PHRASEOLOGICAL EVALUATION OF AUTOMATIC INTERPRETING ASSISTED BY YANDEX: STUDY AND ANALYSIS OF VERBAL IDIOMS

Pablo Ramírez Rodríguez[*]
Universidad de Granada, Spain

ABSTRACT**:** Verbal idioms pose a significant challenge for automatic translation and interpreting systems due to their idiomatic nature and contextual variability. In this context, and with the aim of analyzing the effectiveness of automatic interpreting of phraseology using the integrated Yandex corpus, this research seeks to evaluate the quality of automatic interpreting through examples of verbal idioms extracted from Canal Sur Noticias 24 horas on YouTube using a voice recognition and interpreting application provided by Yandex. To conduct the study, 50 frequently used verbal idioms were selected from Spanish news and their automatic interpreting into Russian were analyzed. Finally, the results were compared in both languages to determine the accuracy and naturalness of the resulting automatic interpreting.
KEYWORDS**:** Automatic Interpreting; Phraseology; Verbal Idioms; Interpreting Evaluation

1. Introduction

The analysis of human interpreting quality is a complex and multidisciplinary field of research, encompassing areas such as linguistics, translation studies, psychology, and sociology (García Becerra, 2013). This analysis aims not only to establish standards and quality criteria but also to enhance interpreting practices through continuous training and the development of new technologies, including AI-assisted tools and voice recognition software (Han & Lu, 2023). It also examines the differences between simultaneous and consecutive interpreting, as well as variations in specialized contexts such as medical, legal, and technical settings. Recent studies have incorporated both quantitative and qualitative methods to evaluate accuracy, fluency, and naturalness in interpreting, providing a comprehensive and detailed view of the essential competencies and skills for effective interpreting (Collados Aís, 2016; Christen & Gracia, 2017; Kolchugina, 2021).

Throughout history, theoretical and methodological studies on interpreting quality have laid a solid foundation for contemporary research. One of the earliest approaches to evaluating interpreting is error analysis, which focuses on identifying and classifying errors during interpreting. This method is used to assess accuracy and fidelity by considering errors such as omissions, additions, and distortions. However, while effective in detecting specific issues, its fragmented approach may not capture essential aspects of overall interpreting, such as pragmatic adequacy. Additionally, this method can be subjective, as the identification of errors often depends on the evaluator's perspective and criteria (Pradas Macías, 2003).

The impact of contextual factors on interpreting is crucial for understanding the quality of the process. These factors include both cultural and situational elements, which can significantly influence the execution and perception of interpreting. Cultural

[*] ramires-rodriges-p@rudn.ru

differences notably affect the interpreting of phraseological units, which require a deep understanding of the cultures involved. According to Corpas Pastor's classification (1997), phraseological units are divided into collocations, idioms, and phraseological statements, and they present a challenge due to their fixed and sometimes idiomatic nature, as is the case with idiomatic expressions. These expressions have meanings that cannot be easily inferred from the sum of their components, necessitating detailed cultural and contextual understanding (Ruiz Gurillo, 1997). Cultural differences can greatly alter the interpreting of these phraseological units, as their meanings can vary significantly between different languages and cultures (Mogorrón Huerta, 2020).

In this context, we believe that phraseology in general, and verbal idioms in particular, represent a largely unexplored quality parameter due to their linguistic complexity and specificity. Research in this field demonstrates how interpreters manage not only conventional linguistic structures but also deeply rooted idiomatic expressions found in everyday uses of a language (Ramírez Rodríguez, 2024). Verbal idioms possess not only idiomatic meanings but also cultural and pragmatic connotations that can vary depending on the context. Additionally, studying the interpreting of verbal idioms provides an opportunity to better understand how semantic and pragmatic nuances are conveyed in different linguistic and cultural contexts, thereby enhancing the effectiveness and accuracy of interpreting in various communicative situations (Tormo, 2021).

The evolution of technology has led to the incorporation of computational tools in simultaneous interpreting (Corpas Pastor, 2024) and statistical analyses to complement human evaluation, improving the objectivity and reliability of interpreting assessments (Defrancq et al., 2024). Today, automatic tools utilize AI algorithms and models to analyze and compare translations, providing quick and objective assessments (Ramírez Rodríguez, 2022). Additionally, combining automated evaluations with human review has become a common practice, where AI conducts a preliminary analysis that is then reviewed and adjusted by human experts to ensure accuracy and cultural appropriateness (Prandi, 2023).

The motivation driving this study lies in the critical importance of phraseology and verbal expressions for accurate and faithful interpreting of the original discourse, particularly within the context of media and AI. Verbal expressions, which include idioms and specific technical terms, not only convey meanings deeply rooted in cultural contexts but also influence how information is perceived and interpreted by the audience.

On the other hand, advances in AI within the field of interpreting present both significant challenges and opportunities concerning the handling of phraseology and verbal expressions. The ability of AI systems to understand and accurately convey these complex linguistic structures in real time is crucial for their effective integration into media environments, where precision and coherence are essential. Additionally, the study considers how advances in language models are enhancing AI's capacity to interpret idiomatic expressions through contextual learning and deep semantic analysis.

2. The role of AI in evaluating phraseological interpreting

In recent years, two events have significantly altered the field of interpreting and the interaction of interpreters with technology: the COVID-19 pandemic and the introduction of generative AI (Wang & Fantinuoli, 2024). Prior to the pandemic, remote communication tools were gradually gaining popularity, with remote interpreting considered more of a future vision than a current trend. However, the global pandemic made the use of remote communication technologies an urgent necessity, leading to a surge in demand for remote interpreting services. Even though the pandemic has now subsided, remote communication tools continue to be widely used because virtual and hybrid meetings and events remain common. By the end of 2022, coinciding with the end of the COVID-19 crisis, generative AI emerged, transforming various professions, including interpreting (Dong & Zhao, 2019). This technology employs advanced natural language processing (NLP) methods to create conversational agents capable of performing various text-related tasks. These developments have had a significant impact on interpreters, software developers, and researchers in the field, exploring different perspectives on interpreting technologies and presenting a promising future for both interpreters and innovators in this sector.

Since the inception of translation studies focusing on how digital technologies and computer tools impact and enhance the translation process and multilingual communication, efforts have been made to optimize the management of corpora and reference resources (Pöchhacker, 2024). Current tools primarily extract multilingual concordances from parallel texts, creating frequency-ordered term lists paired with their equivalents. In the field of interpreting, the adoption of specific tools has been slower and less widespread due to concerns about their reliability and potential distractions (Giustini, 2023). In this context, machine interpreting (MI) can be seen as an extension of machine translation, utilizing speech recognition for input and synthetic speech for output. It involves software that decodes and transcribes spoken language, generating an automatic translation and reproducing it through a speech synthesizer (Gaber, 2023). Currently, most companies and institutions developing MI projects use neural machine translation engines. However, one of the main challenges for MI is its anticipatory capacity, especially when dealing with less common languages (Fantinuoli, 2016, 2017; Ramírez Rodríguez, 2023).

In general, the primary goal of quality assessment has been and continues to be the development of accurate and reliable methods to measure the effectiveness of interpreting in terms of fidelity, coherence, and cultural appropriateness. Over the decades, various models and theoretical approaches have been proposed to address this challenge, each with its own advantages and limitations (Defrancq & Fantinuoli, 2021). These approaches include both quantitative and qualitative assessments, combining detailed linguistic analyses with advanced technological tools such as natural language processing (NLP) algorithms and automated evaluation systems. Additionally, research has considered contextual and psychological factors affecting interpreting, such as stress and mental fatigue, as well as the impact of bilingual proficiency and interpreter experience on performance quality.

Recent scientific advancements have further enriched this field by incorporating insights from cognitive psychology and neurolinguistics. Studies have utilized neuroimaging techniques like fMRI and EEG to explore the neural underpinnings of simultaneous interpreting, revealing the activation of specific brain regions associated with bilingual language processing and cognitive control (Hervais-Adelman et al., 2015). This has provided a more comprehensive understanding of the cognitive load and mental processes involved in interpreting. Moreover, experimental designs in psycholinguistics have been employed to investigate the effects of stress and mental fatigue on interpreter performance. For example, research has demonstrated that high-stress conditions can significantly impair working memory capacity and multitasking abilities, leading to reduced accuracy and increased errors in interpreting (Gile, 2009).

The integration of big data analytics has also allowed for the examination of large corpora of interpreted texts, facilitating the identification of patterns and trends in interpreter behavior and translation choices across different languages and contexts. These methodologies enable a more robust and nuanced analysis of interpreter performance, contributing to the development of more sophisticated evaluation metrics and training programs. By leveraging these interdisciplinary approaches, researchers have been able to provide a deeper understanding of the multifaceted nature of interpreting. This comprehensive approach has identified specific challenges faced by interpreters in different contexts and communicative situations, thereby informing the development of targeted strategies to enhance interpreting quality and efficacy.

Phraseology, understood as the set of fixed and occasionally idiomatic expressions unique to a language, plays a crucial role as an intraparameter in the accurate transmission of the source discourse and the terminology used. In this context, the ability of AI to accurately process and manage specific phraseology and specialized terminology plays a crucial role in determining the quality of interpreting when compared to human interpreters. Unlike humans, AI systems rely on pre-trained models and large databases to recognize and reproduce fixed expressions, technical terms, and idiomatic phrases within a given domain. While this allows for consistent and precise handling of terminology in specialized fields such as medicine, law, or engineering, it may lack the flexibility and contextual awareness that human interpreters naturally possess.

Human interpreters, on the other hand, excel at adapting to nuanced linguistic and cultural contexts, detecting speaker intent, and managing ambiguous or context-dependent expressions. They can dynamically adjust their output based on non-verbal cues, tone, or situational changes elements that AI systems may struggle to fully interpret.

Several advances in NLP and machine learning have significantly improved the capacity of AI systems to recognize and translate phraseological expressions. However, challenges remain due to the inherent complexity and variability of idiomatic language. Research has shown that phraseology poses a unique challenge to AI due to its contextual dependence and cultural specificity (Sag et al., 2002). One scientific approach to addressing this issue involves the development of large, annotated corpora that include a wide range

of idiomatic expressions. These corpora serve as training data for machine learning models, enabling them to better understand and generate appropriate translations for phraseological units (Ramisch, 2015). Additionally, advanced techniques such as deep learning and transformer models like BERT and GPT have shown promise in capturing the nuances of idiomatic language by leveraging context-aware embeddings and transfer learning (Vaswani et al., 2017).

To evaluate the performance of AI in handling phraseology, researchers employ both intrinsic and extrinsic evaluation methods. Intrinsic methods focus on the linguistic accuracy of translations, assessing aspects such as semantic fidelity, syntactic appropriateness, and contextual relevance. Extrinsic methods, on the other hand, examine the practical impact of translations in real-world tasks, such as user comprehension and task success rates in multilingual communication settings (Manning, 2008). Moreover, psycholinguistic studies have explored how human interpreters process phraseological expressions, shedding light on cognitive strategies employed during interpreting. These insights have informed the design of AI systems by highlighting the importance of context and the need for adaptive translation strategies that can dynamically adjust to different communicative scenarios (Colson, 2005).

By integrating these scientific perspectives, the study aims to provide a comprehensive analysis of the effectiveness of AI in interpreting phraseology compared to human interpreters. This includes examining factors such as accuracy, fluency, and cultural appropriateness, and identifying specific areas where AI systems may require further enhancement. Through this investigation, the study contributes to the broader understanding of the potential and limitations of AI in the field of interpreting, offering valuable insights for future research and development in NLP and computational linguistics.

3. Methodology

The objective of this study is to analyze and evaluate the quality of automatic interpreting using examples of verbal idioms extracted from the Canal Sur Noticias 24 Horas YouTube channel, performed by Yandex's AI. To achieve this goal, 50 verbal idioms were selected and analyzed, using Yandex's intelligence service corpus as the primary method. For this analysis, a bilingual corpus was used consisting of Spanish audio transcriptions and their corresponding automatic interpreting into Russian, generated by Yandex's voice recognition system. This approach allows for the assessment of verbal idioms interpreting within a colloquial spoken language context, which may present challenges different from written language. The audios were transcribed using automated speech recognition (ASR) tool provided by Yandex and then manually reviewed to ensure accuracy. Non-verbal elements, such as facial expressions, gestures, and tone of voice were removed, and transcription errors corrected. Subsequently, the transcriptions were segmented into phrases and aligned with their automatic interpreting generated by Yandex.

To ensure a representative sample of verbal idioms, Spanish texts were selected from various sources on the Canal Sur Noticias 24 Horas YouTube channel. Two main types of

content were chosen to guarantee the necessary variety and richness for this study: live news broadcasts and interviews. Live news broadcasts provide a rich source of natural and spontaneous language. In this context, live broadcasts covering a wide range of topics, including politics, economics, culture, and sports, were transcribed. These transcriptions offer a varied context where verbal idioms may naturally appear in different language registers. Interviews with various personalities provide dialogues rich in idiomatic expressions and verbal idioms. These transcriptions allow for observation of how verbal idioms are used in more informal and often contextual conversations. The combination of these two types of content enables a thorough and representative collection of verbal idioms, covering both the formal and structured language of live news and the more colloquial and expressive language of interviews. This diversity of sources ensures that the resulting ad hoc corpus reflects a broad range of uses and contexts, providing a solid basis for linguistic analysis and the study of contemporary Spanish verbal idioms. Additionally, random sampling was applied within the selected content types to avoid over-representation of certain topics or expressions. This random sampling allowed for equitable coverage of various themes within the corpus.

In this study, a technology developed by Yandex was used to translate a live broadcast from Spanish into Russian, specifically focusing on certain verbal idioms. Translating a live broadcast presents a significant challenge due to the need to process and translate information in real-time, which has been addressed through the development of an advanced technique based on neural networks. This technique relies on deep neural networks, which have proven effective in capturing complex contexts and generating accurate translations in machine translation applications. In this context, our study focuses on evaluating this new technique and its ability to preserve the nuances of semantic meaning and cultural connotations inherent in phraseological expressions across different linguistic contexts. Verbal idioms, as key components of phraseology, present challenges for translation due to their idiomatic meanings and context dependence. The evaluation is based on analyzing how the system handles these expressions in terms of semantic fidelity and contextual appropriateness, with the aim of identifying strengths and limitations in preserving linguistic and cultural nuances. Although our study does not delve into the specific technical aspects of this tool, such as the architecture of the neural networks or the training algorithms used, these aspects are considered due to their potential impact on translation quality. Neural network-based machine translation models require extensive training with multilingual data and the integration of attention mechanisms to manage the complexity of phraseological expressions. Understanding these technical elements can provide valuable insights for the optimization and continuous improvement of the translation tool.

In the initial phase, the system captures the audio stream and transcribes it into plain text using ASR technology provided by Yandex. This process faces several challenges due to potential unwanted sounds, such as background noise and music, and the variability in individuals' speaking styles, including different accents, speeds, and diction styles.

Additionally, the presence of multiple speakers further complicates the task. To address these challenges, the technology must ensure that the context and coherence of the speech are maintained during translation. The algorithm processes the input as sequences of audio fragments, extracting relevant acoustic features such as intonation, rhythm, and phonetic characteristics. These acoustic features are fed into a deep neural network, which has been trained to identify patterns and relationships in the language. The neural network then generates a set of word sequences that the language model uses to select the most coherent and plausible interpreting. This stage is crucial to ensure that the transcribed text accurately reflects the original audio content, considering both semantic accuracy and speech fluency.

The next phase involves using a machine translation model to convert the transcribed text into the target language. Several challenges arise during this phase: translating literally, word by word or phrase by phrase, may compromise quality, and waiting for a long pause to ensure the end of a sentence can lead to significant delays. To address these issues, the technology is designed to group words into complete sentences while maintaining the original meaning, avoiding excessively long or fragmented sentences. A critical aspect of this stage is speaker identification, which is essential for accurately assigning dialogue lines and reproducing the voice appropriately. This identification process allows the translation to be adapted to be consistent with the original speaker's identity and style. Once individual sentences and lines have been selected and contextualized, translation proceeds using Yandex's translator, which employs advanced NLP techniques and neural translation models to ensure accurate and fluent conversion of text into the desired language. This technology also incorporates deep learning algorithms to continually improve translation quality based on real-time feedback and data analysis.

After translation, the text is converted into audio using text-to-speech technology, designed to produce natural and coherent sound. This technology adjusts the synthesis according to the speaker's gender identified in the previous stage, ensuring the generated voice is appropriate and consistent with the original speaker. The algorithm then synchronizes the translated speech with the corresponding segments of the live video, ensuring that the audio aligns correctly with the video frames. During this phase, the neural network also faces the challenge of managing discrepancies in speech speed; for instance, the speaker in the video may deliver a sentence at a very high speed, or the translated sentence may be twice as long as the original. To resolve this, the synthesized audio is temporally adjusted, compressed, or expanded as needed to match the video’s pacing. Finally, the translated audio is overlaid onto the live video feed as a new audio track. This audio is integrated into the overall video stream and embedded in the viewer's browser page, allowing real-time listening of the translation while viewing the content. This integration is crucial for ensuring that the translated audio perfectly synchronizes with the video, providing a smooth and coherent viewing experience for the end-user.

The methodology used for evaluating the machine translation of phraseological expressions in live broadcasts includes the following steps: 1. The audio of the original live

broadcast (audio recording #1) and the translated version (audio recording #2) are simultaneously recorded, ensuring that both recordings are captured in parallel and aligned with identical time intervals for precise comparison. In the study, two separate devices were used to capture these recordings. 2. Instances of phraseological expressions in audio recording #1 used in the live broadcast in the original language (Spanish, in this case) are identified and located. 3. These are compared with the translations of these phraseological expressions in audio recording #2 (in Russian). 4. The accuracy of the translations are analysed and assessed.

To ensure the robustness of the methodology, several scientific considerations were incorporated. Firstly, a rigorous sampling method was employed, where recordings were sampled at various times of the day. This approach was designed to account for different types of news content and linguistic variability, ensuring a comprehensive analysis. Secondly, high-fidelity recording devices were utilized to capture the audio accurately. This technical precision was crucial to minimize potential distortions that could affect the identification of phraseological expressions.

For the phraseological evaluation, NLP utilized a range of methods to process and interpret human language, which is fundamental for the analysis of verbal idioms in a corpus assisted by speech recognition. Among the techniques used is part-of-speech tagging, which assigns grammatical labels to each word in a text. This technique facilitated the identification and classification of verbal idioms by distinguishing grammatical components that form these phraseological units, such as verbs and their complements. Additionally, syntactic analysis provided a structural decomposition of sentences, allowing for a detailed understanding of how words are assembled into verbal idioms, contributing to a more accurate identification of these units within the corpus.

- **Semantic network analysis** focused on representing and visualizing the semantic relationships between words within verbal idioms. This technique involves constructing semantic networks where each node represents a word or expression, and connections between nodes illustrate semantic relationships. The visualization of these networks allowed for the identification of recurrent patterns and linguistic structures that characterize verbal idioms, providing a comprehensive perspective on how words combine to form phraseological units.
- **Speech recognition-assisted transcription analysis** was crucial for ensuring the accuracy of automatically generated transcriptions. This technique includes precision review, where automatic transcriptions are compared with the original audio to identify errors. Adaptation and correction of transcriptions are performed to improve accuracy by adjusting mis-transcribed words and ensuring that verbal idioms are represented accurately.
- **Transcript comparison** was employed to validate the accuracy of verbal idioms identified in speech recognition-generated transcriptions. This technique involves comparing manually automatic transcriptions of verbal idioms, allowing for the

identification of discrepancies and errors. Additionally, comparison with supplementary linguistic resources, such as dictionaries and phraseological databases, helps ensure the correct classification of verbal phrases, guaranteeing that the analysis is rigorous and precise.

- **Speech recognition quality assessment** analyzed how imperfections in speech recognition affect the identification of verbal idioms. This assessment considers the impact of transcription errors on the representation of phrases, such as recognition errors, omissions, and additions. Based on these analyses, adjustments can be made to speech recognition algorithms or correction methods to improve the overall quality of transcriptions and ensure a more accurate representation of verbal idioms in the corpus.

For the comparative analysis, the accuracy of translations was assessed based on a predefined set of criteria, encompassing semantic equivalence, contextual appropriateness, and fluency. These criteria aimed to ensure that the translated content preserved the original meaning, aligned with the intended context, and maintained natural readability in the target language. The comparative analysis framework utilized was carefully adapted from well-established translation evaluation models within the field of computational linguistics. This approach not only provided a structured and systematic methodology but also ensured that the evaluation process adhered to rigorous academic standards, making it both reliable and relevant to contemporary research practices. The audio recordings were digitized and stored manually in a structured database, facilitating efficient retrieval and analysis. Metadata such as time stamps, speaker identification, and context notes were meticulously documented to support the accuracy and reproducibility of the analysis.

4. Results and discussion

This study has outlined that AI employs various techniques for oral interpreting of information, including NLP, machine learning, speech recognition, and semantic analysis. Within this framework, Yandex's automatic interpreting AI proposed for this study not only analyzes and decomposes the meaning of words and phrases in their specific context but also applies machine learning algorithms to continually enhance its oral translation capabilities. This technology relies on speech recognition to capture and transcribe data in real time and uses neural networks along with sequence processing models to refine translation accuracy. Moreover, these structures enable the AI to understand temporal and contextual relationships between words in a sentence, resulting in more coherent and contextually appropriate translations.

From a corpus linguistics perspective, Yandex benefits significantly from analyzing large linguistic datasets to adjust and refine its translation models. The use of Yandex’s corpus for automatic interpreting of verbal idioms has provided a substantial empirical basis that allows the AI to identify patterns and trends in language use, thus facilitating

more precise and contextually adapted translations. Consequently, Yandex's corpus serves as a valuable tool for continuously evaluating and adjusting translation models, improving their adaptability to linguistic and contextual variations present in real data transcriptions. In the case of the phraseological interpreting of the analyzed verbal idioms, this capability to analyze extensive and varied data has been crucial for capturing nuances and ensuring that translations accurately reflect the meaning and context of phraseological expressions across different languages.

In this context, after analyzing Yandex's integrated corpus, the AI employed NLP techniques to conduct an initial analysis of the message texts studied and subsequently extract key information related to phraseology. The application of machine learning enhanced speech recognition for transcribing and understanding human speech, while semantic analysis focused on deciphering the underlying meaning of verbal idioms. When analyzing these phraseological units, it was crucial to consider various factors for a comprehensive interpretative evaluation. These include ambiguity that may vary between literal and figurative senses; understanding the cultural and linguistic context in which they occur, as many of these phraseologisms are deeply rooted in Hispanic culture; the ability to identify them as integral semantic units; and the capacity to translate them accurately into other languages, in this case, from Spanish into Russian.

However, after interpreting the content and analyzing the verbal idioms, several factors affecting the accuracy of the translation of these expressions were identified. According to the results obtained, out of the 50 units analyzed, only 15 were correctly translated by Yandex's AI, representing less than half of the cases. This finding indicates that the ability of automatic translation systems to handle and translate idiomatic expressions reflected in corpora remains a significant challenge in the advancement of translation technologies. The intricate nature and diversity of idiomatic expressions across languages make their accurate and contextual translation a complex process that requires more thorough analysis and the refinement of specialized translation algorithms.

Among the crucial factors affecting the accuracy of Yandex's translations were the ambiguity in the meaning of idiomatic expressions, the understanding of the cultural context in which these expressions are used, and attention to the type of linguistic register. The inherent ambiguity of idioms presents a notable challenge for AI, as the presence of multiple possible interpretations complicates the selection of the most appropriate translation according to the context. This lack of clarity in the meaning of idioms led to incorrect, sometimes literal translations, or even omissions of parts of the original message. Additionally, the cultural context in which these idiomatic expressions occur is essential for their precise interpreting. Yandex faced difficulties in identifying and capturing the cultural connotations and implicit meanings in these expressions, resulting in translations that did not adequately reflect the original sense of the idioms in question. Furthermore, the nature of the linguistic register impacted the quality of translations. The presence of specific colloquialisms with no direct equivalents in the target language complicated the correct interpreting of these expressions, leading to imprecise or incomplete translations. The

most representative examples of the translations of the analyzed idioms are presented below, illustrating the difficulties encountered and areas for improvement in automatic interpreting.

When it comes to ambiguity, many of the verbal idioms studied, such as *pisar fuerte, abrir puertas, poner en pie, abrir camino, or ajustar cuentas* can have multiple meanings, complicating their accurate translation. In this context, both AI in general and Yandex may struggle to select the appropriate translation based on the usage context of the expression. In our case, Yandex had difficulty determining the correct meaning of certain idioms due to contextual ambiguity, resulting in inaccurate or incorrect translations. This was the case for idioms such as *pegar ojo, salir rana, caérsele la casa encima, dar tela, echar una mano, chuparse los dedos,* or *tener mala leche.* All these idioms were translated literally, resulting in confusing and incomprehensible translations in Russian. Therefore, it is crucial to consider these potential ambiguities when using automatic translation tools. This can be addressed by supplementing machine translations with post-editing, where a human translator reviews and refines the output for accuracy. Additionally, for complex cases involving idiomatic expressions or culturally nuanced content, human translators may need to be employed from the start to ensure a precise and contextually appropriate interpreting.

The interpreting of the mentioned verbal idioms did not work correctly due to several limitations inherent to the corpus integrated in Yandex. First, the linguistic corpus used may not have included enough contextual examples of these idioms in various communicative situations, which is crucial for AI to learn and recognize contextual and cultural variations of idiomatic expressions. Without adequate representation of these idioms in the corpus, AI cannot develop a deep understanding of their use and meaning. Additionally, it is possible that the corpus was not sufficiently updated or enriched with data from diverse and contemporary sources. Idioms often evolve and may acquire new meanings over time, especially in specific cultural contexts. If the corpus does not reflect these changes, AI may fail to interpret the most current or colloquial idiomatic expressions correctly.

Another limitation is the corpus's ability to capture the semantic and pragmatic richness of verbal idioms. Idioms not only depend on individual words but also on their usage in specific situations and the accompanying cultural and social nuances. If the corpus lacks these dimensions, AI will struggle to make accurate and contextually appropriate interpreting.

On the other hand, the integrated corpus may not be designed to handle semantic ambiguity effectively. Verbal idioms, by their nature, are prone to multiple and often figurative meanings, requiring a deep and nuanced understanding that goes beyond simple literal translation. Without a specialized approach to handling ambiguity, AI may resort to literal translations that are inadequate or incomprehensible in the target language.

In addition to linguistic and semantic aspects, the cultural context in which idiomatic expressions are rooted significantly influences their translation into Russian. Many of these expressions are specific to a particular culture and reflect values, beliefs, and traditions

unique to that culture. Consequently, a literal translation may result in a loss of their original meaning and connotation. In our study, Yandex opted to translate some of these idioms in a more general or contextualized manner, considering their underlying meaning in the source culture, in this case, Spanish. Thus, Yandex attempted to preserve the essence and original intent of the expressions by adapting them to the cultural context of the target language. However, this approach has its problems, as the resulting translations are not always equivalent or congruent in meaning. Examples of failed translations by Yandex include: *estar en una nube, meter salsa, ser la guinda del pastel, dar con la tecla,* and *dar cosa.* Although the translation into Russian of these idioms was not literal, Yandex did not consider the specific cultural context in which they are embedded. This highlights the importance of accounting for not only linguistic aspects but also cultural ones to achieve a precise and meaningful translation of idiomatic expressions.

For instance, *estar en una nube* was translated as *быть на облаке* ('being on a cloud'), an expression in Russian that refers to someone who has passed away, contrasting with the Spanish usage where it denotes someone excited or distracted by something positive. In the case of *meter salsa*, this expression was translated as *водить фальшивки* ('to cheat'), whereas in Spanish it is used to describe someone who meddles in other people's lives to criticize them. The Russian translation refers to someone cheating in a game, which does not fit the original context.

An interesting expression is *ser la guinda del pastel*, which has a literal equivalent in Russian, *вишенка на торте* ('the cherry on the cake'). However, Yandex chose to translate it as *глазурь на торте* ('the icing on the cake'), a comprehensible but non-idiomatic version. Regarding *dar con la tecla*, generally used as a synonym for 'hit the mark,' it was translated into Russian as *опредеделить деталь* ('to determine the detail'). On the other hand, the expression *me da cosa*, meaning 'it makes me uneasy,' was translated as *это много мне дает* ('it gives me a lot') due to a lack of context. Such limitations, including ambiguity and polysemy, lack of cultural or situational context, errors in semantic analysis, or constraints in machine learning, highlight the need to enrich the corpora used to train machine translation systems with more examples of idiomatic expressions in various contexts and to consider both linguistic and cultural aspects. Collaboration with human translators and the integration of specific cultural knowledge could also improve translation accuracy.

In terms of linguistic register, it is crucial to note that some idioms include colloquialisms specific to a particular level of formality. These aspects also presented an additional challenge for machine translation, as they may not have direct equivalents in the target language or may be incomprehensible to native speakers. In these situations, Yandex faced the challenge of deciding how to translate such expressions, sometimes choosing to omit the expression if an adequate equivalent could not be found or, alternatively, seeking an approximate equivalent that preserved the general sense of the original expression. Examples of omissions include: *echarle morro, meterse en un fregao, estar al pie del cañón,* and *meter la pata*. In these cases, Yandex did not identify their specific meaning or context

in Spanish because they are very colloquial idiomatic expressions with no direct translation into Russian, leading the tool to omit them to avoid possible interpreting errors.

For finding an approximate equivalent that preserves the general sense of the original expression, the following examples can be cited: *quitarle hierro a un asunto* (*снять напряжение* – 'to relieve tension'), *meterse en un lío* (*втянуть в беспорядок* – 'to get caught in a mess'), *estar en auge* (*процветать* – 'to flourish'*), tener entre algodones a alguien* (*держать в объятиях* – 'to keep in one's arms'), *plantar cara* (*противостоять* – 'to stand up to'), or *estar en un pozo sin fondo* (*быть в яме* – 'to be in a hole'). All these examples retain the metaphorical meaning of the source text, and their translations, though not always idiomatic expressions, integrate well into Russian discourse.

It is also interesting and positive to analyze the idiomatic expressions that Yandex translated correctly. This was the case with expressions like: *tener en el bolsillo* (*есть в кармане* – 'to have in one's pocket'), *dar alas* (*дать крылья* – 'to give wings'), *hacerse la boca agua* (*слюнки текут* – 'to make one's mouth water'), *tocar madera* (*постучать по дереву* – 'to knock on wood'), and *no levantar cabeza* (*не поднимать головы* – 'to not lift one's head'). In these cases, Yandex was able to capture the contextual meaning of the expressions and provide an appropriate translation based on their Russian equivalents. This demonstrates that, on certain occasions, machine translators can correctly translate idiomatic expressions if they have a broad and up-to-date database that allows them to recognize and understand the meaning of these expressions in different languages. It is notable that most correctly translated expressions are idioms with counterparts in Russian. This clearly indicates that both Yandex's corpus and other machine translation systems still exhibit a bias towards a literal approach, offering more accurate translations primarily in cases where there are complete equivalents between the two languages.

5. Conclusions

In general, while AI and machine learning have advanced significantly, automated translation systems still face limitations in understanding and accurately translating idiomatic expressions. This highlights the critical need for human intervention and linguistic knowledge in translating texts that are complex and culturally rich. Therefore, it is essential to continue research and refinement of automated translation systems to handle idiomatic expressions and improve the quality and accuracy of the translations produced more effectively.

Yandex's corpus includes an extensive database of texts in multiple languages, enabling it to successfully recognize and translate some idiomatic expressions that have direct equivalents in Russian. This is evident in expressions like *tener en el bolsillo* (*есть в кармане*) and *dar alas* (*дать крылья*), where the system was able to preserve the meaning and essence of the original expressions. The presence of cultural equivalents in the corpus allows Yandex to provide translations that reflect the context and appropriate use of expressions in the target language. Additionally, the use of neural machine translation models based on deep neural networks in Yandex's corpus contributes to greater accuracy

in translating idiomatic expressions. These models help capture the full context of a sentence and consider the syntax and semantics of words, resulting in more natural and contextually appropriate translations. This demonstrates that Yandex's corpus, by incorporating advanced techniques, is well-positioned to offer accurate translations when clear equivalences exist between languages.

However, despite these advances, Yandex's corpus faces challenges with highly colloquial idiomatic expressions or those specific to a particular linguistic register. Idioms such as *echarle morro* or *meterse en un fregao* were omitted or incorrectly translated due to the lack of direct equivalents in Russian. The corpus's inability to handle these colloquial idioms reflects a limitation in understanding contexts and linguistic registers that do not have an exact match in the target language. Idiomatic expressions with specific cultural meanings may be misinterpreted or inaccurately translated if the corpus lacks sufficient information about the cultural context. For example, *estar en una nube* was translated as *быть на облаке* ('to be in the cloud'), an expression used to refer to death in Russian, rather than the positive emotion or distraction implied in Spanish. The lack of specific cultural data in Yandex's corpus can lead to translations that do not capture the appropriate meaning, resulting in incorrect interpreting. Although Yandex has attempted to adapt idiomatic expressions to the cultural context of the target language, some translations do not accurately reflect the original meaning. For instance, *dar con la tecla* was translated as *опредеделить деталь* ('to determine the detail'), which fails to capture the metaphorical meaning of the expression in Spanish. The lack of a deep understanding of cultural and linguistic context in Yandex's corpus limits the AI's ability to provide culturally appropriate translations.

In conclusion, Yandex's corpus demonstrates both strengths and weaknesses in translating idiomatic expressions. While the inclusion of a broad database and advanced techniques improves accuracy in some cases, limitations in understanding colloquial idioms and specific cultural contexts highlight the need to continue developing and refining automated translation systems. Integrating more detailed cultural and colloquial information into Yandex's corpus is crucial to overcoming these challenges and offering translations that are both linguistically precise and culturally appropriate.

REFERENCES

Christen y Gracia, L. M. (2017) *Incidencia del género de la voz del intérprete en la percepción de la calidad de la interpretación simultánea de conferencias*. Universitat de Vic - Universitat Central de Catalunya.

Collados Aís, Á. (2016) 'La entonación monótona y la calidad de la interpretación simultánea: frecuencia, conceptualizaciones y efectos', *Meta*, 61(3), pp. 675–691.

Colson, J. P. (2005) 'Cross-linguistic phraseological studies' in *Conference Proceedings Phraseology*. Belgium: Mouton de Gruyter.

Corpas Pastor, G. (1997) 'Grados de equivalencia translémica de las locuciones en inglés y español', *XVIII Congreso de AEDEAN: Alcalá de Henares*. Universidad de Alcalá, 15–17 diciembre 1994. Editorial Universidad de Alcalá, pp. 329-334.

Corpas Pastor, G. (2024) 'Towards the "tech" revolution in interpreting', *Routledge Encyclopedia of Technology and the Humanities*, p. 183.

Defrancq, B., & Fantinuoli, C. (2021) 'Automatic speech recognition in the booth: Assessment of system performance, interpreters' performances and interactions in the context of numbers', *Target*, 33(1), pp. 73-102.

Defrancq, B., Snoeck, H., & Fantinuoli, C. (2024) 'Interpreters' performances and cognitive load in the context of a CAI Tool', *Translation, Interpreting and Technological Change: Innovations in Research*, *Practice and Training*, p. 37.

Dong, Y., Li, Y., & Zhao, N. (2019) 'Acquisition of interpreting strategies by student interpreters', *The Interpreter and Translator Trainer*, 13(4), pp. 408–425.

Fantinuoli, C. (2016) 'Revisiting corpus creation and analysis tools for translation tasks', *Cadernos de Tradução*, 36, pp. 62-87.

Fantinuoli, C. (2017) 'Computer-assisted interpreting: Challenges and future perspectives', *Trends in E-tools and Resources for Translators and Interpreters*, pp. 153-174.

Fantinuoli, C. (2023) 'Towards AI-enhanced computer-assisted interpreting' in *Interpreting technologies–current and future trends*. John Benjamins, pp. 46-71.

Gaber, M. (2023) 'Soluciones tecnológicas para intérpretes: reconocimiento automático del habla e interpretación remota/technological solutions for interpreters: automatic speech recognition and remote interpreting', PhD diss., Universidad de Málaga.

García Becerra, O. (2013) 'La incidencia de las primeras impresiones en la evaluación de la calidad de la interpretación simultánea: un estudio empírico', PhD diss., Universidad de Granada.

Gile, D. (2009) *Basic concepts and models for interpreter and translator training*. John Benjamins publishing company

Giustini, D. (2023) 'COVID-19 and the configuration of materiality in remote interpreting: Is technology biting back?' in *Translation and Interpreting in the Age of COVID-19*. Singapore: Springer Nature Singapore, pp. 197-213.

Han, C., & Lu, X. (2023) 'Can automated machine translation evaluation metrics be used to assess students' interpretation in the language learning classroom?', *Computer Assisted Language Learning*, 36(5-6), pp. 1064–1087.

Hervais-Adelman, A., Moser-Mercer, B., Michel, C. M., & Golestani, N. (2015) 'fMRI of simultaneous interpretation reveals the neural basis of extreme language control', *Cerebral Cortex*, 25(12), pp. 4727–4739 .

Kolchugina, N. (2021) 'La repercusión del uso de estrategias en la interpretación bilateral y en la evaluación de su calidad en los ámbitos de entretenimiento deportivo y artístico-musical: propuesta de un catálogo didáctico de estrategias' PhD diss., Universidad del País Vasco-Euskal Herriko Unibertsitatea.

Manning, C. D. (2008) *Introduction to information retrieval*. Syngress Publishing.

Mogorrón Huerta, P. (2020) 'Análisis multidisciplinar del fenómeno de la variación fraseológica en traducción e interpretación', *MonTI. Monografías de traducción e interpretación*, pp. 36–64.

Pöchhacker, F. (2024) 'Is machine interpreting interpreting?', *Translation spaces*. John Benjamins Publishing Company.

Pradas Macías, E. (2003) 'Repercusión del intraparámetro pausas silenciosas en la fluidez: Influencia en las expectativas y en la evaluación de la calidad en interpretación simultánea' PhD diss., Universidad de Granada.

Prandi, B. (2023) *Computer-assisted simultaneous interpreting: A cognitive-experimental study on terminology*. Language Science Press.

Ramírez Rodríguez, P. (2022) 'Fraseologismos en el aula de ELE: la problemática de la traducción automática', *Tradumàtica*, 20, pp. 77-95.

Ramírez Rodríguez, P. (2023) 'La cognición fraseológica en traducción: El caso de las locuciones verbales relacionadas con la alimentación', *Hikma*, 22(2), pp. 207-233.

Ramírez Rodríguez, P. (2024) 'La adquisición fraseológica cognitiva de las locuciones verbales somáticas en español por estudiantes rusos', *Círculo de lingüística aplicada a la comunicación*, 98, pp. 263–273.

Ramisch, C. (2015) *Multiword expressions acquisition. A Generic and open framework.* Cham: Springer International Publishing.

Ruiz Gurillo, L. (1997) 'Relevancia y fraseología: la desautomatización en la conversación coloquial', *Español Actual: Revista de Español Vivo*, 68, pp. 21-30.

Sag, I. A., Baldwin, T., Bond, F., Copestake, A., & Flickinger, D. (2002) 'Multiword expressions: A pain in the neck for NLP', *Computational Linguistics and Intelligent Text Processing: Third International Conference*. Mexico City, 17-23 February. Springer Berlin Heidelberg, pp. 1-15.

Tormo, C. U. (2021) 'Perspectiva teórica y aplicada en el análisis de locuciones verbales que expresan emociones', *Journal of Spanish Language Teaching*, 8(1), pp. 48–62.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser L., & Polosukhin, I. (2017) 'Attention is all you need', *Advances in Neural Information Processing Systems*, p. 30.

Wang, X., & Fantinuoli, C. (2024) 'Exploring the correlation between human and machine evaluation of simultaneous speech translation', *arXiv*, 14 [online]. Available at: https://doi.org/10.48550/arXiv.2406.10091.

About the author: Pablo R. Rodríguez studied Translation and Interpreting at the University of Granada. He holds a master's degree in European Studies from the University of Seville and a postgraduate degree in Russian as a Foreign Language from the HSE University in Moscow. In 2022 he completed a PhD in Comparative Linguistics at the RUDN University in Moscow, Faculty of Foreign Languages, where he works as a doctoral assistant teaching ELE and translation.