New challenges and opportunities in technology-assisted phraseology interpreting: the case of Yandex live stream translation

Novos desafios e oportunidades na interpretação da fraseologia assistida pela tecnologia: o caso da tradução de vídeo ao vivo do Yandex

Pablo Ramírez Rodríguez [iD][*1] **and Evgeniy Antonov** [iD][†2]

[1]*Universidad de Granada, Granada, Spain.*
[2]*National Research Nuclear University, MEPhI, Moscow, Russia.*

Abstract

Computer-assisted interpreting (CAI) has become an essential element in an increasingly globalized and interconnected world. With the growing demand for instant online communication and the need to overcome language limitations in a global context, CAI has become a valuable tool for a wide range of applications. However, the CAI is not without its challenges. The interpretation of idiomatic expressions remains a significant barrier, as these linguistic constructs can be particularly difficult to interpret accurately due to their culturally embedded nature. In this context, the objective of this article is to address the problem of interpreting idiomatic expressions, in our case 52 verbal idioms within the framework of CAI, focusing on the Spanish-Russian language combination. The aim is to analyze how this technology meets the challenges of idiomatic phraseology and how it influences accurate and effective intercultural communication. To achieve this goal, a comprehensive methodology has been applied, combining linguistic analysis with practical observations of real-time technology-assisted interpreting situations using the live on-air speech translator and multimedia content provided by Yandex. The results of this research provide a deeper understanding of how translation and interpreting technology addresses the challenges of idiomatic expression, while also providing critical insight into the effectiveness and limitations of technological solutions in the field of intercultural communication.

Keywords: CAI. Phraseology. Verbal idioms. Live stream translation. Automatic interpreting.

Resumo

A interpretação assistida por computador (IAC) tornou-se um elemento essencial em um mundo cada vez mais globalizado e interconectado. Com a crescente demanda por comunicação instantânea online e a necessidade de superar as limitações linguísticas em um contexto global, a IAC se tornou uma ferramenta valiosa para uma ampla gama de aplicações. No entanto, a IAC não está isenta de dificuldades. A interpretação de expressões idiomáticas continua sendo um obstáculo significativo, pois essas construções linguísticas podem ser especialmente difíceis de interpretar com precisão devido à sua natureza culturalmente enraizada. Nesse contexto, o objetivo deste artigo é abordar o problema da interpretação de expressões idiomáticas, no nosso caso de 52 locuções verbais no âmbito da IAC, focando na combinação linguística espanhol-russo. O objetivo é analisar como essa tecnologia responde aos desafios da fraseologia idiomática e como influencia uma comunicação intercultural precisa e eficaz. Para alcançar esse objetivo, foi aplicada uma metodologia exaustiva que combina a análise linguística com observações práticas de situações de interpretação assistida por tecnologia em tempo real, utilizando o tradutor de voz ao vivo e os conteúdos multimídia fornecidos pela Yandex. Os resultados desta pesquisa proporcionam uma compreensão mais profunda de como a tecnologia de tradução e interpretação enfrenta os desafios da expressão idiomática, ao mesmo tempo que oferecem uma visão crítica da eficácia e das limitações das soluções tecnológicas no âmbito da comunicação intercultural.

Palavras-chave: IAC. Fraseologia. Locuções verbais. Tradução em tempo real. Interpretação automática.

[*]Email: pablorarod@gmail.com
[†]Email: eantonov@kaf65.ru

1 Introduction

Technology-assisted interpreting has experienced exponential growth in recent years, reflected in the dizzying progress in the development of information and communication technology (ICT) tools and resources (Gutiérrez Artacho; Olvera Lobo; Hunt Gómez, 2016; Mezcua, 2019). These innovative technologies have greatly facilitated the interpretation and comprehension of texts in different linguistic contexts (Olalla-Soler; Vert Bolaños, 2015). In this sense, computer-assisted interpreting (CAI) has become crucial in an increasingly globalised and connected society, playing a pivotal role in breaking down language barriers and facilitating effective communication in an environment characterised by cultural and linguistic diversity (Mellinger, 2019; Li, 2021). In response to the growing demand for instant online communication and the need to overcome language limitations in a global environment, CAI has established itself as an infallible tool for a wide range of applications, from interpreting international business meetings to translating multimedia content in real time (Fantinuoli, 2017a; Alcaide Martínez, 2021).

The integration of technology has revolutionised the way individuals interact with language, opening new possibilities for increasing the efficiency and accuracy of text interpretation, both in real time and asynchronously (Gaber, 2023b; Ramírez Rodríguez, 2023). The development of ICT tools has led to improvements in natural language processing, machine translation, speech recognition and other related technologies, resulting in significant advances in text interpretation (Valero-Garcés, 2024). These tools have been particularly useful in overcoming language barriers in multicultural environments, facilitating communication between speakers of different languages and promoting intercultural understanding (Pérez *et al.*, 2020).

The use of CAI does, however, present certain difficulties. In the phraseological domain, understanding idiomatic expressions remains a significant challenge, as these linguistic constructions can be particularly difficult to interpret accurately due to their cultural embeddedness (Corpas Pastor; Gaber, 2021; Ramírez Rodríguez, 2022). Cultural and contextual differences can lead to inaccurate or ambiguous translations of these expressions, resulting in misunderstandings and communication problems. In addition, CAI is often based on pre-defined algorithms and databases, which can limit its ability to adapt to new idioms and emerging expressions (Ortigoza; Morillo-Montoya; Monpué, 2024). This problem identified in the interpretation of idiomatic expressions by CAI adds to the previously mentioned challenges in technology-assisted phraseology, where accuracy and fluency in the interpretation of texts in different linguistic contexts are key to effective communication. The complexity of idiomatic expressions, such as idioms, and their culturally contextual nature require new approaches and specialised tools to improve the interpretation of such expressions, which represents a relevant area of research in the development of technology-assisted phraseology.

2 Theoretical Framework

2.1 The phraseology-technology binomial

Today, phraseology and technology have significantly converged with the advent of the internet and social networks (Corpas Pastor, G., 2013). Phraseology constitutes a specialized field of linguistic study that focuses on fixed or semi-fixed expressions within a language, encompassing idiomatic expressions, collocations, and other multi-word units. Idiomatic expressions are characterized by meanings that cannot be directly inferred from the literal interpretation of their constituent elements, as their comprehension relies heavily on cultural and contextual factors. These expressions pose significant challenges in interpretation and translation due to their language-specific, figurative nature and nuanced usage. Thus, phraseology systematically explores the structural, semantic, and functional dimensions of these expressions, emphasizing their linguistic and cultural specificity.

In parallel, within the realm of technology, the concept of CAI pertains to the application of technological tools and specialized software to augment the interpreting process, enhancing both efficiency and precision. CAI encompasses a range of resources, including terminology management systems, speech recognition software, and real-time transcription tools. These technologies play a pivotal role in fostering terminological consistency and streamlining the workflow of professional interpreters, thereby contributing to the overall quality and effectiveness of interpreting practices.

Digital platforms allow users to share phrases, expressions, and thoughts instantly and globally, creating new forms of interaction and communication (Piccioni; Pontrandolfo, 2017). In this context, artificial intelligence (AI) and natural language processing are revolutionising the way we interact with technology. In this sense, virtual assistants such as Siri, Alexa and Google Assistant are understanding and responding to voice commands in increasingly sophisticated ways, changing the way we use language in our daily lives. In other words, the combination of phraseology and technology represents a dynamic and fruitful interaction capable of transforming the way we understand and use idiomatic expressions in different linguistic contexts. Phraseology, as a linguistic discipline focused on the study of phraseological units and their use in discourse, has been greatly enriched by technological advances that have allowed the development of increasingly sophisticated language analysis and processing tools (Arnáiz, 2015).

The integration of technology in the study of phraseology opens new possibilities for the study and analysis of idiomatic expressions in different languages (Mogorrón Huerta, 2012). Thus, using digital linguistic corpora and natural language processing tools, patterns, and regularities in the use of phraseological units can be identified, which has enriched our understanding of how these expressions are used in different discursive contexts (Fantinuoli, 2017b; Corpas Pastor; Rubio, 2023; Gaber, 2023a). However, in this context, the pairing of phraseology and technology also presents challenges and limitations (Sevilla Muñoz, 2012). In this sense, the development of CAI has been catalysed by significant advances in fields such as AI, machine learning and computational linguistics. These technologies have enabled the development of increasingly sophisticated interpreting systems capable of instantly and accurately translating real-time conversations and written texts in multiple languages (Koponen; Nunes Vieira; Spinol, 2021; Guo; Han; Anacleto, 2023).

Idiomatic expressions encapsulate cultural identity, reflecting a community's values, traditions, and collective experiences. For example, an English expression like "spill the beans" (to reveal a secret) might be meaningless or misinterpreted if translated literally into a language without a similar metaphorical framework. Such cultural nuances demand interpretive skills that transcend linguistic competence, requiring interpreters to access a deep understanding of both the source and target cultures (Ramírez Rodríguez, 2024). When these nuances are ignored or mistranslated, the resulting communication can lack authenticity, coherence, or even lead to misunderstandings. In practice, the challenge lies not only in decoding the meaning of idiomatic expressions but also in determining how to convey their intended effect whether by using an equivalent idiom in the target language, paraphrasing the underlying concept, or providing additional contextual information.

Current CAI systems are ill-equipped to handle the complexity of idiomatic expressions, largely because their underlying algorithms are designed to prioritize literal, data-driven translations rather than context-sensitive or culturally nuanced interpretations. Speech recognition software often struggles with regional accents or informal speech, resulting in errors at the input stage. Similarly, machine translation engines, though increasingly sophisticated, rely on probabilistic models that may fail to capture the figurative meanings of idioms or their cultural connotations. Moreover, most CAI tools lack the ability to dynamically adjust to the contextual or pragmatic needs of a conversation (Corpas Pastor, G., 2020; Corpas Pastor, Gloria, 2022).

2.2 The role of technology in interpreting: the case of Yandex browser

As mentioned above, technology has played a key role in the development of interpreting, improving the quality of interpreting through the availability of resources such as terminology databases and online dictionaries (Cifuentes-Férez, 2015; Rockwell; Sinclair, 2022). These enable interpreters to find the accurate and up-to-date information they need to do their job effectively. A prominent example in this area is Yandex browser, a platform developed by the Russian company Yandex that combines AI and speech recognition technology to provide real-time interpreting services. Yandex browser represents a significant advance in CAI, as it harnesses the ability of AI to process large amounts of linguistic data efficiently and quickly (Jibreel, 2023).

This tool is designed to be easily integrated into different platforms and devices, allowing users to access real-time interpretation services from anywhere, at any time. This system uses natural language

processing and machine learning algorithms to automatically translate conversations into different languages (Erbsen; Põldre, 2023). In addition, Yandex browser includes interpreting capabilities in various contexts, such as business meetings, international conferences, video conferences and online events. The use of this technology in real-time situations demonstrates the ability of AI to adapt to different circumstances and provide a practical and efficient interpreting experience (Novozhilova et al., 2020).

Recently, new technological tools have been explored to improve interpreting, such as the integration of AI for context analysis and the detection of cultural nuances in discourse (Defrancq; Fantinuoli, 2021). These advances are revolutionising the way interpreting challenges are addressed, allowing professionals to adapt more effectively to users' needs and preferences. Moreover, the application of machine learning algorithms significantly contributes to improving the accuracy and fluency of interpreting, opening new possibilities for human-technology collaboration in this field, playing a key role in improving intercultural communication (Alotaibi, 2020).

In this context, the speech recognition technology built into Yandex browser is another key aspect of its functionality. This system allows spoken conversations to be instantly converted to text, which is then analysed and translated by AI algorithms to provide seamless interpretations (Shadiev; Liu, 2023). The ability to translate both speech and text in real time is essential to ensure effective communication in different situations and contexts. Furthermore, Yandex browser is characterised by its ability to adapt to different accents, intonations and speaking styles, which in theory improves the accuracy of translations and ensures smooth communication between interlocutors. This versatility is particularly important in multicultural environments where significant linguistic variations can occur. In this context, such adaptability to possible variations is based on algorithms that can detect patterns in speech and dynamically adjust the interpretation to accurately reflect the original meaning of the message, as well as on advanced acoustic modelling and speech analysis techniques (Kim; Kweon, 2020).

To further enhance its capabilities, the Yandex Live Multimedia platform also benefits from regular updates and continuous improvements to its AI algorithms. This allows it to evolve and provide increasingly advanced and efficient interpretation solutions by exploring the use of new technologies based on deep language modelling and recurrent neural networks. These techniques enable the platform to better understand the context, tone, and subtleties of human language, resulting in more natural interpretation. In addition, the integration of reinforcement learning systems into Yandex Live Multimedia's AI algorithms allows the platform to improve its ability to adapt in real time to new dynamic situations and contexts (Tao; Busso, 2021). This innovative approach is redefining the way intercultural communication challenges are addressed today, while helping to position Yandex as a leader in the field of CAI and setting new standards in the quality and accuracy of machine translation services.

However, despite advances in the field aimed at overcoming existing barriers to automated interpreting, there still appear to be significant challenges to the accuracy and fluency of real-time text interpretation. Phraseology, understood as the study of language-specific idiomatic expressions endowed with meanings, is an elusive area in the context of CAI due to its complexity and variability. In the interpreting process, phraseology poses a challenge to AI systems, as literal translations of idiomatic expressions may not convey the correct meaning in the target language. The variability of idiomatic expressions and the presence of regionalisms and jargon make it difficult for CAI systems to accurately recognise and translate the implied meaning of these fixed constructions. The lack of standardisation of phraseology across languages and the constant evolution of colloquial expressions also make it difficult to develop automated interpreting systems that can accurately handle these linguistic aspects.

3 Methodology

In this paper, a technology developed by Yandex in 2021 that translates a live stream from English, Spanish, French, Italian, German, or Chinese into Russian has been used. The translation of a live stream is a challenging task that has been addressed by the development of a novel technique based on

neural networks. Our study is devoted to the evaluation of this new technique and focuses on its ability to preserve the subtleties of semantic meaning and cultural connotations inherent in phraseological expressions in different linguistic contexts. Although our study does not address the technical details underlying this tool, we address these issues as they may help to inform translation. This technology incorporates advanced machine learning algorithms to enable the instantaneous translation of language during audiovisual broadcasts. In essence, this algorithm comprises five fundamental steps, each executed by a distinct neural network model.

Initially, the audio stream is captured and transcribed into plain text using automatic speech recognition. The video may contain extraneous sounds such as noise and music, people may speak with different accents, speeds and diction, and there may be many speakers, so the technology must ensure that context and coherence are maintained during the translation process. Therefore, the algorithm takes a sequence of audio chunks as input, extracts acoustic features, and passes them into the neural network. The neural network in turn produces a set of word sequences from which the language model selects the most plausible hypothesis.

Subsequently, a machine translation model is employed to translate the text into the desired target language. There are several problems here: if you translate word-by-word or phrase-by-phrase, the quality will suffer, and if you wait for a long pause to guarantee the end of a sentence, there will be a long delay. So, the technology groups words into sentences without losing meaning or making sentences too long. For correct translation at this stage, it is also necessary to determine the gender of the speaker to determine to whom a particular line belongs and to reproduce the voice correctly. After selecting individual sentences and lines, the translation is performed, for which Yandex uses its own translator.

Once the translation is completed, the translated text is processed through text-to-speech technology, which converts the written content into spoken audio. This step ensures that the generated speech sounds natural and coherent, considering various linguistic features such as tone, rhythm, and inflection. Additionally, the gender of the speaker, which was previously identified during the initial stages of the process, is incorporated into the synthesis to ensure the voice matches the intended speaker's profile. This level of customization helps improve the overall quality and authenticity of the synthesized speech, making it more relatable and contextually appropriate for the target audience.

Furthermore, the algorithm ensures that the translated speech is accurately synchronized with the corresponding segments of the video stream, aligning seamlessly with the visual content, and maintaining synchronization with the video frames. This process is crucial to ensure that the audio matches the timing of the speaker's lip movements and actions in the video. Additionally, the neural network addresses several challenges in this stage, such as when the speaker delivers a sentence rapidly, or when the translated sentence is significantly longer than the original. In these cases, the system must dynamically adjust the synthesized audio by compressing or shortening it to fit within the allotted time frame, ensuring smooth and natural speech flow that aligns with the visual context.

Finally, the translated speech is seamlessly integrated into the live video stream, replacing the original audio with the newly generated translated audio. This newly created audio is then encapsulated into an audio stream, which is embedded directly into the browser interface of the viewer, allowing them to hear the translated speech while watching the video in real-time. The technology used for this process is currently functional exclusively within the Yandex browser, which was the platform selected for the study. Figure 1 displays a detailed screenshot of the Yandex browser interface, clearly highlighting the key components involved in the translation and audio synchronization process. It provides a visual representation of how the translated audio is integrated into the video stream, showing the alignment between the original video content and the overlaid translated speech. The screenshot also highlights the user interface elements that facilitate the viewer's interaction with the translated video, such as volume controls, language options, and playback features.

To enhance the study's transparency and validity, it is crucial to provide comprehensive details regarding the sample selection, data collection methods, and the analytical techniques employed. The live video functionality within the interface is activated via a clearly visible and easily accessible button, allowing for straightforward interaction with the system. A notable feature of the translation

Figure 1. Screenshot of the Yandex browser interface when utilizing the automatic translation features.
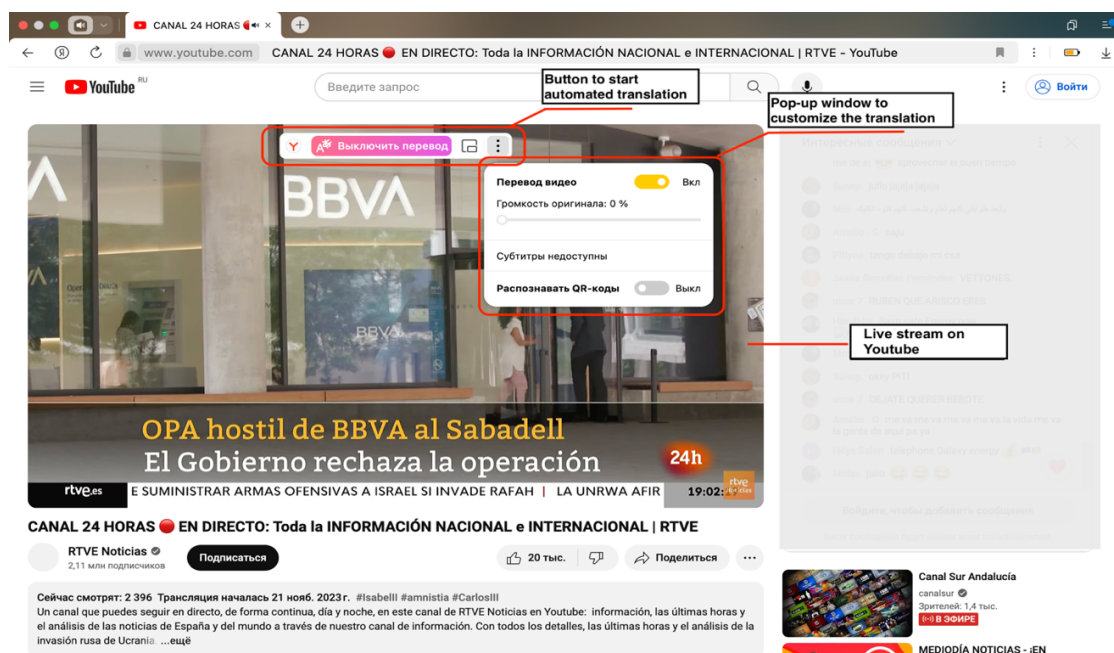
Source: Author's own work

process is the 40-50 second temporal delay between the original live video stream and its translated counterpart in Spanish. This delay, as derived from a detailed evaluation of the system's operation, is purposefully incorporated to allow sufficient time for the system to process the content contextually, which is essential for delivering an accurate translation in real-time, especially when dealing with live broadcasts. Such a delay also enables the system to manage the complexities inherent in real-time translation, such as adjusting for idiomatic expressions, speech nuances, and varying speaking speeds.

Moreover, the Yandex browser's functionality extends beyond mere passive translation by allowing users to actively customize their experience. This includes features like adjusting the volume of the original audio track, which could be crucial in environments where background noise or other factors might interfere with the audio quality. Additionally, the availability of subtitles provides an extra layer of understanding and flexibility for users, especially in cases where visual cues or context may not fully suffice to ensure a clear understanding of the translated content. These customizable features add a significant level of adaptability to the translation process, catering to various user preferences and enhancing the overall user experience. Such details are integral to understanding the technical infrastructure of the study and ensure its findings can be accurately interpreted and replicated in future research.

This study employed the Yandex browser and a range of accessible technological tools to explore the automated translation of live news streams. The analysis focused on two YouTube channels, "RTVE Noticias" and "Canal Sur Andalucía," both of which provide continuous news coverage. These channels were selected as the primary subjects of investigation due to their ongoing news broadcasts, providing a robust sample for examining the translation of live content. The methodology for the study, as outlined in Figure 2, was structured to facilitate a systematic comparison of the automated translation of phraseological expressions in real-time news streams.

The approach for comparing the automated translation of phraseological expressions involved several key stages:

1. *Audio recording:* The first step in the methodology involved recording both the original live stream (audio recording #1) and its corresponding translation (audio recording #2) simultaneously, ensuring that both recordings occurred at the same intervals during the live broadcast. Two separate devices were used to capture these audio recordings concurrently, ensuring precise alignment of the original and translated content.

2. *Detection of phraseologisms (verbal idioms):* In the second stage, instances of phraseologisms, commonly used idiomatic expressions, were identified in audio recording #1, which captured the live stream in the original language (Spanish). The identification process was carefully carried out to ensure that the expressions detected were contextually relevant and representative of the language used in the broadcast.

3. *Translation matching:* Once the phraseologisms were identified in the original audio, the next step involved matching the corresponding translations of these expressions in audio recording #2, which was in Russian. This step was crucial for establishing whether the automated translation system accurately rendered the idiomatic expressions in a culturally appropriate and linguistically accurate manner.

4. *Comparison and analysis:* The final stage entailed a detailed comparison of the accuracy and correctness of the translations. A comprehensive comparison table was developed to facilitate this process, allowing for a side-by-side evaluation of the original and translated expressions. The total duration of the audio recordings in both Spanish (the original language) and Russian (the translated language) amounted to 243 minutes for each language. These recordings were made across different intervals, ranging from 15 to 60 minutes, and were captured on multiple occasions throughout the day over four non-consecutive days. This sampling strategy was employed to ensure a diverse representation of news topics covered during the broadcasts. As part of the analysis, 52 verbal idioms were identified and thoroughly examined within the context of the live news stream, providing insights into the effectiveness of the automated translation system in handling phraseological expressions in a dynamic, real-time setting.

Figure 2. General outline of the methodology presented.
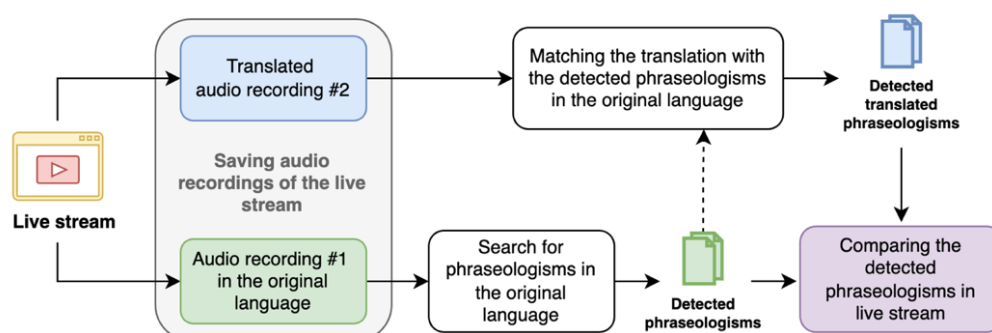
Source: Author's own work.

4 Results and discussion

As we have mentioned in this paper, AI uses various interpretation techniques to orally translate any kind of information, such as natural language processing, machine learning, speech recognition and semantic analysis. In this sense, AI not only analyses and understands the meaning of words and phrases in the context in which they are used, but also uses machine learning algorithms to continuously improve its ability to translate orally, as well as using speech recognition technologies to capture and transcribe in real time. Importantly, AI also relies on the use of neural networks and sequence processing models to improve translation accuracy. These architectures allow AI to understand the temporal and contextual relationships between words in a sentence, resulting in a more coherent and contextual translation.

In this case, Yandex used natural language processing to firstly analyze the text of the analyzed messages and secondly to extract relevant information related to phraseology. It used machine learning to improve speech recognition to transcribe and understand human speech, and semantic analysis to understand the meaning behind the verbal idioms displayed. When interpreting phraseological units,

especially idioms, several factors need to be considered. Among them are the ambiguity of the meaning of the idioms between literal and figurative meaning; the understanding of their cultural and linguistic context, since most of the proposed phraseologisms are immersed in Hispanic culture; the ability to recognize them as a single semantic unit; and the ability to translate them appropriately into different languages, in our case Russian. All of this is achieved through good use of contextual resources and good adaptation to natural conversation. Therefore, it is extremely important for AI to have access to contextual resources, such as linguistic databases or text corpora, which allow it to understand the use and meaning of verbal idioms in different contexts and communication situations. Moreover, idioms are often used in natural and colloquial conversational contexts, so AI must be able to interpret them in a fluent and natural way, maintaining their original tone and style.

Nevertheless, after the speech interpretation of the speeches taken from the news and the analysis of verbal idioms, it was observed that several factors influenced the way in which these expressions were interpreted. According to the data obtained, of the 52 units analyzed, only 20 were correctly translated by Yandex, i.e., less than half. This suggests that the ability of machine translation systems to interpret and translate idiomatic expressions remains a major challenge in the development of advanced translation technologies. The complexity and diversity of idiomatic expressions in different languages make their accurate and contextual translation a complicated process that requires further analysis and development of specialized translation algorithms.

The presence of ambiguity in the meaning of idiomatic expressions, consideration of the cultural context in which these expressions are used, and attention to the type of linguistic register were crucial aspects that affected the accuracy of the translations performed by Yandex. The ambiguity in the meaning of idiomatic expressions posed a significant challenge to AI, as the presence of multiple possible interpretations made it difficult to select the most appropriate translation based on the context in which these expressions were used. The lack of clarity in the meaning of idioms led to incorrect, literal translations or even the omission of part of the original message.

Furthermore, the cultural context in which the idiomatic expressions were rooted was critical to their accurate interpretation. Yandex had difficulty identifying and understanding the cultural connotations and meanings implicit in these expressions, resulting in translations that did not adequately reflect the original meaning of the idiomatic expressions. On the other hand, the nature of the linguistic register of the idioms also had an impact on the quality of the translations produced. The presence of colloquialisms for certain idioms, which had no direct equivalent in the target language, made it difficult to interpret these expressions correctly, resulting in incorrect or incomplete translations. The following are the most representative examples of translations of the idioms analyzed.

In the case of ambiguity, many of the verbal idioms analyzed, such as: "pisar fuerte" (to make a strong impression), "abrir puertas" (to open doors), "poner en pie" (to get on one´s feet), "abrir camino" (to pave the way) o "ajustar cuentas" (to settle scores), may have other meanings or interpretations, which undoubtedly makes it difficult to translate them accurately. In this context, AI in general, and Yandex in particular, may have difficulties in choosing the appropriate translation depending on the context in which the expression is used. In our case, Yandex had difficulties in determining the correct meaning of certain idioms based on the context in which they are used due to ambiguity, that led to inaccurate or incorrect translations. This was the case with the following idioms: "pegar ojo" (to get some sleep), "salir rana" (to turn out badly), "caérsele la casa encima" (to feel overwhelmed), "dar tela" (to give a hard time), "echar una mano" (to lend a hand), "chuparse los dedos" (to lick your fingers) or "tener mala leche" (to be in a bad mood). All these verbal idioms have been translated literally, resulting in confusing and incomprehensible translations in Russian. It is therefore important to take these potential ambiguities into account when using machine translation tools, and in some cases, it may be necessary to use human translators to achieve an accurate interpretation of idiomatic expressions.

In addition to linguistic and semantic aspects, the Russian translation of the idiomatic expressions was also influenced by the cultural context in which they are rooted. This is because many of the idiomatic expressions presented are unique to a particular culture and reflect values, beliefs, and traditions specific to that cultural context. Therefore, a literal translation of these expressions may

result in a loss of their original meaning and connotation. In our case, Yandex chose to translate some of these idioms in a more general or contextualized way, considering the underlying meaning they convey in the culture of origin, i.e., in Spanish. In this way, Yandex tried to preserve the essence and original intention of the expression, adapting it to the cultural context of the target language. The problem with this is that the two translations are not always equivalent or congruent in meaning. Some examples of unsuccessful translations by Yandex are: "estar en una nube" (to be on cloud nine), "meter salsa" (to spill the beans), "ser la guinda del pastel" (to be the icing on the cake), "dar con la tecla" (to hit the nail on the head) or "dar cosa" (to feel uneasy). Although the Russian translation of these idioms was not literal, Yandex did not consider the specific cultural context in which they are embedded.

Thus, "estar en una nube" was translated as быть на облаке ("to be in a cloud"), an expression used in Russian when a person dies and not, as in the Spanish case, when one is excited or distracted by something positive that has happened. In the case of meter salsa, this expression was translated as водить фальшивки ("to cheat"). The Spanish context was that of a person who interferes in other people's lives to criticize them, while the Russian translation as a person who cheats in the game does not make sense in the given context. The following expression is curious in that its Russian counterpart exists in its literal form вишенка на торте ("to be the icing on the cake"). However, Yandex chose to change one of the components of the idiom, translating it as глазурь на торте (to be the frosting of the cake), which is understandable but not idiomatic. As for the expression "dar con la Tecla", which is usually used as a synonym for "dar en el blanco" (to hit the target), it was translated into Russian as опредеделить деталь (to determine in detail), while the expression "me da cosa" in the sense of "me da apuro" (I feel embarrassed) has been translated into Russian as это много мне дает (it gives me a lot) due to lack of context.

In terms of linguistic register, it is important to note that some idiomatic phrases contain colloquialisms that are specific to a particular register or level of linguistic formality. These elements also posed an additional challenge for machine translation, as they may not have a direct equivalent in the target language or may be incomprehensible to native speakers of the target language. In these cases, Yandex faced the difficulty of deciding how to approach the translation of these expressions, choosing to omit the expression in question if it could not find a suitable equivalence or, on the contrary, to look for an approximate equivalence that preserved the general meaning of the original expression. Examples of omissions include: "echarle morro" (to be cheeky), "meterse en un fregao" (to get into a mess), "estar al pie del cañón" (to hold the fort) and "meter la pata" (to mes up). In these cases, Yandex did not recognize their specific meaning or context in Spanish because they are very colloquial idiomatic expressions that do not have a direct translation in Russian, so the tool decided to omit them directly to avoid possible errors of interpretation. Regarding the search for an approximate equivalence in order to preserve the general meaning of the original expression, the following examples of idioms could be given: "quitarle hierro a un asunto" – to defuse a situation (снять напряжение – quitar presión – to relieve pressure), "meterse en un lío" – to ge tinto trouble (втянуть в беспорядок – arrastrarse al desorden – to descend into chaos), "estar en auge" – to be on the rise (процветать – florecer – to prosper), "tener entre algodones a alguien" - to treat someone with kid gloves (держать в объятиях – mantener en abrazos - to envelop someone in hugs), "plantar cara" – to face up (противостоять – oponerse – to oppose) o "estar en un pozo sin fondo" - to be in a never-ending hole (быть в яме – estar en un hoyo – to be in a hole). All the above examples perfectly preserve the metaphorical meaning of the source text, and their translations, even though in many cases they are not phraseologisms, fit perfectly into Russian speech.

It is also interesting and positive to analyze those verbal idioms that Yandex translated correctly. This was the case with the expressions: "tener en el bolsillo" – to have somene under control (есть в кармане), "dar alas" – to inspire (дать крылья), "hacerse la boca agua" - to make one's mouth water (слюнки текут), "tocar madera" - to knock on wood (постучать по дереву) у "no levantar cabeza" - to be overwhelmed (не поднимать головы). In these cases, Yandex may have recognized the contextual meaning of the expressions and provided an appropriate translation based on their Russian equivalents. This shows that in some cases machine translators may be able to correctly

translate idiomatic expressions if they have a large and up-to-date database that allows them to recognize and understand the meaning of such expressions in different languages. It is also noticeable that most of the correctly translated expressions are idioms whose counterparts exist in the Russian language. This is a clear indication that both Yandex and other machine translators still process a certain bias towards literal language, showing more correct translations only in cases where there are full equivalences in both languages.

The findings reveal that CAI systems like Yandex face considerable challenges in accurately interpreting idiomatic expressions and culturally specific language. This observation aligns with theoretical perspectives emphasizing the limitations of current natural language processing systems in capturing the contextual and pragmatic dimensions of language. Such limitations reinforce the theoretical argument that AI-driven systems lack the nuanced understanding required to manage linguistic and cultural subtleties, a domain where human interpreters excel.

Additionally, the study highlights a critical gap in the scalability of CAI tools for real-time applications, particularly in scenarios that demand high fidelity in the translation of phraseological units. Existing theories often posit that AI can significantly augment human capabilities in real-time interpretation. However, the study's findings challenge this assumption by demonstrating that the technology struggles with temporal constraints and the need for cultural and contextual alignment. This discrepancy underscores the necessity for advancements in natural language processing algorithms that incorporate larger, more diverse training datasets, particularly those focused on colloquial and idiomatic language.

Furthermore, the study advances the theoretical discourse by illustrating the importance of hybrid approaches that integrate human expertise with AI capabilities. While CAI tools can enhance efficiency and accessibility, their limitations necessitate human oversight to address ambiguities, cultural nuances, and the interpretive depth required for idiomatic language. This reinforces theories advocating for a collaborative framework in which AI serves as an assistive, rather than autonomous, tool.

The study's results highlight that while CAI tools like Yandex offer notable advantages in terms of efficiency and accessibility, they fall short in handling the complexities of idiomatic expressions and culturally nuanced language. This limitation underscores the theoretical argument that AI systems, even with advanced natural language processing capabilities, currently lack the cognitive and contextual understanding necessary to navigate the intricacies of phraseological translation effectively. These insights align with established theories suggesting that the interpretive depth required for accurate and culturally appropriate translations remains beyond the scope of autonomous systems.

Moreover, the study challenges theories advocating for full automation in language interpretation, revealing that the nuanced decision-making inherent to human interpreters is indispensable for managing linguistic ambiguity and cultural specificity. It supports the proposition that hybrid models, where human judgment complements AI''s speed and processing capacity, offer the most effective framework for real-world applications of CAI. This collaborative approach is particularly crucial in high-stakes scenarios, such as live news translation, where real-time accuracy and cultural resonance are imperative.

5 Conclusion

Overall, despite advances in AI and machine learning, there are still limitations in the ability of machine translation systems to understand and faithfully translate idiomatic expressions. This underlines the importance of and need for human intervention and linguistic knowledge in the translation of complex and culturally rich texts. Therefore, further research and improvement of machine translation systems is needed to deal more effectively with the translation of idiomatic expressions and to improve the quality and accuracy of the interpretations generated. It is also worth noting that advances in the field of AI have enabled the development of machine translation systems that incorporate more sophisticated techniques, such as the use of neural translation models based on deep neural networks. These models could capture the full context of a sentence more effectively and to consider the syntax and semantics of words in context, resulting in more accurate and natural translations of idiomatic expressions.

However, AI, and CAI developed by Yandex in particular, faces challenges such as natural language processing, understanding the context and semantics of utterances, and a lack of sufficient data to train translation models.

The findings of this study carry considerable practical implications for professionals in the fields of translation, interpretation, and education. For translators and interpreters, particularly those engaged with idiomatic expressions and culturally specific language, the study underscores key challenges, such as the limitations of AI and CAI systems in grasping context and semantics, that highlight the continued necessity of human expertise to ensure both linguistic accuracy and cultural appropriateness in translations. While AI-driven tools like Yandex's CAI can enhance operational efficiency in certain scenarios, they are not yet capable of replicating the nuanced comprehension and interpretative decisions that skilled human translators and interpreters can provide. As such, practitioners should view CAI systems as complementary tools, leveraging their capabilities to support the translation process, but not as standalone solutions. Human expertise remains essential for navigating ambiguities, managing cultural subtleties, and accurately interpreting idiomatic expressions, which continue to present significant challenges for current machine translation systems.

Moreover, for educators in the fields of translation and interpretation, the study's insights offer valuable guidance for curriculum development. Emphasizing the limitations and potential of CAI systems in interpreting complex language features like idioms can help future translators and interpreters become more proficient in utilizing technology effectively. Educators could integrate hands-on training with machine translation tools, while also teaching students how to navigate the cultural subtleties and contexts that technology may miss. Understanding the need for continuous advancement in AI systems could be woven into training, preparing students to bridge the gap between human understanding and technological capabilities.

References

ALCAIDE MARTÍNEZ, M. Crítica de Postigo Pinazo (2020): La interpretación en un mundo cambiante: nuevos escenarios, tecnologías, retos formativos y grupos vulnerables. *Babel*, John Benjamins, v. 67, n. 2, p. 245–248, 2021. ISSN 0521-9744. DOI: 10.1075/babel.00216.mar. Available from: https://www.jbe-platform.com/content/journals/10.1075/babel.00216.mar.

ALOTAIBI, Hind M. Computer-Assisted Translation Tools: An Evaluation of Their Usability among Arab Translators. *Applied Sciences*, v. 10, n. 18, p. 62–95, 2020.

ARNÁIZ, Marta Saracho. *La fraseología del español*: Una propuesta de didactización para la clase de ELE basada en los somatismos. 2015. Doctoral dissertation – Departamento de Lengua Española, Facultad de Filología, Universidad de Santiago de Compostela, Santiago de Compostela.

CIFUENTES-FÉREZ, P. Herramientas en línea para la traducción de UF: diccionarios monolingües y bilingües. *In:* TARRÍO, Germán Conde; HUERTA, Pedro Mogorrón; GARCÍA-SECO, David Prieto (eds.). *Enfoques actuales para la traducción fraseológica y paremiológica: ámbitos, recursos y modalidades.* [*S. l.*]: Centro Virtual Cervantes, 2015. p. 97–110.

CORPAS PASTOR, G. Detección, descripción y contraste de las unidades fraseológicas mediante tecnologías lingüísticas. *In:* MORENO, Inés Olza; RICHARD, Elvira Manero (eds.). *Fraseopragmática.* [*S. l.: s. n.*], 2013. p. 335–374.

CORPAS PASTOR, G. Language Technology for Interpreters: the VIP project. *Translating and the Computer*, v. 42, p. 36–48, 2020.

CORPAS PASTOR, G.; GABER, M. Extracción de fraseología para intérpretes a partir de corpus comparables compilados mediante reconocimiento automático del habla. *In:* CORPAS PASTOR, Gloria; BAUTISTA ZAMBRANA, María Rosario; HIDALGO TERNERO, Carlos Manuel (eds.). *Sistemas fraseológicos en contraste: enfoques computacionales y de corpus.* [*S. l.: s. n.*], 2021. p. 271–289.

CORPAS PASTOR, G.; RUBIO, E. G. Fraseología computacional y de corpus aplicadas al español. *Romanica Olomucensia*, v. 1, p. 1–8, 2023.

CORPAS PASTOR, Gloria. Interpreting Tomorrow? How to Build a Computer-Assisted Glossary of Phraseological Units: in (Almost) No Time. *In:* CORPAS PASTOR, Gloria; MITKOV, Ruslan (eds.). *Computational and Corpus-Based Phraseology*. Cham: Springer International Publishing, 2022. p. 62–77. ISBN 978-3-031-15925-1.

DEFRANCQ, B.; FANTINUOLI, C. Automatic speech recognition in the booth: Assessment of system performance, interpreters' performances and interactions in the context of numbers. *Target*, v. 33, n. 1, p. 73–102, 2021.

ERBSEN, Heidi; PÕLDRE, Siim. Is all Russian news the same? Framing in Russian news media generated by the Yandex news algorithm for the United States, Estonia, and Russia. *Journalism*, v. 24, n. 8, p. 1789–1816, 2023. DOI: 10.1177/14648849211069237.

FANTINUOLI, C. Computer-assisted interpreting: Challenges and future perspectives. *In:* TRENDS in E-tools and resources for translators and interpreters. [*S. l.: s. n.*], 2017. p. 153–174.

FANTINUOLI, C. Computer-assisted preparation in conference interpreting. *Translation & Interpreting: The International Journal of Translation and Interpreting Research*, v. 9, n. 2, p. 24–37, 2017.

GABER, M. Cómo dominar la fraseología y automatizar el proceso de documentación: una solución tecnológica para la formación de intérpretes en la combinación español<> árabe. *Romanica Olomucensia*, v. 35, n. 1, p. 55–70, 2023.

GABER, M. *Soluciones tecnológicas para intérpretes:* reconocimiento automático del habla e interpretación remota. 2023. PhD thesis – Universidad de Málaga. Technological solutions for interpreters: automatic speech recognition and remote interpreting.

GUO, M.; HAN, L.; ANACLETO, M. T. Computer-Assisted Interpreting Tools: Status Quo and Future Trends. *Theory and Practice in Language Studies*, v. 13, n. 1, p. 89–99, 2023.

GUTIÉRREZ ARTACHO, J.; OLVERA LOBO, M. D.; HUNT GÓMEZ, C. I. Incorporación de las nuevas tecnologías de la innovación y un modelo didáctico adaptado a la formación en Traducción e Interpretación. *In:* MENESES, Eloy López; SANCHIZ, David Cobos; PADILLA, Antonio Hilario Martín; GARCÍA, Laura Molina; MARTÍNEZ, Alicia Jaén (eds.). *III Congreso Internacional Virtual innovación pedagógica y praxis educativa: INNOVAGOGÍA 2016*. [*S. l.: s. n.*], 2016. p. 1030–1039.

JIBREEL, I. Online Machine Translation Efficiency in Translating Fixed Expressions Between English and Arabic: Proverbs as a Case-in-Point. *Theory and Practice in Language Studies*, v. 13, n. 5, p. 1148–1158, 2023.

KIM, Jong-Bae; KWEON, Hye-Jeong. The Analysis on Commercial and Open Source Software Speech Recognition Technology. *In:* LEE, Roger (ed.). *Computational Science/Intelligence and Applied Informatics*. [*S. l.: s. n.*], 2020. p. 1–15.

KOPONEN, M.; NUNES VIEIRA, L.; SPINOL, N. Introduction to the Dossier Issue Studying human-computer interaction in translation and interpreting: software and applications. *Tradumàtica*, n. 19, p. 66–74, 2021.

LI, C. A. A Comparative Analysis of Korean and Chinese Studies on the Technological Turn in Interpreting. *The Journal of Translation Studies*, v. 22, n. 3, p. 221–253, 2021.

MELLINGER, Christopher D. Tecnologías de interpretación asistida por ordenador y cognición de intérpretes: una perspectiva orientada hacia el producto y el proceso. *Revista Tradumàtica: tecnologies de la traducció*, n. 17, p. 33–44, 2019.

MEZCUA, A. R. Competencia digital y TICs en interpretación: «renovarse o morir». *Edmetic*, v. 8, n. 1, p. 55–71, 2019.

MOGORRÓN HUERTA, P. Explotación informática de una base de datos multilingüe de unidades fraseológicas. *In:* REY, María Isabel González (ed.). *Unidades fraseológicas y TIC*. [*S. l.*]: Instituto Cervantes, 2012.

NOVOZHILOVA, A. A.; KOROLKOVA, S. A.; MITYAGINA, V. A.; NAUMOVA, A. P. SMART Technologies in Translation: Globalization as a Factor of Progress. *In:* POPKOVA, E. G.; SERGI, B. S. (eds.). *"Smart Technologies" for Society, State and Economy*. Cham: Springer International Publishing, 2020. p. 228–237.

OLALLA-SOLER, C.; VERT BOLAÑOS, O. Traducción y tecnología: herramientas del proceso traductor como actividad profesional. El punto de vista de los estudiantes. *Tradumàtica*, n. 13, p. 623–641, 2015.

ORTIGOZA, M. S.; MORILLO-MONTOYA, R.; MONPUÉ, G. Desentrañando el lenguaje: impacto de la PNL en la era de la inteligencia artificial. *Revista Científica Saperes Universitas*, v. 7, n. 1, p. 30–47, 2024.

PÉREZ, Á. G.; RODRÍGUEZ, A. M. P.; MERINO, C. S.; SÁEZ, F. T. Tecnología para la Enseñanza y el Aprendizaje de Lenguas Extranjeras: La Enseñanza de Lenguas Asistida por Ordenador. Pasado, presente y futuro. *Pragmalingüística*, n. 28, p. 238–254, 2020.

PICCIONI, S.; PONTRANDOLFO, G. Competencia traductora y recursos informáticos: Por qué las tecnologías no sustituyen la formación en traducción. *Revista de Lingüística y Lenguas Aplicadas*, v. 12, p. 87–101, 2017.

RAMÍREZ RODRÍGUEZ, P. Fraseologismos en el aula de ELE: La problemática de la traducción automática. *Revista Tradumàtica: Traducció i Tecnologies de la Informació i la Comunicació*, n. 20, p. 77–95, 2022.

RAMÍREZ RODRÍGUEZ, P. La fraseología a partir de las TIC: Propuesta metodológica en el aula de RLE. *Archivum: Revista de la Facultad de Filosofía y Letras*, n. 73, p. 359–384, 2023.

RAMÍREZ RODRÍGUEZ, P. La TAV como recurso didáctico digital: El caso de las paremias en ruso. *Transfer: Revista electrónica sobre Estudios de Traducción e Interculturalidad / e-Journal on Translation and Intercultural Studies*, v. 19, n. 1-2, p. 112–138, 2024.

ROCKWELL, G.; SINCLAIR, S. *Hermeneutica*: Computer-assisted interpretation in the humanities. [*S. l.*]: MIT Press, 2022.

SEVILLA MUÑOZ, M. Utilización de recursos en línea en la enseñanza/aprendizaje de traducción de unidades fraseológicas. *In:* REY, María Isabel González (ed.). *Unidades fraseológicas y TIC*. [*S. l.*]: Instituto Cervantes, 2012.

SHADIEV, R.; LIU, J. Review of research on applications of speech recognition technology to assist language learning. *ReCALL*, v. 35, n. 1, p. 74–88, 2023.

TAO, F.; BUSSO, C. End-to-End Audiovisual Speech Recognition System With Multitask Learning. *IEEE Transactions on Multimedia*, v. 23, p. 1–11, 2021. DOI: 10.1109/tmm.2020.2975922.

VALERO-GARCÉS, C. Análisis de aperturas en interpretación telefónica: estudio de caso. *Cadernos de Tradução*, v. 44, p. 1–19, 2024.

Author contributions