



# Integrating Generative AI in Language Education: A Systematic Review of Pedagogical, Ethical, and Technological Themes

Wang Cong , Lim Seong Pek\* , Zhang Yuchen   
INTI International University  
Negeri Sembilan, Malaysia

Imratul Najwa Abdul Latif   
Universiti Teknologi Mara (Kampus Dengkil)  
Selangor, Malaysia

Zhou Bo   
Meishan Pharmaceutical College  
Sichuan Province, China

Li Ming   
North China University of Water Resources and Electric Power  
Henan Province, China

**Abstract.** Artificial Intelligence (AI) has emerged as a transformative force in language education offering innovative solutions for personalized instruction, real-time feedback, and inclusive learning environments. While several systematic reviews have explored AI in language learning broadly, few have specifically targeted generative AI using a combined Systematic Literature Review (SLR) and Bibliographic Coupling Analysis (BCA). This study fills this gap by synthesizing findings from 19 peer-reviewed articles published between 2020 and 2024 in the Web of Science database, using a dual approach that combines a SLR and Bibliographic Coupling Analysis to explore the educational potential and thematic development of AI in language learning. The review examines how AI facilitates adaptive learning paths, enhances language skills development, supports assessment and teacher roles, improves accessibility for diverse learners, and raises critical ethical and cross-cultural considerations. Using the PRISMA framework to guide the selection and synthesis process, and bibliographic coupling to identify intellectual linkages, the analysis reveals six main research clusters: personalized and adaptive learning,

---

\*Corresponding author: Lim Seong Pek; [seongpek.lim@newinti.edu.my](mailto:seongpek.lim@newinti.edu.my)

language skill enhancement, AI-driven assessment, inclusivity and accessibility, ethical and critical engagement, and system usability. According to the literature, AI supports learner autonomy, promotes engagement, and addresses various learner needs, although challenges such as digital inequality, algorithmic bias, and over-reliance on technology persist. In alignment with Sustainable Development Goal (SDG) 4: Quality Education, this study underscores the importance of inclusive, ethical, and learner-centred AI integration. Future research should address the long-term impacts of AI in education, ensure equitable access, and balance technological advancement with pedagogical integrity. This review provides practical recommendations for integrating generative AI into language classrooms, highlights the pedagogical opportunities and challenges associated with AI adoption and outlines future research directions related to long-term learning outcomes and equitable AI implementation.

**Keywords:** Ethical AI; Generative AI; Inclusivity; Language Education; Personalized Learning

## 1. Introduction

The rapid evolution of Artificial Intelligence (AI) has brought about transformative changes in language education, offering new opportunities for personalized learning and improved student engagement. AI-driven technologies, including large language models (LLMs) such as ChatGPT, have begun reshaping traditional approaches to language instruction by facilitating real-time feedback, personalized learning pathways, and adaptive educational tools (Karataş et al., 2024; Jia et al., 2022). These technologies have proved effective in enhancing various aspects of language learning, including grammar correction, vocabulary acquisition, and pronunciation practice (Goodman et al., 2024; Song & Song, 2023), with particular promise in promoting learner autonomy and self-regulation (Wei, 2023).

The integration of AI into language education holds significant potential for improving educational equity. By offering tailored support to learners with diverse backgrounds, including those with disabilities or from underprivileged contexts, AI has the capacity to democratize access to high-quality language instruction (Szabó & Szoke, 2024; Ou et al., 2024). Additionally, AI systems can facilitate the development of language skills by dynamically adapting to individual learners' proficiency levels and providing continuous, personalized feedback (Qiao & Zhao, 2023; Alqahtani et al., 2023). This aligns with global educational goals, such as Sustainable Development Goal 4, which emphasizes the importance of inclusive and quality education for all.

Despite its promising applications, the widespread adoption of AI in language learning raises several critical challenges. Ethical concerns surrounding data privacy, algorithmic bias, and academic integrity have emerged as key issues in the discourse on AI-enhanced education (Lund et al., 2023; Petrillo et al., 2024). Moreover, while AI technologies offer significant advantages in terms of accessibility and scalability, there are risks associated with over-reliance on these

tools, potentially hindering the development of critical thinking and creativity among learners (Munn & Henrickson, 2024). While several reviews have examined AI in language learning more broadly, few have systematically synthesized research specifically targeting generative AI through a methodological design that combines Systematic Literature Review (SLR) and Bibliographic Coupling Analysis (BCA). Existing reviews tend to focus either on technological affordances or learner outcomes but rarely analyse the intellectual structure and thematic evolution of the field using bibliometric network approaches.

This study aims to address these challenges by conducting a comprehensive Systematic Literature Review (SLR) combined with Bibliographic Coupling Analysis (BCA) to synthesize the current state of research on AI in language education. Focusing on studies published between 2020 and 2024, the review seeks to (1) identify and categorize the key themes and sub-themes of AI applications in language learning, (2) map intellectual linkages and research clusters using bibliographic coupling, and (3) explore the pedagogical, technological, and ethical implications of AI integration in this field. Through the analysis of 19 peer-reviewed articles, this study offers a structured framework that highlights both the potential benefits and challenges of AI-enhanced language education. Based on the above methodological procedures, the following section presents the results generated from the bibliographic coupling analysis and systematic literature review.

## 2. Methodology

This study employs a mixed-methods design combining quantitative bibliometric indicators (bibliographic coupling strength, cluster analysis via VOSviewer) with qualitative thematic synthesis following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol. The integration of numerical coupling links and interpretive coding of themes enables a comprehensive understanding of both structural and substantive developments in the field.

To rigorously explore the application of Generative Artificial Intelligence (AI) in language education, this study adopts a dual-method design, integrating a Systematic Literature Review (SLR) with Bibliographic Coupling Analysis (BCA). Bibliographic coupling is a bibliometric approach that examines the degree of relatedness between academic publications by calculating the number of shared references (Kessler, 1963). When two documents cite the same third work, they are considered coupled, with a greater number of shared references indicating a closer thematic relationship.

This analysis was supported using VOSviewer, a specialized tool for constructing and visualizing bibliometric networks (Zou et al., 2025). Specifically, bibliographic coupling by author was performed to identify and map clusters of researchers whose works exhibit strong intellectual connections. The resulting visualizations provide an intuitive overview of the structural composition of the research field and help to identify dominant themes and co-authorship patterns

(Zupic & Čater, 2015). Although Kessler (1963) is an early foundational study, it remains the seminal work that formally introduced bibliographic coupling, and contemporary bibliometric research continues to build on this principle (Zupic & Čater, 2015; Zou et al., 2025).

For the systematic review component, this study follows the PRISMA guidelines to ensure methodological transparency and reproducibility (Moher et al., 2009). A predefined protocol based on the PICO (Population, Intervention, Comparison, Outcome) framework was used to establish inclusion criteria, define the scope of inquiry, and guide the search and selection of relevant studies. Articles were sourced from the Web of Science database and were limited to peer-reviewed journal publications from 2020 to 2024 that explicitly address the role of generative AI tools in language teaching and learning. By combining bibliographic coupling and systematic review, this approach enables both a thematic synthesis and a structural mapping of the research landscape, offering insights into emerging trends, intellectual linkages, and methodological patterns within the domain of AI-enhanced language education.

### 2.1. PICO Framework

The PICO (Population, Intervention, Comparison, and Outcome) framework is used to efficiently structure the research question and inclusion criteria in order to narrow the study's scope (Schardt et al., 2007). The population for this study consists of language learners of all ages who engage with generative AI tools in educational settings. The intervention focuses on the use of generative AI technologies, such as large language models (e.g., ChatGPT), for personalized language instruction, real-time feedback, and adaptive learning pathways. Research that contrasts generative AI interventions with traditional language teaching methods or other non-AI-based approaches is part of the comparison. Outcomes evaluated in this review include improvements in language skills (e.g., grammar, vocabulary, pronunciation), learner autonomy, engagement, accessibility for diverse learners, and the ethical implications of AI integration.

According to Higgins et al. (2019), this methodological approach guarantees that the study stays focused on pertinent research that directly assesses the effect of generative AI on language education. The following research questions have been developed to direct this review:

- In what ways does bibliographic coupling analysis highlight the intellectual growth and thematic structure of studies on generative AI in language education?
- What effects does generative AI in language learning have on learners' language skills, autonomy, and engagement?
- What difficulties does its implementation present, particularly regarding ethical concerns and accessibility?
- To what extent can the findings, assessment tools, and feedback generated by generative AI in language education be considered valid and reliable, and what methodological and pedagogical safeguards are required to ensure their trustworthiness?

## 2.2. PRISMA Framework

It is commonly acknowledged that the PRISMA framework improves the reproducibility and transparency of systematic reviews. By explicitly recording the number of studies found, screened, eliminated, and ultimately included in the review, it offers a methodical approach to choosing pertinent studies (Page et al., 2021) thus guaranteeing that only studies that satisfy the inclusion criteria are examined and that the selection process is impartial. This procedure is shown graphically using the PRISMA flow diagram, which ensures traceability and clarity in the study selection. This study provides a methodologically sound review that adds to the body of knowledge already available in generative AI in language education research by following PRISMA guidelines. Based on the methodological procedures displayed in Figure 1, the following section presents the results generated from the bibliographic coupling analysis and systematic literature review.

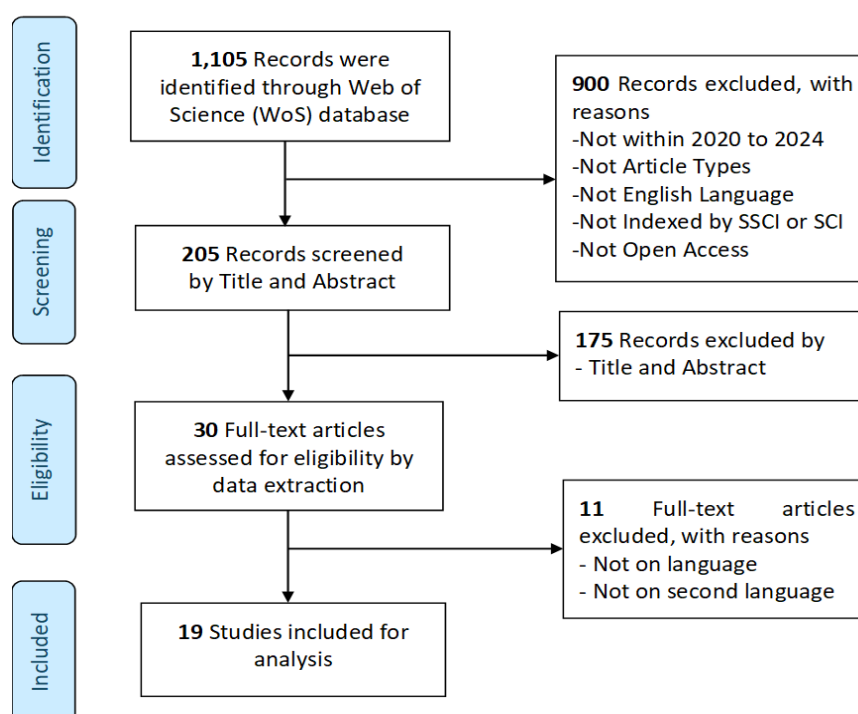


Figure 1: PRISMA Framework

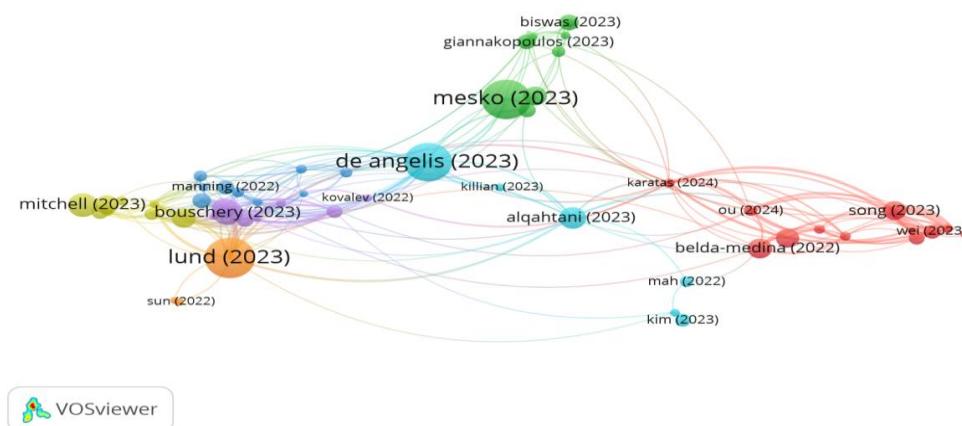
## 3. Results

### 3.1. Bibliographic Coupling Analysis

The analysis found seven clusters (Fig. 2) based on the Bibliographic Coupling that was done. Each cluster represents a thematic focus within the corpus of 55 articles. As shown in Table 1, five clusters collectively highlight the multifaceted character of AI-enhanced language learning as well as the growing use of generative AI technology to promote learners' language skills, autonomy, engagement, ethical considerations, and technical integration. The relationship between generative AI, personalized learning, and learners' language outcomes is the focus of Cluster 1 (Red). This cluster of studies examines the effects of various forms of AI-driven interventions, including adaptive learning paths,

student modelling, and mobile/contextual learning on learners' language skill development, autonomy, and engagement (Wei, 2023; Qiao & Zhao, 2023; Kim & Su, 2024). Among the subjects covered are the long-term impacts of AI customization on language proficiency, the multifaceted impacts of adaptive feedback, and the function of teacher support in AI-mediated environments (Jia et al., 2022; Ou et al., 2024). This cluster emphasizes how crucial it is becoming to understand how to use generative AI in a balanced way to promote learners' healthy language development and educational equity.

Studies that concentrate on how generative AI environments foster language skills development and writing support are grouped under Cluster 2 (Green) which focuses on creating AI-driven tools that encourage writing proficiency, pronunciation enhancement, and academic writing support (Makrygiannakis et al., 2024; Giannakopoulos et al., 2023; Ye et al., 2024). The research investigates how elements like real-time feedback, adaptive writing platforms, and personalized assignments improve learning opportunities and foster the growth of linguistic competence and motivation (Zaretsky et al., 2024; Panthier et al., 2023). This cluster encapsulates the expanding relationship between generative AI design principles, language learning, and educational technology.



**Figure 2: Bibliographic Coupling by Clusters (VOSviewer visualisation)**

Studies that concentrate on how generative AI environments enhance assessment, feedback, and teacher support are grouped under Cluster 3 (Blue) with a focus on creating AI-driven systems that encourage automated assessments, real-time feedback, and teacher role adjustment (Binder et al., 2022; Fang et al., 2024; Jackson et al., 2024). The research looks at how elements like learning analytics, intelligent tutoring systems, and data-driven insights improve learning outcomes and foster the growth of teaching efficiency and student progress tracking (Snoswell et al., 2025; Ong et al., 2024). This cluster encapsulates the expanding relationship between generative AI design principles, educational assessment, and teacher support.

Studies that concentrate on how generative AI environments promote accessibility, inclusivity, and equity are grouped under Cluster 4 (Yellow) which focuses on creating AI-driven tools that encourage support for neurodiverse learners, digital divide bridging, and inclusive design (Pellert et al., 2024; Son et al., 2023; Mitchell & Krakauer, 2023). The research examines how elements like speech-to-text features, multilingual support, and self-paced learning improve learning opportunities and foster the growth of educational equity and learner autonomy (Hohenstein et al., 2023; Hubert et al., 2024). This cluster encapsulates the expanding relationship between generative AI design principles, inclusive education, and equitable technology.

Studies that concentrate on how generative AI environments address ethics, critical thinking, and cross-cultural AI use are grouped under Cluster 5 (Purple) with a focus on creating AI-driven frameworks that encourage AI literacy, ethical considerations, and cross-cultural communication (Bauer et al., 2023; Magee et al., 2023). The research looks at how elements like data privacy, algorithmic bias mitigation, and intercultural competence improve learning opportunities and foster the growth of responsible AI use and critical thinking (Luitse & Ziewitz, 2021; AlShaikh et al., 2024). This cluster encapsulates the expanding relationship between generative AI design principles, ethical education, and cross-cultural technology.

**Table 1: Bibliographic Coupling Cluster Table**

Cluster & Colour	Cluster Label	Items	References
Cluster 1 (Red)	Personalized Learning Innovations	12	Wei (2023); Qiao & Zhao (2023); Jia et al. (2022); Karataş et al. (2024); Ou et al. (2024); Song & Song (2023);
Cluster 2 (Green)	Language Skills Enhancement	9	Giannakopoulos et al. (2023); Makrygiannakis et al. (2024); Panthier et al. (2023); ); Ye et al. (2024); Zaretsky et al. (2024)
Cluster 3 (Blue)	Assessment and Teacher Empowerment	9	Binder et al. (2022); Fang et al. (2024); ); Ong et al. (2024); ; Snoswell et al. (2025);
Cluster 4 (Yellow)	Inclusive Education Solutions	8	Hohenstein et al. (2023); Hubert et al. (2024); Jakesch et al. (2023); Pellert et al. (2024); Son et al. (2023); )
Cluster 5 (Purple)	Ethical and Critical Frameworks	8	AlShaikh et al. (2024); Magee et al. (2023); Bauer et al. (2023); Luitse & Ziewitz (2021)

### 3.2. Systematic Literature Review

The goal of the Systematic Literature Review (SLR) is to investigate how generative AI affects language education. Through an organized and methodologically guided review of 19 academic articles, this SLR identifies five key thematic areas, each of which represents crucial facets of how generative AI influences different aspects of language learning, according to its contribution to Personalized & Adaptive Language Learning; Language Skills Development & Writing Support; Assessment, Feedback & Teacher Support; Accessibility, Inclusivity & Equity; and Ethics, Critical Thinking & Cross-Cultural AI Use. Each article is categorized using a thematic mapping matrix (Table 2). The review offers a clear visual representation of how individual studies relate to these overarching themes by methodically extracting and organizing data.

**Table 2: Main Themes and Article Mapping (Systematic Literature Review Matrix)**

Article	Personalized & Adaptive Language Learning	Language Skills Development & Writing Support	Assessment, Feedback & Teacher Support	Accessibility, Inclusivity & Equity	Ethics, Critical Thinking & Cross-Cultural AI Use
Wei (2023)	✓	✓			
Qiao & Zhao (2023)	✓	✓			
Alqahtani et al. (2023)			✓		
Rahman et al. (2024)		✓	✓		
Jovanović & Campbell (2023)		✓			✓
Tang (2024)					✓
Karataş et al. (2024)		✓			
Lund et al. (2023)					✓
Szabó & Szoke (2024)	✓			✓	✓
Goodman et al. (2024)		✓		✓	
Munn & Henrickson (2024)					✓
Xia et al. (2024)				✓	✓
Akhavan & Jalali (2024)					✓
Ou et al. (2024)			✓		
Jia et al. (2022)	✓		✓	✓	
Petrillo et					✓

al. (2024)			
Kang & Kang (2022)	✓		✓
Song & Song (2023)		✓	
Zhang & Dong (2024)	✓		

---

### 3.2.1. Personalized and Adaptive Language Learning

AI-powered language learning tools provide highly personalized experiences that are tailored to the individual learner's needs, preferences, and learning pace. The use of AI in language education enhances adaptive learning systems that adjust to the proficiency level and learning style of students. Personalized learning paths based on AI algorithms enable students to engage with content that matches their current knowledge, while offering opportunities to advance at their own pace (Wei, 2023; Jia et al., 2022). This customization leads to greater learner autonomy, as students can set their own learning goals and monitor progress (Qiao & Zhao, 2023).

Additionally, tools like AI-based chatbots and intelligent tutoring systems help in providing real-time feedback on language skills, ensuring that learners receive immediate correction on their mistakes, fostering continuous improvement (Jia et al., 2022). Personalized vocabulary acquisition, real-world contextual learning, and language output are significantly enhanced through AI applications such as image recognition and adaptive grammar correction systems (Zhang & Dong, 2024). However, AI's impact is not without challenges; concerns about data privacy, over-reliance on technology, and the unequal distribution of technological resources remain significant barriers to its widespread adoption (Lund et al., 2023). Nevertheless, the potential of AI to transform language learning experiences by making them more interactive, individualized, and flexible is clear.

### 3.2.2. Language Skills Development and Writing Support

AI tools have notably transformed language education by improving learners' writing and grammar proficiency. These tools assist with writing tasks by providing real-time corrections for grammar, spelling, and stylistic issues, helping learners improve the accuracy and fluency of their written work (Goodman et al., 2024). By utilizing large language models (LLMs) such as ChatGPT, learners receive instant feedback, enabling them to revise their texts efficiently (Karataş et al., 2024). Furthermore, AI-driven systems like adaptive learning platforms personalize writing assignments to suit learners' levels and specific needs (Song & Song, 2023). Karataş et al. (2024) found that learners using ChatGPT reported improvements in writing, grammar and vocabulary acquisition, along with increased motivation and engagement.

The authors highlight ChatGPT's capacity to provide immediate corrective feedback, which suggests potential for richer language-task involvement. Studies show that motivation also increases as learners feel more confident of

their writing abilities owing to the supportive role of AI tools (Qiao & Zhao, 2023). However, while AI provides support in writing quality, some scholars raise concerns about over-dependence on AI tools, suggesting that such reliance might reduce the development of critical thinking and originality in student writing (Lund et al., 2023). Moreover, issues like academic integrity and plagiarism are prevalent, with AI systems being used for unauthorized content generation (Petrillo et al., 2024). Despite these concerns, AI's contribution to writing development is transformative, particularly in language skills enhancement.

### 3.2.3. *Assessment, Feedback and Teacher Support*

AI's role in automating assessment and providing personalized feedback has reshaped foreign language education. AI systems like intelligent tutoring systems (ITS) can assess student performance continuously and offer feedback that is both immediate and tailored to individual learning needs (Alqahtani et al., 2023). These systems provide insights into areas where students struggle and offer targeted practice, which leads to more efficient learning processes (Rahman et al., 2024). Additionally, AI facilitates more comprehensive assessments through learning analytics, which help track progress and identify trends in student behaviour (Song & Song, 2023). Teachers benefit from AI's ability to analyse large sets of data to gauge student progress, allowing them to focus more on strategic instructional planning rather than grading (Ou et al., 2024).

This shift in focus from administrative tasks to more meaningful engagement with students enhances teaching effectiveness and reduces educator burnout (Jovanović & Campbell, 2023). Furthermore, AI tools can support formative assessment, wherein students engage in peer feedback and self-assessment through AI-mediated interactions (Xia et al., 2024). However, challenges remain, such as ensuring the reliability and fairness of AI assessments and addressing concerns about privacy and algorithmic biases that could impact the accuracy of feedback (Petrillo et al., 2024).

### 3.2.4. *Accessibility, Inclusivity and Equity*

AI has the potential to make language education more inclusive, particularly for students with disabilities or those from disadvantaged backgrounds. AI systems can be designed with accessibility features such as speech-to-text, text-to-speech, and adaptive fonts to assist learners with visual, auditory, or learning impairments (Goodman et al., 2024). For example, dyslexic students benefit from AI tools that allow for easy text customization and provide real-time feedback on writing and reading tasks (Goodman et al., 2024).

Additionally, AI-powered systems can offer multilingual support, allowing students from diverse linguistic backgrounds to engage in language learning without the language barrier being a significant hindrance (Szabó & Szoke, 2024). This makes AI-powered language learning more equitable, as it accommodates various learning styles and needs (Zhang & Dong, 2024). Moreover, AI tools enable self-paced learning, allowing students to access resources at their convenience and promoting autonomy (Wei, 2023). However, the digital divide remains a challenge, with students in regions with limited internet access or

technological resources being at a disadvantage (Rahman et al., 2024). Therefore, while AI offers significant promise in advancing inclusive education, it is critical to ensure that the technology is accessible to all students, regardless of their socioeconomic status.

### 3.2.5. *Ethics, Critical Thinking and Cross-Cultural AI Use*

As AI becomes an integral part of foreign language education, it is essential to address the ethical implications surrounding its use. Data privacy, algorithmic bias, and academic integrity are significant concerns that arise with the adoption of AI tools in educational settings (Ou et al., 2024). AI's reliance on large datasets means that students' personal data could be at risk, and biases inherent in the algorithms could affect the fairness of AI-driven assessments (Munn & Henrickson, 2024). Additionally, educators and learners need to develop AI literacy to navigate these ethical challenges effectively and to understand the limitations of AI technologies (Akhavan & Jalali, 2024). Critical thinking plays a vital role in ensuring that AI is used responsibly, encouraging learners to question the content generated by AI and to assess its validity (Lund et al., 2023).

Furthermore, AI's role in cross-cultural communication is increasingly recognized. AI tools can be used to introduce students to cultural contexts, fostering intercultural competence through immersive experiences (Szabó & Szoke, 2024). However, scholars warn that AI tools must be designed to respect cultural diversity and avoid reinforcing stereotypes (Karataş et al., 2024). These ethical, social, and cultural considerations are crucial in ensuring that AI is used in ways that are equitable, transparent, and supportive of global learning environments. Building on these findings, the next section interprets the results in light of existing scholarship and discusses their pedagogical, ethical, and technological implications for AI-enhanced language education.

## 4. Discussion

The findings of this review demonstrate that the integration of generative artificial intelligence (AI) in language education has significant pedagogical, technological, and ethical implications. The analysis of 19 peer-reviewed articles published between 2020 and 2024 reveals that generative AI enhances learners' autonomy, engagement, and language skills by offering personalized feedback, adaptive learning pathways, and real-time interaction. These interventions, powered by LLMs such as ChatGPT, support individualized instruction that caters to diverse learner needs and facilitates self-directed learning (Wei, 2023; Qiao & Zhao, 2023).

Additionally, generative AI tools have been found effective in providing writing and pronunciation support, grammar correction, and context-specific learning scenarios. These benefits not only improve linguistic competence but also foster confidence and motivation among language learners (Karataş et al., 2024; Song & Song, 2023). However, challenges such as over-reliance on AI, reduced critical thinking, and academic integrity issues – particularly plagiarism – pose serious concerns. This highlights the importance of establishing clear usage guidelines

and incorporating AI literacy into language education programs (Lund et al., 2023; Petrillo et al., 2024).

A comparative synthesis of the reviewed studies reveals areas both of convergence and divergence in the impact of generative AI on language learning. Several studies report consistent improvements in writing accuracy, vocabulary development, and learner motivation when AI-generated feedback is used (Goodman et al., 2024; Karataş et al., 2024). However, other studies highlight more moderate gains, noting that the effectiveness of generative AI depends heavily on learners' prior proficiency levels and the quality of human-AI interaction (Qiao & Zhao, 2023; Song & Song, 2023). For example, while adaptive learning systems were found to significantly enhance autonomous learning in higher-proficiency learners (Wei, 2023), studies involving beginner learners reported slower progress and greater reliance on teacher mediation.

In terms of ethical and pedagogical implications, some researchers emphasize the role of AI in promoting learner autonomy and inclusive participation, whereas others caution that algorithmic bias and the risk of plagiarism may undermine pedagogical integrity (Lund et al., 2023; Petrillo et al., 2024). These contrasting findings suggest that the influence of generative AI is not uniform but shaped by contextual variables such as learner demographics, instructional design, and institutional readiness. Therefore, AI should be integrated in a context-sensitive manner that accounts for learner needs, technological access, and cultural expectations.

This review also highlights the inclusive potential of generative AI. Features like speech-to-text, multilingual support, and customizable content provide accessibility for neurodiverse learners and those from under-resourced contexts. By supporting diverse learning styles and enabling self-paced progression, generative AI contributes to educational equity and aligns with Sustainable Development Goal 4 (Szabó & Szoke, 2024; Goodman et al., 2024). Still, the persistent digital divide must be addressed to avoid deepening educational inequality between regions with differing access to technology and infrastructure.

This study shows that generative AI is a promising tool for transforming language education; to maximize its benefits and reduce associated risks, its implementation must be thoughtful, balanced, and learner centred. Responsible AI integration requires transparent guidelines, ethical policy frameworks, and culturally sensitive design approaches. Educational stakeholders can utilize AI to create more personalized, accessible, and ethical learning environments – ultimately fostering a more equitable and innovative educational future. Drawing upon the above discussion, the concluding section synthesizes the key contributions of this review and outlines future research directions.

#### **4.1. Critical Interpretation of Findings**

A critical comparison of the reviewed studies reveals both converging and diverging perspectives regarding the role of generative AI in language education.

While multiple studies (e.g., Wei, 2023; Qiao & Zhao, 2023) consistently demonstrate that AI-powered tools enhance learner autonomy and self-regulation, other research raises concerns about the potential erosion of critical thinking and creativity due to students' over-reliance on automated feedback (Lund et al., 2023; Petrillo et al., 2024). This divergence suggests that the effectiveness of AI tools may depend heavily on instructional design and the extent of teacher mediation.

Regarding writing improvement, a strong consensus indicates that AI-driven feedback enhances grammatical accuracy and writing fluency (Goodman et al., 2024; Karataş et al., 2024). However, some studies caution that AI-generated scaffolding may inadvertently mask learners' linguistic gaps, resulting in superficial proficiency gains (Munn & Henrickson, 2024). This contrast highlights the need to balance AI-assisted writing with traditional pedagogical strategies that cultivate analytical and rhetorical skills.

The review also identifies a tension between the promise of inclusivity and the persistent structural challenges that limit equitable access. While studies such as Szabó and Szoke (2024) and Goodman et al. (2024) show that AI supports neurodiverse learners and provides adaptive, accessible learning pathways, others emphasize the widening digital divide and unequal technology distribution that restrict adoption in resource-limited settings (Rahman et al., 2024). These contradictory findings underscore the need for context-sensitive implementation strategies.

Moreover, cross-cultural studies (e.g., Xia et al., 2024) argue that AI can enhance intercultural communication, whereas other scholars warn against algorithmic biases that may reinforce cultural stereotypes (Magee et al., 2023). Therefore, culturally responsive AI design remains a crucial area for further inquiry.

## **5. Conclusion**

The results of this systematic literature review indicate that the integration of generative AI into language education holds transformative potential for enhancing learner autonomy, engagement, and language proficiency. AI-powered tools, particularly large language models such as ChatGPT, offer dynamic and personalized learning experiences that support real-time feedback, adaptive instruction, and tailored language practice. These affordances significantly enhance grammar accuracy, writing fluency, and speaking confidence across diverse learner populations.

Educators should be equipped with professional training to leverage generative AI tools in ways that enrich classroom instruction while preserving critical thinking and learner creativity. Institutions should adopt balanced models that blend AI-enhanced learning with teacher guidance and human interaction. Furthermore, policies must be established to ensure the transparent use of AI, protect student data privacy, and guarantee fair access for learners from varied socioeconomic and cultural contexts.

The adoption and impact of AI vary considerably across cultural and national contexts, influenced by technological infrastructure, educational values, and cultural attitudes toward automation. Countries with strong digital ecosystems report more positive learning outcomes, while regions with limited resources face barriers that shape AI acceptance and efficacy. Future research should explore the long-term effects of generative AI on language acquisition, learner identity, and intercultural competence. Comparative studies across global regions are essential to understand how context-specific variables influence AI adoption and efficacy. There is also a growing need to investigate how human-AI collaboration can foster equitable, ethical, and sustainable education in the age of intelligent technology.

### **5.1. Limitations, Opportunities, and Challenges**

This review is subject to several limitations. First, the analysis is restricted to peer-reviewed journal articles published between 2020 and 2024, which may exclude relevant grey literature or emerging preprint studies. Second, the reliance on the Web of Science database may limit the coverage of publications from other regions or languages. Third, the relatively small number of included studies limits the generalizability of the findings, particularly across diverse cultural and educational settings. Finally, variations in research design, sample size, and methodological rigor across the reviewed studies introduce potential bias in synthesizing outcomes.

At the same time, the use of artificial intelligence presents both significant opportunities and notable challenges for language education. Opportunities include personalized learning pathways, adaptive feedback, increased accessibility for diverse learners, and more efficient formative assessment. However, these benefits must be balanced against challenges such as algorithmic bias, data privacy concerns, risks to academic integrity, unequal access to AI resources, and limited teacher preparedness for AI-integrated instruction. Understanding and addressing these opportunities and challenges is essential for ensuring that generative AI contributes ethically, equitably, and effectively to future language learning environments.

### **5.2. Implications for Research, Practice, and Policy**

This review contributes to the theoretical understanding of AI-enhanced language learning by synthesizing evidence across writing support, learner autonomy, inclusivity, and ethical challenges. Future studies should conduct longitudinal research to examine long-term learning outcomes, explore discipline-specific AI applications, and investigate the interplay between human-AI collaboration and learner identity formation. More cross-context comparative studies are also needed to understand variations across cultural and institutional settings.

For practitioners, the review highlights the importance of integrating AI tools as complementary - not substitutive - components of instruction. Teachers should provide guidance on critical AI use, support self-regulated learning, and design activities that encourage creativity rather than dependence on automated solutions. Institutions should offer professional development programmes that

equip educators with skills to implement AI effectively, especially for supporting diverse learner groups.

At the policy level, the findings call for the establishment of comprehensive guidelines to govern ethical AI use in education, including data privacy protection, transparency of AI algorithms, and standards for responsible use. Policymakers should also prioritize digital equity initiatives to ensure that all learners - including those in underserved regions - benefit from AI-enhanced learning tools. Investment in infrastructure, teacher training, and accessibility standards is essential for realizing the inclusive potential of generative AI.

### **5.3. Research recommendations and practical recommendations**

Based on these findings, several research and practical recommendations can be proposed. Future research should adopt more rigorous experimental, longitudinal, and mixed method designs to examine how generative AI supports learners across different proficiency levels, cultural contexts, and instructional environments. Studies should further explore the mechanisms underlying AI-mediated learning, including metacognition, learner agency, and the interaction between human and AI feedback.

From a practical perspective, educators and institutions should integrate AI tools within well-designed pedagogical frameworks rather than treating them as isolated technologies. Teachers need targeted training in AI literacy and ethical guidelines to help students use AI responsibly and creatively. Additionally, institutions should establish clear policies on data privacy, academic integrity, and equitable access to AI resources. Considering local and national cultural contexts is essential to ensuring that AI-enhanced language learning remains inclusive, culturally responsive, and pedagogically meaningful.

In conclusion, the analytical and critical examination of the conceptual model underscores that the integration of artificial intelligence into language education is shaped by a multi-layered interplay among pedagogical, ethical, technological, and theoretical dimensions. By revisiting the foundational components of the research through contemporary learning theories - such as sociocultural learning theory, personalized learning frameworks, and human-AI interaction models - the analysis demonstrates that AI does not function as a detached technological tool but as an active pedagogical agent embedded within complex instructional ecosystems.

The opportunities emerging from the use of AI in foreign language teaching - such as adaptive learning pathways, multimodal feedback, and enhanced learner autonomy - are inherently linked to challenges involving ethical accountability, privacy, academic integrity, and the risk of over-reliance on automated systems. This conceptual synthesis highlights the need for theoretically grounded, ethically aligned, and pedagogically informed AI integration, offering a comprehensive framework that can guide future empirical research and support the development of responsible and context-sensitive AI-enhanced language education.

Taken together, the dual-method approach adopted in this study contributes original insights by mapping both the thematic terrain and intellectual structure of the emerging field of generative AI-assisted language education. This study suggests that generative AI, when thoughtfully and inclusively implemented, can serve as a powerful tool for reimagining language education. By addressing technical, pedagogical, and ethical dimensions holistically, educational stakeholders can harness AI's potential to build a more personalized, inclusive, and future-ready learning ecosystem. These findings provide a conceptual foundation that can guide future research, policy development, and instructional design.

## 6. Acknowledgments

The authors wish to acknowledge the use of ChatGPT in the writing of this paper. This tool was used solely to help improve the language, grammar, and clarity of the manuscript. No generative content was used in a manner that would compromise the originality or intellectual contribution of the research. The authors also express their sincere appreciation to INTI International University for providing a supportive academic environment that enabled the successful completion of this study. The paper remains an accurate representation of the authors' own work and intellectual contributions.

## 7. References

- Akhavan, A., & Jalali, M. S. (2024). Generative AI and simulation modeling: how should you (not) use large language models like ChatGPT. *System Dynamics Review*, 40(3), e1773. <https://doi.org/10.1002/sdr.1773>
- Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., Bin Saleh, K., ... & Albekairy, A. M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in Social and Administrative Pharmacy*, 19(8), 1236–1242. <https://doi.org/10.1016/j.sapharm.2023.05.016>
- AlShaikh, R., Al-Malki, N., & Almasre, M. (2024). The Implementation of the Cognitive Theory of Multimedia Learning in the Design and Evaluation of an AI Educational Video Assistant Utilizing Large Language Models. *Heliyon*, 10(3), e25361. <https://doi.org/10.1016/j.heliyon.2024.e25361>
- Bauer, E., Greisel, M., Kuznetsov, I., Berndt, M., Kollar, I., Dresel, M., & Fischer, F. (2023). Using Natural Language Processing to Support Peer-Feedback in the Age of Artificial Intelligence: A Cross-Disciplinary Framework and a Research Agenda. *British Journal of Educational Technology*, 54(5), 1222–1245. <https://doi.org/10.1111/bjet.13336>
- Binder, M., Heinrich, B., Hopf, M., & Schiller, A. (2022). Global Reconstruction of Language Models with Linguistic Rules: Explainable AI for Online Consumer Reviews. *Electronic Markets*, 32(4), 2123–2138. <https://doi.org/10.1007/s12525-022-00612-5>
- Fang, X., Che, S., Mao, M., Zhang, H., Zhao, M., & Zhao, X. (2024). Bias of AI-Generated Content: An Examination of News Produced by Large Language Models. *Scientific Reports*, 14, 4588. <https://doi.org/10.1038/s41598-024-55686-2>
- Giannakopoulos, K., Kavadella, A., Salim, A. A., Stamatopoulos, V., & Kaklamanos, E. G. (2023). Evaluation of the Performance of Generative AI Large Language Models ChatGPT, Google Bard, and Microsoft Bing Chat in Supporting Evidence-Based Dentistry: Comparative Mixed Methods Study. *Journal of Medical Internet Research*, 25, e51580. <https://doi.org/10.2196/51580>

- Goodman, S. M., Buehler, E., Clary, P., Coenen, A., Donsbach, A., Horne, T. N., ... & Morris, M. R. (2024). LaMPost: AI writing assistance for adults with dyslexia using large language models. *Communications of the ACM*, 67(9), 80–89. <https://doi.org/10.1145/3626952>
- Higgins, J. P., López-López, J. A., Becker, B. J., Davies, S. R., Dawson, S., Grimshaw, J. M., ... & Caldwell, D. M. (2019). Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Global Health*, 4(Suppl 1), e000858. <https://doi.org/10.1136/bmjgh-2018-000858>
- Hohenstein, J., Kizilcec, R. F., DiFranzo, D., Aghajari, Z., Mieczkowski, H., Levy, K., Naaman, M., Hancock, J. T., & Jung, M. F. (2023). Artificial Intelligence in Communication Impacts Language and Social Relationships. *Scientific Reports*, 13, 5487. <https://doi.org/10.1038/s41598-023-30938-9>
- Hubert, B., Awa, E., & Zabelina, D. L. (2024). The Current State of Artificial Intelligence Generative Language Models Is More Creative Than Humans on Divergent Thinking Tasks. *Scientific Reports*, 14, 3440. <https://doi.org/10.1038/s41598-024-53303-w>
- Jackson, I., Saenz, M. J., & Ivanov, D. (2024). From Natural Language to Simulations: Applying AI to Automate Simulation Modelling of Logistics Systems. *International Journal of Production Research*, 62(4), 1434–1457. <https://doi.org/10.1080/00207543.2023.2276811>
- 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), 15527. <https://doi.org/10.3390/su142315527>
- Jovanović, M., & Campbell, M. (2023). Connecting AI: Merging large language models and knowledge graph. *Computer*, 56(11), 103–108. <https://doi.org/10.1109/MC.2023.3305206>
- Kang, B., & Kang, S. (2022). Construction of Chinese language teaching system model based on deep learning under the background of artificial intelligence. *Scientific Programming*, 2022(1), 3960023. <https://doi.org/10.1155/2022/3960023>
- Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, 1–24. <https://doi.org/10.1007/s10639-024-12574-6>
- Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14(1), 10–25. <https://doi.org/10.1002/asi.5090140103>
- Kim, A., & Su, Y. (2024). How Implementing an AI Chatbot Impacts Korean as a Foreign Language Learners' Willingness to Communicate in Korean. *System*, 122, 103256. <https://doi.org/10.1016/j.system.2024.103256>
- Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, 74(5), 570–581. <https://doi.org/10.1002/asi.24750>
- Luitse, D., & Ziewitz, M. (2021). The Great Transformer: Examining the Role of Large Language Models in the Political Economy of AI. *Big Data & Society*, 8(2), 1–15. <https://doi.org/10.1177/20539517211047734>
- Makrygiannakis, M. A., Giannakopoulos, K., Kaklamanos, E. G., & Stamatopoulos, V. (2024). Evidence-Based Potential of Generative Artificial Intelligence Large Language Models in Orthodontics: A Comparative Study of ChatGPT, Google Bard, and Microsoft Bing. *European Journal of Orthodontics*, 46, cjae017. <https://doi.org/10.1093/ejo/cjae017>

- Magee, L., Arora, V., & Munn, L. (2023). Structured Like a Language Model: Analysing AI as an Automated Subject. *Big Data & Society*, 10(2), 1–12. <https://doi.org/10.1177/20539517231210273>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group\*, T. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), 264–269. <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>
- Mitchell, M., & Krakauer, D. C. (2023). The Debate Over Understanding in AI's Large Language Models. *Proceedings of the National Academy of Sciences*, 120(13), e2215907120. <https://doi.org/10.1073/pnas.2215907120>
- Munn, L., & Henrickson, L. (2024). Tell me a story: a framework for critically investigating AI language models. *Learning, Media and Technology*, 1–17. <https://doi.org/10.1080/17439884.2024.2327024>
- Ong, J. C. L., Seng, B. J. J., Law, J. Z. F., Low, L. L., Kwa, A. L. H., Giacomini, K. M., & Ting, D. S. W. (2024). Artificial Intelligence, ChatGPT, and Other Large Language Models for Social Determinants of Health: Current State and Future Directions. *Cell Reports Medicine*, 5(1), 101356. <https://doi.org/10.1016/j.xcrm.2023.101356>
- Ou, A. W., Stöhr, C., & Malmström, H. (2024). Academic communication with AI-powered language tools in higher education: From a post-humanist perspective. *System*, 121, 103225. <https://doi.org/10.1016/j.system.2024.103225>
- Panthier, C., Roustit, J., Cahuzac, M., & Cochereau, I. (2023). Success of ChatGPT, an AI Language Model, in Taking the French Language Version of the European Board of Ophthalmology Examination: A Novel Approach to Medical Knowledge Assessment. *Journal Français d'Ophthalmologie*, 46(7), 706–711. <https://doi.org/10.1016/j.jfo.2023.05.006>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372. <https://doi.org/10.1136/bmj.n71>
- Pellert, M., Lechner, C. M., Wagner, C., Rammstedt, B., & Strohmaier, M. (2024). AI Psychometrics: Assessing the Psychological Profiles of Large Language Models Through Psychometric Inventories. *Perspectives on Psychological Science*, 19(5), 808–826. <https://doi.org/10.1177/17456916231214460>
- Petrillo, L., Martinelli, F., Santone, A., & Mercaldo, F. (2024). Toward the Adoption of Explainable Pre-Trained Large Language Models for Classifying Human-Written and AI-Generated Sentences. *Electronics*, 13(20), 4057. <https://doi.org/10.3390/electronics13204057>
- Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, 14, 1255594. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Rahman, A., Raj, A., Tomy, P., & Hameed, M. S. (2024). A comprehensive bibliometric and content analysis of artificial intelligence in language learning: Tracing between the years 2017 and 2023. *Artificial Intelligence Review*, 57(4), 107. <https://doi.org/10.1007/s10462-023-10643-9>
- Schardt, C., Adams, M. B., Owens, T., Keitz, S., & Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC medical informatics and decision making*, 7, 1–6. <https://doi.org/10.1186/1472-6947-7-16>
- Son, Y., Clouston, S. A. P., Kotov, R., Eichstaedt, J. C., Bromet, E. J., Luft, B. J., & Schwartz, H. A. (2023). World Trade Center Responders in Their Own Words: Predicting PTSD Symptom Trajectories With AI-Based Language Analyses of Interviews. *Psychological Medicine*, 53(3), 918–926.
- wSong, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of

- ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Snoswell, C. L., Snoswell, A. J., Kelly, J. T., Caffery, L. J., & Smith, A. C. (2025). Artificial Intelligence: Augmenting Telehealth with Large Language Models. *Journal of Telemedicine and Telecare*, 31(1), 150–154. <https://doi.org/10.1177/1357633X231169055>
- Szabó, F., & Szoke, J. (2024). How does generative AI promote autonomy and inclusivity in language teaching? *ELT Journal*, 78(4), 478–488. <https://doi.org/10.1093/elt/ccae052>
- Tang, K. S. (2024). Informing research on generative artificial intelligence from a language and literacy perspective: A meta-synthesis of studies in science education. *Science Education*, 108(5), 1329–1355. <https://doi.org/10.1002/sce.21875>
- Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14, 1261955. <https://doi.org/10.3389/fpsyg.2023.1261955>
- Xia, Y., Shin, S. Y., & Kim, J. C. (2024). Cross-cultural intelligent language learning system (CILS): Leveraging AI to facilitate language learning strategies in cross-cultural communication. *Applied Sciences*, 14(13), 5651. <https://doi.org/10.3390/app14135651>
- Ye, C., Zweck, E., Ma, Z., Smith, J., & Katz, S. (2024). Doctor Versus Artificial Intelligence: Patient and Physician Evaluation of Large Language Model Responses to Rheumatology Patient Questions in a Cross-Sectional Study. *Arthritis & Rheumatology*, 76(3), 479–484. <https://doi.org/10.1002/art.42737>
- Zaretsky, J., Kim, J. M., Baskharoun, S., Zhao, Y., Austrian, J. S., Aphinyanaphongs, Y., Gupta, R., Blecker, S. B., & Feldman, J. (2024). Generative Artificial Intelligence to Transform Inpatient Discharge Summaries to Patient-Friendly Language and Format. *JAMA Network Open*, 7(3), e240357. <https://doi.org/10.1001/jamanetworkopen.2024.0357>
- Zhang, Y., & Dong, C. (2024). Exploring the Digital Transformation of Generative AI-Assisted Foreign Language Education: A Socio-Technical Systems Perspective Based on Mixed-Methods. *Systems*, 12(11), 462. <https://doi.org/10.3390/systems12110462>
- Zou, R., Jiang, L., & Wider, W. (2025). Bibliometric Insights into the Open Education Landscape. *International Review of Research in Open and Distributed Learning*, 26(1), 283–309. <https://doi.org/10.19173/irrodl.v26i1.7953>
- Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, 18(3), 429–472. <https://doi.org/10.1177/1094428114562629>