

Evaluating Artificial-Intelligence Translation Tools' Ability to Imitate Human Translators in Translating Synonymous Words in Specialized Texts

Omar LAABED¹ *

¹University of Batna 1 (Algeria)
omar.laabed@univ-batna.dz

Nadjat BADI²

²University of Batna 1 (Algeria)
hist.rev.alg@gmail.com

DOI: 10.33705/1111-019-001-023

Received: 15/02/2026

Accepted: 01/06/2026

Published: 30/06/2026

*Corresponding Author

Citation :
LAABED,O.& BADI,N (2026).
Evaluating Artificial-Intelligence
Translation Tools' Ability to Imitate
Human Translators in Translating
Synonymous Words in Specialized
Texts
Maalim
I(1), 55-74

Abstract:

In a world of fast-paced technological development, the idea that Artificial-Intelligence translation tools can imitate human translators in the English-Arabic language combination may still seem not thoroughly covered. This suggests that there are some aspects that still need more testing and evaluation. This paper aims at evaluating two AI translation tools' ability to imitate human translators in translating synonymous words from English into Arabic in specialized texts. Using content analysis as the primary research tool, the results showed that the ability of AI translation tools to imitate human translators is limited to the choice of the translation procedures, and that machine-translation post-editing is necessary to validate AI output translations.

Keywords: AI translation tools; synonymous words; MT evaluation; human translator imitation; historical text.

Maalim

© 2026 The Author(s).

Published by the High council of the Arabic
language.

This is an open access article
under the [CC BY license](#)



Résumé :

Dans un monde caractérisé par un développement technologique rapide, l'idée que les outils de traduction de l'intelligence artificielle peuvent imiter le traducteur humain dans la combinaison linguistique Anglais-Arabe ne semble pas suffisamment étudiée. Ceci veut dire qu'il y a des aspects qui nécessitent toujours une évaluation. Cette étude vise à évaluer l'aptitude de deux outils de traduction de l'intelligence artificielle à imiter le traducteur humain dans la traduction des synonymes de l'anglais vers l'arabe. En se basant sur l'analyse de contenu comme méthode de recherche principale, les auteurs ont conclu que l'aptitude des outils de traduction de IA à imiter le traducteur humain est limitée au choix des procédés de traduction, et que la relecture de la traduction automatique est nécessaire pour valider la traduction produite par l'IA.

Mot clés : outils de traduction de l'intelligence artificielle, synonymes, évaluation de la traduction automatique, imitation du traducteur humain, texte historique

تقييم قدرة أدوات الترجمة بالذكاء الاصطناعي على محاكاة المترجم البشري في ترجمة المترادفات في النصوص المتخصصة

أ. عمر العابد¹

¹ جامعة باتنة 1 (الجزائر)

أ. نجات بادي²

² جامعة باتنة 1 (الجزائر)

الملخص:

في عالم يتميز بتطور تكنولوجي سريع، تبدو فكرة قدرة أدوات الترجمة بالذكاء الاصطناعي على محاكاة أداء المترجم البشري في الأزواج اللغوي عربي-إنكليزي بحاجة إلى دراسة أكثر تعمقا، ولذلك، يسعى هذا البحث إلى تقييم مدى قدرة أداتين من أدوات الترجمة بالذكاء الاصطناعي على ترجمة المترادفات في النصوص المتخصصة، وتوصل البحث بالاعتماد بشكل أساسي على منهجية تحليل المحتوى إلى أنَّ قدرة الأداتين على محاكاة المترجم البشري تنحصر في اختيار إجراءات الترجمة فقط، وضرورة القيام بعملية المراجعة من طرف المراجع البشري من أجل اعتماد الترجمات التي تُقدِّمها أدوات الذكاء الاصطناعي.

الكلمات المفتاحية: أدوات الترجمة بالذكاء الاصطناعي؛ المترادفات، تقييم الترجمة الآلية؛ محاكاة المترجم البشري؛ النص التاريخي.

1. Introduction: When Gutenberg invented the printing press back in the 15th century, it was massively welcomed and celebrated throughout Europe and beyond. However, not everyone was happy with this invention. The scribes, who used to make their living with their bare hands by copying people's words onto manuscript books, did not seem to like it because it meant to them the beginning of the end of their 'craft'. In fact, the use of the word 'craft' is not innocent; this word was used by Newmark (1981/1988) to describe what translation is all about; "the attempt to replace a written message and/or statement in one language by the same message and/or a statement in another language" (p. 07), but the idea of making a living by translating people's discourses may well sound synonymous to the word 'craft' and, by extension, expresses similarity between the scribes and the translators.

Six centuries after the invention of the press, technology seems to have again made significant developments. In a world witnessing a fast-paced development of computer-aided translation (CAT) tools and Artificial Intelligence (AI) tools, translation of texts has become a matter of just a few clicks away. As a result of this, the demand for human translation seems to have been negatively affected. Hence, similarity between the 15th century scribes and today's translators may be well noticed.

However, in the midst of the seemingly endless machine-man fight, there is also room for collaboration. In translation, the machine-man relation may even be tighter than it appears to be.

Comparative studies carried out on the evaluation of machine translation (MT) as compared to human translation, in language combinations including Arabic, show that the ability of the machine translation tools to imitate human translators may have not been thoroughly and sufficiently studied. This could be an indicator of a research gap that this particular research paper attempts to fill. Based on these grounds, this study aims to identify whether AI translation tools, as a new facet of machine translation, have the ability to imitate human translators in the language combination English-Arabic, specifically in handling the translation of synonymous words (juxtaposed or separated by few words) in historical texts. Following this aim, the main research questions that this study aims to answer may be formulated as follows: 1) To what extent can AI tools can imitate

human translators? 2) Are there specific cases (of synonymous-word translation) where AI tools cannot imitate human translators?

Accordingly, the following hypotheses were formulated: 1) Although, AI translation tools can imitate human translators in translating synonymous words, this ability is relative and the imitation may be limited to particular cases; 2) There are cases where AI tools may not be able to imitate human translators.

To answer the two questions above, a descriptive analytical method was followed, and content analysis was used as a primary research instrument. The corpus of this study consists of sixteen (16) passages non-randomly selected from Mathew Carey's 1794 historical book *A Short Account of Algiers*. These passages contain thirty-two (32) synonymous words (juxtaposed in pairs or separated by 1 to 3 words only). The passages were then translated using two AI translation tools; namely DeepL and ChatGPT, in order to compare them to the human reference translation provided by Ali Tablit, outed back in 2013. The procedure followed in this research adheres to the recommendations of previous studies, mainly in terms of machine translation evaluation. The use of the human translation as a reference translation has been applied for many years now; this may be traced back to the emergence of machine translation first metrics (BLUE, METERO, Etc.) using human translation as a reference to evaluate outputted machine translation.

2. Literature Review:

2.1. Translation and Technology: Translation has been defined by many scholars with diverging perspectives. It was viewed as a linguistic transfer of messages from one language into another (Catford, 1965). It was also regarded as a "craft" (Newmark, 1981), and more recently as a negotiation not only between two languages, but also between two cultures (House, 2009). This denotes the relation between translation and language, between translation and careers, and between translation and culture. In more recent times, the word 'translation' seems even to have more connotations. These have to do with technology, and with users understanding by translation the use of some machines and technological devices to convey wording from one language into another. In fact, this rhymes well with the idea that technology refers to the application of scientific

knowledge to the practical aims of human life (Britannica.com). Based on this, the idea of considering translation among the practical aims of human life denotes a tight relationship between translation and technology.

The translation-technology relationship idea has further been strengthened by the outing of an encyclopedia titled "*The Routledge Encyclopedia of Translation Technology*", with a first edition released back in 2014 and a second released just recently in 2023. The encyclopedia contains as many as 47 chapters; this indicates, by extension, the development that machine translation has recently reached. This development has resulted in the emergence of new concepts such as: Human-aided machine translation (HAMT), Machine-aided human translation (MAHT), Computer-aided translation (CAT) tools, knowledge-based machine translation (KBMT), example-based machine translation (EBMT), online translation, deep learning, etc.

2.2. A Historical Overview of Machine Translation:

2.2.1. Emergence and Beginnings: In general terms, machine translation (MT) is defined as "the use of computer programs to translate texts from one natural language into another automatically" (Ping, 2009, p. 162). From a more scholarly perspective, this concept refers to "a sub-field of computational linguistics (CL) or natural language processing (NLP) that investigates the use of software to translate text or speech from one natural language to another" (Qun & Xiaojun, 2023, p. 112).

Based on the literature related to machine translation (for instance, Mounin, 1964; Hutchins, 1999; Hutchins, 2001, Hutchins, 2006; Ping, 2009), the emergence of MT may be owed to the technological developments that the world saw particularly in the aftermath of WWII. However, it is also mentioned that its emergence may be traced back to the 1930s, when Alan Turing, a British mathematician, suggested that machines may mimic humans to the extent that it becomes difficult to distinguish between them. Going back to the idea of the post WWII developments, Hutchinson & Waters (1987, p. 6) described these developments as "the demands of a brave new world" characterized by an enormous and unprecedented expansion in scientific, technological and economic activity on an international scale. In this regard, Hutchins (2006) mentioned the central

role of the invention of the electronic computer in the 1940s in paving the way for the emergence of machine translation. Lederer (2006, p. 142) mentioned that computers were used during WWII to help decipher messages transmitted in foreign languages. Mounin (1964, p. 5) mentioned that the year 1946 saw the beginnings of what he called 'new research area' of machine translation, and the result was that one year later (i.e. 1947) the idea of the possibility of using computers to translate natural languages was confirmed. This is when two scientists, the American Warren Weaver and the British Andrew Booth, suggested that natural languages may be translated automatically. Two years later (1949), Weaver wrote a memorandum entitled "Translation", the last of his four proposals on mechanical translation: 1) examination of immediate contexts may be useful to solve the problem of multiple meaning; 2) the possibility that all languages share certain logical features; 3) the cryptographic methods concerned with the basic statistical properties of communication may be applied in mechanical translation; and 4) there may be linguistic universals (Qun & Xiaojun, 2023, p. 112). This latter proposal is argued to trigger research on MT in the USA, leading later to significant achievements.

From the 1930s up until the 1950s, the idea that machine translation can mimic human translators, became, more and more, true through significant developments. These developments resulted in more important achievements starting from the 1950s, when the first artificial intelligence translation method was applied, through until today.

2.2.2. Milestones in the Development of AI Translation Methods: The development of AI translation methods may be broken down into three main stages: the rule-based translation approach stage (late 1950s - the 1980s), the statistics-based translation approach stage (late 1980s –1990s), and the neural-network breakthrough in the 2010s.

The rule-based translation algorithms stage constitutes the emergence of AI-powered translation systems, and, therefore, its role in the subsequent developments is regarded as significant. In this stage, language rules were stored in the software to produce basic translation. A remarkable event in this stage is the 1954 joint project between the members of the Institute of Linguistics at George Town University and, the computer company, IBM. This project resulted in the first public live

demonstration in which 49 sentences were selected and then translated from Russian into English using 250 words and only 6 grammatical rules (Hutchins, 1995: 436). This demonstration is argued to mark machine translation, and, by extension, AI translation, as a reality after seven years of Weaver's idea of the possibility of using computers to translate natural languages (Qun & Xiaojun, 2023, p. 113). The years that followed this demonstration saw interesting developments, and interest in machine translation went beyond the US boundaries to several countries in Europe and Asia (Qun & Xiaojun, 2023, p. 113).

However, the rule-based approach to machine translation was limited because language simply started to be more flexible, and rules were not used in the manner that machines were instructed to apply them in their language processing. As a result, the late 1980s saw the emergence of a new approach to AI translation; this is usually referred to as statistics-based approach.

At this point, it is also worth noting that machine translation focused in this period of time on the English-Russian language combination, and, by extension, interest in machine translation was mainly limited to the US. However, the Cold War between the US and the Soviet Union resulted in shifting the interest away from MT in the US. This explains why Hutchins (2001, p. 16) named the decade from 1967 to 1976 as "The Quiet Decade".

As an alternative to the criticized rule-based translation approach, the 1980s saw the emergence of a new approach called 'the statistics-based approach' to machine translation. Soon, this approach gained more and more space especially. It is based on the idea of using massive bilingual text datasets to find matching words or phrases across languages without using pre-stored rules; this resulted in machines "learning" real-life language, and, consequently in more accurate translation, especially in languages with large digitalized archives such as English and French (Al-Aqad, 2025, p. 67). These systems used neural networks to imitate human cognition. Neural networks, on the other hand, could process sentences Al-Aqad, (2025, p. 67), rather than words or phrases, and context was primarily considered in the translation process. As human need for more development increased in this direction, the concept of "deep learning" emerged, allowing for more human-like attitudes to

translation. Today, OpenAI translation programs allow for more accurate, context-aware and creative translations; let alone the fact that they are time efficient and cost-effective.

To summarize the above, today's AI translation programs have seen, in the space of 75 years, a revolution in natural language processing (NLP). Although, the rule-based approach had many weaknesses, but it was the first idea from which stemmed all the subsequent developments.

2.3. Machine Translation Tools' Ability to Imitate Human Translators: The idea that the machine can imitate humans in the tasks of translation is not recent. In fact, this may be traced back to the 1930s with Alan Turing's test designed to validate his own thoughts of the possibility that machines may mimic humans; this test was later named after him "the Turing Test".

The test consists of the following: in a closed room there is a machine or a human. Outside the room, there is an interrogator who does not know whether the being in the room is the machine or the human. Both the machine and the human communicate with the interrogator through a screen. The interrogator's job consists of asking questions. The machine or the human processes the questions and provides answers. Based on these answers, the interrogator decides who is in the closed room. The yielded results were promising, and today's status of AI proves the validity of Turing's idea. What further this test validates is the ability of the machine to mimic humans in both understanding the language and making responses accordingly.

2.3.1. Machine Translation Evaluation: Broadly speaking, MT evaluation may be carried out using one of the two main types of methods: the methods using metrics that consist, in general, of comparing the outputted machine translation to a human reference translation; and the methods that completely rely on human evaluation consisting of appointing to a native speaker, usually of the target language (TL), with the task of evaluating the machine outputted translation. With reference to the former type, the birth of the pioneering metric called BLEU (Bilingual Evaluation Understudy), developed in 2002 by Papineni and his colleagues, is argued to have led to intensifying efforts to develop more metrics aiming at MT evaluation. Among these, literature mentions the following: NIST (National Institute for Statistics and Technology) by Doddington in 2002, METEOR (Metric for Evaluation of Translation with Explicit Ordering) by Banerjee and Lavie in 2005, TER

(Translation Edit Rate) by Snover and his team back in 2006. This leads to understand that the distinctive line that should be drawn between the former and latter types of evaluation (automated and human) lays in the idea that the former type requires itself constant testing and evaluation.

2.3.2. English-Arabic, Arabic-English MT Evaluation: Since this paper focuses on English-Arabic combination, it is worth taking a look at some existing literature that falls within this particular frame. The interest of Arab translation specialists in MT evaluation systems seems to be very recent compared to that of non-Arab specialists. For instance, Fahad al-Salhi (2019, p. 277) provided a summary of the studies carried out by Arab scholars. The range of the studies he mentioned is from the very recent year 2000 to 2013. The BLEU metric was widely used in almost all the studies that he mentioned, suggesting the need to test the other metrics in this particular language combination. The following are studies carried out by Arab researchers on the theme of MT evaluation:

In 2000, Chalabi (2000, cited in al-Salhi, 2019) tested a free-machine translation system to translate from English into Arabic; he reported an achievement of 60% accuracy level. In 2004, Guessoum and Zantout provided a collaborative work titled « A Methodology for Evaluating Arabic Translation Systems”. The authors suggested a method to assess Arabic translation systems. They focused on evaluating lexical and grammatical coverage, semantic correctness, and pronoun resolution correctness. In 2009, applying the BLEU evaluation method, Adly and al-Ansary carried out a comparative study of the Universal Network Language (UNL) translation and widely used translation systems; namely Google Translate and Babylon. The study results suggested that Google Translate performed better than the other systems included in the study. However, three years later, the use of Google Translate as a translation system was advised against by al-Sheikh (2012), who tested Google Translate performance in the translation of three texts of three different types, his study focused on the English-Arabic combination. In 2014, al-Samawi (2014, cited in al-Salhi, 2019), focusing also on the English-Arabic combination, and applying human -rather than automated- evaluation, carried out a study in which he examined the translation of a 100 sentences belonging to 10 different types of texts based on error and content analysis. He reported an average of 3.66

errors per sentence. In 2013, Al-Kabi (2013, cited in al-Salhi, 2019) and his colleagues carried out a study comparing between Google Translate and Babylon MT systems in English-Arabic translation. The authors applied the BLEU method in their study. They reported that Google Translate MT system performed better compared to Babylon MT system. In the same year (2013), another study comparing the same systems, and also applying the BLEU method, was carried out by Kadhim and his colleagues. They reported that Google Translate performed better in terms of accuracy, while Babylon performed better in terms of style, and that both systems showed the same clarity level.

To summarise, in most cases of MT evaluation studies done by Arab scholars, the metric BLEU was used. However, the reliability of BLEU has been widely criticised since it does not provide precise scores. Therefore, other methods seem to be interestingly required.

2.4. Machine Translation Post-Editing (MTPE): Machine translation post editing may be regarded as a similar process to proofreading in human translation post editing. This process consists of rereading the outputted machine translation in order to correct potential errors relating mainly to accuracy, fluency, and other aspects. However, most importantly, this process is carried out by a human editor. This suggests the centrality of the human contribution to the translation of texts.

2.5. Synonymy in Specialized Texts: Synonymy is considered as one type of lexical relations. Originally, the word 'synonymy' derives from ancient Greek, and is divided into two parts: 'syn' meaning 'with', and 'onama' (name); this leads to think of synonymy as a state of 'sameness of meaning'. A synonym, according to Webster's New Dictionary of Synonyms (1984), is one of two or more words which have the same or very nearly the same essential meaning (p. 24a). The author added, however, that this similarity is not superficial, and that distinguishing between synonyms may be achieved by adding an implication or a connotation to the synonymous words, or by identifying a difference in their idiomatic use or their application (p. 25a). The author stresses that the only thin distinguishing line between the synonymous words is their denotation; in other words, words with similar denotation may have identical meaning. An example of this is the word 'nice' in the dictionary that has one fixed denotation (exacting or displaying exacting standards) with other

words like 'dainty' or 'fastidious' in different contexts (such as in selection, judgment, or workmanship) (p. 25a)

On the other hand, a specialized text is a text that uses wording that belongs to a specific jargon (historical, military, legal, religious, medical, etc.), and that is associated with a special context. This study does not concentrate on the specificity of the jargon of the specialized texts, but, rather on the use of synonyms in such specialized texts.

3. Study Corpus and Methodological Design:

3.1. Corpus: This study uses Carey's 1794 'A Short Account of Algiers'. It contains 46 pages, which, as its title may indicate, is short and concise. The Arabic translation, outed in 2013, was carried out by Ali Tablit. It contains 63 pages with an a 20-page appendix containing several letters in relation with the book theme. From linguistic, cultural and translation perspectives, the book is loaded with a good amount of items that may be highly interesting as research topics. The author used a formal language to describe, according to his own views, the events, population, customs and traditions that constituted the living of the Algerians and surrounding regions before the arrival of the French occupier in 1830. The translator, Ali Tablit, on the other hand, offered a translation that was rather based on critical judgments of some of the information provided by the author.

3.2. Methodological Design: This paper followed a qualitative approach. It used content analysis as its primary research tool. By definition, content analysis refers to "a research method that provides a systematic and objective means to make valid inferences from verbal, visual, or written data in order to describe and quantify specific phenomena" (Downe-Wambolt, 1992, p. 314). Based on these grounds, this paper is descriptive analytical since the gathered data were described and analyzed.

3.2.1. Procedure: Upon completion of the reading of the English text, sixteen (16) passages, containing thirty-two (32) juxtaposed synonymous words and synonymous words separated by one to three words, were extracted. The synonymous words and phrases were written in bold as to highlight them. An online dictionary was used for the purpose of deciding whether the words are synonymous. In the extraction process, context was primarily considered, and the words and

phrases, put in their context, were given to ChatGPT and DeepL for translation. The human translation provided by Ali Tablit is used as a reference translation. Upon this step, the translation procedures applied by the two AI-translation tools were identified. The aim is to identify the degree at which the two AI translation tools may imitate the provided human translator. A secondary aim is to know whether machine translation post-editing (MTPE) may be used to evaluate the translations outputted by AI-translation tools. The evaluation was completely human and did not rely on any automated evaluation system; this is to adhere to the recommendations of previous studies suggesting that human evaluation methods are more accurate than the metrics developed for this purpose.

3.2.2. Gathered Data: List of Synonymous Words: The authors counted thirty-two (32) synonymous words/phrases throughout the book. This count was based on two criteria: a) the words are synonymous or related to each other as verified by the dictionary, and b) the words are in pairs (juxtaposed) or separated by 1 to 3 words only. In order to verify the two criteria mentioned above, the synonymous words (3 occurrences with 1 to 3 words separating the two synonymous words from each other), and synonym pairs (13 occurrences with no word separating the two words) were listed using the Merriam-Webster online dictionary & thesaurus; first to check whether they are considered synonymous (or related), and second to check their definitions and how they relate to each other. In order to ensure understanding of the synonymous words by the two AI-translation tools, sufficient context was added to each pair of synonymous words. The idea is to check how ChatGPT and DeepL handle the translation of the mentioned synonymous words as compared to the human reference translator.

3.3. Results Discussion and Analysis:

3.3.1. Discussion: The results showed that, in order to translate the synonymous words, DeepL, ChatGPT, and Ali Tablit, the human reference translator, used four main procedures as shown in the table below:

Table 1 : Choice of the translation procedure

Procedure	Double-used a word to translate the synonym pair		Naturalized one of the two synonymous words		Used one functional equivalent to translate the synonym pair		Translated the SL synonym pair using a TL synonym pair	
	Cases	Rate %	Cases	Rate %	Cases	Rate %	Cases	Rate %
DeepL	5	31.25	3	18.75	00	00	7	43.75
ChatGPT	00	00	05	31.25	01	6.25%	10	62.5
HRT	00	00	01	6.25	01	6.25	14	87.5

Table 1. above shows that DeepL, ChatGPT, and the human translator used four translation procedures; these include: 1) double-use of the same word to translate the synonym pair (DeepL only); 2) Naturalization of one of the two synonymous words; 3) Use of one functional equivalent to translate the synonym pair; and 4) Use of a TL synonym pair to translate the SL synonym pair.

Table 1. also shows that, in terms of the ability to imitate human translators, ChatGPT is more likely to imitate human translators in that it had the same score as the human reference translator in two columns (1 & 3), and a close score in the fourth column. On the other hand, the scores of DeepL are significantly far from the human reference translator's scores.

3.3.2. Analysis: This section focused on the evaluation of the ability of the two AI translation tools to imitate human translators. However, semantic and syntactic evaluation was also considered since the synonymous words in question were put in their context and, by extension, required a look at how context was also conveyed. As mentioned previously, four translation procedures were identified. The following is a demonstration of how each procedure was applied.

2.3.2.1. Double-use of a word to translate the synonym pair:

This procedure was used only by DeepL in five cases as shown in the examples below:

Table 2. one word is used to translate the two words of the synonym pair

Synon. words	DeepL Tr.	ChatGPT Tr.	HRT
Since it ceased to be the seat of an independent government, it has shrunk into ruins and desolation . (p. 6)	ومنذ أن توقفت عن أن تكون مقراً لحكومة مستقلة، انكمشت إلى خراب وخراب. wa mundhu an tawwaqafat 'an an takūna maqrān li-ḥukūma mustaqilla, inkamashat ilā kharāb wa-kharāb	منذ أن توقفت عن أن تكون مقراً لحكومة مستقلة، تقلصت إلى خراب ودمار. mundhu an tawwaqat 'an an takūna maqrān li-ḥukūma mustaqilla, taqallaṣat ilā kharāb wa-damār.	ومنذ توقفها كمرکز لحكومة مستقلة فإنها انكمشت لتصبح آثارا وخرابا. (ص. 16) wa mundhu tawwaqūfiha qamārqazin li-ḥukūma mustaqilla, fa inaha inkamashat athāran wa-kharāb
Their principal arms are a short lance and a scymitar , though they are likewise acquainted with the bow and the musket. (p. 13)	وأسلحتهم الرئيسية هي الرمح القصير والرمح القصير، على الرغم من أنهم يعرفون أيضاً القوس والبنديقية. Wa asliḥatuhum ar-ra'isiyya hiya arrumḥu qaṣīr wa- arrumḥu qaṣīr, 'ala rughmi min annahum ya'rifūna ayḍan al-qaws wa-l-bunduqiyya.	أسلحتهم الرئيسية هي رمح قصير وسيف مقووس، مع أنهم يعرفون أيضاً القوس والبنديقية. asliḥatuhum ar-ra'isiyya hiya rumḥ qaṣīr wa-sayf maqwās, ma'annahum ya'rifūna ayḍan al-qaws wa-l-bunduqiyya.	فأسلحتهم الرئيسية هي رمح قصير وقوس. ولو أنهم ملمون أيضاً باستعمال القوس والمسكيت (بنديقية صغيرة). (ص. 24) Fa asliḥatuhum ar-ra'isiyya hiya rumḥ qaṣīr wa-qaws, wa law anahum mulimuna ydan, bist'imali l-qaws walmusqit.

The two examples explained in the table above how DeepL used one word "خراب" [kharāb] to translate the synonym pair 'ruins and desolation', and the phrase 'الرمح القصير' [arumḥ alqaṣīr] to translate the related words 'lance and a scymitar' whereas ChatGPT used two different words to translate the two synonym pairs respectively 'خراب ودمار' [kharāb wa-damār], 'رمح قصير وسيف' [rumḥ qaṣīr wa-sayf maqwās]. Following the human reference translation, ChatGPT applied a similar translator procedure as the human translator. This suggests that ChatGPT is more likely to imitate human translators in handling the translation of the mentioned synonym pairs in that it provided two TL synonymous words as equivalents to the SL synonymous words. This does not necessarily mean that ChatGPT does not require a machine-translation post-editing, but it rather means that if it does require such a step, it should be at minor level compared to DeepL that is likely

to necessitate more focused machine-translation post-editing (MTPE) to make the translation more accurate.

3.3.2.2. Naturalization of one of the two synonymous words:

Table 3. Naturalization of one of the two words to translate the two words of the synonym pair:

Syn. words	Deepl Tr.	ChatGPT Tr.	HRT
Many of the troops were cut to pieces by the Moors and Arabs; while several gallies and other vessels , endeavouring to gain some neighbouring creeks along the coast, were immediately plundered, and their crews massacred by the inhabitants. (p. 23)	وقطع المغاربة والعرب العديد من الجنود إرباً؛ بينما نُهبت على الفور عدة سفن غالية وسفن أخرى كانت تحاول الاستيلاء على بعض الخلجان المجاورة على طول الساحل، وذبح السكان طواقمها. wa qatṭa'a l-maghāriba wa-l-'arab l'adida mina mina l-junūd irban irban; baynamā nuhibat 'alā l-fawr 'idda sufun ghalīa wa-sufun ukhrā kānat tuḥāwīlu l-istila 'alā ba'd al-khuljān al-mujawira 'alā tuli s-sāḥil, wa-dhabaha s-sukkān ṭawāqumuhā	قُطع كثير من الجنود إرباً على يد المغاربة والعرب؛ بينما نُهبت على الفور عدة غوالي وسفن أخرى كانت تحاول الوصول إلى بعض الخلجان القريبة على الساحل، وقُتل طواقمها على يد السكان. quṭṭi'a kathīr mina l-junūd irban 'alā yadi l-maghāriba wa-l-'arab; baynamā nuhibat 'alā l-fawr 'idda ghawālī wa-sufun ukhrā kānat tuḥāwīlu l-wuṣūl ilā ba'd al-khuljān al-qarība 'alā s-sāḥil, wa-qutīla ṭawāqumuhā 'alā yadi s-sukkān.	لقد قطعوا اربا من قبل البربر والعرب، في حين حاولت سفن شراعية عديدة ومراكب أخرى الوصول إلى خلجان مجاورة على امتداد الساحل، لكنها تعرضت فوراً للسلب وقتل الأهالي جميع طواقمها. (ص. 36) Laqād kotī'o irban min kibali lbārbari wa-l-'arab, fi hin hawalat sufun shira'iya 'adidā wa marakib ukhrā l-wuṣūl ilā khuljān mujawira 'alā mtidad s-sāḥil, lakinaha tā'aradhat fawran lis-salb wa-qatala l-ahali jami' ṭawāqumihā
He then kept firing among their tents, while some well manned galliot and brigantine were dispatched to attack their shipping. (p. 32)	ثم استمر في إطلاق النار بين خيامهم، في حين تم إرسال بعض الغاليات والبوارج المدججة بالرجال لمهاجمة سفنهم. thumma istamarra fi iṭlāq al-nār bayna khayāmihi, fi ḥīn tumm irsāl ba'd al-ghāliyat wa-al-bawārij al-mudajja bi-al-rijāl li-muhājamati sufunihi.	ثم واصل إطلاق النار بين خيامهم، بينما أرسلت بعض الغليوتات والبريغنتينات المأهولة جيداً لمهاجمة سفنهم. Thumma wasala iṭlāq al-nār bayna khayāmihi, baynama ursilat ba'd al-ghilyūtāt wa-lbrigantināt lma'hulati jayidan limuhajamati sufunihi	عندها قام بإطلاق النار على خيامهم، في حين أعدت مجموعة غليوطات ومراكب شراعية للقيام بمهاجمة سفنهم (ص. 45) 'indahā qāma bi-iṭlāq al-nār 'alā khayāmihi, fi ḥīn a'addat majmū'at ghilyūtāt wa-marākib sharā'iyya lil-qiyām bi-muhājamati sufunihi.

Both DeepL and ChatGPT used naturalization to translate the word 'galliot'; whereas DeepL used 'الغاليات' [al-ghāliyāt], ChatGPT used 'الغليوتات' [al- ghilyūtāt]. However, from the transcription of the words, as compared to the human reference translation 'غليوطات' [ghilyūtāt], it appears that ChatGPT is more likely to imitate the human translator in that it used a word that is almost similar to that used by the human translator; only the first sound 'ت' [t] needs to be replaced by the emphatic sound 'ط' [ṭ]. On the other hand, the word 'brigantines' was translated by DeepL as 'بوارج' [bawārij], and was naturalized by ChatGPT as 'البريغنتينات' [lbrigantināt]. The human translator offered a the two word equivalent 'مراكب شرعية' [marākib sharā'iyya]. Therefore, in this particular case, it is DeepL that imitated the human translator in the choice of the translation procedure.

3.3.2.3. Use of one functional equivalent to translate the two synonymous words:

Table 4. Use of one functional equivalent to translate the two words of the synonym pair:

Synon. words	DeepL Tr.	ChatGPT Tr.	HRT
This being refused, the Frenchman without ceremony carried off the Turkish viceroy, and his cadi or judge ,... (p. 31)	ولما رُفِضَ هذا الطلب، قام الفرنسي دون احتفال بخلع الوالي التركي وقاضيه أو قاضيه،.... wa-lammā rufiḍa hadha t-talab, qāma l- faransī duna ihtifial bi khal'l l-wali at-turkī wa-qāḍiyahu wa- qāḍiyahu,	ولما رُفِضَ ذلك، أخذ الفرنسي بلا تردد نائب الملك التركي وقاضيه، wa-lammā rufiḍa dhālik, akhadha l- faransī bilā taraddud nā'ib al- malik at-turkī wa- qāḍiyahu,	غير أنّ طلبه هذا وجد رفضاً قاطعاً من الجزائريين. قام الأميرال الفرنسي وبدون مراسيم احتفالية باختطاف نائب الداي التركي، ومعه قاض ... (ص. 44) Ghayra anna talabhu hādha wajada rafḍhan qāti'an mina ljazayriynā, qāma l-amiral lfiransi wa biduni marāsima htifily bikhtitāfi na –ibi dayi turqi, wa ma'ahu qādhi

The synonym pair 'cadi and judge' refers actually to one valid equivalent in this case which is 'judge'; the first word being of Arabic origin and was, for some reason, originally transcribed by Mathew Carey, author of the book from which the passage was extracted. To translate this pair, DeepL double used the Arabic word 'قاضيه' [qāḍiyahu], whereas ChatGPT used one word 'قاضيه' [qāḍiyahu] to translate the pair. The human translator also used one word 'قاض' [qāḍi] to translate the pair. By extension, ChatGPT imitated the human translator in his choice of the translation procedure, and, therefore, minor post-translation editing is needed with the DeepL translation.

2.3.2.4. Translating the SL synonym pair by a TL synonym pair:

Table 5. Translating a SL synonym pair by a TL synonym pair:

Synonymous words	DeepL Tr.	ChatGPT Tr.	HRT
who were just arrived from the porte, with all their equipage and retinue . (p. 31)	الذين كانا قد وصلا لتوهما من الميناء مع كل عتادهما وحاشيتهما. alladhayn kana qad wasalā litwwihima tawwi mina l-mina, ma'a kull 'atadihima wa- ḥāshiyatihimā.	الذين قدما للتو من الباب العالى، مع كل أمتعتهم وحاشيتهم. alladhayn qadimā li- tawwi mina l-bāb al-ālī, ma'a kull amti'atihimā wa-ḥāshiyatihimā.	وصلا توا من الباب العالى مع طاقمهما وحاشيتهما. (ص. 44) wasalā tawwan mina mina l-bāb al-ālī ma'a tāqimihima wa- ḥāshiyatihimā

The synonym pair 'equipage and retinue' was translated by DeepL as 'عتادهما وحاشيتهما' [atadihima wa-ḥāshiyatihimā], and by ChatGPT as 'أمتعتهم وحاشيتهم' [amti'atihimā wa-ḥāshiyatihimā], whereas it was translated by the human translator as 'طاقمهما وحاشيتهما' [tāqimihima wa-ḥāshiyatihimā]. This means that the three translations overlapped in the translation of the first word of the pair 'equipage'. The meaning with which this latter word was associated by DeepL and ChatGPT is 'the set of articles and physical items that may be used by people as equipment'. Archaically, the word 'equipage' is synonymous with the word 'retinue', both referring to "the group of retainers or attendants" (<https://www.merriam-webster.com>, accessed 6 Sep. 2025) that used to accompany important personalities in the past; this is the meaning that was considered by the human translator, who translated the pair as 'طاقمهما وحاشيتهما' [tāqimihima wa-ḥāshiyatihimā]. The two latter Arabic words are synonymous or close to each other; the first referring to the 'retainers and attendants of important personalities' and the second to their 'associates and advisors'. Both DeepL and ChatGPT imitated the human translator in his choice of the translation procedure, but the validation of their translations requires post-editing.

4. Conclusion: As shown in the study results, the ability of AI translation tools to imitate human translators is limited to the choice of the translation procedures, and not necessarily the word choice. It has also been shown that machine-translation post-editing, however minor or major, is necessary to validate AI translation tools' outputted equivalents to synonymous words. In addition, four main translation procedures were used by AI translation tools to render the synonymous words; these are:

a) Use of one word to translate the two synonymous words; b) Naturalization of one of the two words constituting the synonym pair; c) Use of a single functional equivalent to translate the two words of the synonym pair; and d) Use of a TL synonym pair to translate the SL synonym pair. The authors think that the provided procedures are listed from the least appropriate to the most appropriate, and that the machine-translation post-editing is required even if the AI translation tool's choice of the translation procedure is similar to the human translator's.

The results showed that the two AI tools, mentioned in this study, could not imitate the human translator in the word choice. Also, among the cases discussed in this study, two particular cases showed that the mentioned AI tools could not imitate the human translator; the first concerns DeepL, while the second concerns both AI tools: a) the case where the relationship between the words is very tight that the difference between them is slight, as in the case of 'ruins and desolation' with both denoting "destruction", where DeepL translated both words using one word twice; b) the case of the use of archaic English, where the AI tools confused the meaning of the word 'equipage' with that of 'equipment', when it rather meant 'servants' who accompanied an important personality.

In a nutshell, testing the AI ability to imitate a human translator's cognitive skills, in the English–Arabic language combination, seems still in need of constant work, testing, and evaluation. Based on the above, both study hypotheses were validated.

Bibliographical references:

1. Al-Aqad M. (2025). *Role of AI in Translation and Interpreting*. USA: IG Global Scientific Publishing
2. Al-sahli F. S. (2019). "Evaluating Neural Machine Translation Using Error Analysis In English - Arabic Texts", *Journal of Aswan Faculty of Arts*, 5, 1, 2019, 270-289. doi: 10.21608/mkasu.2019.212252
3. Catford J. C. (1965). *A Linguistic Theory of Translation*. UK: Oxford University Press
4. Doddington G. (2002). Automatic Evaluation of Machine Translation Quality Using N-gram Co-Occurrence Statistics. In: *Proceedings of the Human Language Technology Conference (HLT-2002)* (San Diego, USA, March, 2002). pp. 128-132
5. Downe-Wambolt B. (1992). Content Analysis: Method, Application and Issues. *Health Care Women Int.* Jul-Sep;13(3):313-21. doi: 10.1080/07399339209516006.
6. House J. (2009). *Translation*. UK: Oxford University Press
7. Hutchins J. (2001). "Machine Translation over fifty years". In: *Histoire Épistémologie Langage*. Tome 23, fascicule 1, 2001. Le traitement automatique des langues. pp. 7-31
8. Hutchinson T. & Waters A. (1987). *English for Specific Purposes*. UK: Cambridge University Press
9. Lederer M. (2006). *La traduction aujourd'hui, le modèle interprétatif*. Nouvelle édition. Caen : Lettres Modernes Minard
10. Mounin G. (1964). *La machine à traduire: Histoire des problèmes linguistiques*. La Haye: Mouton
11. Newmark P. (1988). *A Textbook of Translation*. UK: Prentice Hall
12. Papineni K., Roukos S., Ward T., & Zhu W. J. (2002). Bleu: a Method for Automatic Evaluation of Machine Translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.
13. Ping K. (2009). Machine Translation. In: BAKER Mona. *The Routledge Encyclopedia of Translation Studies*. USA: Taylor and Francis, pp. 162-168

14. Qun L. & Xiaojun Z. (2023). Machine Translation: General. In: Sin-Wai C. *The Routledge Encyclopedia of Translation Technology*. USA: Taylor & Francis. pp. 112-127
15. Webster M. (1984). Webster's New Dictionary of Synonyms. MA: Merriam-Webster Inc.
16. Online encyclopaedia: www.britannica.com (Sep. 3rd, 2025, at 08:23 am)
17. Online dictionary and thesaurus: www.merriam-webster.com (Sep. 3rd, 2025, at 08:23 am)