THE EFFECTIVENESS OF AI TOOLS IN EVALUATING IDIOMATIC EXPRESSIONS USAGE AMONG LEARNERS

Aamal Alfaqih [1*]

[1]Study Program of Technology in Education and Second Language Acquisition, College of Education, University of South Florida, United States

Article Info	Abstract
	This paper analyzes the potential of Artificial Intelligence interventions for assessing idiomatic expressions among language learners. Idiomatic expressions are a challenging aspect of language proficiency as their meanings are figurative and culture specific. This paper aspires to explore how current AI technologies like natural language processing, adaptive learning platforms, and automated speech evaluation systems m easure learners' understanding of idiomatic expressions. It reviews the strengths and limitations of AI in providing real-time feedback and where AI fails to understand idioms regarding cultural or contextual nuances accurately. It emphasizes the importance of developing more sophisticated AI systems that can consider this cultural understanding. It proposes hybrid models combining AI with human evaluation for fine-grained judgments. Ultimately, it suggests that AI is a promising way of assisting learners in learning an idiomatic language, and future directions are for improving AI's role in idiomatic expression evaluation.

*✉ Correspondence:
Email Address: aamalalfaqih@usf.edu

INTRODUCTION

Artificial Intelligence (AI) has recently been reshaping the field of second language acquisition. As AI evolves, it assesses learners in multiple linguistic areas, from grammar and pronunciation to vocabulary and idiomatic expressions (Woo & Choi, 2021; Jia et al., 2022). Especially idiomatic expressions are problematic because they have a symbolic meaning and a long history of cultural references (Cieślicka, 2006). Students who have learned and are proficient with idioms can have high proficiency levels (Laufer, 2005; Erman & Warren, 2000). However, this is not critical in this context; idiomatic expressions are notoriously complex to learn, teach, and test (Baldwin & Kim, 2010). The main question of this paper is whether AI tools can assess idiomatic phraseology learners to the same degree as human analysts and, if so, to what extent it helps them develop their idiomatic fluency.

The recent advancements in AI, specifically in natural language processing (NLP) and

machine learning, show strong potential for tackling these issues (Vaswani et al., 2017). It is possible to evaluate idiomatic expressions at scale using artificial intelligence and provide learners with real-time feedback regarding idiom usage (Yalamati & Batchu, 2024). These tools can analyze vast amounts of linguistic data and catch learners' idiomatic usage while adjusting learning materials for the individual's needs (Jia et al., 2022; Zou et al., 2023). Using AI would allow for idiomatic expression testing among many language learning applications (Xia et al., 2024). This is one step towards using data-driven, personalized technology to help people learn languages better than ever (Zou et al., 2023). However, there are concerns about AI systems' abilities to comprehensively understand various idiomatic languages in cultural and contextual nuances necessary for accurate assessment (Son et al., 2023).

This paper explores how effective AI tools are in measuring idiomatic expression usage among learners, mainly, it is intended to address the following questions:

1. How accurately do AI tools assess idiomatic expression usage compared to traditional human evaluation?
2. To what extent can AI tools provide meaningful feedback on idiomatic expressions' cultural and contextual appropriateness?
3. What are the strengths and limitations of AI-powered platforms in promoting the acquisition of conversational fluency?

This analysis is worth pursuing because AI tools are becoming an increasingly integral part of language learning, and their potential to represent complex linguistic phenomena, such as idiomatic expressions, may bring quality gains to language instruction. This research also contributes to developing more effective, culturally aware AI systems for language learning by outlining the benefits and drawbacks of current AI-driven tools to establish better culturally aware systems for language learning. The ability of AI to deliver scalable, consistent, and specific feedback makes it a promising tool for tackling one of the most complex parts of language acquisition, which is the usage of idiomatic expressions.

The following section outlines the methodological approach used to identify, select, and synthesize recent literature relevant to AI-based assessment of idiomatic expressions. This methodological foundation supports the theoretical framework that follows, which explores the significance of idiomatic expressions in language learning and the growing role of AI in language assessment. Subsequent sections examine key AI tools and techniques–such as natural language processing (NLP), adaptive learning platforms, and automated speech evaluation systems–used to assess idiomatic usage. The paper then evaluates these tools in terms of their accuracy, scalability, and ability to provide culturally and contextually meaningful feedback. It also discusses the main challenges in assessing idioms through AI, including limitations in cultural understanding and idiom variability. Finally, the paper outline future directions for enhancing AI's capacity to evaluate idiomatic expressions, such as hybrid models combing human insight with machine assessment, concludes with implications for second language acquisition (SLA) research and practice.

METHODS

This study adopts a conceptual review design focusing on recent scholarly works (2019-2024) that examine the role of Artificial Intelligence (AI) in evaluating idiomatic expressions in language learning. The sources were identified through academic databases such as Scopus, Google Scholar, and ResearchGate, using keywords such as "AI in idiom assessment," "language learning AI," and "idiomatic expressions evaluation AI". A total of 85 peer-reviewed articles were selected based on relevance, publication quality, and alignment with the study's objectives. The literature was synthesized thematically across four domains: (1) the role of idiomatic expressions in SLA, (2) AI tools and platforms used in language assessment, (3) strengths and weakness of AI in idiom analysis, and (4) future prospects and hybrid evaluation models. No empirical data was collected directly; rather, findings were derived from comparative analysis and critical synthesis of secondary data.

THEORETICAL FRAMEWORK

Idiomatic expressions are not merely linguistic elements; they also convey cultural and social nuances, reflecting the values and historical contexts of native speakers (Boers & Lindstromberg, 2008). It can be seen that the effective use of idioms is a critical aspect of reaching language proficiency, as it demonstrates a learner's ability to comprehend and engage with the cultural dimensions of a language (Liu, 2023; Samani, 2012). It is, however, difficult to assess the usage of idiomatic expressions, especially when learners are required to navigate figurative meanings and cultural significance (Fornaciari et al., 2024). AI-driven platforms can evaluate various linguistic features, providing real-time feedback on pronunciation, grammar, and idiomatic expressions (Mykhalevych, 2024; Opeton, 2024). While AI tools have made progress, they still have difficulty dealing with the cultural and contextual complexity of idioms (Lingo Vista, n.d.; Tongue Tactic, n.d.). This section delves into the importance of idiomatic expressions in language learning and the role of integrating AI in language assessment.

The Importance of Idiomatic Expressions in Language Learning

Idiomatic expressions are considered essential to language learning toward fluency or cultural literacy. They also carry a meaning beyond the literal interpretation of words, requiring learners to understand their linguistic and cultural norms (Boers et al., 2007). Since idioms are used so often in everyday conversation, learning to understand and use them correctly can be essential for students aiming to interact effectively with native speakers. They typically reflect the language's speakers' values, history, and culture. In other words, idiomatic expressions such as 'to break the ice' and 'to spill the beans' are examples of a phenomenon that is not merely linguistic constructs but also carries social behavior and practices within it (Goshkheteliani, 2013). Due to their cultural specificity, idioms often prove to be one of the most challenging genres for non-native speakers as they have to translate the words and understand the contextual implications that ideational meaning depends on social conventions (Charteris-Black, 2002). As Bardovi-Harlig (2002) notes, idioms are vital for interpreting humor, social behaviors, and most informal speech, all of which are necessary for native-level proficiency in a language.

Regarding idiomatic competence, the language assessment and proficiency field often view

this type of facility with language as a hallmark of advanced language use (Cieślicka, 2015; Leung, 2022). The ability to use idiomatic phrases is traditionally regarded as an indicator that students have progressed beyond some of the more basic structural aspects of language and reached a higher realm of nuanced, native-like expression (Liu, 2017). As idiomatic expressions are spontaneously part of everyday informal conversation, the successful choice of an idiom or collocation reveals language knowledge (Alharthi, 2019; Cieślicka, 2015). It indicates competence in navigating proper social and cultural settings (Liontas, 2002). In this sense, learners need to develop idiomatic skills in order to sound 'native' or speak with some approximation of fluency, at least in informal or colloquial situations (Zyzik, 2011).

Idiomatic expressions competence is also connected with pragmatic competence and the knowledge of showing awareness of the social rules that govern language dynamics across various situations (Taguchi, 2015). Idioms often go beyond their literal meaning and can relate to other emotional tones, attitudes, or even relationships with others (Halenko & Wang, 2022). Learners must grasp these pragmatic dimensions to use language effectively in everyday communication (Wray, 2000). This aligns with communicative language teaching, which stipulates that language learning should be realistic and practical, rooted in average communication skills, as context is vital for the functional use of language (Richards & Rodgers, 2014).

Additionally, idiomatic expressions promote emotional engagement and depth of learning in teaching some languages (MacIntyre & Gregersen, 2012). Idioms frequently evoke imagery and metaphors to appeal to the learner's emotional side and memory systems more than plain language forms (Boers & Lindstromberg, 2008). Moreover, using colorful and memorable phrases can add fun to learning and, as a result, promote motivation and retention of new language items (Siyanova-Chanturia, 2015).

Overall, idiomatic expressions are not an insignificant aspect of language learning but an essential component of reaching fluency by attaining communicative competence and cultural understanding. As idiomatic expressions are complex and emerge from culture, they play an indispensable role in language use, where language learners gain vital aspects for communicating, in this case, informal settings or social interactions. The demand for high-quality tools to help master idiomatic patterns will likely rise as language learning moves increasingly toward technology-mediated approaches.

The Role of AI in Language Learning and Assessment

One of the most critical aspects that AI has brought is its power to offer automated assessments for language learning (Shi & Aryadoust, 2024). In traditional learning environments, assessments such as speaking and writing tests are time-consuming to administer and evaluate, often requiring the expertise of human instructors (Almusharraf & Alotaibi, 2023). Tools such as auto-writing evaluation (AWE) systems and automatic speech recognition (ASR) tools powered by AI can quickly assess language analysis through linguistic features, such as grammar, pronunciation accuracy, fluidity of speaking well, and coherence (Chen et al., 2018). In addition, ETS' Speech Rater and Pearson's Versant measure spoken language skills by analyzing features such as pronunciation, intonation, and pace (Chen et al., 2018). As a result of their deep learning algorithms, these

tools can mimic human scoring with accurate metrics and objective feedback (Zechner et al., 2009; Zhang et al., 2022). Being highly scalable, AI can quickly evaluate a large amount of data, reducing the workload for teachers to assess every piece meticulously and helping monitor learner progress more extensively over time (Mollick et al., 2024). It is undeniable that artificial intelligence reliably evaluates surface-level linguistic elements. Nevertheless, it has yet to develop its capacity to evaluate more nuanced aspects of language, such as idioms and cultural implications (Amini et al., 2024).

Another significant aspect in which AI can be used for learning languages is having learners' work reviewed immediately with customized feedback (Guest, 2023). Machine Learning algorithms on AI-powered platforms like Duolingo and Memrise, for example, track learner progress over time by analyzing user data to personalize learning paths catering to areas of strengths and weaknesses (Bicknell et al., 2023; Duvall, 2024). These platforms track the learner's performance with each exercise and task; they adapt exercises to provide targeted practice in specific areas a learner struggles with (Guan et al., 2021; Zou et al., 2023). AI-powered chatbots and virtual tutors help support the feedback loop by allowing learners to run through conversational practice exercises independently (Labadze et al., 2023; Sensei, 2023). These systems leverage Natural Language Understanding (NLU) to emulate a dialogue and provide corrective feedback on grammar errors, vocabulary misuse, or idiomatic expressions (Labadze et al., 2023; Nassaji & Kartchava, 2017). In other words, chatbots can correct mistakes instantly and help learners practice language in a low-stakes environment. Such systems provide reliable, non-judgmental feedback, which can reduce the anxiety learners might feel when making mistakes during conversational practice (Chen, 2024; Hawanti & Zubaydulloevna, 2023).

AI generates adaptive learning environments for learner-specific requirements. Capturing and evaluating data in an AI system can modify educational materials as learners progress, providing them with further engagement while remaining challenging at the right stage (Benkhalfallah & Laouar, 2023; Chen et al., 2020). Adaptive platforms utilize machine learning and data analytics to tailor content delivery according to learners' performance, enabling a focused approach to challenging language skills such as idiomatic expressions (Molenaar, 2022; Zawacki-Richter et al., 2019). AI algorithms implemented in platforms like Mosa Lingua and Busuu can check user data and then adapt the lesson plans to more idioms or other challenging topics in language learning. Sophisticated machine learning techniques combined with spaced repetition algorithms enable systems to reinforce complex idiomatic expressions by presenting them to learners at carefully optimized intervals, such as every week to two weeks, ensuring effective retention (Son et al., 2023). By providing learners with a structured and tailored learning path based on their cognitive capacity, these platforms prevent learners from being overloaded with information (Imhof et al., 2020).

AI TOOLS AND TECHNIQUES FOR IDIOMATIC EXPRESSION EVALUATION

Evaluating the usefulness of idiomatic expressions in language learning is an exciting challenge, particularly when attempting to automate the process (Adewumi et al., 2022; Knietaite et al., 2024). Human intuition and a sense of empathy are often appealed to in traditional assessment methods, but AI tools can overcome these barriers as they allow for more modern methodologies (Zeng et al., 2023). This section analyzes some AI

technologies and methodologies that can be used to evaluate idiomatic expression utilization. It discusses their current scope of work, limitations, and possible further developments.

One of the basic AI techniques used when learning a language, specifically in connection with idioms, is Natural Language Processing (NLP) (Selvi et al., 2023). Deep learning algorithms, such as transformers (BERT, GPT) in NLP tools, understand beyond the literal meaning of language. They can also use non-literal language, as these models can identify idiomatic phrases by analyzing contextual clues, making them helpful in detecting non-literal language and idiomatic usage (Knietaite et al., 2024; Vaswani et al., 2017). The difficulty arises, however, because many idiomatic expressions are based on cultural and pragmatic knowledge in addition to syntactic recognition. Despite some improvements in transformer models for context, NLP might only sometimes adequately account for idiomatic expressions when cultural subtleties are involved (Zeng et al., 2023). With the help of approaches (like BERT and GPT), NLP models can partially understand figurative language. However, its performance relies heavily on data diversity and quality, especially in a multi-culturally dissimilar training corpus where a significant variation exists (Liu et al., 2019). In addition, the rise of NLP models trained to detect literal and metaphorical meanings has strengthened idiomatic expression evaluation (Zeng & Bhat, 2021). Semantic compatibility models are employed to assess idioms based on semantically implausible context, similar to the route this system might have used (Liu & Hwa, 2019). The approach is deep learning, which uses a continuous bag-of-words (CBOW) model and has been successful in recognizing idioms at the sentence level by measuring semantic cohesion (He et al., 2024).

Learning platforms powered by artificial intelligence, like Duolingo, Memrise, and Busuu, use gamification and machine learning algorithms to customize each learner's learning experience. In particular, these platforms are effective in offering personalized learning experiences via the use of AI, which tracks errors made by users and tailors future exercises in order to target weaknesses, including idiomatic expressions (Dandachi, 2024; Khakpour & Colomo-Palacios, 2021; Zawacki-Richter et al., 2019; Zeng & Bhat, 2022). Adaptive systems can provide learners with a variety of sentence contexts for idiomatic expressions, which may assist learners in determining the meaning of idiomatic expressions from these different contexts (Liu & Zu, 2024). Even though adaptive learning platforms excel at presenting idiomatic content, their feedback mechanisms for idiomatic expression accuracy still need to be developed (Wu et al., 2023). In many cases, they do not provide learners with the nuanced, real-time feedback they are entitled to in order to understand subtle cultural differences (Xia et al., 2024). Consequently, these platforms focus primarily on more common idiomatic phrases, with little attention to more culturally specific or uncommon phrases (He et al., 2024).

Automated Speech Recognition (ASR) systems have made significant strides in evaluating spoken language proficiency, including the correct usage of idiomatic expressions (Loukina et al., 2015). Systems like ETS's Speech Rater and Pearson's Versant analyze spoken language features such as pronunciation, fluency, and rhythm and are increasingly incorporating idiomatic language assessments (Zechner et al., 2014). ASR systems evaluate how idiomatic expressions are used in speech by recognizing pronunciation and intonation patterns that differ from literal speech (Cheng, 2011). While ASR systems can

provide feedback on whether an idiomatic expression is pronounced correctly or used in a grammatically correct context, they often struggle with assessing the appropriateness of idiom usage in culturally sensitive situations (Vishwakarma, 2023). The challenge arises when idioms carry nuanced meanings depending on social and situational factors, which ASR systems cannot fully comprehend (Ngueajio & Washington, 2022; Tilmatine et al., 2021). Additionally, the integration of speech synthesis models (such as those used by Google and Amazon Alexa) has shown promise in simulating conversational contexts where idiomatic expressions are used (Akuzawa et al., 2018). With this technology, students can practice using idiomatic expressions in real-time conversations and receive feedback on their usage, thus bridging the gap between practice and real-life language application (Adewumi et al., 2022).

To sum up, several AI tools and techniques for evaluating idiomatic expressions are being developed, and these tools and techniques may contribute substantially to improving language learning. Some new technologies are expected to contribute to advancing idiomatic expression assessment, including natural language processing, machine learning algorithms, automatic speech recognition, and adaptive learning platforms. While artificial intelligence has made remarkable advancements, it still faces significant challenges in understanding subtle cultural and contextual elements of idiomatic expressions, vital for accurate recognition. In the future, cultural competence should be a high priority in developing artificial intelligence. Using hybrid models in which AI is combined with human oversight may provide a more accurate assessment of idiomatic expressions.

EVALUATION OF AI TOOLS FOR IDIOMATIC EXPRESSION USAGE

A growing interest in evaluating the effectiveness of AI technologies in assessing idiomatic expression usage has resulted from the continued integration of AI technologies into language learning. In automated assessment systems, idiomatic expressions present a unique challenge due to their complex cultural and contextual nuances (Shen et al., 2024). This section focuses on the accuracy, scalability, and consistency of AI tools in this domain, as well as the cultural and contextual limitations of AI's understanding of idiomatic expressions.

Accuracy in Idiomatic Expression Assessment

AI-based tools such as NLP and machine learning are demonstrating impressive results regarding identifying idiomatic expressions in context. BERT and GPT can recognize non-literal language patterns, distinguish idiomatic expressions from their literal counterparts, and determine appropriate contexts due to their utilization of large linguistic datasets (Liu, 2019). A study conducted by Liu et al. (2020) demonstrated that transformer-based models can accurately identify and classify idiomatic expressions in written and spoken languages. Nevertheless, even the most advanced models may find it challenging to interpret idioms in complex or culturally rich contexts, resulting in errors that may cause learners to be confused (Madabushi et al., 2021). While AI can detect idioms, it may need help assessing their level of understanding or appropriate use in various contexts (Tedeschi et al., 2022). It is important to note that idioms are not only fixed expressions; they can be used in various ways based on tone, audience, and formality (Liu & Hwa, 2019). Thus, although

idiomatic phrases are initially identified with a high degree of accuracy, there is still room for improvement in the comprehensive assessment of idiomatic usage.

Scalability and Consistency

Unlike traditional assessment methods, AI tools provide consistent evaluation at scale, avoiding subjective interpretation by individual evaluators that can lead to inconsistencies. All assessments are graded and feedbacked uniformly by AI systems, regardless of the number of students or instances (Teachflow AI, 2022). This feature is particularly beneficial in blended and online learning environments, where various learners need consistent feedback regarding idiomatic expression usage. A study by Singla et al. (2021) on automated scoring systems highlights that AI consistently outperforms human graders in providing standardized assessments across large groups. This capability is crucial for idiomatic expressions, which often require precise interpretation. However, the standardized approach can also be a deficiency, as idiomatic expressions vary in use depending on regional dialects, speaker intent, and cultural context factors that AI systems need help adapting to dynamically (Kabra et al., 2023).

Cultural and Contextual Limitations

Another significant challenge for AI in idiomatic assessment is its limited ability to interpret and evaluate idioms in culturally nuanced ways (Prabhakaran et al., 2022). Idioms are deeply tied to cultural contexts and may require a specific understanding of social and historical backgrounds to be used correctly. Current AI tools, trained primarily on generalized datasets, may need more specificity to recognize culturally unique idioms or those that hold dual meanings depending on social context. For instance, AI systems often misinterpret idioms with cultural double messages, resulting in incorrect evaluations of a learner's idiomatic proficiency (Buddemeyer et al., 2021; Prabhakaran et al., 2022). Moreover, idioms are often interpreted differently depending on the region, age group, and even the subculture within a particular language. As a result of this limitation, AI can face a challenge since idioms understood in one part of the world may be entirely misunderstood in another (Li et al., 2024). Although some AI models are now trained on multilingual and multicultural corpora to address this issue, these models still struggle to account for regional variations and cultural subtleties (Helm et al., 2024; Li et al., 2024). Considering this limitation, AI cannot yet provide genuinely personalized language instruction for idiomatic expressions, especially for learners seeking native proficiency (Hernandez, 2024; Xia et al., 2024).

CHALLENGES IN EVALUATING IDIOMATIC EXPRESSIONS

With the advancement of AI and its integration with language learning, idiomatic expressions face unique challenges in terms of evaluating them accurately (Fornaciari et al., 2024). In addition to their figurative nature, cultural particularity, and context dependency, idioms present a wide variety of challenges (Glucksberg & McGlone, 2001). This section explores the complexity of idioms due to their figurative and cultural meanings, cultural awareness in AI's idiom assessment, variability in idiomatic forms and expressions, bias in training data, and limited cultural diversity.

Complexity of Idioms Due to Figurative and Cultural Meanings

It can be seen that idioms are challenging to evaluate because they are non-literal, culture-specific, hard to translate, depend on context, and involve pragmatic subtleties (Vula & Tyfekçi, 2024). This difficulty becomes even more pronounced when AI uses are layered on top, as these components demand more sophisticated linguistic signal processing than simple pattern recognition (GeeksforGeeks, 2024). Evaluating idiomatic expressions can be more challenging due to nuances of culture and context. They frequently reflect a historical or social background that AI may need the mechanisms to fully comprehend, as idioms are traditionally culture-based (Škvorc et al., 2022). For instance, the English idiom 'kick the bucket' has a meaning far removed from its literal interpretation, requiring cultural understanding to decipher (Gibbs & Colston, 2012, p. 1). Idiomatic expressions are complex to translate directly into other languages, and the AI system could unintentionally misunderstand or inaccurately predict whether an idiom is appropriate for a different culture (Zhou et al., 2021).

Cultural Awareness in AI's Idiom Assessment

Cultural awareness is critical here since the interpretation of an idiom can vary from one region, social group, or specific context to another (Goshkheteliani, 2013; Prabhakaran et al., 2022). In contrast, AI systems are predominantly based on data-driven learning algorithms and are frequently trained on massive datasets without the necessary depth of cultural background (Li et al., 2024). As a result, this limitation reduces the effectiveness of AI in determining how acceptable a given use of idiomatic language may be across various cultural landscapes (Knietaite et al., 2024; Bukovčanová, 2024). With this understanding, AI might decide if an idiom suits a specific register or context (i.e., formal versus informal communication).

Variability in Idiomatic Forms and Expressions

One aspect of idiomatic expressions is that they have various forms or variations, complicating the way of assessing them (Carrol & Segaert, 2024). So, for instance, an expression like 'gave him a piece of my mind' might appear in different forms like 'giving her a piece of their mind.' Adaptable parsing and deep linguistic pattern recognition are required to identify these variations, which may be beyond the capabilities of many AI systems dependent on rigid word patterns or specific phrase structures (Nedumpozhimana et al., 2022). Also, considering that students modify the idioms many times or partially use idioms, creating hybrid formulations that still convey the intended meaning but only differ from the conventional form (Adewumi et al., 2022). From an AI perspective, where patterns must match exactly, such variation can be troublesome as penalizing learners for pragmatically or contextually acceptable variations is easy.

Bias in Training Data and Limited Cultural Diversity

Training AI systems to assess idioms accurately requires access to large and varied datasets with various idioms in different contexts (Adewumi et al., 2021; Dunn, 2022). Nevertheless, the corpora used to train NLP models often rely on formal language or might need more examples of conversational idiom usage. So, real-world data is required to

generate idiomatic assessments that generalize across diverse communication contexts, such as informal conversations, digital communications, and cross-cultural exchanges (Fornaciari et al., 2024). AI systems are trained on existing data, and this existing data may include inherent biases or underrepresented linguistic and cultural diversity (Behzad et al., 2022). Therefore, these systems may give overbroad generalization or, in other cases, fail to capture idiomatic use across different dialects or language varieties. For example, an idiom commonly used in American English might be less widely used or mean the same thing in British English; however, AI systems mainly trained on American English data may be incapable of capturing such nuances (Luo & Puett, 2024; Myers, 2023). This bias can result in inaccurate evaluations, particularly for learners from different linguistic and cultural backgrounds.

Ultimately, the challenge of assessing idiomatic expressions demonstrates the limitations of current AI systems in handling more intricate forms of figurative language. In spite of AI's potential for scalability and efficiency when involving language learning and assessment, it is still necessary to develop the cultural, contextual, and pragmatic aspects of idioms. Solving these challenges will necessitate more advanced AI models, with greater cultural understanding, as well as pragmatic awareness and more contextual flexibility, potentially achieving this through hybrid systems that leverage human judgment alongside automation.

FUTURE DIRECTIONS

Even though current AI tools can detect and provide feedback on idiomatic usage, many limitations must be addressed, specifically in the area of cultural comprehension, context sensitivity, and subjective feedback (Gamage et al., 2022; Li & Sporleder, 2010). Future development must address these gaps to enhance the effectiveness of AI in evaluating idioms. This section addresses specific areas for improvement and innovation, starting with the benefits of hybrid models that combine AI-driven and human-based evaluations, examine how emotionally intelligent AI can provide nuanced feedback, and suggest integrating VR and AR technologies to create immersive environments for idiomatic practice, envisioning a comprehensive, contextually rich approach to AI-supported idiom evaluation.

Hybrid Models of Assessment

Undoubtedly, AI cannot capture the idiomatic nuances of a language, which is why hybrid assessment models combine AI-driven and human-based evaluations to offer a more comprehensive assessment solution. Using such a model, AI would perform initial, scalable assessments to identify patterns and flag potential problems for further investigation. Human evaluators could then analyze these flagged instances for cultural, pragmatic, and contextual appropriateness, especially in complex cases (Bukovčanová, 2024; Smith, 2023). This dual approach would enable large-scale assessments without sacrificing the qualitative depth required for accurate evaluation, particularly in advanced or nuanced language learning contexts (Ouyang et al., 2022). Additionally, hybrid systems could adapt over time, allowing the AI to learn from human feedback and gradually improve its ability to assess idiomatic usage independently (Ouyang et al., 2022).

Expanding Idiomatic Expression Databases and Contextual Variety

Current AI systems primarily focus on interpreting widely recognized English idioms, limiting their adaptability across diverse cultural and linguistic contexts. As AI technology advances, enhancing these systems to recognize idiomatic expressions across different regional dialects, professional languages, and informal speech will improve their effectiveness and broaden their applicability (Mykhalevych, 2024). Expanding the idiomatic database to include contextually rich examples across formal, informal, and cross-cultural communication would enhance AI's ability to recognize and evaluate idioms across real-life scenarios. This expanded database could also help AI systems differentiate between literal and figurative usage, improving context sensitivity and reducing misinterpretations (Beck & Weber, 2020; Liu & Hwa, 2019).

Developing Emotionally Intelligent AI for Nuanced Feedback

Developing emotionally intelligent AI could also significantly enhance feedback on idiomatic expressions and cultural awareness. This emotional intelligence system would analyze the correctness, tone, and intent of idiomatic expressions (Bishop et al., 2024). For instance, certain idioms may carry connotations of humor, sarcasm, or affection, which are critical for learners to understand to use idioms appropriately (Liu, 2017). Emotionally intelligent AI could help learners develop pragmatic competence, which is essential in language learning, by providing feedback on tone and suggesting alternative expressions that may better convey the intended emotion or social nuance (Taguchi, 2024).

Integrating AI Idiom Training with Virtual and Augmented Reality (VR/AR)

Incorporating AI systems with virtual reality (VR) and augmented reality (AR) technology could enhance the immersion and engagement of idiomatic learning. It would be possible for students to practice idioms in a virtual environment by simulating real-life situations such as professional meetings, casual conversations, or cultural exchanges. Through this experiential learning approach, learners will be able to gain a deeper understanding of idiomatic expressions as well as their contextual applications while participating in a safe learning environment (Lin & Wang, 2021; Peixoto et al., 2021). VR/AR applications, supported by AI, could offer real-time feedback on idiomatic usage, helping learners refine their skills and gain confidence in using idioms appropriately within varied contexts.

CONCLUSION

It can be concluded that integrating AI into the evaluation of idiomatic expressions offers promising results, but the process is still challenging. Using AI tools, language learners can obtain rapid and scalable feedback and identify common idiomatic patterns. It should be noted that AI cannot fully assess cultural nuances, pragmatic subtleties, and contextual variations essential to mastering idioms. Hybrid models that combine AI-driven assessments with human oversight are required to address cultural and contextual issues related to language learning. Further consideration should be given to AI models that incorporate idiomatic databases drawn from a variety of cultural backgrounds in order to enhance their awareness of regional dialects and social contexts. Developing an emotion-aware AI that provides context-appropriate feedback would significantly improve the

learning process for idiomatic expressions. Aside from acknowledging the limitations of AI, educators, and developers must invest in culturally sensitive technologies and continuously improve the capability of AI to recognize and analyze idiomatic expressions across diverse contexts since these implications are significant. This approach could foster more meaningful and contextually accurate language learning environments, better preparing learners for real-world communication.

REFERENCES

Adewumi, T. P., Vadoodi, R., Tripathy, A., Nikolaidou, K., Liwicki, F., & Liwicki, M. (2021). Potential idiomatic expression (PIE)-English: Corpus for classes of idioms. *arXiv*. https://doi.org/10.48550/arxiv.2105.03280

Adewumi, T., Liwicki, F., & Liwicki, M. (2022). Vector representations of idioms in conversational systems. *Sci, 4*(4), 37. https://doi.org/10.3390/sci4040037

Akuzawa, K., Iwasawa, Y., & Matsuo, Y. (2018). Expressive speech synthesis via modeling expressions with variational autoencoder. *In Proceedings of the Interspeech 2018 Conference* (pp. 3067–3071). ISCA. https://doi.org/10.21437/interspeech.2018-1113

Alharthi, T. (2019). The learner's levels of collocation knowledge and idiomatic expressions: Exploring the relationship of acquisition of two items of phraseology. *International Journal of Applied Linguistics and English Literature, 8*(5), 39–48.

Almusharraf, N., & Alotaibi, H. (2023). An error-analysis study from an EFL writing context: Human and automated essay scoring approaches. *Technology, Knowledge and Learning, 28*(3), 1015–1031. https://doi.org/10.1007/s10758-022-09592-z

Amini, M., Ravindran, L., & Lee, K. F. (2024). Implications of using AI in translation studies: Trends, challenges, and future direction. *Asian Journal of Research in Education and Social Sciences, 6*(1), 740–754.

Baldwin, T., & Kim, S. N. (2010). Multiword expressions. *In Handbook of Natural Language Processing* (2nd ed., pp. 267–292). Chapman and Hall/CRC.

Bardovi-Harlig, K. (2002). A new starting point? Investigating formulaic use and input in future expression. *Studies in Second Language Acquisition, 24*(2), 189–198. https://doi.org/10.1017/S0272263102002036

Beck, S. D., & Weber, A. (2020). Context and literality in idiom processing: Evidence from self-paced reading. *Journal of Psycholinguistic Research, 49*, 837–863. https://doi.org/10.1007/s10936-020-09719-2

Behzad, S., Sakaguchi, K., Schneider, N., & Zeldes, A. (2022). Elqa: A corpus of metalinguistic questions and answers about English. *arXiv preprint arXiv:2205.00395*. https://doi.org/10.48550/arXiv.2205.00395

Benkhalfallah, F., & Laouar, M. R. (2023). Artificial intelligence-based adaptive e-learning environments. In K. Kabassi, P. Mylonas, & J. Caro (Eds.), *Proceedings of the 3rd International Conference on Novel & Intelligent Digital Systems (NiDS 2023)* (Lecture Notes in Networks and Systems, Vol. 677, pp. 62–66). Springer. https://doi.org/10.1007/978-3-031-44097-7_6

Bicknell, K., Brust, C., & Settles, B. (2023). How Duolingo's AI learns what you need to learn: The language-learning app tries to emulate a great human tutor. *IEEE Spectrum, 60*(3), 28–33. https://doi.org/10.1109/MSPEC.2023.10061631

Bishop, J., Hassan, W., Bilsland, R., & Alexander, E. (2024). Enhancing emotionally intelligent responses in AIML systems through idiom-emoticon integration and analysis. In B. Christiansen & F. F. Adedoyin (Eds.), *Generative AI and multifactor productivity in business* (pp. 133–159). IGI Global.

Boers, F., & Lindstromberg, S. (Eds.). (2008). *Cognitive linguistic approaches to teaching vocabulary and phraseology*. De Gruyter Mouton. https://doi.org/10.1515/9783110199161

Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research, 11*(1), 43–62. https://doi.org/10.1177/1362168806072460

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*. https://doi.org/10.48550/arXiv.2005.14165

Buddemeyer, A., Walker, E., & Alikhani, M. (2021). Words of wisdom: Representational harms in learning from AI communication. *arXiv Preprint arXiv:2111.08581*. https://doi.org/10.48550/arXiv.2111.08581

Bukovčanová, A. (2024). *The cultural nuances AI can't capture: Why human translators are irreplaceable*. Translata. https://www.translata.eu/blog/the-cultural-nuances-ai-cant-capture-why-human-translators-are-irreplaceable

Carrol, G., & Segaert, K. (2024). As easy as cake or a piece of pie? Processing idiom variation and the contribution of individual cognitive differences. *Memory & Cognition, 52*(2), 334–351. https://doi.org/10.3758/s13421-023-01463-x

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access, 8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chen, L., Zechner, K., Yoon, S.-Y., Evanini, K., Wang, X., Loukina, A., Tao, J., Davis, L., Lee, C. M., Ma, M., Mundkowsky, R., Lu, C., Leong, C. W., & Gyawali, B. (2018). Automated scoring of nonnative speech using the SpeechRater v. 5.0 engine. *ETS Research Report Series, 2018*(1), 1–31. https://doi.org/10.1002/ets2.12198

Chen, M. R. A. (2024). The AI chatbot interaction for semantic learning: A collaborative note-taking approach with EFL students. *Language Learning & Technology, 28*(1), 1–25. https://hdl.handle.net/10125/73586

Cheng, J. (2011). Automatic assessment of prosody in high-stakes English tests. *In Proceedings of Interspeech 2011* (pp. 1589–1592). International Speech Communication Association. https://doi.org/10.21437/Interspeech.2011-477

Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research, 22*(2), 115–144. https://doi.org/10.1191/0267658306sr263oa

Cieślicka, A. B. (2015). Idiom acquisition and processing by second/foreign language learners. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 208–244). Cambridge University Press. https://doi.org/10.1017/CBO9781139342100.012

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics, 32*, 45–61. https://doi.org/10.1017/S0267190512000074

Dandachi, I. E. (2024). AI-powered personalized learning: Toward sustainable education. In R. R. Heredia & A. B. Cieślicka (Eds.), *Navigating the intersection of business,*

sustainability and technology (pp. 109–118). Springer Nature. https://doi.org/10.1007/978-981-99-8572-2_5

Dunn, J. (2022). *Natural language processing for corpus linguistics*. Cambridge University Press. https://doi.org/10.1017/9781009070447

Duvall, J. (2024). *Memrise vs. Duolingo: Which is better for language learning?* Duolingo. https://duolingoguides.com/memrise-vs-duolingo/

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text & Talk*, *20*(1), 29-62. https://doi.org/10.1515/text.1.2000.20.1.29

Fornaciari, F. D. L., Altuna, B., Gonzalez-Dios, I., & Melero, M. (2024). A hard nut to crack: Idiom detection with conversational large language models. *arXiv Preprint arXiv:2405.10579*. https://doi.org/10.48550/arXiv.2405.10579

Gamage, G., De Silva, D., Adikari, A., & Alahakoon, D. (2022, July). A BERT-based idiom detection model. In *Proceedings of the 15th International Conference on Human System Interaction* (pp. 1–5). IEEE. https://doi.org/10.1109/HSI55341.2022.9869485

GeeksforGeeks. (2024). *Signal processing and artificial intelligence (AI)*. GeeksforGeeks. https://www.geeksforgeeks.org/signal-processing-and-artificial-intelligence-ai/

Gibbs, R. W., & Colston, H. L. (2012). *Interpreting figurative meaning*. Cambridge University Press. https://doi.org/10.1017/CBO9781139168779

Glucksberg, S., & McGlone, M. S. (2001). *Understanding figurative language: From metaphor to idioms* (Oxford Psychology Series No. 36). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195111095.001.0001

Goshkheteliani, I. (2013). Idioms in cross-cultural communication. In J. Szerszunowicz, B. Nowowiejski, K. Yagi, & T. Kanzaki (Eds.), *Research on phraseology across continents* (Vol. 2, pp. 19–36). University of Bialystok Publishing House.

Guest, C. (2023, December 15). *The impact of AI on language learning*. Pearson. https://www.pearson.com/languages/community/blogs/2023/12/ai-and-language-learning.html

Halenko, N., & Wang, J. (Eds.). (2022). *Pragmatics in English language learning*. Cambridge University Press. https://doi.org/10.1017/9781108894241

Hawanti, S., & Khudoiberdieva, M. Z. (2023). AI chatbot-based learning: Alleviating students' anxiety in English writing classroom. *Bulletin of Social Informatics Theory and Application, 7*(2), 182–192. https://doi.org/10.31763/businta.v7i2.659

He, W., Idiart, M., Scarton, C., & Villavicencio, A. (2024). Enhancing idiomatic representation in multiple languages via an adaptive contrastive triplet loss. *arXiv Preprint arXiv:2406.15175*. https://doi.org/10.48550/arXiv.2406.15175

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, *24*, 470-497. https://doi.org/10.1007/s40593-014-0024-x

Helm, P., Bella, G., Koch, G., & Giunchiglia, F. (2024). Diversity and language technology: How language modeling bias causes epistemic injustice. *Ethics and Information Technology, 26*(1), Article 8. https://doi.org/10.1007/s10676-023-09742-6

Hernandez, M. (2024, May 17). Build proficiency for language learners by assessing AI outputs: Insights from a UChicago instructor's experimentation. *Academic Technology Solutions*. https://academictech.uchicago.edu/2024/05/17/build-

proficiency-for-language-learners-by-assessing-ai-outputs-insights-from-a-uchicago-instructors-experimentation/

Imhof, C., Bergamin, P., & McGarrity, S. (2020). Implementation of adaptive learning systems: Current state and potential. In P. Isaias, D. G. Sampson, & D. Ifenthaler (Eds.), *Online teaching and learning in higher education* (pp. 93–115). Springer. https://doi.org/10.1007/978-3-030-48190-2_6

Jia, F., Sun, D., Ma, Q., & Looi, C.-K. (2022). Developing an AI-based learning system for L2 learners' authentic and ubiquitous learning in English language. *Sustainability, 14*(23), Article 15527. https://doi.org/10.3390/su142315527

Kabra, A., Liu, E., Khanuja, S., Aji, A. F., Winata, G. I., Cahyawijaya, S., ... & Neubig, G. (2023). Multi-lingual and multi-cultural figurative language understanding. *arXiv Preprint arXiv:2305.16171.* https://doi.org/10.48550/arXiv.2305.16171

Khakpour, A., & Colomo-Palacios, R. (2021). Convergence of gamification and machine learning: A systematic literature review. *Technology, Knowledge and Learning, 26*(3), 597–636. https://doi.org/10.1007/s10758-020-09456-4

Knietaite, A., Allsebrook, A., Minkov, A., Tomaszewski, A., Slinko, N., Johnson, R., ... & Villavicencio, A. (2024). Is less more? Quality, quantity and context in idiom processing with natural language models. *arXiv Preprint arXiv:2405.08497.* https://doi.org/10.48550/arXiv.2405.08497

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, *20*(1), 56. https://doi.org/10.1186/s41239-023-00426-1

Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, *5*(1), 223-250. https://doi.org/10.1075/eurosla.5.11lau

Leung, C. (2022). Language proficiency: From description to prescription and back. *Educational Linguistics*, *1*(1), 56-81. https://doi.org/10.1515/eduling-2021-0006

Li, C., Chen, M., Wang, J., Sitaram, S., & Xie, X. (2024). Culture LLM: Incorporating cultural differences into large language models. *arXiv Preprint* arXiv:2402.10946. https://doi.org/10.48550/arXiv.2402.10946

Li, L. (2023). From idioms to constructions. *In Quadri-syllabic schematic idioms in Chinese: Description and acquisition* (pp. 7–28). Springer. https://doi.org/10.1007/978-981-19-7202-7_2

Li, L. (2023). Idiom acquisition and usage-based language learning theory. In *Quadri-syllabic schematic idioms in Chinese: Description and acquisition* (pp. 29–48). Springer. https://doi.org/10.1007/978-981-19-7202-7_3

Li, L., & Sporleder, C. (2010, August). Linguistic cues for distinguishing literal and non-literal usages. In C.-R. Huang & D. Jurafsky (Eds.), *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010): Posters* (pp. 683–691). Coling 2010 Organizing Committee. https://aclanthology.org/C10-2078

Li, S., Chen, J., Yuan, S., Wu, X., Yang, H., Tao, S., & Xiao, Y. (2024, March). Translate meanings, not just words: IdiomKB's role in optimizing idiomatic translation with language models. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 17, pp. 18554–18563). https://doi.org/10.1609/aaai.v38i17.29817

Lin, Y.-J., & Wang, H.-C. (2021). Using virtual reality to facilitate learners' creative self-efficacy and intrinsic motivation in an EFL classroom. *Education and Information Technologies, 26*(4), 4487–4505. https://doi.org/10.1007/s10639-021-10472-9

Lingo Vista. (n.d.). *Understanding the cultural context of idioms.* Lingo Vista. https://lingovista.com/language-learning/cultural-insights/understanding-the-cultural-context-of-idioms/

Liontas, J. (2002). Context and idiom understanding in second languages. *EUROSLA Yearbook*, *2*(1), 155-185. https://doi.org/10.1075/eurosla.2.11lio

words: IdiomKB's role in optimizing idiomatic translation with language models. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 17, pp. 18554–18563). https://doi.org/10.1609/aaai.v33i01.33016738

Liu, D. (2017). *Idioms: Description, comprehension, acquisition, and pedagogy*. Routledge. https://doi.org/10.4324/9781315092843

Liu, J., O'Hara, N., Rubin, A., Draelos, R., & Rudin, C. (2020, July). Metaphor detection using contextual word embeddings from transformers. In *Proceedings of the Second Workshop on Figurative Language Processing* (pp. 250–255). Online: Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.figlang-1.34

Liu, N. F., Gardner, M., Belinkov, Y., Peters, M. E., & Smith, N. A. (2019). Linguistic knowledge and transferability of contextual representations. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 1073–1094). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1112

Liu, Y. (2019). Fine-tune BERT for extractive summarization. *arXiv Preprint arXiv:1903.10318.* https://doi.org/10.48550/arXiv.1903.10318

Liu, Y., & Zu, Y. (2024). Design and implementation of adaptive English learning system integrating language contexts. *Journal of Electrical Systems, 20*(9s), 85–91. https://journal.esrgroups.org/jes/article/view/4293

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). ROBERTA: A robustly optimized BERT pretraining approach. *arXiv Preprint arXiv:1907.11692.* https://doi.org/10.48550/arXiv.1907.11692

Loukina, A., Zechner, K., Chen, L., & Heilman, M. (2015, June). Feature selection for automated speech scoring. *In Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 12–19). Association for Computational Linguistics. https://doi.org/10.3115/v1/W15-0602

Luo, Q., & Puett, M. (2024). Anglo-American bias could make generative AI an invisible intellectual cage. *Nature, 629*(8014), 1004. https://doi.org/10.1038/d41586-024-01573-9

MacIntyre, P., & Gregersen, T. (2012). Emotions that facilitate language learning: The positive-broadening power of the imagination. *Studies in Second Language Learning and Teaching, 2*(2), 193–213. https://doi.org/10.14746/ssllt.2012.2.2.4

Madabushi, H. T., Gow-Smith, E., Scarton, C., & Villavicencio, A. (2021). A stitch in language models: Dataset and methods for the exploration of idiomaticity in pre-trained language models. *arXiv Preprint arXiv:2109.04413.* https://doi.org/10.48550/arXiv.2109.04413

Molenaar, I. (2022). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence, 3*, Article 100070. https://doi.org/10.1016/j.caeai.2022.100070

Mollick, E., Mollick, L., Bach, N., Ciccarelli, L. J., Przystanski, B., & Ravipinto, D. (2024). AI agents and education: Simulated practice at scale. *arXiv Preprint arXiv:2407.12796.* https://doi.org/10.48550/arXiv.2407.12796

Myers, A. (2023). *AI detectors biased against non-native English writers.* Stanford Institute for Human-Centered Artificial Intelligence. https://hai.stanford.edu/news/ai-detectors-biased-against-non-native-english-writers

Mykhalevych, N. (2024, April 4). *Does AI understand idioms?* Preply. https://preply.com/en/blog/ai-idioms/

Nassaji, H., & Kartchava, E. (2017). *Corrective feedback in second language teaching and learning: Research, theory, applications, implications*. Routledge.

Nedumpozhimana, V., Klubička, F., & Kelleher, J. D. (2022). Shapley idioms: Analysing BERT sentence embeddings for general idiom token identification. *Frontiers in Artificial Intelligence, 5*, Article 813967. https://doi.org/10.3389/frai.2022.813967

Ngueajio, M. K., & Washington, G. (2022, June). Hey ASR system! Why aren't you more inclusive? Automatic speech recognition systems' bias and proposed bias mitigation techniques: A literature review. In J. Zhou & Q. Tian (Eds.), *HCI International 2022 – Late Breaking Papers: Interacting with eXtended Reality and Artificial Intelligence* (Lecture Notes in Computer Science, Vol. 13517, pp. 421–440). Springer. https://doi.org/10.1007/978-3-031-21707-4_30

Opeton. (2024). *AI-driven language acquisition: Unlocking the power of artificial intelligence in language learning.* Opeton. https://www.opeton.co/blog/ai-driven-language-acquisition-unlocking-power-of-artificial-intelligence-in-language-learning

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., ... Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems, 35*, 27730–27744.

Peixoto, B., Pinto, R., Melo, M., Cabral, L., & Bessa, M. (2021). Immersive virtual reality for foreign language education: A prisma systematic review. *IEEE Access*, *9*, 48952-48962. https://doi.org/10.1109/ACCESS.2021.3068858

Prabhakaran, V., Qadri, R., & Hutchinson, B. (2022). Cultural incongruencies in artificial intelligence. *arXiv Preprint arXiv:2211.13069.* https://doi.org/10.48550/arXiv.2211.13069

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press.

Samani, E. R. (2012). The effect of conceptual metaphors on learning idioms by L2 learners. *International Journal of English Linguistics*, *2*(1), 249-256. https://doi.org/10.5539/ijel.v2n1p249

Selvi, R. S., Mohana, S., Rennis, G., Bharath, S., Boomeswar, A., & Liviyan, C. L. (2023). Idiom sense: Finding figurative meaning of idioms using natural language processing techniques. *Journal of Emerging Technologies and Innovative Research, 10*(7), f169–f175. http://www.jetir.org/papers/JETIR2307522.pdf

Burt, R. (2023, November 20). *Driving education forward: Chatbots as teaching tools*. Sensei LMS. https://senseilms.com/chatbot-for-learning/

Shen, S., Logeswaran, L., Lee, M., Lee, H., Poria, S., & Mihalcea, R. (2024). Understanding the capabilities and limitations of large language models for cultural

commonsense. *arXiv Preprint* arXiv:2405.04655. https://doi.org/10.48550/arXiv.2405.04655

Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, *36*(2), 187-209. https://doi.org/10.1017/S0958344023000265

Singla, Y. K., Krishna, S., Shah, R. R., & Chen, C. (2022, June). Using sampling to estimate and improve performance of automated scoring systems with guarantees. *Proceedings of the AAAI Conference on Artificial Intelligence 36*(11), 12835-12843. https://doi.org/10.1609/aaai.v36i11.21563

Siyanova-Chanturia, A. (2015). On the 'holistic' nature of formulaic language. *Corpus Linguistics and Linguistic Theory*, *11*(2), 285-301. https://doi.org/10.1515/cllt-2014-0016

Škvorc, T., Gantar, P., & Robnik-Šikonja, M. (2022). MICE: mining idioms with contextual embeddings. *Knowledge-Based Systems*, *235*, 107606. https://doi.org/10.1016/j.knosys.2021.107606

Smith, S. (2023). *Combining human and AI translations: A skilful approach.* Day translations blog. https://www.daytranslations.com/blog/combining-human-and-ai-translations/

Son, J. B., Ružić, N. K., & Philpott, A. (2023). Artificial intelligence technologies and applications for language learning and teaching. *Journal of China Computer-Assisted Language Learning*. Advance online publication. https://doi.org/10.1515/jccall-2023-0015

Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology, 14*, Article 1210187. https://doi.org/10.3389/fpsyg.2023.1210187

Taguchi, N. (2015). *Developing interactional competence in a Japanese study abroad context*. Multilingual Matters. https://doi.org/10.21832/9781783093731 Dokumen

Taguchi, N. (2024). Technology-enhanced language learning and pragmatics: Insights from digital game-based pragmatics instruction. *Language Teaching*, *57*(1), 57-67. https://doi.org/10.1017/S0261444823000101

Teachflow AI. (2022). *Revolutionizing assessment: AI's automated grading and feedback—unlocking efficiency, objectivity, and personalized learning*. Teachflow AI. https://teachflow.ai/revolutionizing-assessment-ais-automated-grading-feedback-unlocking-efficiency-objectivity-and-personalized-learning/

Tedeschi, S., Martelli, F., & Navigli, R. (2022). ID10M: Idiom identification in 10 languages. *Findings of the Association for Computational Linguistics: NAACL 2022*, 2715–2726. https://doi.org/10.18653/v1/2022.findings-naacl.208

Tilmatine, M., Hubers, F., & Hintz, F. (2021). Exploring individual differences in recognizing idiomatic expressions in context. *Journal of Cognition, 4*(1), Article 37. https://doi.org/10.5334/joc.183

Tongue Tactic. (n.d.). *Exploring idiomatic expressions across cultures.* Tongue tactic. https://tonguetactic.com/language-learning/cultural-insights/exploring-idiomatic-expressions-across-cultures/

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *arXiv Preprint* arXiv:*1706.03762*. https://doi.org/10.48550/arXiv.1706.03762

Vishwakarma, V. K. (2023, August). Translating cultural nuances: Challenges and strategies. *ELT Voices: International Peer-Reviewed Journal for the Teachers of English, 13*(2), 80–90. https://doi.org/10.5281/zenodo.8268531

Vula, E., & Tyfekçi, N. (2024, March 5). Navigating non-literal language: The complexities of translating idioms across cultural boundaries. *Academic Journal of Interdisciplinary Studies, 13*(2), 284–307. https://doi.org/10.36941/ajis-2024-0049

Woo, J. H., & Choi, H. (2021). Systematic review for AI-based language learning tools. *arXiv Preprint* arXiv: *2111.04455.* https://doi.org/10.48550/arXiv.2111.04455

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, *21*(4), 463-489. https://doi.org/10.1093/applin/21.4.463

Wu, W.-C. V., Hsieh, J. C., & Yang, J.-C. (2023). Personalizing flipped instruction to enhance EFL learners' idiomatic knowledge and oral proficiency. In J. M. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, design, and technology: An international compendium of theory, research, practice, and policy* (pp. 883–905). Springer International Publishing. https://doi.org/10.1007/978-3-319-17461-7_59

Xia, Y., Shin, S.-Y., & Kim, J.-C. (2024). Cross-cultural intelligent language learning system: Leveraging AI to facilitate language learning strategies in cross-cultural communication. *Applied Sciences, 14*(13), Article 5651. https://doi.org/10.3390/app14135651

Yalamati, S., & Batchu, R. K. (2024). Smart data processing: Unleashing the power of AI and ML. In J. M. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Practical applications of data processing, algorithms, and modeling* (pp. 205–221). IGI Global. https://doi.org/10.4018/979-8-3693-2909-2.ch016

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zechner, K., & Hsieh, C.-N. (2024). Automated scoring and feedback for spoken language. In M. D. Shermis & J. Wilson (Eds.), *The Routledge international handbook of automated essay evaluation* (pp. 141–160). Routledge. https://doi.org/10.4324/9781003397618-10

Zechner, K., Evanini, K., Yoon, S. Y., Davis, L., Wang, X., Chen, L., & Leong, C. W. (2014, June). Automated scoring of speaking items in an assessment for teachers of English as a foreign language. *In Proceedings of the Ninth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 134–142). Association for Computational Linguistics. https://aclanthology.org/W14-1816

Zeng, Z., & Bhat, S. (2021). Idiomatic expression identification using semantic compatibility. *Transactions of the Association for Computational Linguistics, 9*, 1546–1562. https://doi.org/10.1162/tacl_a_00442

Zeng, Ziheng, & Bhat, Suma. (2022). Getting BART to ride the idiomatic train: Learning to represent idiomatic expressions. *Transactions of the Association for Computational Linguistics, 10*, 1120–1137. https://doi.org/10.1162/tacl_a_00510

Zeng, Z., Cheng, K. T., Nanniyur, S. V., Zhou, J., & Bhat, S. (2023). IEKG: A commonsense knowledge graph for idiomatic expressions. *arXiv Preprint* arXiv:*2312.06053*. https://doi.org/10.48550/arXiv.2312.06053

Zhou, J., Gong, H., & Bhat, S. (2021, August). PIE: A parallel idiomatic expression corpus for idiomatic sentence generation and paraphrasing. In P. Cook, J. Mitrović, C. Parra Escartín, A. Vaidya, P. Osenova, S. Taslimipoor, & C. Ramisch (Eds.), *Proceedings of the 17ᵗʰ Workshop on Multiword Expressions (MWE 2021)* (pp. 33–48). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.mwe-1.5

Zou, B., Reinders, H., Thomas, M., & Barr, D. (2023). Using artificial intelligence technology for language learning. *Frontiers in Psychology, 14*, Article 1287667. https://doi.org/10.3389/fpsyg.2023.1287667

Zyzik, E. (2011). Second language idiom learning: The effects of lexical knowledge and pedagogical sequencing. *Language Teaching Research, 15*(4), 413–433. https://doi.org/10.1177/1362168811412025