Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multimodal Learning Strategies in English Language Teaching: A Systematic Literature Review of English-Speaking Skills Development

Lilian Guia Burgos

lgbburgos01202501907@usep.edu.ph

University of Southeastern Philippines https://orcid.org/0009-0005-0386-2284

Ben Cuamag

University of Southeastern Philippines https://orcid.org/0009-0003-3999-0013

Marvie Martizano

University of Southeastern Philippines https://orcid.org/0009-0006-8981-5944

Mennard Vigo

University of Southeastern Philippines https://orcid.org/0009-0002-7440-4430

Gwen Leanne Sison

University of Southeastern Philippines

Systematic Review

Keywords: language teaching, multimodal learning strategies, English-speaking skills, Artificial intelligence, systematic literature review

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11739282/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Background

The increasing integration of technology in education has transformed English language teaching, giving rise to multimodal learning strategies that combine text, visuals, audio, and interactive tools. While these approaches are widely adopted, there remains limited synthesized evidence on how they specifically influence students' English-speaking skills across educational levels. This study addresses this gap by examining the effectiveness of multimodal strategies in developing speaking proficiency.

Methods

Using a systematic literature review (SLR) design guided by PRISMA protocols, relevant studies published between 2020 and 2026 were retrieved from ProQuest and ScienceDirect databases. From an initial pool of 2,499 articles, eight studies met the inclusion criteria. These studies included participants from elementary, secondary, and tertiary levels and explored approaches such as translanguaging, augmented reality (AR), artificial intelligence (AI), extended reality (XR), and other multimodal strategies.

Results

Findings reveal that multimodal strategies generally improve students' speaking skills, particularly in terms of fluency, vocabulary, pronunciation, and confidence.

Conclusion

Translanguaging enhances participation and meaning-making, while AI-enhanced tools provide interactive and low-anxiety environments for speaking practice. However, some approaches show uneven effects, particularly in developing grammatical accuracy, and may be less effective when overused.

INTRODUCTION

Technology has transformed educational practices by expanding access to learning resources and enabling more flexible, interactive, and learner-centered instructional approaches. Digital platforms and the Internet have reshaped how knowledge is created, accessed, and shared, making technology an integral component of teaching and learning rather than merely a supplementary tool (Means et al., 2021; Almahasees et al., 2021; Haleem et al., 2022; Kumar et al., 2022). In higher education and other learning contexts, digital repositories, online platforms, and virtual learning environments have become central to knowledge construction, collaboration, and communication (Van Laar et al., 2017; Giannakos et al., 2022).

Within this digital transformation, multimodal learning has emerged as a prominent instructional approach in English Language Teaching (ELT). Multimodal learning integrates multiple modes of communication including text, images, audio, video, gestures, and interactive technologies to facilitate richer learning experiences and improve learner engagement (Blikstein & Worsley, 2016; Di Mitri et al., 2018). Digital technologies enable language teachers to combine these modes in meaningful ways, creating opportunities for learners to process information through multiple sensory channels while enhancing comprehension, retention, and motivation (Cowie & Sakui, 2021; Kessler, 2018; Al Rahmi et al., 2018). As a result, multimodal learning has become an increasingly important pedagogical approach in technology-enhanced language classrooms.

The growing adoption of multimodal learning has generated considerable interest in its potential to improve English language proficiency, particularly speaking skills. In English as a Foreign Language (EFL) and English as a Second Language (ESL) contexts, multimodal strategies such as video-based projects, digital presentations, interactive online tasks, and multimedia-rich instruction have been associated with improvements in speaking fluency, pronunciation, confidence, and learner motivation (Tan et al., 2022; Zhussupova et al., 2022; Cowie & Sakui, 2021; Sanako, 2024). These approaches also encourage active participation, collaborative meaning-making, and peer feedback, all of which contribute to meaningful oral communication practice (Klimova, 2021). Furthermore, technology-enhanced tools—including mobile applications and artificial intelligence (AI)-assisted platforms—provide learners with authentic, interactive, and low-anxiety environments for speaking practice while supporting individualized feedback and continuous language development (Haleem et al., 2022; Yu & Trainin, 2022).

Despite the increasing body of research on multimodal learning in ELT, existing evidence remains fragmented across diverse instructional approaches, educational settings, and learner populations. Individual studies have examined strategies such as translanguaging, augmented reality (AR), artificial intelligence (AI), extended reality (XR), and other multimodal approaches; however, there is limited synthesized evidence explaining how these strategies collectively influence English-speaking skills across different educational levels. A comprehensive synthesis is therefore needed to identify common instructional approaches, evaluate their reported effects on speaking development, and determine areas where evidence remains inconsistent.

Accordingly, this systematic literature review examines how multimodal learning strategies including translanguaging, augmented reality, artificial intelligence, extended reality, and other multimodal approaches are utilized in English Language Teaching to support speaking instruction. Guided by the PRISMA framework, the review synthesizes empirical studies published between 2020 and 2026 to identify the multimodal strategies employed, evaluate their reported effects on learners' English-speaking skills, and highlight implications for future research and instructional practice.

Purpose of the Review

The purpose of this systematic literature review is to examine how multimodal learning strategies are utilized in English Language Teaching (ELT) to support speaking instruction and to synthesize existing evidence on their effectiveness in improving students' English-speaking skills.

The research questions were raised through the following procedures of systematic literature review (SLR):

1. What types of multimodal learning strategies are used in ELT speaking instruction?
2. What findings are reported regarding students' English-speaking skills when using multimodal learning strategies?

Specifically, the study aims to provide a comprehensive understanding of the different multimodal approaches employed in language classrooms and how these approaches influence learners' speaking development across various educational contexts. This review is guided by the following objectives:

1. To identify and describe the different types of multimodal learning strategies used in ELT speaking instruction.
2. To analyze and synthesize the reported findings on students' English-speaking skills when multimodal learning strategies are implemented.

METHODS

Systematic Literature Review (SLR) serves as a method for synthesizing scientific evidence to address a specific research question in a manner that is both transparent and reproducible (Lame, 2019). They aim to encompass all published evidence related to the topic while also evaluating the quality of this evidence. As stated by Carrera-Rivera et al. (2022), its objective is to assess key aspects of existing knowledge regarding a topic related to research questions in order to propose areas for additional investigation. Moreover, SLR seeks to categorize, assess, classify, and produce realistic support that compiles predetermined suitability criteria to address and resolve a given research issue.

Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) is a technical procedure utilized to investigate, collect, and generate comprehensive relevant information from prior research in order to present the current state of the research. PRISMA data enables researchers to improve and expand the scope of paper evaluation and align it with the specific objectives of the research findings (Khan & Qureshi, 2020).

Searching Strategies

A strategy for searching was developed to identify pertinent articles for this systematic review: multimodal learning strategies, enhancing language speaking skills, basic education or multimodal instruction, multimodal strategies and multimodal instruction. The engine searching area: ProQuest (https://www.proquest.com/) and Science Direct (https://www.sciencedirect.com/) were the two

databases employed in search techniques. All searches encompassed journals, reviews, and conferences, which were solely accessible in English and spanned the timeframe from databases launched in 2020 to 2025.

Selection Standards

The primary aim of this systematic literature review was to find out the types of multimodal learning strategies in English Language Teaching and how it its reported findings to educational setting. The exploration was placed from 2020 through 2025. The analysis did not include articles published before 2020. Moreover, the investigation encompassed all nations globally. Consequently, a total of 2, 499 research articles were filtered at this stage. At this juncture, 39 records have been retrieved.

Assessment Quality

The study centers on recent research articles and conference papers. To guarantee the quality and significance of the educational material included in the analytical method, the abstracts of the publications were rigorously checked for assessment and refinement. Subsequently, all research pieces were comprehensively assessed. The subsequent rejection measure was to filter documents that were exclusively published in English. 15 of them were not fully text. Thus, they had been taken out of the study. Additionally, 16 articles were eliminated due to unclear methods used and articles not focusing on speaking skills. Lastly, there were 8 articles which were identified and taken to be reviewed.

Extraction of Data

The results of the study were available only in English and had been published and reviewed between 2020 and 2025. Throughout the review process, 39 papers were identified. These publications were examined to ascertain the objectives of the study. Among other aspects, the following terms were related to multimodal learning strategies, enhancing language speaking skills, basic education or multimodal instruction, multimodal strategies and multimodal instruction, as referenced.

RESULTS AND DISCUSSION

Based on the synthesized findings from the attached table, multimodal learning strategies, particularly translanguaging and AI-enhanced approaches demonstrate a generally positive impact on students' English-speaking skills across educational contexts. Among the eight eligible studies, two involved elementary learners, five focused on secondary students, and one on tertiary level, indicating that most empirical evidence is concentrated in secondary education.

The synthesis of these studies demonstrates a significant paradigm shift toward immersive and embodied learning environments that leverage both high-tech and socio-linguistic tools. High-tech

interventions such as Augmented Reality (AR), Virtual Reality (VR), and Human-Robot Collaboration (HRC) focus on creating authentic, interactive spaces. This aligns with the broader RRL regarding Situated Learning Theory, which posits that knowledge is best acquired when embedded in the activity, context, and culture in which it is used (Oto-Millera et al., 2025; Yan et al., 2025). By creating these digital "contact zones," these tools effectively reduce "inert knowledge" information that can be recalled but not applied and increase student agency. For instance, the use of a "visualized mirror-neuron learning system" in HRC (Liang & Hwang, 2025) provides an empirical bridge to neuro-educational research, suggesting that physical, embodied interaction maps social-cognitive processes more deeply than traditional rote memorization.

Table 1

Multimodal Approaches and Tools used in ELT Speaking Instruction

Study	**Core Multimodal Approach/Tool**	**Description and Application**
Oto-Millera et al. (2025)	Augmented Reality (AR)	Immersive technology used to improve vocabulary, listening, and fluency among gifted ESL/EFL learners.
Fernández García (2023)	Contemporary Music	Uses the highly emotional lyrics and narrative structures of country music to foster narrative and oral skills in adolescent EFL learners.
Larsson et al. (2025)	Sign-Supported Speech (SSS)	Integrating speech and sign simultaneously to support second language development of learners with special needs.
Yan et al. (2025)	AI and Extended Reality (XR)	Integrates AI (natural language processing, conversational agents) with XR (VR, AR, MR) to support speaking and intercultural competence.
Liang & Hwang (2025)	Human-Robot Collaboration (HRC)	Employs robot-assisted language learning (RALL) and digital storytelling to enhance persuasive communication and bodily-emotional awareness.
Rahman & Hu (2025)	Translanguaging & Trans-semiotizing	Leverages multiple linguistic and semiotic resources (gestures, visuals, and various languages) to construct knowledge in English-medium classrooms.
Nguyen (2025)	Transvocality & Decolonial Dialogue	Utilizes minoritized indigenous languages alongside the majority language to facilitate classroom participation and "voice" for indigenous students.
Granqvist (2024)	Literary Book-Group Discussions	Uses student-led discussions of novels to create "safe houses" for linguistic participation and social belonging in diverse classrooms.

Complementing these technological tools are pedagogical frameworks like translanguaging, trans-semiotizing, and the use of authentic cultural materials such as music. The RRL in this area (Rahman & Hu, 2025; Nguyen, 2025) emphasizes that language should be treated as a fluid, multimodal practice rather than a static set of rules. This supports the Sociocultural Theory of Mind, which suggests that

semiotic resources including gestures, emotional resonance, and mother tongues are essential psychological tools for cognitive development. Furthermore, the inclusion of "transvocality" and indigenous languages reflects a growing trend in decolonial RRL, where the goal of language education is not mere "fluency" in a dominant tongue, but the creation of "safe houses" (Granqvist, 2024) where learners can maintain their social belonging while acquiring new linguistic identities. By integrating these diverse resources from digital avatars to local indigenous languages the studies collectively advocate for a move beyond traditional grammar-centered instruction toward a more holistic, embodied, and socially-situated model of language acquisition.

Table 2

Findings on Students' English-Speaking Skills Using Multimodal Strategies

Article Title	Author(s) & Year	Findings on Students' English-Speaking Skills (via Multimodal / Translanguaging Strategies)
Translanguaging and trans-semiotizing in English-medium classrooms: Upholding university's policies or constructing knowledge?	Rahman, M. M. & Hu, G. (2025)	Multimodal strategies, including translanguaging, enhance speaking skills by improving learners' confidence, participation, comprehension, and fluency through linguistic and visual supports.
Teachers' use of sign-supported speech in interaction with multilingual children in Swedish preschools	Karolina Larsson, Polly Björk-Willén, Katarina Haraldsson, Kristina Hansson (2025)	As a multimodal strategy, sign-supported speech (SSS) demonstrates mixed effects, improving communication and participation but offering limited gains in verbal speaking skills and oral interaction.
Voicing decolonial dialogues: Indigenous teachers' translanguaging in the mainstream classroom	Trang Thi Thuy Nguyen (2025)	Translanguaging as a multimodal strategy enhances speaking development by improving comprehension, participation, engagement, confidence, interaction, and meaning-making, fostering more effective classroom communication.
Creating a safe house for active literary book-group discussions in a contact zone classroom	Angelica Granqvist (2024)	Student-led multimodal learning strategies enhance speaking skills by increasing participation, confidence, fluency, dialogic skills, meaningful speech, and overall communication competence.
Augmented Reality in English Language Acquisition Among Gifted Learners: A Systematic Scoping Review (2020–2025)	Millera, Ostin, Bustamante (2025)	Augmented reality (AR) enhances speaking skills by improving pronunciation, fluency, vocabulary, motivation, engagement, and speaking confidence while reducing anxiety, although grammar development remains limited.
Integrating Artificial Intelligence and Extended Reality in Language Education: A Systematic Literature Review (2017–2024)	Weijian Yan, Belle Li, Victoria L. Lowell (2025)	As a multimodal strategy, augmented reality (AR) improves speaking performance by enhancing pronunciation, fluency, vocabulary, motivation, and confidence while reducing speaking anxiety, though grammar gains remain limited.
Fostering EFL Learners' Narrative Skills Through Contemporary Music: The Case of Country Music	Antonio Fernández García (2023)	Contemporary music as a multimodal strategy, contemporary music enhances speaking skills by improving narrative organization, oral expression, engagement, and participation through authentic language exposure, although its effects on fluency and accuracy remain limited.

Article Title	Author(s) & Year	Findings on Students' English-Speaking Skills (via Multimodal / Translanguaging Strategies)
Embodied Social Agency in Human-Robot Collaboration for Promoting EFL Learners' Persuasive Communication and Bodily-Emotional Awareness: A Functional Linguistics Perspective	Jia-Cing Liang; Gwo-Jen Hwang (2025)	HRC-SDST as a multimodal strategy, HRC-SDST enhances speaking skills by improving persuasive communication, idea organization, emotional and nonverbal awareness, motivation, self-efficacy, participation, and speaking confidence.

The table presented implies that translanguaging as a multimodal strategy enhances speaking development by enabling learners to utilize their full linguistic repertoire. For instance, Rahman and Hu (2025) found that combining L1 and English improves learners' confidence, participation, and clarity of expression, while Nguyen (2025) similarly reported increased willingness to communicate and improved classroom interaction. These findings support the argument that translanguaging facilitates meaning making and reduces linguistic barriers, leading to more fluent oral production. This is consistent with broader literature, where translanguaging is shown to strengthen communicative competence and oral fluency in multilingual settings (Wei & García, 2022; Wei, 2017; Zhang et al., 2022).

Moreover, AI-enhanced multimodal approaches including augmented reality (AR), extended reality (XR), and conversational agents provide interactive environments that foster speaking practice. Studies by Millera et al. (2025) and Yan et al. (2025) indicate improvements in pronunciation, vocabulary, and fluency, as well as increased motivation and reduced anxiety.

These findings align with Huang et al. (2023), who emphasize that AI tools support real-time feedback and autonomous learning, contributing to improved speaking performance. However, both the reviewed studies and recent literature note limitations in grammar development and pedagogical integration, suggesting that AI tools are more effective for fluency than for linguistic accuracy.

Another important implication is the role of multimodal resources (e.g., gestures, visuals, music, embodied interaction) in supporting speaking skills. For example, Granqvist (2024) reported that multimodal, discussion-based environments enhance participation and dialogic skills, while Fernández-García (2023) found that music supports narrative speaking and engagement. These findings are supported by multimodality theory, which posits that communication is constructed through multiple semiotic modes that enhance meaning-making and discourse development (Kress, 2019; Jewitt et al., 2016). Recent study further confirms that multimodal instruction improves interactional competence and spoken discourse quality (Lim & Kessler, 2023).

However, the findings also reveal that not all multimodal strategies equally support speaking development. Larsson et al. (2025) found that sign-supported speech (SSS) has limited impact on verbal skills and may even hinder speaking when overused. This suggests that multimodal support should be used as scaffolding tools rather than substitutes for verbal interaction, reinforcing the need for balanced

instructional design. Additionally, affective factors such as confidence, motivation, and anxiety reduction emerge as critical mediators of speaking development. Several studies (e.g., Rahman & Hu, 2025; Millera et al., 2025) highlight that multimodal environments create safe and engaging learning spaces, increasing learners' willingness to communicate. This is supported by Dewaele and MacIntyre (2019), who found that positive emotional experiences significantly enhance speaking performance in second language learning.

Overall, the reviewed studies, together with recent literature from 2020 to 2026, consistently show that multimodal approaches especially those that combine translanguaging and AI-enhanced tools can effectively improve students' English-speaking skills. These approaches do not only support linguistic development, such as fluency and vocabulary, but also engage cognitive processes and address affective factors like confidence and motivation. In this way, speaking development is supported more holistically.

Limitations

This review is limited by the small number of included studies (n = 8), which may affect the generalizability of the findings. The evidence is also concentrated at the secondary level, with limited representation from elementary and tertiary contexts. Additionally, the inclusion of only English-language studies published between 2020 and 2025 may have excluded relevant research. Variations in study designs and multimodal approaches make direct comparisons difficult, while reliance on secondary data means findings depend on the quality and reporting of the original studies.

DISCUSSION

This systematic literature review examined the role of multimodal learning strategies in enhancing English-speaking skills among different educational contexts. Drawing from eight selected studies published from the last five years, the review provides converging evidence that multimodal approaches generally exert a positive influence on learners' speaking development. These strategies ranging from translanguaging practices and AI-enhanced tools to augmented and extended reality, music-based instruction, and embodied interaction collectively support improvements in key speaking components such as fluency, vocabulary, pronunciation, and communicative confidence.

A major pattern emerging from the synthesis is that multimodal learning is most effective when it promotes active engagement, interaction, and meaning-making. In particular, translanguaging practices enable learners to leverage their full linguistic repertoire resulting to reducing language anxiety and enhancing participation and clarity of expression. Similarly, AI and XR-based tools create interactive, low-anxiety environments that facilitate real-time speaking practice and immediate feedback, contributing to increased motivation and communicative competence. These findings suggest that multimodal strategies do not merely supplement instruction but can fundamentally reshape the conditions under which speaking skills are developed.

However, the review also highlights important limitations and uneven outcomes. While multimodal approaches consistently support fluency and engagement, their impact on grammatical accuracy remains less pronounced. Certain strategies, such as sign-supported speech, demonstrate conditional effectiveness and may hinder verbal development when over-relied upon. These findings emphasize the importance of pedagogical balance indicating that multimodal resources should function as scaffolds rather than substitutes for authentic verbal interaction.

CONCLUSION

Overall, the evidence points to a paradigm shift in English language teaching from traditional, form-focused instruction toward a more holistic, communication-oriented, and learner-centered approach. Speaking development is shown to be a multidimensional process shaped not only by linguistic input but also by cognitive engagement and affective factors such as confidence and motivation. Nevertheless, the relatively small number of studies, the concentration on secondary-level learners, and variations in methodological rigor suggest the need for further empirical validation. Despite these limitations, this review contributes to the growing body of literature by offering a synthesized understanding of how multimodal strategies can be effectively harnessed to support speaking proficiency in diverse learning contexts.

Recommendations

In light of the findings, several recommendations are proposed for educators, curriculum developers, and future researchers. For classroom practice, teachers are encouraged to adopt multimodal learning strategies in a purposeful and pedagogically grounded manner. Rather than relying solely on technological tools, educators should integrate multimodal resources with communicative language teaching principles to ensure meaningful interaction and sustained speaking practice. Translanguaging should be recognized as a valuable instructional approach that supports comprehension, participation, and confidence, especially in multilingual classrooms. Additionally, teachers should maintain a balanced use of multimodal supports ensuring that these tools enhance rather than replace opportunities for authentic verbal communication.

At the institutional level, schools and educational stakeholders should consider embedding multimodal approaches into curriculum design and language programs. This includes providing adequate training for teachers in the effective use of digital tools such as artificial intelligence, augmented reality, and interactive platforms. Investments in technological infrastructure and professional development are essential to maximize the potential of multimodal instruction. Furthermore, policies that promote inclusive and culturally responsive teaching practices such as the integration of learners' first languages can further strengthen speaking development.

For future research, there is a need to expand the scope and depth of investigations on multimodal learning in English language teaching. More empirical studies should examine the long-term effects of

multimodal strategies on speaking proficiency particularly in relation to grammatical accuracy and discourse competence. There is also a need for greater representation of tertiary-level learners and diverse educational contexts to enhance the generalizability of findings. Comparative studies exploring the relative effectiveness of different multimodal approaches as well as research employing rigorous experimental and mixed-methods designs are highly recommended. Finally, future studies should aim to establish clearer pedagogical frameworks that guide the optimal integration of multimodal strategies in language instruction.

By addressing these areas, future work can build on the current evidence base and further clarify how multimodal learning can be strategically implemented to support comprehensive and sustainable development of English-speaking skills.

DECLARATIONS

Author Contributions

L. G. Burgos: Conceptualization, Methodology, Writing- Reviewing and Editing; **B. Cuamag:** Writing-Original draft preparation, Writing- Reviewing and Editing; **G. L. Sison:** Data curation, Writing-Original draft preparation, Writing- Reviewing and Editing**; M. Martizano:** Visualization, Investigation, Writing-Reviewing and Editing; **M. Vigo:** Supervision, Conceptualization, Writing-Original draft preparation, Writing- Reviewing and Editing;

Funding

This research received no external funding

Ethical Approval

Not applicable.

Competing interest

The authors declare no conflicts of interest.

Data Availability

Data will be made available by the corresponding author on request.

Declaration of Artificial Intelligence Use

The authors declare that this article was prepared and written solely by human effort.

REFERENCES

1. Almahasees Z, Mohsen K, Amin MO (2021) Faculty's and students' perceptions of online learning during COVID-19. Front Educ 6:638470. https://doi.org/10.3389/feduc.2021.638470
2. Al-Rahmi WM, Alias N, Othman MS, Marin VI, Tur G (2018) A model of factors affecting learning performance through the use of social media in Malaysian higher education. Comput Educ 121:59–72. https://doi.org/10.1016/j.compedu.2018.02.010
3. Blikstein P, Worsley M (2016) Multimodal learning analytics and education data mining: using computational technologies to measure complex learning tasks. J Learn Analytics 3(2):220–238. https://doi.org/10.18608/jla.2016.32.11
4. Carrera-Rivera A, Ochoa W, Larrinaga F, Lasa G (2022) How-to conduct a systematic literature review: A quick guide for computer science research. MethodsX 9:101895. https://doi.org/10.1016/j.mex.2022.101895
5. Castañeda L, Selwyn N (2018) More than tools? Making sense of the ongoing digitizations of higher education. InternSSational J Educational Technol High Educ 15(1). https://doi.org/10.1186/s41239-018-0109-y
6. Cowie N, Sakui K (2021) Teacher and student-created videos in English language teaching. ELT J 75(1):97–102. https://doi.org/10.1093/elt/ccaa054
7. Dewaele JM, MacIntyre PD (2019) The predictive power of multicultural personality traits, learner and teacher variables on foreign language enjoyment and anxiety. In *Evidence-based second language pedagogy*(pp. 263–286). Routledge. https://eprints.bbk.ac.uk/id/eprint/30105/3/30105.pdf
8. Di Mitri D, Schneider J, Specht M, Drachsler H (2018) From signals to knowledge: A conceptual model for multimodal learning analytics. J Comput Assist Learn 34(4):338–349. https://doi.org/10.1111/jcal.12288
9. Fernández-García A (2023) EFL learners' narrative skills: the case of contemporary country music. Miscelánea J Engl Am Stud 67:41–69. https://doi.org/10.26754/ojs_misc/mj.20237385
10. Giannakos M, Spikol D, Di Mitri D, Sharma K, Ochoa X, Hammad R (2022) Introduction to multimodal learning analytics. Multimodal Learn Analytics Handb 3–28. https://doi.org/10.1007/978-3-031-08076-0_1
11. Granqvist A (2024) Creating a safe house for active literary book-group discussions in a contact zone classroom. Linguistics Educ 83:101335. https://doi.org/10.1016/j.linged.2024
12. Haleem A, Javaid M, Qadri MA, Suman R (2022) Understanding the role of digital technologies in education: A review. Sustainable Oper Computers 3:275–285. https://doi.org/10.1016/j.susoc.2022.05.004
13. Huang H, Zheng O, Wang D, Yin J, Wang Z, Ding S, Yin H, Xu C, Yang R, Zheng Q, Shi B (2023) ChatGPT for shaping the future of dentistry: the potential of multi-modal large language model. Int J Oral Sci 15(1). https://doi.org/10.1038/s41368-023-00239-y
14. Jewitt C, Bezemer J, O'Halloran K (2016) *Introducing multimodality*. https://doi.org/10.4324/9781315638027

15. Kessler G (2018) Technology and the future of language teaching. Foreign Lang Annals 51(1):205–218. https://doi.org/10.1111/flan.12318
16. Khan N, Qureshi MI (2020) A systematic literature review on online medical services in Malaysia. Int J Online Biomedical Eng 16(6):107–118. https://doi.org/10.3991/ijoe.v16i06.13573
17. Klimova B (2021) An insight into online foreign language learning and teaching in the era of COVID-19 pandemic. Procedia Comput Sci 192:1787–1794. https://doi.org/10.1016/j.procs.2021.08.183
18. Kress G (2019) Transposing meaning. In *Translation and Multimodality*(pp. 24–48). Routledge. https://doi.org/10.4324/9780429341557-2
19. Kumar P, Garg RK, Kumar P, Panwar M (2022) Teachers' perceptions of student barriers to sustainable engagement in online education. Int J Knowl Learn 15(4):373. https://doi.org/10.1504/ijkl.2022.126273
20. Lame G (2019) Systematic literature reviews: An introduction. *Proceedings of the Design Society: International Conference on Engineering Design*, *1*(1), 1633–1642. https://doi.org/10.1017/dsi.2019.169
21. Larsson K, Björk-Willén P, Haraldsson K, Hansson K (2025) Teachers' use of sign-supported speech in interaction with multilingual children in Swedish preschools. Linguistics Educ 87:101398. https://doi.org/10.1016/j.linged.2025.101398
22. Liang J, Hwang G (2025) Embodied social agency in human-robot collaboration for promoting EFL learners' persuasive communication and bodily-emotional awareness: A functional linguistics perspective. Educ Inform Technol 30(17):24291–24335. https://doi.org/10.1007/s10639-025-13747-7
23. Lim J, Kessler M (2023) Multimodal composing and second language acquisition. Lang Teach 57(2):183–202. https://doi.org/10.1017/s0261444823000125
24. Means B, Peters V, Neisler J, Wiley K, Grifths R (2021) *Lessons from remote learning during COVID-19*. Digital Promise. https://doi.org/10.51388/20.500.12265/116
25. Nguyen TTT (2025) Voicing decolonial dialogues: Indigenous teachers' translanguaging in the mainstream classroom. Linguistics Educ 88:101440. https://doi.org/10.1016/j.linged.2025.101440
26. Oto-Millera N, Pellicer-Ortín S, Bustamante JC (2025) Augmented reality in English language acquisition among gifted learners: A Systematic scoping Review (2020–2025). Appl Sci 15(21):11487. https://doi.org/10.3390/app152111487
27. Rahman MM, Hu G (2025) Translanguaging and trans-semiotizing in English-medium classrooms: Upholding university's policies or constructing knowledge? Linguistics Educ 87:101424. https://doi.org/10.1016/j.linged.2025.101424
28. Tan KH, Rajendran A, Muslim N, Alias J, Yusof NA (2022) The potential of TikTok's key features as a pedagogical strategy for ESL classrooms. Sustainability 14(24):16876. https://doi.org/10.3390/su142416876
29. Van Laar E, Van Deursen AJ, Van Dijk JA, De Haan J (2017) The relation between 21st-century skills and digital skills: A systematic literature review. Comput Hum Behav 72:577–588.

https://doi.org/10.1016/j.chb.2017.03.010

30. Vázquez-Cano E, Bernal-Bravo C, Ballesteros-Regaña C, López-Meneses E (2024) Exploring the role of interactive games in improving Spanish spelling skills: A Mobile App study. Edmetic 14(1):4. https://doi.org/10.21071/edmetic.v14i1.17374
31. Wei L (2017) Translanguaging as a practical theory of language. Appl Linguist 39(1):9–30. https://doi.org/10.1093/applin/amx039
32. Wei L, García O (2022) Not a first language but one repertoire: translanguaging as a decolonizing project. RELC J 53(2):313–324. https://doi.org/10.1177/00336882221092841
33. Yan W, Li B, Lowell VL (2025) Integrating Artificial intelligence and Extended Reality in language Education: A Systematic Literature Review (2017–2024). Educ Sci 15(8):1066. https://doi.org/10.3390/educsci15081066
34. Yu A, Trainin G (2022) A meta-analysis examining technology-assisted L2 vocabulary learning. ReCALL 34(2):235–252. https://doi.org/10.1017/S0958344021000239
35. Zhang Q, Lin X, Osborne C (2022) A think-aloud method of investigating translanguaging strategies in learning Chinese characters. Appl Linguistics Rev 15(4):1453–1476. https://doi.org/10.1515/applirev-2022-0135
36. Zhussupova R, Amrenova A, Ayazbayeva A, Kassimbekova N (2022), October 10 Designing multimodal texts in developing speaking skill for pre-service EFL teachers. *Proceedings of ADVED 2022- 8th International Conference on Advances in Education* (pp. 225–235). OCERINT. https://doi.org/10.47696/adved.202231

Figures

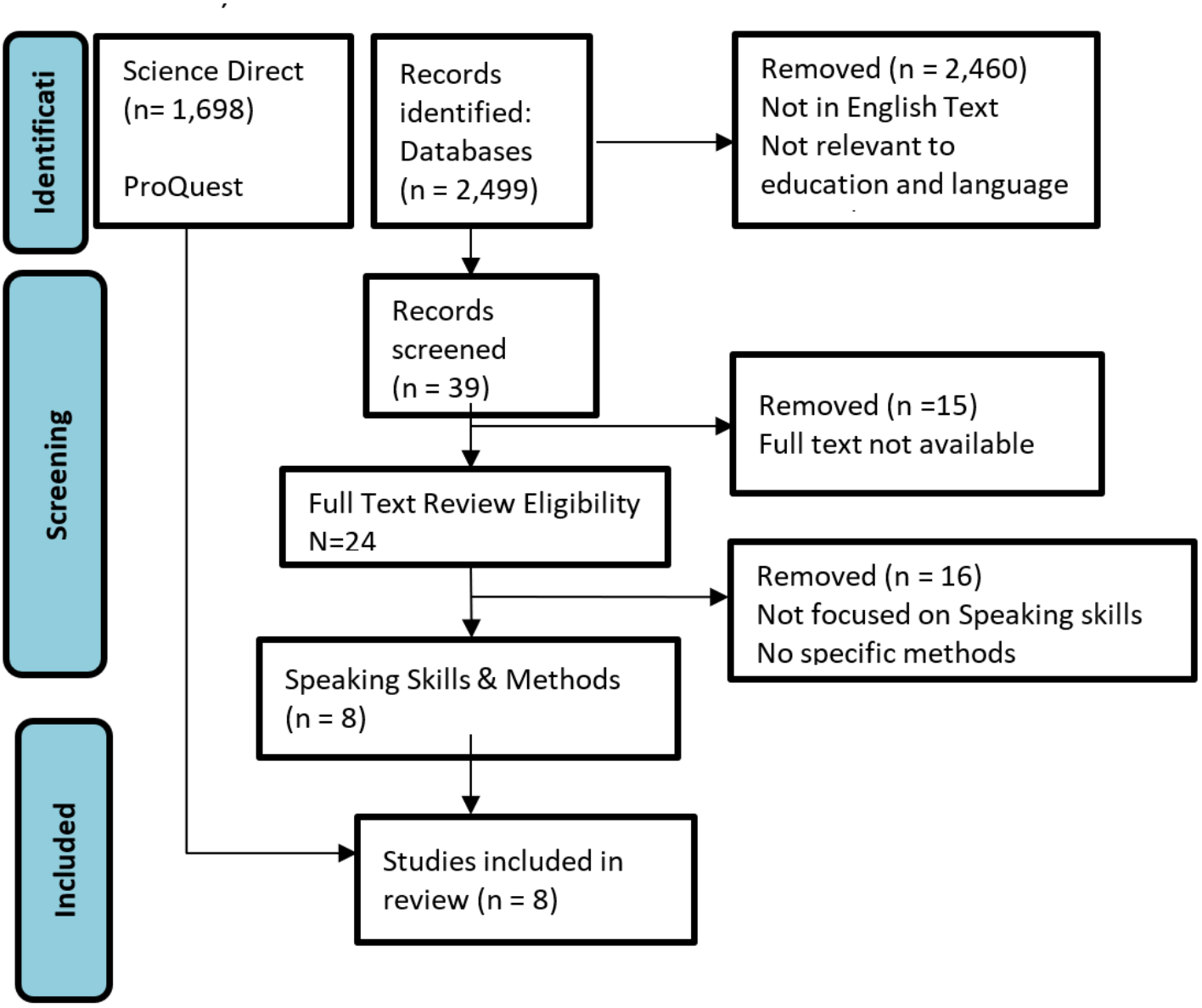


Figure 1

Systematic review of the related literature

Source: Page MJ, et al. BMJ 2021;372:n71. doi: 10.1136/bmj.n71.