

## **Translating (Metaphors) in the Age of AI: Opportunities, Challenges, and Implications for the EFL Classroom**

### **ABSTRACT**

The paper explores the use of AI translation tools in EFL classrooms, focusing on metaphor translation. We investigate the attitudes of first- and third-year English students at the University of Rijeka, Croatia, towards AI tools and evaluate three platforms: Google Translate, ChatGPT, and Glosbe, regarding their ability to accurately translate metaphors. The findings show a generally positive student disposition towards AI tools but also highlight frequent inaccuracies in AI-generated metaphor translations. We discuss the implications of these results for EFL teaching, emphasizing the potential value of error correction as a pedagogical tool. Our analysis suggests that the limitations of AI tools can serve as valuable pedagogical resources for fostering critical engagement, improving students' understanding of culturally and contextually impregnated language, and enhancing their linguistic skills. Our findings underscore the need for an urgent and systematic integration of AI tools into classrooms.

**Keywords:** AI in education, machine translation, metaphors, error correction, English as a Foreign Language (EFL)

### **Prevajanje (metafor) v dobi umetne inteligence: priložnosti, izzivi in posledice za učilnico angleščine kot tujega jezika**

#### **IZVLEČEK**

V članku raziskujemo uporabo prevajalskih orodij, ki temeljijo na uporabi umetne inteligence (UI), pri pouku angleščine kot tujega jezika (EFL) s poudarkom na prevajanju metafor. Preučujemo stališča študentov in študentk prvega in tretjega letnika angleščine na Univerzi na Reki na Hrvaškem do orodij UI ter ocenimo zmožnosti natančnega prevajanja metafor na platformah Google Translate, ChatGPT in Glosbe. Ugotovitve kažejo na splošno pozitiven odnos študentov in študentk do uporabe orodij UI, a hkrati v raziskavi izstopa tudi pogosta netočnost pri prevodnih metaforah, ki jih prevede UI. Razpravljamo o pedagoških posledicah teh rezultatov, pri čemer poudarjamo didaktični potencial popravljjanja napak kot učnega pristopa. Ugotavljamo tudi, da omejitve orodij UI lahko predstavljajo dragocena izhodišča pri pouku, saj spodbujajo kritično razmišljanje, pomagajo pri razumevanju kulturno in kontekstualno zaznamovanega jezika ter prispevajo k izboljšanju jezikovnih spretnosti. Ugotovitve podpirajo nujnost sistematičnega vključevanja orodij UI v pedagoški proces.

**Ključne besede:** umetna inteligenca v izobraževanju, strojno prevajanje, metafore, popravljjanje napak, angleščina kot tuji jezik (EFL)

# 1 Introduction

Contemporary science in general, and cognitive linguistics in particular, embrace the view that human experience – and, by extension, human language – are profoundly conditioned and shaped by the human body, perception, and culture. More broadly, the relativistic view that language both reflects the conceptual structure of the speaker and simultaneously influences cognitive processes to some degree has been well-established and validated within the field (Whorf 1956; Lakoff and Johnson 1980). According to this perspective, language not only mirrors human thought but also actively contributes to shaping it.<sup>1</sup>

In the specific context of how language reflects and influences cognition, metaphors are often regarded as pivotal elements that guide thought processes, serving as “switching points” on the rail junctions of our ideas (Lakoff and Johnson 1980).

Most fluent speakers of English know that the phrases ‘switching point’ and ‘rail junction’ used in the previous sentence are not to be understood literally – i.e., as referring to a pivotal point at a rail junction where trains can go one direction or another. Instead, these two expressions are expressive devices of language that describe metaphors as having the power to direct our ideas into one among the many possible options that our cognitive, or thought processes can create. In fact, in the sentence under scrutiny, we have just resorted to metaphors to explain what they are. In other words, we have compared the thought process to a train journey, and tried to illustrate the capacity of metaphors to direct, redirect and shape our ideas in a certain way (rather than some other possible one) by comparing metaphors to switching points on rail junctions that can direct the train (i.e., thought) along different paths. Our aim was to render – as clearly as possible – the idea that the thought process can be strongly directed by using metaphors. In fact, as illustrated by our example, metaphors are linguistic tools that make us understand one thing (usually a more abstract one – in our case the cognitive process) in terms of another (usually a simpler, more ‘accessible’ one – in our case the train travelling on rail tracks that can and do go in different directions by being directed at switch-points).

If we now try to translate the sentence ‘Metaphors are switching points on the rail junctions of our ideas’ using a few AI tools,<sup>2</sup> we get the following:

a) ChatGPT:

Metafore su “točke prebacivanja” na željezničkim čvorištima naših ideja. (literal translation with inadequate lexical selection of ‘transfer points’ and ‘rail nodes’/‘railway crossing’).

---

<sup>1</sup> See, e.g., the work done in the past few decades by the interdisciplinary Language and Cognition Group (now subdivided into multiple groups) of the Max Planck Institute for Psycholinguistics in Nijmegen, on how language influences perception, categorization, and conceptualization. For more information and references, visit <https://www.mpi.nl>. For a further range of studies on the interplay between linguistic structures and cognitive processes (e.g., the interplay between language and the perception of space and time), see the extensive body of work by Lera Boroditsky (e.g., Boroditsky and Gaby 2010). For recent work, see Maier and Abdel Rahman (2024).

<sup>2</sup> These three AI translation tools – Google Translate, ChatGPT, and Glosbe – were chosen because they are the most commonly used tools in Croatia, among both students and professional translators (for more details, see Section 2 below). Google Translate and Glosbe use machine learning algorithms, and ChatGPT relies on advanced language models (like GPT) that involve deep learning.

## b) Google translate:

Metafore su “sklopne točke” na željezničkim raskrižjima naših ideja. ((literal translation with inadequate lexical selection of ‘switch points’ and ‘rail nodes’/‘railway crossing’)).

## c) Glosbe:

Metafore su prekretnice na željezničkim raskrižjima naših ideja. (literal translation with inadequate lexical selection of ‘turning point / milestone’ and ‘rail nodes’/‘railway crossing’).

All three translation versions are too literal and prove to be contextually and culturally inappropriate in the target language. In all three cases, the machine translation tool yields an attempt to render lexical accuracy by proposing a literal translation of what is recognized as (railroad) technical terms – namely switching point and rail junction, (rendered into Croatian – in the case of the English ‘switching point’ – as ‘transfer point’ in a), ‘switching point’ in b), and ‘turning point / milestone in c), while all three tools render ‘rail junctions’ as ‘rail node’ / ‘railway crossing’). At the same time, neither of the three translation versions manages to render the pragmatic value of the source language metaphoric expressions, thus failing to convey the message. All three versions have problems with the two metaphors in the source language sentence (metaphors seen as ‘switching points’ and viewing ‘the rail junctions’ of ideas).

If we now turn to student translations, we get the following:

d) Translation by students:<sup>3</sup>

Metafore su skretničari na raskrižjima naših misli. (literally, ‘Metaphors are the switchmen at the junctions of our thoughts’).

Immediately, we note a stark contrast between the AI-generated translations and the human translation. While the former are overly literal and culturally inappropriate, the student version, characterized by a degree of creative liberty (i.e., a slight departure from the source language), functions effectively within the context and is perfectly adapted to the target language and culture.

Building on these observations, and within the broader academic discussion surrounding the use of AI tools by English as a Foreign Language (EFL) students and translators (Gašpaović et al. *in prep.*), as well as the challenges faced by AI in idiomatic translation (Gašpaović et al. *in prep.*), in this paper we investigate the possibilities and challenges associated with metaphor translation by AI tools. The issue is explored within the larger, applied context of AI tool usage in the EFL classroom. One of our main aims is to highlight the growing and urgent need to explore and standardize possible applications of AI in the classroom, starting from a detailed understanding of AI and its potential in language learning. The paper is structured as follows: after a brief literature review, we introduce the study, outlining the methodology and results. These results are then discussed, and the implications for AI translation tools in

<sup>3</sup> All the AI translations of metaphors under investigation in this study were also analysed by the students who completed the questionnaire, post questionnaire completion, and the translation proposed as ‘translation by the students’ refers to the version that was agreed upon by the students in class as the best translation option.

general – specifically in the context of metaphor translation – and their integration into EFL classrooms are addressed. The study concludes with a discussion of the current landscape of AI in pedagogy.

## 2 Setting the Stage

It is indisputable that the world is currently undergoing a profound digital transformation across all sectors. In terms of both speed and scale, this change can be described as tectonic. A central driving force behind this transformation is Artificial Intelligence (AI), defined here as the simulation of human intelligence processes by computer systems (see Healey 2020).

The three translation tools under scrutiny in this paper – Google Translate, ChatGPT, and Glosbe – are all powered by information technology (IT) and are thus frequently referred to in the literature as IT translation tools. However, in this study, we refer to them as AI translation tools because our focus is specifically on their capabilities in terms of how they process and generate language – that is, how these systems can perform tasks typically requiring human intelligence, such as language understanding, generation, contextualization, and cultural adaptability.

While language processing is one of the most rapidly advancing fields within AI, and AI is increasingly infiltrating various aspects of life, one surprising area where AI adoption appears to be lagging is in language didactics, with the EFL (English as a Foreign Language) classroom serving as a notable example. A considerable number of teachers still have a double-edged sword attitude toward the implementation and integration of AI tools into the foreign language teaching process (this issue is discussed in detail in Section 4 below).

Numerous recent studies suggest that the integration of AI tools into second/foreign language teaching should be viewed not merely as a passing trend but as an urgent and growing need (Crompton, Edmett, and Ichaporia 2023; Edmett et al. 2023; Vogt and Flindt 2023). As highlighted in our literature review, while the potential of AI is widely recognized by both students and, to a certain extent, teachers, its application in classroom practice remains limited. One region that seems to be reversing this trend is Asia (Crompton, Edmett, and Ichaporia 2023). Meanwhile, most countries within the European Union are still awaiting national policy guidelines on the issue, with notable exceptions such as Sweden (Musk 2022) and the UK (Edmett et al. 2023). A recent study by Vogt and Flindt (2023) demonstrates that even low-threshold AI tools have been integrated into classroom practice in a limited and hesitant manner, with a general tendency for such tools to be “ignored and excluded from language teaching” (2023, 2). Furthermore, the integration of AI tools into EFL classrooms remains underexplored, despite their clear importance (Crompton, Edmett, and Ichaporia 2023; Vogt and Flindt 2023).

The above implies that we are neglecting the vast potential applications of AI tools in foreign language classrooms. AI-supported tools, particularly translation tools, can assist students to improve their language skills by offering instant translations, exposing them to diverse language structures, and providing immediate feedback (Dizon and Gayed 2021; Farrokhnia et al. 2023; Schmidt and Strasser 2022). Moreover, AI tools can provide personalized learning

experiences, adapting to the individual needs of students and offering resources tailored to their specific learning levels (Okolo et al. 2024). AI chatbots like ChatGPT, for instance, can engage students in conversational practice, enhancing their fluency and comprehension (Crompton, Edmett, and Ichaporia 2023; Kazu and Kuvvetli 2023). Finally, AI tools can create innovative and stimulating learning environments and contexts, such as through virtual reality (Chen et al. 2022).

True, the vast potential of integrating AI into EFL teaching comes with certain limitations. Setting aside the many ethical issues, which are outside the scope of this paper, one of the most significant challenges in integrating AI tools into foreign language pedagogy lies in the pragmatic and culturally embedded nature of language, particularly regarding idiomatic expressions and metaphors. As illustrated in the introduction, AI tools are not equipped to provide the meaningful, context-based, and possibly culture-specific interpretations necessary for appropriately translating or explaining culturally and/or contextually loaded phrases (see also Naveen and Trojovský 2024). These tools often rely on literal translations, overlooking important nuances such as tone, cultural implications, and the pragmatic functions of language in real-world contexts. To reiterate our initial point, metaphors – acting as cognitive and cultural connectors in language – present a significant challenge for AI translation tools.

One of the main messages we wish to convey is that this limitation need not be viewed solely in a negative light; instead, it can be regarded as a potentially valuable pedagogical tool. When considering the issue in the context of the critical role of error correction in the EFL classroom (see Khansir and Pakdel 2018), the potential of AI translation tools emerges as far more valuable than initially apparent. These tools are useful pedagogical resources not only when they provide accurate translations, but also when they fail to do so, since these failures can offer powerful opportunities for teaching about culture and context-specific elements, critical evaluation of translations, and overall learning. This is the central argument we will further explore in the discussion below. Before delving into that, in the central part of this study, we will review our research examining student attitudes and habits regarding the use of AI translation tools, as well as the performance of these tools in translating metaphorical language.

### 3 The Study

Given that students, teachers, and translators alike face numerous challenges related to the use of AI translation tools in their everyday work, we decided to explore some of these pressing issues in greater detail and, at the very least, provide a more comprehensive framework for their future investigation. The study presented in this paper is motivated by the following questions:

1. What are the most widely used translation AI tools among Croatian students of English?
2. Which opportunities and challenges do students recognize regarding the use of these tools for translation purposes?
3. Are students actively encouraged to use AI tools in their translation work, and/or provided adequate guidance in this respect?
4. What guidance would they give to users of AI translation tools?

5. How do the most widely used AI tools cope in the translation of metaphors?
6. What are the implications of all the above for EFL classrooms?

The first four questions were investigated using a questionnaire. To address the fifth question, twenty metaphorical expressions – both cross-linguistically transparent and opaque – were run through the most popular AI translation tools (Google Translate, ChatGPT, and Glosbe), and their English-to-Croatian and Croatian-to-English translations were evaluated in terms of accuracy and cultural appropriateness. The final, sixth question is explored in the discussion section of the paper, drawing on the findings from the first five research questions.

### 3.1 Methodology

This study employed a mixed-methods approach, integrating both qualitative and quantitative techniques. The initial stage involved a comprehensive literature review, focusing on existing scholarly work related to the implementation of AI tools in EFL classrooms in general, and specifically for the purposes of metaphor translation. The goal of this stage was not only to synthesize existing scholarly discussions but also to assess the level of awareness and identify current areas of interest. This review served as both a stimulus and a foundation for the subsequent phases of the research. As noted in the previous section, this phase highlighted the fact that the scholarly awareness regarding the need for and ways of integrating AI into EFL classrooms, particularly in metaphor translation, is extremely limited and insufficient.

The next phase involved administering a mixed-format questionnaire (see Appendix), which included both closed and open-ended questions. This questionnaire was used to collect primary data and assess the current situation regarding the use of AI tools by students in their everyday activities, both in and outside the classroom. The survey was completed by seventy-two university students enrolled in the undergraduate English program at the University of Rijeka, from two different years of study: forty-four students were in their first year, and twenty-eight were in their third year. The closed-ended questions aimed to gather quantitative data on the frequency of AI tool usage, exposure to these tools, and preferences for specific AI translation tools. The open-ended questions were designed to collect qualitative responses regarding participants' views on various aspects of AI translation tools in the context of EFL teaching and learning.

The final stage of the study involved a controlled experiment in which a set of metaphors, ranging from cross-linguistically transparent (i.e., having lexical and pragmatic translational equivalents) to cross-linguistically opaque (no straight one-to-one translational matches), were translated from English to Croatian and vice-versa using several AI translation tools (Google Translate, ChatGPT, and Glosbe). The translations were evaluated one by one based on accuracy, fidelity to the source meaning, and cultural appropriateness of the target expression. The results of these translations were subsequently discussed in class.

### 3.2 Results

The results are presented below in the order of their appearance on the questionnaire. The responses to the quantitative questions are reported and displayed in graphs, while the

answers to the qualitative questions are organized and presented based on the frequency of recurring response themes.

### 3.2.1 Use of AI Translation Tools (Question 1)

Both first-year and third-year students reported using AI translation tools in their everyday life and academic work. As a result, the answer to **Question 1**, which asked whether students used AI translation tools, was **100% “yes”** in both groups.

### 3.2.2 Most Widely Used AI Translation Tools (Question 2)

When asked about the most widely used AI translation tools among Croatian students of English at the University of Rijeka, the findings revealed significant differences between first-year and third-year students in their software preferences.

- **First-Year Students:** Among the first-year students ( $n = 44$ ), **Google Translate** emerged as the overwhelmingly preferred translation tool, with 42 students reporting its use. Other tools, such as **ChatGPT** (8), **Glosbe** (5), and **DeepL** (2), were used to a much lesser extent.
- **Third-Year Students:** In contrast, **third-year students** ( $n = 28$ ) favoured **Glosbe** (24), with **Google Translate** (17) as the second most popular option. Additionally, **online dictionaries** (7), **DeepL** (6), and **ChatGPT** (4) were also commonly used, while **Eudict** and **Reverso** each had one user.

We note that the total number of responses for this question exceeds the number of students who completed the questionnaire. This is because many students reported using multiple translation tools, with this behaviour being more prevalent among third-year students.

TABLE 1. Preferred translation tools.

| Translation Tool    | First-Year Students ( $n = 44$ ) | Third-Year Students ( $n = 28$ ) |
|---------------------|----------------------------------|----------------------------------|
| Google Translate    | 42                               | 17                               |
| Glosbe              | 5                                | 24                               |
| ChatGPT             | 8                                | 4                                |
| DeepL               | 2                                | 6                                |
| Online Dictionaries | 0                                | 7                                |
| Eudict              | 0                                | 1                                |
| Reverso             | 0                                | 1                                |

### 3.2.3 Guidance on AI Translation Tools (Questions 3 and 4)

The responses to Questions 3 and 4 reveal a notable contrast between first-year and third-year students regarding guidance on AI translation tools.

- **First-Year Students:** The majority of first-year students ( $n=44$ ) reported that they had not received any guidance on the availability or use of AI translation tools during the

first three months of their university education, or prior to that.<sup>4</sup> This lack of guidance was reflected in their negative responses.

- **Third-Year Students:** In contrast, the third-year students (n = 28) were divided, with 7 students indicating that they had received guidance on AI translation tools at the university level. This was an unexpected finding, but further investigation revealed that all seven of these students had German as their second major. As part of their German courses, they were introduced to **Glosbe**, which they subsequently shared with their classmates, explaining the high popularity of Glosbe among third-year students.

### 3.2.4 Benefits of AI Translation Tools (Question 5)

Regarding the benefits of using AI translation tools, students from both first and third years highlighted similar advantages, with **speed** and the **multiplicity of options** being the most frequently mentioned. Other noted benefits include the following:

- Simplicity and ease of use
- The availability of alternatives, allowing users to choose the best solution
- Tools that sometimes remind users of the appropriate translation
- Vocabulary expansion
- Potential as a learning tool

Since the advantages identified by first-year and third-year students did not differ significantly, the results have been combined into one graph below.

### 3.2.5 Challenges in Using AI Translation Tools (Question 6)

When it comes to challenges that may hinder the use of AI translation tools, the respondents were almost unanimous in identifying **accuracy** and issues related to **(de)contextualization** as major concerns. Specific challenges mentioned included the following:

- Problems with idioms and metaphors
- Issues with collocations
- Difficulty handling cultural allusions
- Struggles with the pragmatic aspects of language, such as colloquial expressions, proverbs, and personal perspectives

Additionally, some students mentioned less common challenges, such as the following:

- Problems with syntax
- Insufficient data for Croatian, resulting in poorer performance when translating from or into Croatian, compared to “major” languages

---

<sup>4</sup> The questionnaire was administered in late December 2024, by which time first-year students had completed three months of lectures, as the academic year begins in October. In terms of guidance on the use of AI translation tools, respondents were instructed to consider any form of guidance they had received, including both during their time at university and prior to enrolment, at any level of education.

- A lack of habit in using physical dictionaries and other literature
- A potential narrowing of creativity and autonomy owing to reliance on IT tools

No significant differences were noted between the first- and the third-year respondents.

### 3.2.6 Use of IT Tools for Translating Metaphors (Question 7)

Regarding whether students would use IT tools for translating metaphors, the responses revealed a clear ‘no’ from the first-year students, and a more divided response from the third-year students.

- **First-Year Students:** All 44 first-year students (100%) stated that they would not use IT tools for translating metaphors.
- **Third-Year Students:** Among the 28 third-year students, 21 (75%) preferred not to use IT tools for translating metaphors, while 7 (25%) were in favour of using them.

It is important to note that while first-year students simply answered ‘no’ to the question of using IT tools for metaphor translation, third-year students often elaborated on both their ‘yes’ and ‘no’ responses. Positive responses were frequently qualified with phrases such as “Yes, but...” or “Yes, if...”. Common elaborations included the following:

- “Yes, but with caution/guidance/care/a pinch of salt.”
- “Yes, if I cannot think of an equivalent myself / if I am unfamiliar with the source language expression.”
- “Yes, but always double-checking the proposed translation with authentic target language data, teachers, or native speakers.”

On the other hand, most negative responses were followed by explanations that reiterated the limitations of AI translation tools, particularly when dealing with the idiomatic, metaphoric, pragmatic, and cultural aspects of translation – issues already highlighted in the responses to Question 6 (above).

### 3.2.7 Guidance for Using AI Translation Tools as Future EFL Teachers (Question 8)

Finally, when it comes to the guidance that first- and third-year students, as future English as a Foreign Language (EFL) teachers, would give to their students regarding the use of AI translation tools, the responses were grouped based on common themes, which were then ranked according to their frequency. The most frequent themes identified are listed below, from most to least common:

#### **First-Year Students:**

- Use IT tools as a help rather than relying exclusively on them
- Always double-check the solutions proposed by IT tools, as they can make many mistakes

- Compare different sources
- First, study and understand the language, and only then use IT tools
- Use with great caution or moderation; do not overuse them (they will make you “dumb”)

Responses that were unique or only appeared once include:

- Do not use for sentences, only for words
- Do not use them, as they are above your level
- Use them because they can help you better understand a text in a foreign language
- Use them only as an exercise in recognizing mistakes commonly made by others

### **Third-Year Students:**

- Use IT tools, but do not abuse them; they are just a tool, a help, support
- Use more than one IT tool and check the different solutions; triple fact-check; compare and contrast different solutions; always remember that these tools make mistakes all the time
- Do not use IT tools if your knowledge is poor
- Learn to think critically about the proposed translations
- First, learn to work without their help
- I would discourage their use; use as a last resort
- If you get paid for a translation, you must do it yourself – anyone can feed a source language into a tool, but that is not the point
- Best for specific/technical vocabulary

We immediately observe that the responses from third-year students reveal more pedagogically oriented comments compared to those from first-year students. While first-year students mainly focus on cautionary advice, such as using the tools sparingly and cross-checking results, third-year students demonstrate a deeper understanding of the broader educational implications of IT tool usage. They emphasize critical thinking, the importance of building foundational knowledge before relying on tools, and the need to use these resources as a last resort or for specific tasks. This shift toward more pedagogically sound advice reflects their growing awareness of the role of an EFL teacher and their understanding of how to guide students effectively in the classroom. This topic will be discussed in more detail in the next section.

### **3.2.8 Accuracy of AI Translation Tools in Translating Metaphors**

In order to explore the level of accuracy demonstrated by a selected set of AI tools in translating metaphors, we created two sets of metaphorical expressions: one containing 10 cross-linguistically transparent metaphoric expressions (i.e., metaphorical phrases that have lexical and pragmatic equivalents in both languages), and the other set containing 10 cross-linguistically opaque metaphoric expressions (i.e., metaphorical phrases with no direct one-to-one translational equivalents). These two sets are exemplified below.

### Cross-linguistically transparent metaphorical phrases:

1. *A double-edged sword* – *Mač s dvije oštrice* (meaning: something that has both positive and negative consequences or effects);
2. *The tip of the iceberg* – *Vrh ledenog brijega* (meaning: what is visible or known is just a fraction of the whole);
3. *A wolf in sheep's clothing* – *Vuk u janjećoj koži*<sup>5</sup> (lit. 'Wolf in lamb's skin' – meaning: someone who appears to be harmless or trustworthy is actually dangerous and deceitful)
4. *To be on cloud nine* – *Biti na devetom nebu* (meaning: extremely happy or in a state of bliss).

### Cross-linguistically opaque metaphorical phrases:

1. *A silver lining* – *U svakom zlu neko dobro* (lit. 'In every evil, there is some good') (meaning: a positive or hopeful aspect to a generally negative or difficult situation)
2. *Caught between a rock and a hard place* – *Između čekića i nakovnja* (lit. 'Between the hammer and the anvil') (meaning: be caught in a position where one is pressed from two sides, with no easy way out);
3. *Like two peas in a pod* – *Kao dvije kapi vode* (lit. 'Like two drops of water') (meaning: two things or people that are extremely similar or identical)
4. *Miss the boat / the ship has sailed* – *prošla baba s kolačima* (lit. 'the grandmother with the cakes has passed/left') / *Prošao voz* (literally, 'the train has passed') (meaning: an opportunity has been missed, and it is too late to act now)

All expressions from both sets were translated from English to Croatian and vice-versa using three AI translation tools: Google Translate, ChatGPT, and Glosbe. All translations were first evaluated by the teacher for accuracy, fidelity to the source meaning, and cultural appropriateness of the target expression. The results of these translations were subsequently discussed with the students in class.<sup>6</sup>

The results we obtained when translating our two sets of metaphors align closely with the findings presented by Gašparović, Brala-Vukanović, and Brkić-Bakarić (in prep). In fact, the accuracy of translations of expressions that contain metaphorical language (either metaphorical descriptions or conventional metaphors embedded in idioms, proverbs, etc.) seems to be influenced by three key factors: a) the degree of equivalence between source and target metaphoric expressions; b) the source and target language pair; and c) the translation platform.

More specifically, in terms of platform performance, ChatGPT demonstrates a higher level of translation accuracy compared to Google Translate, which, in turn, outperforms Glosbe. It

<sup>5</sup> To give a good example of overliteral translation, let us note here that albeit both English and Croatian use this idiom, and the metaphorical mapping between good and evil is clear and frequently used, ChatGPT translates the English expression 'wolf in sheep's clothing' as 'wolf in sheep's skin', mistaking the Croatian culturally usual 'lamb's skin' with the literal translation from English, i.e. 'sheep'.

<sup>6</sup> The evaluations of the translations in class were done after the students had completed the questionnaire.

is worth noting that the study by Gašparović, Brala-Vukanović, and Brkić-Bakarić (in prep) indicates that Microsoft Copilot surpasses ChatGPT in translation accuracy. However, since our participants did not use Microsoft Copilot, it was not included in this study.

Additionally, translation accuracy tends to be higher when translating from Croatian to English rather than from English to Croatian. This is unsurprising, as AI translation tools are trained on large datasets. Translating from a language with moderate data resources, such as Croatian, into a language with extensive data resources, like English, poses less of a challenge for AI systems compared to translating in the reverse direction.

Finally, as anticipated, metaphorical expressions with close translational equivalents in the target language were translated more accurately than those without such close (lexical, semantic, and pragmatic) equivalents.

Considering the variability in translation quality – from highly accurate and contextually appropriate translations to completely inaccurate or unacceptable ones – the role of the teacher in this process becomes crucial. This issue is discussed in detail in the next section of this paper.

## 4 Discussion

The data presented in Section 3 reveal interesting trends in the use of AI translation tools among Croatian students of English, with notable differences between first-year and third-year students. These differences reflect both the development of various linguistic and pedagogical skills, as well as the evolving relationship students have with these tools as they progress through their studies. The findings emphasize a (potential) critical link between linguistic competence and reliance on IT (translation) tools. Additionally, they underscore the urgent and growing need for more structured guidance on the use of AI translation tools in the EFL classroom.

At this point, an important observation is that while the use of AI translation tools is widespread among students, this practice is not always mirrored in teaching practices or reflected in the national curriculum. In Croatia, there seems to be a disconnect between the rapid development of these tools and the pace at which educators are integrating them into the EFL classroom. One plausible explanation for this disconnect lies in the generational gap between students, who are more inclined to use technology in general, and teachers, who may be less familiar with or receptive to these tools. Interestingly, while everyday practice shows that teachers frequently use AI translation tools outside the classroom for their own purposes, our data suggest that they remain reluctant to incorporate these technologies into their teaching practices. This reluctance may stem from concerns about AI tools replacing traditional teaching methods or teachers feeling inadequately prepared to integrate them among their pedagogical tools (Edmett et al. 2023; Vogt and Flindt 2023).

Furthermore, this lack of integration of AI tools in language education is at least in part due to the lack of structured guidance or curricular support for AI integration in the EFL classroom. Given students' natural inclination towards IT, and given that research has demonstrated

the potential benefits of AI translation tools in language learning, including vocabulary improvement, stylistic refinement, and even anxiety control (Crompton, Edmett, and Ichaporia 2023), lack of structured guidance is undermining the many potential pedagogical benefits these tools could provide for both teachers and students.

In particular, deeper analysis and guidance is needed for clearly understanding and describing the tasks that AI tools can handle effectively, how these tasks can be incorporated into the classroom, and what exactly remains the role of the teacher and teacher expertise (for a comprehensive discussion regarding this point, see Edmett et al. 2023). In the context of the current topic, i.e. AI translation of metaphors, understanding how metaphor translation tasks can be incorporated into EFL teaching, and how students and teachers could benefit from this, would allow AI to serve as a useful pedagogical tool. In fact, while our study has shown that AI translation tools may often underperform with metaphorical expressions, teachers should be made aware of the fact that AI's failures in accuracy should be seen as valuable teaching opportunities. In cases of inaccurate AI translation output (e.g., results too literal or contextually inappropriate), a detailed error correction process can serve not just to improve students' language skills, but also to foster students' cultural sensitivity and their ability to critically evaluate AI translation tools and their output from diverse perspectives.

The lack of curricular guidance on the use of IT (translation) tools results in their pedagogical integration in Croatia remaining largely informal and individualized. At the same time, as our data show, when guidance is provided, it is generally well received by students. In fact, while first-year students from our study primarily rely on Google Translate as a convenient tool, third-year students exhibit more varied use, with Glosbe emerging as the most popular choice. This shift is likely due to guidance they received in their German classes, highlighting how even a minor introduction to AI tools can significantly influence students' perceptions and usage. The success of Glosbe among third-year students underscores the importance of structured guidance in AI tool integration, as it can influence how students navigate and evaluate the performance of different translation tools. Moreover, since AI tools develop at varying rates (some showing decline in performance results over time – cf. Gašparović, Brala-Vukanović, and Brkić-Bakarić (*in prep*)), it is crucial that students are taught to critically assess and reassess these tools continuously.

Our data also shows an interesting evolution in students' critical attitudes toward AI tools as they advance from the first to the third year, even without structured, formal pedagogical guidance in this respect. While first-year students often approach AI tools with caution, third-year students show a more nuanced understanding of these tools' limitations, especially in the translation of culturally specific content, such as metaphors.<sup>7</sup> This progression seems to reflect their trial-and-error experiences with translation tools, and only sporadic, informal guidance from individual teachers. The third-year students, despite lacking formal pedagogical training, demonstrate a critical approach, acknowledging the inherent limitations of these tools and the necessity of double-checking translations. In this regard, integrating cultural awareness training into the use of AI translation tools, through error analysis and correction of Idioms

<sup>7</sup> For an interesting study on how students working in groups negotiate limitations in the use of Google Translate, see Rowe (2022).

and metaphors, for example, constitutes a step in the right direction. The current situation, where the development of critical thinking and critical evaluation skills regarding these tools is left to the individual trial-and-error experiences of students, and their intuitive awareness, is certainly not the most effective pedagogical method.

Given all the above, it would be fair to say that AI translation tools can undoubtedly be useful for basic translation tasks, and students should be encouraged to use them as a complement to their own language skills, rather than a substitute. Instead of simply proposing IT translation tools as potential help, teachers should actively guide students in the practical application of these tools, crucially focusing on their potential to be used as excellent – and de-personalized – error-correction pedagogical opportunities. Notably, the de-personalized nature of AI's inadequate or inaccurate translations eliminates the personal failure moments that typically occur in a real classroom, thus removing the potential for demotivating students or exposing them to negative and stressful emotions. This aspect is a crucial factor that should not be overlooked when analysing the future role of AI (translation) tools in the EFL classroom.

Providing opportunities for students to critically assess the quality of (a variety of) translations, especially when dealing with idiomatic or metaphorical language, can help improve not only their linguistic and translation skills, but also their wider understanding of cultural and contextual nuances of language and communication. Error correction as a pedagogical tool has a valid ally in IT. While AI-driven tools can and do make mistakes, the human ability to critically analyse and correct these mistakes, as well as to learn from them, is an efficient way for deeply and productively engaging with language and the translation process in the context of AI tools.

## 5 Concluding Remarks

In light of the discussion above, it is evident that the integration of AI translation tools into the EFL curriculum is essential and urgently needed. This integration would not only reflect students' current habits, interests, and needs, but also allow for a more structured and effective use of these tools in language learning. Several key points need to be addressed in this regard: 1) the structured inclusion of AI tools in the curriculum and teacher training on their pedagogical use; 2) continuous monitoring of the development of AI translation tools; and 3) fostering critical thinking in students, particularly in relation to translating culturally specific language.

While AI translation tools offer both opportunities and challenges, their current lack of integration in the classroom limits their potential. By fostering critical thinking and promoting a balanced use of these tools, educators can enhance students' translation abilities, especially when dealing with the complexities of culturally embedded language, such as metaphors. Even though AI translation tools often struggle with culture-specific expressions, these shortcomings provide excellent opportunities for error-correction-based pedagogy. As AI technology continues to evolve, it is crucial that educators – and before them education policy makers – adopt a proactive, structured approach to integrating these tools into the curriculum. The future of translation and language teaching is inevitably intertwined with AI, and it is crucial that both teachers and students are equipped to

navigate this evolving landscape and use it in an informed and guided way, so as to make full use of their potential and avoid possible traps.<sup>8</sup>

Further research is needed to explore how AI translation tools can be more effectively integrated into language learning curricula. Given their rapid development, investigating just how exactly AI tools can be tailored to address specific challenges – such as translating metaphors and other culturally embedded language features – will be crucial. Understanding which pedagogical tasks are best suited for AI and which should remain the (sole) responsibility of the teacher will provide valuable insights into how educators can best prepare students for the challenges and opportunities AI tools present.

Ultimately, the gap between AI's potential and its actual integration in EFL classrooms is not only a technological issue but also a pedagogical one. Future research and teacher training should address both the practical and theoretical aspects of AI tool use in language education. Training programs should not only focus on how to use these tools effectively but also emphasize fostering students' critical engagement with them. Teaching students to understand the limitations of AI tools and using these limitations as a pedagogical opportunity will enhance their language skills and ensure that AI becomes a valuable, guided resource in the EFL classroom.

In conclusion, while many of the issues discussed remain at an intuitive level, their importance and relevance demand immediate scholarly attention and structured curricular guidance. The time to act is now – further delays in integrating AI translation tools into EFL teaching would mean missing a critical opportunity to significantly enhance language learning and teaching in the digital age.

## References

- Boroditsky, Lera, and Alice Gaby. 2010. "Remembrances of times east: Absolute spatial representations of time in an Australian Aboriginal community." *Psychological Sciences* 21 (11): 1635–39. <https://doi.org/10.1177/0956797610386621>.
- Crompton, Helen, Adam Edmett, and Neenaz Ichaporia. 2023. *Artificial Intelligence and English Language Teaching: A Systematic Literature Review*. British Council.
- Dizon, Gerald, and Jamal M. Gayed. 2021. "Examining the impact of grammarly on the quality of mobile L2 writing." *JALT CALL Journal* 17 (2): 74–92. <https://doi.org/10.29140/jaltcall.v17n2.336>.
- Edmett, Adam, Neenaz Ichaporia, Helen Crompton, and Ross Crichton. 2023. *Artificial Intelligence and English Language Teaching: Preparing for the Future*. British Council. [https://www.teachingenglish.org.uk/sites/teacheng/files/2024-08/AI\\_and\\_ELT\\_Jul\\_2024.pdf](https://www.teachingenglish.org.uk/sites/teacheng/files/2024-08/AI_and_ELT_Jul_2024.pdf).
- Farrokhnia, Mohammad, Sima K. Banihashem, Omid Noroozi, and Allan Wals. 2023. "A SWOT analysis of ChatGPT: Implications for educational practice and research." *Innovations in Education and Teaching International* 61 (3): 460–74. <https://doi.org/10.1080/14703297.2023.2195846>.
- Gartner, Smiljana, and Marjan Krašna. 2023. "Artificial intelligence in education – ethical framework." *12th Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro, 1–7. <https://doi.org/10.1109/MECO58584.2023.10155012>.

<sup>8</sup> Among the many challenges discussed in this study, one has remained outside its primary focus – namely, the ethical considerations. Although the present research does not explicitly address the ethical dimensions of incorporating AI technologies into the classroom, this exclusion should not be interpreted as suggesting that the issues explored here can be considered in isolation from the significant ethical concerns that accompany them. For a comprehensive and recent discussion of ethical considerations related to the integration of AI tools into pedagogical processes, see Gartner and Krašna (2023).

- Gašparović, Marijana, Marija Brala-Vukanović, and Marija Brkić-Bakarić. *in prep.* “Idioms and machine translation: Assessing translation accuracy in the age of AI.” Paper submitted to *CALICO Journal*.
- Healey, Justin. 2020. *Artificial Intelligence*. Volume 450. The Spinney Press.
- Kazu, Ibrahim Yasar, and Murat Kuvvetli. 2023. “The influence of pronunciation education via artificial intelligence technology on vocabulary acquisition in learning English.” *International Journal of Psychology and Education Studies* 10 (2): 480–93. <https://ijpes.com/index.php/ijpes/article/view/1044>.
- Khansir, Ali A., and Farhad Pakdel. 2018. “Place of error correction in English language teaching.” *Educational Process: International Journal* 7 (3): 189–99. <https://doi.org/10.22521/edupij.2018.73.3>.
- Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. University of Chicago Press.
- Maier, Martin, and Rasha Abdel Rahman. 2024. “Transient and long-term linguistic influences on visual perception: Shifting brain dynamics with memory consolidation.” *Language Learning* 74 (1): 157–84. <https://doi.org/10.1111/lang.12631>.
- Musk, Nicholas. 2022. “Using online translation tools in computer-assisted collaborative EFL writing.” *Classroom Discourse* 13 (2): 119–44. <https://doi.org/10.1080/19463014.2021.2025119>.
- Naveen, Palanichamy, and Pavel Trojovský. 2024. “Overview and challenges of machine translation for contextually appropriate translations.” *iScience* 27 (10): 110878. <https://doi.org/10.1016/j.isci.2024.110878>.
- Okolo, Chinwe Jane, Chinyere Grace Ezeonwumelu, Chioma Ihuoma Barah, and Ugwu Nnenna Jovita. 2024. “Language education in the age of AI: Opportunities and challenges.” *Newport International Journal of Research in Education* 4 (1): 39–44. <https://doi.org/10.59298/NIJRE/2024/41139448>.
- Rowe, Lindsey W. 2022. “Google Translate and biliterate composing: Second-graders’ use of digital translation tools to support bilingual writing.” *TESOL Quarterly* 56 (3): 883–906. <https://doi.org/10.1002/tesq.3143>.
- Schmidt, Thomas, and Thomas Strasser. 2022. “Artificial intelligence in foreign language learning and teaching: A CALL for intelligent practice.” *International Journal of English Studies* 33 (1): 165–84.
- Vogt, Kerstin A., and Lars Flindt. 2023. “Artificial intelligence and the future of language teacher education: A critical review of the use of AI tools in the foreign language classroom.” In *The Future of Teacher Education: Innovations Across Pedagogies, Technologies, and Societies*, edited by P. Hohaus and J.-F. Heeren, 179–99. Brill.
- Whorf, Benjamin Lee. 1956. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. Edited by John B. Carroll. MIT Press.

## APPENDIX – Questionnaire Administered to Students

1. Do you use IT/computer-assisted translation tools in your work?
2. If yes, which AI translation tool do you use most frequently?
3. Have you ever been exposed to AI translation tools in your English classes?
4. Have you received any guidance from your university lecturers regarding the use of AI translation tools?
5. In your view, what are the benefits of using AI translation tools?
6. In your view, what are the limitations of using AI translation tools?
7. Do you rely on IT tools for the translation of metaphors?
8. In your view, how efficient are IT tools when translating metaphors?
9. As future EFL teachers, what guidelines would you give your students regarding the use of AI translation tools?