكم" by Google Translate, and "حاسوب" and "أجهزة التحكم" by MateCat. Generally, this is acceptable because the terms are often used interchangeably in Arabic. The same goes for "sea level" since it appeared as "مستوى سطح البحر" in TT1 and "منسوب مياه البحر" in TT2.

6. Discussion

As previously highlighted, it is essential to extrapolate that in some work situations, and depending on the language pair, the format needs to be kept the same in the TT. However, translators are not always obliged to stick to the original format, specifically when the job instructions require that. MateCat, like most translation memory systems, performed well in this area.

Additionally, Punctuation marks are of utmost importance when translating or in any other type of writing. Mogahed (2012) argued that punctuation is crucial to how a text must be understood and misusing it results in faulty translations. Likewise, regarding punctuation, languages are different, and what works for one tongue may not necessarily work for another. Thus, according to Mogahed (2012) “There are many differences between languages with regard to punctuation marks, particularly Arabic and English, which have to be taken into consideration by translators” (p. 2). MateCat gives translators the liberty to adjust the punctuation marks on the spot and change them to what suits the TL.

Furthermore, Translation quality and errors in CAT tools can be controlled by the previously stored TMs or human intervention. One of MateCat’s shortcomings is the lack of a spellchecker. While it is easy to solve this problem via a post-editing phase at the end, a spellcheck addition can save more time. In addition, to deal with translation quality issues, Xiu & Xeauyin (2018) further suggest that “The translation quality of MT systems may be improved either, most obviously, by developing more sophisticated methods or by imposing certain restrictions on the input” (p. 17).

Ultimately, with TMs and glossaries installed in Machine Translation Enhanced Computer-Assisted Translation tools, translators can swiftly deal with vocabulary and terminological challenges. These two features help in saving time and increasing productivity. It should be noted that translators can also update these glossaries regularly by adding new terms. Consequently, they can have a renewable repertoire, each time they translate.

7. Conclusion

Machine Translation Enhanced Computer-Assisted Translation tools are amongst the most sophisticated ways that help translators improve their work on many scales. MateCat, although free, is a good example of a reliable open-source system with a variety of features. It can be suggested that, despite not having the same criteria as paid TM software, free TM systems can be presented as an alternative in translation classes where paid software is unavailable. In the case of MateCat, for instance, students can master using it without needing special training because of its simple interface, accessible group work features, and the ability to add different types of resources. Nevertheless, MateCat lacks an integrated spellchecker for languages, like Arabic, although it uses an external tool, the browser's spellchecker. The developers will likely address this minor shortcoming in future updates.

For future research, and starting from this paper, it is suggested that empirical studies about free CAT tools, such as MateCat, should be conducted in contexts where there is no accessibility to paid and more sophisticated programs. Instructors can expose students to such software to see their attitudes towards them and the degree of usability when performing translation tasks on different types of documents. In particular, Herget (2021) tested MateCat with MA students in post-editing using a product-based approach. The study focused on the importance of training students in MT post-editing using the MateCat tool in an educational setting.

8. References

- Ahrenberg, L. (2017). *Comparing Machine Translation and Human Translation: A Case Study*. Proceedings of the Workshop Human-Informed Translation and Interpreting Technology, 21–28.
- Apriliana, F., Kurniawan, A., Ferianda, S., & Kastuhandani, F. C. (2016). Introducing a Cat Tool to Translate: Wordfast. *Indonesian Journal of English Language Studies (IJELS)*, 2(1), 60–75. <https://doi.org/10.24071/ijels.v2i1.351>
- Bertoldi, N., Cettolo, M., & Federico, M. (2013). *Cache-based Online Adaptation for Machine Translation Enhanced Computer Assisted Translation*. Machine Translation Summit XIV, 35–42. Nice, France.
- Bowker, L., & Corpas Pastor, G. (2015). Translation Technology. *The Oxford Handbook of Computational Linguistics* (2nd ed., pp. 871–905). Oxford, United Kingdom: Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199573691.001.0001>
- Bowker, L., & Fisher, D. (2010). Computer-Aided Translation. *Handbook of Translation Studies* (Vol. 1, pp. 60–65). John Benjamins.
- Bussey, S. (2020, February 20). *Translation Memory eXchange (TMX)*. Retrieved August 11, 2022, from: <https://blog.andovar.com/translation-memory-exchange-tmx>
- Christensen, T., & Schjoldager, A. (2017). Translation-Memory (TM) Research: What Do We Know and How Do We Know It? *HERMES - Journal of Language and Communication in Business*, 23, 1–13.
<https://doi.org/10.7146/hjlc.v23i44.97268>
- Doherty, S. (2016). The impact of translation technologies on the process and product of translation. *International Journal of Communication*, 10, 947–969.
- Federico, M., Bertoldi, N., Cettolo, M., Negri, M., Turchi, M., Trombetti, M., ... Germann, U. (2014). *The Matecat Tool*. COLING, 129–132.
- Federico, M., Koehn, P., Schwenk, H., & Trombetti, M. (2013). *Machine Translation Enhanced Computer Assisted Translation*. Nice MT Summit XIV, 1, 225.
- Herget, K. (2021). *Post-editing Machine Translation in MateCat: A classroom experiment*. 7th International Conference on Higher Education Advances (HEAd'21). Universitat Politècnica de València.
<https://doi.org/10.4995/HEAd21.2021.12789>

- Milad, K. M. B. (2021). *Comparative Evaluation of Translation Memory (TM) and Machine Translation (MT) Systems in Translation between Arabic and English* (Doctorate Thesis). Swansea University, Swansea.
<https://doi.org/10.23889/SUthesis.57439>
- Mogahed, M. M. (2012). Punctuation Marks Make a Difference in Translation: Practical Examples, **1-16**.
- Pearse, B. (2019). *CAT tools in the content economy: A necessity, not a luxury*. Retrieved August 8, 2022, from Translation & Localization Blog website:
<https://www.smartcat.com/blog/cat-tools-in-the-content-economy-a-necessity-not-a-luxury/>
- Peng, H. (2018). *The Impact of Machine Translation and Computer-aided Translation on Translators*. IOP Conference Series: Materials Science and Engineering, 322, 1–5. <https://doi.org/10.1088/1757-899X/322/5/052024>
- Precup-Stiegelbauer, L.-R. (2013). *Automatic Translations Versus Human Translations in Nowadays World*. Procedia - Social and Behavioral Sciences, 70, 1768–1777. <https://doi.org/10.1016/j.sbspro.2013.01.252>
- Temming, M. (2022, June 27). *Scientists Say: Virtual reality*. Retrieved August 4, 2022, from Science News Explores website:
<https://www.snexplores.org/article/scientists-say-virtual-reality>
- Wang, W. (2014). Analysis of characteristics and strength of computer aided translation compared with machine translation. *BioTechnology :An Indian Journal*, 10(19), **11166–11170**.
- What Are CAT Tools?* (n.d.). Retrieved August 8, 2022, from Memsource website:
<https://www.memsource.com/what-are-cat-tools/>
- Xiu, P., & Xeauyin, L. (2018). Human Translation vs Machine Translation: The Practitioner Phenomenology. *Linguistics and Culture Review*, 2(1), **13–23**.
<https://doi.org/10.21744/lingcure.v2n1.8>