Research article

Multimodal immersion in English language learning in higher education: A systematic review

I Wayan Eka Dian Rahmanu [a,*], Gyöngyvér Molnár [b]

[a] *Doctoral School of Education, University of Szeged, Hungary*
[b] *Institute of Education, MTA–SZTE Digital Learning Technologies Research Group, University of Szeged, Hungary*

A B S T R A C T

The development of multimodality has led to increased research on its use to improve students' English language competency. However, no recent review has analyzed multimodality in English language learning in higher education. This systematic review examines 34 research articles published from 2013 to 2024. The primary focus of the study is to explore the application of multimodal pedagogies in higher education, the methods and materials used to assist learners in acquiring English language skills, the English language skills acquired through the usage of multimodality, and the main results of using multiple modes. This systematic review employs the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) standards. It adopts a thorough search strategy across electronic databases, which include Web of Science and Scopus. We found (1) the implementation of multimodality contributes to learners' English language proficiency in English for academic purposes (EAP) and English for specific purposes (ESP) education; (2) there is a predominance of digital multimodality and nonverbal communication use in the higher education classroom, for example, gesture, kinesics, spatial position, facial expression, and gaze the use of the 3D (dimensional) environment, virtual reality (VR); (3) there is an advantage of a multimodal approach in improving higher education learners' vocabulary, reading, speaking, and writing skills and a positive connection between the implementation of multimodality and the development of learners' communicative ability. This systematic review highlights existing research gaps and outlines potential avenues for future investigation aimed at conceptualizing and assessing learners' skills through multimodal approaches.

1. Introduction

Educators create learning media by integrating various resources with a multimodal approach [1,2]. The use of multimodality contributes to significant improvement in English language competency. The implementation of multimodality has noteworthy results on English language learners' mastery of vocabulary [3], their writing skills [4], and their oral comprehension [5]. Using learning resources that combine visual, aural, and interactive elements provides students more opportunities to absorb and retain new words. Students are able to acquire a more profound understanding of the norms and subtleties of written English, which in turn improves their capacity to successfully express themselves in written form. Learners are given the opportunity to practice their spoken English in

---

* Corresponding author.
  E-mail addresses: i.wayan.eka.dian.rahmanu@edu.u-szeged.hu (I.W.E.D. Rahmanu), gymolnar@edpsy.u-szeged.hu (G. Molnár).

a variety of circumstances.

Prior studies have employed multimodality in developing learners' language competency. Multiple modalities have significantly enhanced English language learners' development [6]. Use of multimodality is beneficial to improving aspects of English language learners' competency, for example, vocabulary mastery [7], writing skills [8], and speaking skills [9]. The integration of visual aids and auditory skills has been aimed at enriching vocabulary competency for English language learners [10]. Including interactive exercises expands their lexical development [11]. Similarly, students' English language oral ability can be significantly developed by integrating audio-visual media [3], role play, and interactive dialogues, which draw on the semiotic resources of facial expressions, gestures, and proxemics [12]. In addition, type of educational media has evolved to boost the use of multiple modes through numerous tools, for example, YouTube [13], PowerPoint slides [14], a 3D environment and VR [15], virtual worlds [16], and digital games [7].

The analysis of multimodality in English language learning is limited, especially in the context of higher education. Previous reviews have only focused on multimodal literacy analysis and the use of a multimodal approach in primary and secondary schools [17, 18]. Although studies have elaborated on multimodality, including the visual, gestural, spatial, and linguistic modes [19], it was only a concern in younger learners' development. In addition, prior review studies have failed to provide a thorough treatment of the most recent developments in technology, such as VR, and how it may be integrated into multimodal literacy with the potential of enhancing learners' communication skills. To address the broadened concept of literacy, educational researchers and teachers in higher education in many countries have studied different multimodal pedagogies for multimodal literacy learning in the English language classroom. This study aims to synthesize evidence on the nature of multimodal pedagogy and multimodal literacy learning use in higher education in various places around the world and to suggest implications and research directions to advance the integration of multiple modalities in literacy education. Our systematic review concerns how multimodality is implemented in higher education and is guided by the following research questions.

RQ1. Which courses or programs in higher education use multimodality to enhance English language learning?

RQ2. What modes and digital media have been used to assist learners in acquiring English language skills?

RQ3. What aspects of English language learning have been improved through multimodal teaching?

2. Previous systematic reviews

Multimodal immersion integrates many modes of communication and media to improve learning experiences. It includes visual, aural, gestural, spatial, and linguistic modalities, resulting in a rich and dynamic learning environment [19]. Several systematic literature reviews have previously explored multimodality in teaching and learning activities. According to research conducted by Kulju et al. (2018) and Lim et al. (2022), previous systematic literature reviews have involved multimodal pedagogy in primary and secondary classrooms. The focus was the nature of multimodal literacy and multimodal pedagogy. Earlier studies stressed text production with information and communication technology (ICT) and multimodal assessment within an educational framework.

The first systematic review was conducted by Kulju et al. (2018) [17], who analyzed studies on multimodal pedagogies in elementary classrooms published from 1997 to 2014. The focus of the study was reading and multimodal literacy in primary classrooms. The researchers narrowed the number of 338 articles to a sample of 67 manuscripts. The main topic of the manuscripts was divided into several classifications involving writing and text production, ICT, reading and analysis texts, classroom practices, and significant aspects for further discussion. Additionally, learners generated a range of multimodal texts, including digital videos, television advertisements, and websites using ICT. The findings also revealed that most of the reviewed studies included classroom practices with instructional support. In the second systematic review, similar findings, such as a range of multimodal texts used by teachers, were reported by Lim et al. (2022) [18]. The study was based on manuscripts published between 2010 and 2021. The researchers restricted the number of 890 manuscripts to a sample of 98 articles. In addition, their study covers a broader range of primary and secondary classrooms. The research is concerned with the engagement of multimodal texts. Furthermore, it scrutinized critical, creative, culturally responsive multimodal pedagogies and multimodal assessment. Moreover, the explicit teaching of multimodal literacy, a focus on the effects of multimodal learning, and assessment were also analyzed. The authors showed that multimodal texts, for example, digital games, digital books, poetry, comics, graphic novels, digital photo stories, internet websites, and advertisements, have been used by teachers to engage learners' different interests in classroom activities. Further, critical, creative, and culturally responsive multimodal pedagogies are used by teachers in the English language classroom. In the explicit teaching of multimodal literacy, teachers implemented whole-class, small cohort, and independent activities to teach how to grasp the text type. According to the authors' findings on the attention effect in multimodal learning, the text from learners' life world, for example, comics, trending videos, social media, and advertisements, are frequently integrated to involve learners in their literacy learning. Additionally, assessing multimodal learning is pivotal for the maturation of learners' literacy in classrooms.

Previous reviews have predominantly focused on primary and secondary education, overlooking the challenges and opportunities present in higher education. For instance, Kulju et al. (2018) and Lim et al. (2022) concentrated on multimodal literacy in younger learners, leaving a gap in understanding how these pedagogies translate to adult learners in higher education. In addition, these reviews have failed to provide extensive coverage of the latest advancements in technology, for example, VR, how they can be integrated into multimodal literacy, and what their potential is for improving communication skills [15]. The potential of VR to enhance communication skills through immersive and interactive experiences represents a critical area that remains underexplored in the context of higher education. By incorporating VR and other emerging technologies into multimodal literacy practices, educators can create more engaging and effective learning experiences that cater to the diverse needs of learners [20]. Addressing these shortcomings

will aid in improving the field and promoting the implementation of more effective, evidence-based teaching approaches.

3. Method

This systematic review follows the PRISMA guidelines [21]. The study is also guided by the stages of research synthesis [22], including (a) inclusion and exclusion criteria, (b) specifying search terms, (c) thematic analysis, data extraction, and risk of bias, and (d) synthesis.

3.1. Inclusion and exclusion criteria

The paper selection method for this study follows strict inclusion criteria to guarantee a thorough and focused investigation of the pertinent literature. Manuscripts needed to be published between 2013 and 2024 to be evaluated in order to provide current perspectives. Selected publications were required to focus primarily on implementing multimodality in higher education. Thus, this study focuses on English-language publications, which must present empirical findings and highlight the importance of evidence-based research in studying multimodality in higher education. Both qualitative and quantitative studies are included to provide a comprehensive understanding of multimodal immersion in higher education. The studies are set within higher education institutions to ensure relevance to the targeted educational context. These criteria are crucial benchmarks that have directed the selection process and guaranteed the relevance and thoroughness of the chosen material.
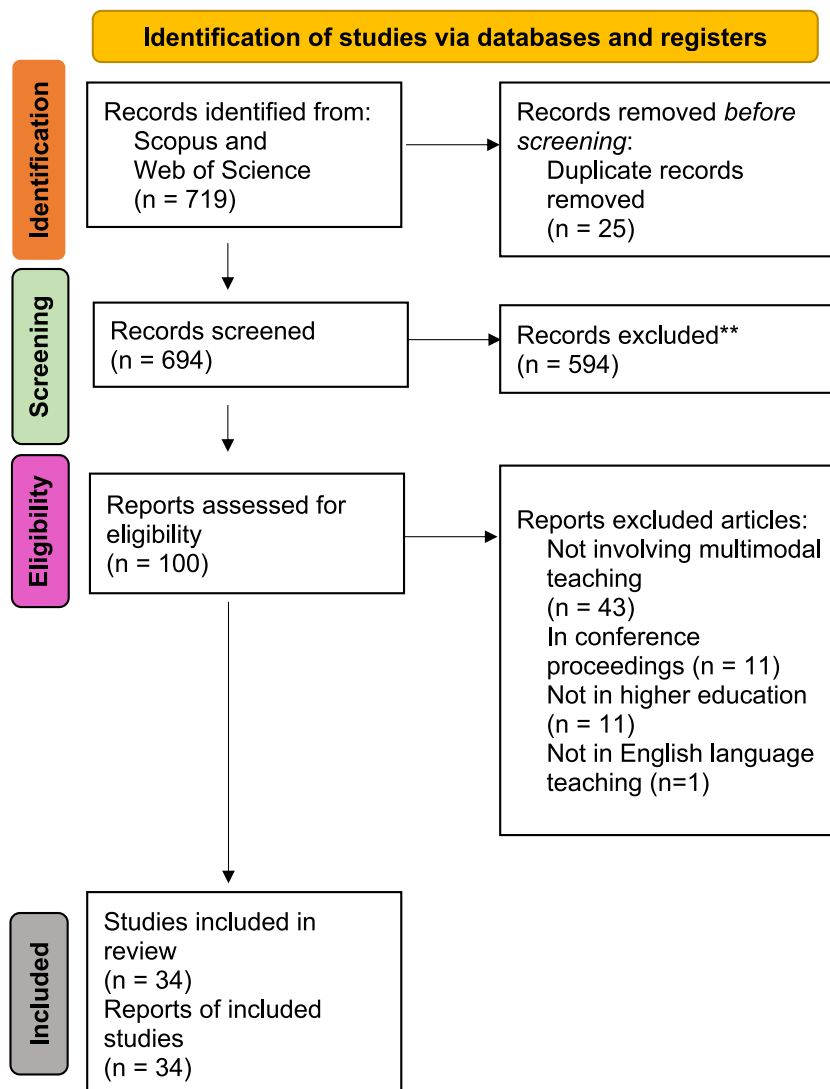

Fig. 1. Stages of the systematic review.

The exclusion criteria used in the selection process have been designed to narrow the scope of the literature review by removing publications that do not match the particular focus on multimodal pedagogies in higher education. Such manuscripts are intentionally not included to ensure the research is thematically consistent and relevant to the objectives. Manuscripts in languages other than English are not included to maintain linguistic consistency and facilitate a more cohesive analysis. Studies conducted in preschool and elementary school settings have been removed to focus on the distinct dynamics and problems of multimodal implementations in higher education environments. Articles from conference proceedings are not considered in order to prioritize peer-reviewed studies, thus enhancing the rigor and credibility of the selected material. In addition, reviews, editorials, opinion pieces, and theoretical papers have been excluded, as they do not provide empirical data on multimodal immersion. Further, duplicate studies have been removed to avoid redundancy and ensure each study is only considered once.

3.2. Search strategy and filtering process

Potentially relevant studies were identified by the Scopus and Web of Science databases. The keywords we used were "multi-modality," "multimodal," "higher education," "English," and "English language." The searches were combined with a Boolean approach that included "AND" and "OR" for greater flexibility in our search. The search thread for Scopus was "(multimodal OR multimodality) AND (English language) AND (higher education) NOT dentistry NOT neuroscience NOT biochemistry NOT genetics NOT molecular biology." The search thread for Web of Science was "(multimodal OR multimodality) AND (education) AND (English language) AND (higher education) NOT bioscience NOT behavioral science."

With regard to the filtering process for titles and abstracts, the identification conducted by the two authors resulted in 719 documents, which were then downloaded from the databases into the RIS type. The RIS document was imported by the authors to Rayyan as the online platform used in this study. Rayyan was employed to screen the titles and abstracts. During the process, Rayyan detected 25 duplicates and deleted them. Subsequently, 694 manuscripts were reviewed by the authors independently based on the inclusion and exclusion criteria and eliminated irrelevant topics. 594 unrelated studies were excluded during the title and abstract filtering process. The eligibility step involved 100 articles, which were then processed according to the inclusion criteria. The remaining manuscripts were independently reviewed based on the qualification criteria. As a result, 66 articles were excluded, as they consisted of irrelevant data (see Fig. 1). Following this, 34 manuscripts met the criteria for inclusion in the last review.

3.3. Thematic analysis, data extraction, and risk of bias

The systematic review involved a rigorous process of extracting data from the chosen articles, which was then systematically organized along specific dimensions in order to facilitate a thorough analysis. The initial dimension concerns geographic dispersion, specifically emphasizing the nation or region of origin for every chosen article. This classification offers valuable perspectives on integrating multimodality into tertiary education worldwide. The second dimension entails categorizing the articles according to the domains of higher education, thereby providing a more comprehensive view of the various academic fields in which multimodal pedagogies have been implemented.

The third dimension pertains to the different modes used in multimodality, providing clarification on the wide range of approaches and methodologies that are included in this educational framework. Simultaneously, the fourth dimension covers the various forms of educational media employed in tandem with multimodal approaches, providing an opportunity for an in-depth analysis of the technological and media-driven instruments incorporated into practices used in higher education.

In addition, the systematic review examined the aspect of acquiring English language skills via multimodal instruments in an effort to determine the way in which multimodality contributes to the improvement of language proficiency. The final dimension consists of the primary outcomes that arise from the implementation of multimodal approaches as learning media. These outcomes offer valuable insights into the educational advantages and effectiveness of multimodal methodologies. By incorporating these dimensions into the analysis of the data extracted from the chosen articles, a nuanced and comprehensive examination has been achieved, providing a multidimensional perspective on the numerous facets of multimodality in higher education.

A risk-of-bias assessment was conducted by considering various factors, including but not limited to disclosure, blinding, results with evaluators being blinded, incomplete data results, selective data reporting, and others [23]. The identified factor in this study is selective reporting, specifically, the tendency of journals to publish only those articles that are deemed significant. A concern arises that a systematic review may overestimate because of the higher probability of including significant studies in a systematic review compared to their unpublished counterparts [24].

3.4. Synthesis

A summary in Table 1 presents the results of the synthesis based on the categories in the coding process and data extraction. As for the synthesis process, the results were painstakingly discussed by the researchers. The summary highlighted the essential patterns and developments we identified in our analysis.

4. Results

The reviewed research has been systematized based on the following assertions: (1) instructors in various fields in higher education use multimodality to enhance English language learning, (2) modes and digital media are used to assist learners in acquiring English

Table 1
List of studies reviewed on multimodality in English language learning in higher education.

	Author/year	Year	Country	Fields of higher education	Modes used	Type of educational media	English language aspects improved	Main outcomes
1	Xiaodong Zhang [12]	(2021)	China	University level (argumentative writing course)	Linguistics (written text) and audio	Drama-based pedagogy and role play	Argumentative writing, including grammar accuracy	Students were able to re-conceptualize writing as a meaning-making process instead of merely engaging with writing at the structural level.
2	Valeria Galimberti, Joan C. Mora, Roger Gilabert [3]	(2023)	Spain	University level (English language program)	Linguistics (lexical), visual modes (videos), and audio	Television series	English vocabulary pronunciation (Phonolexical)	Students enhanced their pronunciation positively.
3	Juan Wang, Yuan Gao, Yaqiong Cui [14]	(2023)	China	University level (Science program)	Linguistics (lexical), gestures, and visual modes (videos)	PowerPoint or slides	Speaking in an academic presentation	The introduction of gesture improved learners' EAP in presentations.
4	Ya-Chun Shih [15]	(2014)	Taiwan	University level (TESOL, language and literature, and arts and design)	Linguistics (text-based), audio, video, and virtual objects	3D virtual environment and VR	Speaking skills	Using a virtual environment helped EFL learners to achieve their communication goals in real time.
5	You Jin Kim, Diane Belcher, Carter Peyton [25]	(2023)	Korea	University level (EFL program)	Linguistics (text-based) and images	DMC tasks and storyboards	Writing skills	Learners writing developed significantly.
6	Francesca Coccetta [9]	(2018)	Italy	University level (English language program, civilization and science of language)	Linguistics (procedural text), visual, audio, image, and animation	LEGO instructional manual	Academic speaking and daily conversation	Students were better able to develop texts in English.
7	Lianjiang Jiang [8]	(2018)	China	University level (engineering, educational technology, Japanese studies with an intermediate proficiency in English)	Linguistics (script-writing), audio and video recordings	DMC, storyboarding, Corel video studio and Movie maker	Writing skills	The study contributed to the learners' motivation to write in English.
8	Ali Alalem [26]	(2023)	USA	University level (Research-focused composition course)	Linguistics (script-writing), audio, visual	Digital storytelling, computer, digital projector, speaker and whiteboard	Writing and reading skills	Learners enhanced their multimodal awareness through digital storytelling.
9	Yiqiong Zhang, Kay L. O'Halloran [27]	(2019)	Asia	University level (English language and literature)	Linguistic, symbolic and visual	Research article and PowerPoint slides	Academic writing	A deeper understanding of written academic texts can be achieved by transforming a verbal text into multimodal form.
10	Vicent Beltran-Planques, Mercedes Querol-Julian [5]	(2018)	Spain	University level (ESP in video game design and Development)	Semiotic resources (spoken language, gesture, facial expression, head movement and gaze)	Role play	Speaking skills	Multimodality stimulates learners' eagerness to use semiotics in activities.
11	Maria Sabat e-Dalmau, Balbina Moncada-Comas [2]	(2023)	Spain	University level (agriculture and food engineering)	Visual, images, kinetics (gestures or gaze), positionality (face making), speech, and semiotics	Face-to-face, PowerPoints and Google Images	Speaking and writing skills	Multimodality acknowledged learners' linguistic diversity, and the tools motivated learners to engage in the EMI classroom.
12	Helen Lee, Regine Hampel, Agnes Kukulska-Hulme [28]	(2019)	UK	University level (English language program)	Visuals (video) and gestures	Skype video conferencing	Speaking and lexical skills	Learning media have the potential to stimulate learners' oral expression.

Table 1 (*continued*)

	Author/year	Year	Country	Fields of higher education	Modes used	Type of educational media	English language aspects improved	Main outcomes
13	Xiao Tan [13]	(2023)	USA	University level (first-year composition course)	Linguistic (citations and written text), visual modes (video and images)	Vox videos, YouTube	Writing	Students were able to use attribution, generate examples and establish links between sources.
14	Olga Kozar [29]	(2015)	Russia	Adult non-formal language education	Visuals and audio	Skype audio and video conferencing	Speaking and grammar skills	Comprehensive teaching and learning activities had the potential to stimulate learners' English language speaking and sentence structure.
15	I-Chung Ke, Hilda Cahyani [6]	(2014)	International (Taiwan and Indonesia)	University level (reading and general English course)	Linguistics (text through e-mail), image and video	Online communication (e-mail, MSN messenger and Facebook)	Writing and grammar skills	Social media could be beneficial for teachers as a tool to enhance learners' English as a lingua franca (ELF) ability.
16	Yo-An Lee, John Hellermann [30]	(2020)	Korea	University level (English language course)	Video and audio	Storytelling	Speaking skills	Although participants were able to speak, speaking assessment was conducted through storytelling.
17	Katarzyna Bromberek-Dyzman, Katarzyna Jankowiak, Paweł Chełminiak [31]	(2021)	Poland	University level (English language course)	Linguistics (text), audio, audio-video	Movies, TV, podcasts, and YouTube	Speaking skills	Multimodality impacted learners' speed and accuracy.
18	Maria Grazia Sindoni [32]	(2021)	Italy	University level (foreign language and literature, public and political communication)	Video calls, spoken and written turns, gaze, kinesics, proxemics and lexis	Skype video conferences, video conversation and academic text	Speaking and writing skills	Video-mediated interaction affected learners' ability to engage in conversation and write an academic text.
19	Teresa Morell [33]	(2018)	Spain	University level (economics, computer technology, architecture, anatomy, nursing, organic chemistry, civil engineering, ecology, educational psychology, history, translation, and English studies)	Linguistics (written text), gaze and gesture	Face-to-face and PowerPoint	Speaking and writing skills	Gaze and gesture had a significant impact on L2 students' comprehension.
20	Lianjiang Jiang, Jasmine Luk [34]	(2016)	China	University level (law, education, engineering and tourism management)	Linguistics, gesture and gaze	Digital video, documentaries and TV programs	Speaking and writing skills	Attractive linguistic meaning, easy-to-spot pronunciation through the video, exciting writing activities through multimodal composing.
21	Mercedes Querol-Julian [35]	(2023)	Spain	University level (business administration program)	Gesture and communicative modes	Website conferencing system (Adobe Connect)	Speaking skills	A digital environment affected learners' participation.
22	Mimi Li [36]	(2020)	USA	University level (TESOL program)	Linguistics (phonetics/phonology, morphology, syntax, semantics and pragmatics), audio, and video	Desire2Learn Brightspace, face-to-face, and textbook (Introducing English Linguistics)	TESOL education	Multimodality enhanced learners' content and collaborative learning.
23	Frantisek Tuma, Tamah Sherman [37]	(2022)	Czech Republic	University level (economics and fine arts)	Utterances, gaze, and gestures	Nil	Speaking skills	Seating arrangement through gesture played an important role in managing students' speaking activities.

Table 1 (*continued*)

	Author/year	Year	Country	Fields of higher education	Modes used	Type of educational media	English language aspects improved	Main outcomes
24	Carolina Girón-García, Inmaculada Fortanet-Gómez [38]	(2023)	Spain	University level (legal English, English for finance, English for science and technology, English for psychology, and English for medicine)	Colors, images, gestures, sound and music; visual prompts: visual signs	YouTube: WHO: influenza, an unpredictable threat	ESP skills	The tutor tended to use video to engage learners' prior knowledge on the topic presented in the course.
25	Amirul Mukminin, Akhmad Habibi, Muhaimin Muhaimin, Marzul Hidayat [39]	(2023)	Indonesia	University level (pre-service English teachers)	Text, video, and audio	Social media (Line, Facebook, i-Map, Instagram, YouTube, and Twitter)	Writing skills	Social media usage was beneficial in assisting learners and teachers in conducting the writing lesson.
26	Xiao Tan [40]	(2023)	USA	University level (first-year composition course)	Images, text vocabulary and audio-visual	Video project	Writing skills	The use of a video project led learners to be more innovative in finishing a writing task.
27	Julian ChengChiang Chen, Sarah Kent [16]	(2020)	Australia	University level (English support program)	3D environment, image, texts and gamification	3D MUVE virtual world and English language task-based	Speaking and writing skills	3D MUVE virtual stimulated learners' interest in learning, and it improved their speaking ability.
28	Matt Kessler [41]	(2020)	USA	University level (applied linguistics and teacher education programs)	Text, color font, color underlining, and color highlighting	Shift-it, Wikipedia, YouDao, Byu's COCA, and digital notes	Writing skills	Digital tools affected L2 learners' writing process.
29	Regina Kaplan-Rakowski, Barbara Loranc [10]	(2019)	Poland	University level (English studies Program)	Text, image, and audio	PowerPoint (slides), pronunciation, and sound effects	Vocabulary skills	Sound effects had a significant positive impact on learners' vocabulary learning.
30	Jian-E Peng [42]	(2019)	China	University level (clinical medicine, journalism and communication, mechanical and electrics, business management)	Audio, video, visual design, voice/facial expression, gesture, and spatial position	Audio/video, PowerPoint (slides)	Communication skills	The use of audio and video effectively has the potential to improve learners' willingness to communicate.
31	Jian-E Peng [43]	(2019)	China	University level (multidisciplinary program)	Kinesthetic (facial expression, gestures, movements), auditory and visual	English songs, videos or visual images, and PowerPoint slides	Communication skills	Modes contribute to learners' active speaking in English.
32	Anisa Cheung [44]	(2022)	Hong Kong	University level (EAP program)	Video, images, auditory and text	Audio/video and Wikipedia	Writing skills	Multimodal adoption was beneficial to open-ended discussions on the content of learners' presentations.
33	So-Yeon Ahn [45]	(2021)	Korea	University level (economics, English linguistics, Chinese education, business, and materials engineering programs)	Linguistics (narrative) and images	Drawings and short descriptions	Speaking skills	Drawing and narrative were employed as effective classroom activities in terms of achieving English language learning.
34	Teresa Morell, Vicent Beltra n-Palanques, Natalia Norte [11]	(2022)	Spain	University level (EMI teaching teacher training program)	Linguistics (writing), speech, nonverbal communication kinetics and spatial mode (position)	Face-to-face and classroom presentation	Verbal and nonverbal communication skills	The learners were effectively engaged through the use of multimodal interaction in the classroom.

language skills, and (3) aspects of English have been improved through multimodal teaching.

4.1. Higher education courses or programs use multimodality to enhance English language learning

Particular higher education courses and programs have implemented multimodal pedagogy in teaching and learning activities. A plethora of previous researchers or lecturers at universities employ multimodality in developing their learners' language skills (see Table 1).

Fig. 2 indicates that multimodality is used beneficially in EAP and ESP instruction. Specifically, a multimodal analysis has been conducted in TESOL, language and literature, and arts and design [15]. Multimodal pedagogy has also been adopted in English as a foreign language (EFL) teaching and learning in higher education [25]. Multimodal instruction has occurred in particular departments, such as civilization and science of language [9] and engineering and educational technology [8]. Multimodal pedagogy has been employed in English for specific purposes (ESP), particularly in video game design and development [5], English-medium instruction (EMI) at an agriculture and food engineering department [2], adult non-formal language education [29], EFL at law, education, engineering, and tourism management departments [34], and EFL at economics and fine arts departments [37]. Additionally, previous research [38] has also investigated the use of multimodality in numerous programs, for example, legal English (in a bachelors' program in law), English for finance (in a bachelors' program in finance and accounting), English for science and technology (in bachelors' programs in agri-food and rural engineering and in industrial design and product development engineering), English for psychologists (in a bachelors' program in psychology), English for medicine (in a bachelors' program in medicine), and a pre-service English teachers program [39].

4.2. Modes and digital media have been used to assist learners in acquiring English language skills

Linguistics and technology have been significantly employed to develop English language learning in the past decade. Implementing written text, video, audio, animation, and images [9,10,12,13,25,26,31,39,44,45] has not only enhanced the learning process but also accommodated diverse learning styles. According to the findings, a combination of several modes, for example, lexis, videos, video calls, gaze, kinesics, and audio [3,14,32], has been employed to develop learners' language skills. The use of gaze and kinesics demonstrates that body language is part of multimodality. In other words, nonverbal communication has potential in language pedagogy. Prior studies have also investigated many forms of nonverbal communication, for example, gesture, spatial position, facial expression, head movement, and gaze, which have been used in language pedagogy [2,11,28,33–35,37,42,43]. Linguistics and technology have been integrated during classroom activities. The use of phonetics, phonology, morphology, syntax, semantics and pragmatics, audio, and video [36] has also been evaluated. Advanced technology has been developed as a teaching tool to expand learners' language ability. The optimization of video, gamification, the three-dimensional (3D) environment, and virtual objects [7,15, 16] has broadened media as a teaching and learning tool in education. Color and symbols have been integrated as teaching and learning media during classroom activities. Text, color font, gesture, music, color underlining, color highlighting, symbols, and visual media [6,27,38,41] have been used and improved to boost learners' language skills.

A wide range of platforms has been methodically used in integrating instructional media. Integrating various technologies and platforms from past studies has enriched the instructional media framework. Previous research has employed Vox videos, YouTube, and podcasts [13,31,38], drama-based pedagogy and role play, television series [3], PowerPoint or slides [2,10,14,27,33,42], storyboards [8,25], digital storytelling, a computer, a digital projector, speakers and a whiteboard [26], website conferencing systems, for example, Adobe Connect and Skype video conferencing [28,29,32,35], Wikipedia, YouDao, Byu's COCA, Shift-it, and digital notes [41, 44], online communication and social media, including e-mail, MSN messenger, Facebook, Line, i-Map, Instagram, and Twitter [6,39], Desire2Learn Brightspace platform and E-Book video [36], drawings and short descriptions [45], digital video, documentaries and TV programs [34], and a LEGO instructional manual [9], which is the optimization of gamification, such as digital game groups (DGGs) [7]. Furthermore, the 3D virtual environment and VR [15,16] have also been employed to facilitate learners in the process of teaching and learning. The system provides EFL teachers and learners with a variety of communication channels and instruments, such as
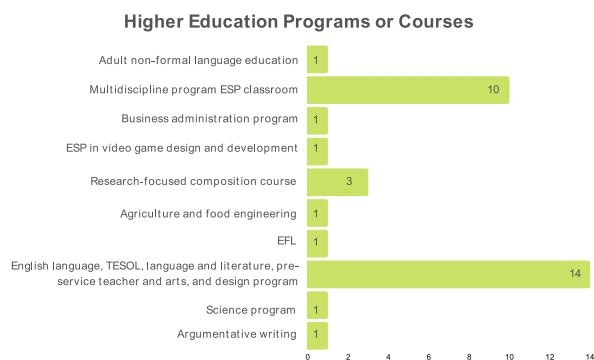

Fig. 2. Higher education courses and programs that use multimodality.

text-based and audio-visual synchronous computer-mediated conversation, user-controlled avatars, and virtual objects.

4.3. Aspects of English have been improved through multimodal teaching

Based on this systematic literature review, multimodality has significantly impacted learners' language competency. The adoption of multimodality to develop English language skills in higher education has been studied by past researchers. Multimodality has been implemented to improve learners' writing skills, emphasizing critical thinking, rhetorical knowledge, and composing processes [13]. The use of multimodality in developing learners' writing skills is beneficial due to its focus on promoting critical thinking skills, developing rhetorical knowledge, and elucidating the complexities of the composing process [5,8,9,13,25,27,39,41].

Similarly, previous research has found that learners' speaking fluency is developed through multimodality. Multimodality improves learners' pronunciation [3,34], thus benefiting oral proficiency. The advantage of multimodality is that it boosts students' development of oral skills [15,16,28–32,37,43,44], which then has the potential to assist learners in developing their presentation skills [14]. A multimodal drama-based program has also been used to measure learners' argumentative writing with a focus on grammar accuracy [12]. By integrating dramatic components, such as visual, aural, and kinesthetic factors, learners have been prompted to make persuasive arguments and motivated to convey their views with grammatical accuracy. Video has also been employed to enhance learners' pronunciation [3]; implementing video and audio in the research demonstrates the significant role of multimodality as an improvement tool for students' pronunciation. Additionally, multimodality has been used to expand TESOL pre-service teachers' practices [36]. Encouraging pre-service teachers to use multimodality as a teaching technique in their lesson plans, resources, and reflections results in the potential to improve students' classroom performance. Similarly, EMI teachers have been involved in previous research. Multimodality has been implemented in EMI micro-teaching, including verbal and nonverbal communication [11]. Specifically, semiotics has been a main concern in assessing EMI teachers' professional development. Another aspect of multimodality, for example, video, subtitles, and transcription, has been implemented to develop ESP skills, which relates to vocabulary [38]. Multiple modes, such as digital storytelling, have been employed to improve learners' reading skills [26]. Digital storytelling as a multimodal tool has been beneficial to active reading. Adapting digital tools broadens the researchers' chance to employ advanced instruments in assessing learners' language proficiency. Another sophisticated device, VR, has been integrated to develop learners' English language abilities. VR implementation aims to provide a multimodal communicative environment to develop speaking skills [15]. Additionally, students have been able to improve their lexical skills through the use of multimodality [7,10].

Furthermore, this study has also found that multimodality not only contributes to learners' development in English but also motivates learners to learn in the classroom [2,11,35]. This heightened motivation is crucial because it both fosters a positive attitude towards learning and encourages persistence and resilience in the face of challenges. Success in developing learners' language acquisition aligns closely with improved motivation, creating a synergistic effect that enhances overall educational outcomes. This motivational boost positively impacts collaborative learning as well, as motivated students are more likely to participate in group activities, share ideas, and support their peers, leading to a richer and more effective learning experience [36].

5. Discussion

5.1. Higher education courses or programs use multimodality to enhance English language learning

Prior reviews by Kulju et al. (2018) and Lim et al. (2022) mostly addressed primary and secondary school multimodal literacy. In contrast, this analysis focuses on higher education, bridging a considerable gap by evaluating the use of multimodal immersion approaches among adult learners. This distinction is critical since learning dynamics and technology engagement in higher education differ significantly from those in younger age groups. Using multimodality is not only beneficial for younger learners' language skill development [1,46–54] but also suitable for higher education. In addition, multimodality has been employed in the EAP and ESP classroom in the previous ten years. Multimodal pedagogy is used not just in conventional language-oriented disciplines and in the EAP classroom, such as TESOL and language and literature but also in diverse domains and in ESP instruction, for example, engineering [2, 8,33], video game design [5], agriculture [2], the law [34], economics [37,45], and medicine [38,42]. Integrating multimodal elements demonstrates an adaptability to different academic domains. Using multimodal pedagogy demonstrates its capacity to adapt to the individual demands and goals of learners in diverse academic fields. The wide-ranging use of multimodal pedagogy highlights its adaptability as an instructional method that can be customized to address the distinct requirements of students in different academic disciplines. Furthermore, integrating multimodal elements in teaching enhances comprehension and retention of complex information and fosters critical thinking and creativity. It aligns with the broader educational goals of higher education institutions, which aim to prepare students for the complexities of the modern world. The widespread implementation of multimodal pedagogies highlights their potential to provide customized learning experiences that effectively address the specific needs of students in diverse academic disciplines.

5.2. Modes and digital media have been used to assist learners in acquiring English language skills

The implementation of numerous modes has underpinned the classroom pedagogy process. This systematic review demonstrates that the multimodality dimension not only consists of text, images, media, video, and audio but also verbal and nonverbal modes as well as advanced technology. Nonverbal cues, such as eye contact (gaze) and physical movements (kinesics), improve communicative understanding and the effectiveness of communication between learners and educators [55]. By recognizing and integrating nonverbal

elements into language instruction, educators may enhance the overall learning process and cultivate a more profound grasp of linguistic concepts among learners. Essentially, this stresses the whole aspect of learning a language, acknowledging that successful communication extends beyond mere words and requires a subtle awareness of nonverbal cues. These studies highlight the possible influence of nonverbal signals on the process of acquiring language, introducing a novel aspect to conventional teaching approaches. In line with this, the study found the advantages of using text, font, and colour text.

Previous researchers have used various platforms and technologies to enhance the instructional media framework. This is aligned with the trend of generating digital multimodal composing (DMC) for teaching and learning pedagogy [4,40,56,57]. This includes a range of multimedia platforms and techniques, such as Vox videos [13] YouTube [38], podcasts [31], role play [12], television shows [3], PowerPoint [14], storyboards [25], digital storytelling [26], and website conferencing systems like Adobe Connect [35] and Skype [28]. This variety demonstrates a comprehensive approach, recognizing the changing nature of instructional technology. However, the advantages of the latest technology, VR, for instance, have not been explored in previous reviews. With regard to the advance technology used, earlier assessments of multimodal literacy have not addressed the latest technological innovations, such as VR. In addition to analyzing the use of video, audio, and text as a multimodal tool, our study fills the gap by reviewing research that uses VR as a multimodal device in the teaching and learning process, providing a more current perspective on the possibilities of immersive media in education. In terms of advanced technology, VR, for instance, integrates modern technology and reflects a progressive approach to teaching language [58]. It is used as an immersive technology to promote language learning and cultural understanding by presenting learners with virtual situations. This technology enables users to visually explore genuine settings, enhancing a more interactive and immersive learning experience. Importantly, VR allows students to gain practical experience in a virtual environment [59]. VR has the potential to enhance social aspects of online education [60]. Using VR contributes to improving learning outcomes [20]. This is consistent with our finding about the use of VR in previous research, which observed that VR has the potential to improve learners' communication skills [15]. Integrating VR into tertiary education can greatly improve the level of engagement and practical utilization of language abilities among university students [61]. The implementation of VR in nurse and hospitality education, for instance, enables nursing students to engage in virtual patient interviews, and it enhances their proficiency in effective communication within clinical environments [62]. Similarly, students studying hospitality might gain advantages from VR scenarios that replicate real customer service encounters, assisting in the development of vital linguistic and interpersonal abilities [58]. The alignment with the continuing digital transformation in higher education highlights the significance of integrating novel technologies such as VR into the curriculum.

5.3. Aspects of English are acquired through the use of multimodality

A combination of linguistics and technology as a multimodality tool has facilitated learners' development in the English language. Using written texts, video, aural stimuli, animation, and graphics has not only enhanced the language development process but also accommodated diverse learning styles. Written text is central in expanding language ability, grammatical competence, and writing skills. Using multimodality in the development of writing, structure, and written argumentation abilities indicates a shift towards a more comprehensive approach to learning. This method seeks to include a thorough treatment understanding of learners' linguistic proficiencies by taking into account many aspects. Simultaneously, the use of video features has improved learners' abilities in oral communication and in understanding spoken English in real-life situations [63]. This is consistent with the overarching objective of language instruction, which is to provide learners with the necessary skills for everyday communicative situations. Integrating video, audio, digital narrative, and VR highlights a clear preference for using technology in the development of learners' language. This indicates an acknowledgment of the possible advantages of sophisticated techniques in improving language proficiency. The emphasis on drama-based multimodality, digital storytelling, and VR points to a need to provide captivating and dynamic educational encounters. By integrating these multimodal media, instructors strive to augment the language acquisition process. The use of multimodal technologies in teaching ESP and EAP signifies a shift towards customizing the learning process to cater to the unique linguistic requirements of various domains. This approach acknowledges the varied linguistic requirements in numerous academic and professional settings.

Based on these study outcomes from the past ten years, the use of multimodal techniques has a positive effect on learners' English language skills and has the potential to stimulate students to learn. This corresponds well with the overall advantages of multimodal pedagogy, which is a versatile and adaptable teaching strategy. Employing many forms of communication in education, multimodal pedagogy has the potential to be customized to address the distinct requirements of students in different academic fields. Through the use of multimodal pedagogy, educators can tailor learning experiences by integrating a variety of modalities, including visual aids, aural components, and kinesthetics.

With regard to the advantages of multimodality, students effectively showcased their capacity to attribute, exemplify, and develop correlations between sources. They altered their viewpoint on writing, seeing it as a process of generating significance rather than only addressing structural components [39,41]. There was a noteworthy advancement in the learning of writing skills by the learners, and the study played a role in cultivating their enthusiasm for writing in English. Likewise, the role of multimodality is important to increase learners' speaking fluency. Proficiency in pronunciation and communication, for instance, demonstrates significant improvement in speaking abilities through the use of multimodality. Additionally, integrating gestures had a beneficial effect on learners' mastery of English, especially in using academic language in presentations. Using advanced technology, for example, VR and 3D tools, is essential in assisting EFL learners in accomplishing their communication goals.

5.4. Limitations

Despite the valuable contributions made by this review, several limitations must be acknowledged. Firstly, there is a potential language and publication bias, as only English-language studies have been included, possibly excluding relevant research published in other languages and skewing towards studies with positive results. Secondly, the heterogeneity among the included studies in terms of design, methodology, and outcomes measured presents challenges in synthesizing findings and drawing generalized conclusions. Future research should strive for methodological standardization to facilitate more direct comparisons. Additionally, while efforts have been made to conduct a comprehensive search, the review may have missed relevant studies not indexed in the chosen databases, suggesting a need to expand search criteria to include additional databases and grey literature. Lastly, while the review covers studies published between 2013 and 2024 to capture recent developments, this time frame may overlook earlier foundational work crucial for a contextual understanding. Future reviews could consider a broader temporal range to encompass long-term trends and shifts in the field.

Further study should prioritize examining the precise impacts of multimodal immersion on several aspects of English instruction and acquisition, including vocabulary acquisition, pronunciation, auditory comprehension, and conversational proficiency. It is essential to determine the ideal circumstances in which these effects are maximized, including the required technology infrastructure, efficient educational tactics, and suitable frequency of use. Furthermore, it is essential for research to investigate the impact of individual variations, such as pre-existing language skills, preferred learning methods, and motivation levels, on the results of multimodal immersion. Moreover, it is crucial to consider learners' attitudes towards technology, cognitive capacities, classroom atmosphere and teacher expertise, to accurately assess VR's overall efficacy in language acquisition. Additionally, it is important to evaluate the necessity of distinguishing other than English language lesson, as well as the potential impact of this differentiation on the utilization of multimodal immersion technologies. Future research can strengthen the understanding and execution of multimodal immersion in English teaching by focusing on these areas and providing specific insights based on empirical evidence.

6. Conclusion

This study aims to investigate the implementation of multimodal pedagogies in higher education, identify specific modes employed, determine the effect of multimodality on English language development, and its influence on learning outcomes in the past ten years. In terms of the advantages of multimodality in higher education, multimodality is not only used in English for academic purposes (EAP) in the classroom but also in English for specific purposes (ESP), for example, in health, business, industry, tourism, and the law. This sheds light on the wide range of levels of higher education engaged in multimodality. The widespread use of multimodality in numerous fields indicates a more significant change in educational paradigms towards more comprehensive and effective teaching approaches. As higher education institutions increasingly acknowledge and use multimodal techniques, students are more prepared to navigate through and thrive in various professional environments where strong communication and information-processing skills are crucial. Although a combination of video, audio, and text followed by traditional media, such as gesture, gaze, and movement, is used frequently, eagerness to develop advanced multimodality has been noted. The virtual environment, 3D media, and gamification games have been developed as multimodal tools. Numerous platforms, for example, YouTube, Skype, social media, and VR with numerous modes, have been integrated in previous studies. In addition, the use of nonverbal communication, for example, gesture, kinesics, spatial position, facial expression, head movement, and gaze, have been used as a multimodal approach in the higher education classroom. Multimodality has contributed to the enhancement of English lexis, grammar, writing and speaking. It has also assisted ESP teachers and learners as learning materials. This means that a variety of academic and professional programs have used multimodality. The primary outcome from previous studies was that multimodality has a positive impact on English proficiency, for example, in speaking and writing skills. Additionally, multimodality successfully enhances learners' enthusiasm to continue the teaching and learning process in higher education. While this review provides valuable insights, future research should take a broader temporal perspective to understand long-term trends in multimodal immersion. Interdisciplinary approaches that integrate insights from education, cognitive science, and technology are crucial for advancing this field. Longitudinal studies are necessary to assess the lasting impact of multimodal immersion on learning outcomes.

Data availability statement

The data associated with this study has been deposited into a publicly available repository (Mendeley data). Mendeley Data Reserved URL ID: https://data.mendeley.com/datasets/rnf4m4dg58.

CRediT authorship contribution statement

I Wayan Eka Dian Rahmanu: Writing – original draft, Software, Methodology, Investigation, Data curation, Conceptualization. **Gyöngyvér Molnár:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing

References

[1] L. Unsworth, K.A. Mills, English language teaching of attitude and emotion in digital multimodal composition, Special Issue on Multimodal Composing in Multilingual Learning and Teaching Contexts 47 (2020) 100712, https://doi.org/10.1016/j.jslw.2020.100712.

[2] M. Sabaté-Dalmau, B. Moncada-Comas, Exploring the affordances of multimodal competence, multichannel awareness and plurilingual lecturing in EMI, System 118 (2023) 103161, https://doi.org/10.1016/j.system.2023.103161.

[3] V. Galimberti, J.C. Mora, R. Gilabert, Audio-synchronized textual enhancement in foreign language pronunciation learning from videos, System 116 (2023) 103078, https://doi.org/10.1016/j.system.2023.103078.

[4] L. Jiang, S. Yu, I. Lee, Developing a genre-based model for assessing digital multimodal composing in second language writing: integrating theory with practice, L2 writing assessment in the digital age 57 (2022) 100869, https://doi.org/10.1016/j.jslw.2022.100869.

[5] V. Beltrán-Planques, M. Querol-Julián, English language learners' spoken interaction: what a multimodal perspective reveals about pragmatic competence, Special Issue: Multimodal Perspectives on English Language Teaching in Higher Education 77 (2018) 80–90, https://doi.org/10.1016/j.system.2018.01.008.

[6] I.-C. Ke, H. Cahyani, Learning to become users of English as a Lingua Franca (ELF): how ELF online communication affects Taiwanese learners' beliefs of English, System 46 (2014) 28–38, https://doi.org/10.1016/j.system.2014.07.008.

[7] P. Sundqvist, P. Wikström, Out-of-school digital gameplay and in-school L2 English vocabulary outcomes, System 51 (2015) 65–76, https://doi.org/10.1016/j.system.2015.04.001.

[8] L. Jiang, Digital multimodal composing and investment change in learners' writing in English as a foreign language, J. Sec Lang. Writ. 40 (2018) 60–72, https://doi.org/10.1016/j.jslw.2018.03.002.

[9] F. Coccetta, Developing university students' multimodal communicative competence: field research into multimodal text studies in English, Special Issue: Multimodal Perspectives on English Language Teaching in Higher Education 77 (2018) 19–27, https://doi.org/10.1016/j.system.2018.01.004.

[10] R. Kaplan-Rakowski, B. Loranc, The impact of verbal and nonverbal auditory resources on explicit foreign language vocabulary learning, System 85 (2019) 102114, https://doi.org/10.1016/j.system.2019.102114.

[11] T. Morell, V. Beltrán-Palanques, N. Norte, A multimodal analysis of pair work engagement episodes: implications for EMI lecturer training, J. Engl. Acad. Purp. 58 (2022) 101124, https://doi.org/10.1016/j.jeap.2022.101124.

[12] X. Zhang, Assessing EFL students' writing development as they are exposed to the integrated use of drama-based pedagogy and SFL-based teaching, Assess. Writ. 50 (2021) 100569, https://doi.org/10.1016/j.asw.2021.100569.

[13] X. Tan, An exploratory study of English as a Second Language students' 'citation' patterns in multimodal writing, J. Engl. Acad. Purp. 66 (2023) 101294, https://doi.org/10.1016/j.jeap.2023.101294.

[14] J. Wang, Y. Gao, Y. Cui, Classroom gesture instruction on second language learners' academic presentations: evidence from Chinese intermediate English learners, J. Engl. Acad. Purp. 66 (2023) 101304, https://doi.org/10.1016/j.jeap.2023.101304.

[15] Y.-C. Shih, Communication strategies in a multimodal virtual communication context, System 42 (2014) 34–47, https://doi.org/10.1016/j.system.2013.10.016.

[16] J. ChengChiang Chen, S. Kent, Task engagement, learner motivation and avatar identities of struggling English language learners in the 3D virtual world, System 88 (2020) 102168, https://doi.org/10.1016/j.system.2019.102168.

[17] P. Kulju, R. Kupiainen, A.M. Wiseman, A. Jyrkiäinen, K.-L. Koskinen-Sinisalo, M. Mäkinen, A review of multiliteracies pedagogy in primary classrooms, Language and Literacy 20 (2) (2018) 80–101.

[18] F.V. Lim, W. Toh, T.T.H. Nguyen, Multimodality in the English language classroom: a systematic review of literature, Ling. Educ. 69 (2022) 101048.

[19] J. Bezemer, G. Kress, Multimodality, Learning and Communication: A Social Semiotic Frame, Routledge, 2015.

[20] Z. Merchant, E.T. Goetz, L. Cifuentes, W. Keeney-Kennicutt, T.J. Davis, Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: a meta-analysis, Comput. Educ. 70 (2014) 29–40, https://doi.org/10.1016/j.compedu.2013.07.033.

[21] M.J. Page, et al., The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, bmj 372 (2021).

[22] H. Cooper, Research Synthesis and Meta-Analysis: A Step-by-step Approach, vol. 2, Sage publications, 2015.

[23] F.B. Schuch, D. Vancampfort, J. Richards, S. Rosenbaum, P.B. Ward, B. Stubbs, Exercise as a treatment for depression: a meta-analysis adjusting for publication bias, J. Psychiatr. Res. 77 (2016) 42–51.

[24] B. Keen, A. Blaszczynski, F. Anjoul, Systematic review of empirically evaluated school-based gambling education programs, J. Gambl. Stud. 33 (2017) 301–325.

[25] Y. Kim, D. Belcher, C. Peyton, Comparing monomodal traditional writing and digital multimodal composing in EAP classrooms: linguistic performance and writing development, J. Engl. Acad. Purp. 64 (2023) 101247, https://doi.org/10.1016/j.jeap.2023.101247.

[26] A. Alalem, Digital storytelling for cultivating a participatory culture in first-year composition, Comput. Compos. 69 (2023) 102792, https://doi.org/10.1016/j.compcom.2023.102792.

[27] Y. Zhang, K.L. O'Halloran, Empowering the point: Pains and gains of a writer's traversals between print-based writing and multimodal composing, Ling. Educ. 51 (2019) 1–11, https://doi.org/10.1016/j.linged.2019.04.003.

[28] H. Lee, R. Hampel, A. Kukulska-Hulme, Gesture in speaking tasks beyond the classroom: an exploration of the multimodal negotiation of meaning via Skype videoconferencing on mobile devices, System 81 (2019) 26–38, https://doi.org/10.1016/j.system.2018.12.013.

[29] O. Kozar, Language education via audio/videoconferencing (LEVAC): a discursive investigation, Ling. Educ. 31 (2015) 86–100, https://doi.org/10.1016/j.linged.2015.05.007.

[30] Y.-A. Lee, J. Hellermann, Managing language issue in second language storytelling, System 93 (2020) 102311, https://doi.org/10.1016/j.system.2020.102311.

[31] K. Bromberek-Dyzman, K. Jankowiak, P. Chełminiak, Modality matters: testing bilingual irony comprehension in the textual, auditory, and audio-visual modality, J. Pragmat. 180 (2021) 219–231, https://doi.org/10.1016/j.pragma.2021.05.007.

[32] M.G. Sindoni, Mode-switching in video-mediated interaction: integrating linguistic phenomena into multimodal transcription tasks, Digital language practices: media, awareness, pedagogy 62 (2021) 100738, https://doi.org/10.1016/j.linged.2019.05.004.

[33] T. Morell, Multimodal competence and effective interactive lecturing, Special Issue: Multimodal Perspectives on English Language Teaching in Higher Education 77 (2018) 70–79, https://doi.org/10.1016/j.system.2017.12.006.

[34] L. Jiang, J. Luk, Multimodal composing as a learning activity in English classrooms: inquiring into the sources of its motivational capacity, System 59 (2016) 1–11, https://doi.org/10.1016/j.system.2016.04.001.

[35] M. Querol-Julián, Multimodal interaction in English-medium instruction: how does a lecturer promote and enhance students' participation in a live online lecture? J. Engl. Acad. Purp. 61 (2023) 101207 https://doi.org/10.1016/j.jeap.2022.101207.

[36] M. Li, Multimodal pedagogy in TESOL teacher education: students' perspectives, System 94 (2020) 102337, https://doi.org/10.1016/j.system.2020.102337.

[37] F. Tůma, T. Sherman, Recruiting help in word searches in L2 peer interaction: a multimodal conversation-analytic study, Ling. Educ. 67 (2022) 100999, https://doi.org/10.1016/j.linged.2021.100999.

[38] C. Girón-García, I. Fortanet-Gómez, Science dissemination videos as multimodal supporting resources for ESP teaching in higher education, Engl. Specif. Purp. 70 (2023) 164–176, https://doi.org/10.1016/j.esp.2022.12.005.

[39] A. Mukminin, A. Habibi, M. Muhaimin, M. Hidayat, Social media use for English writing (SMU-EW): preservice English teachers, Ampersand 10 (2023) 100112, https://doi.org/10.1016/j.amper.2023.100112.

[40] X. Tan, Stories behind the scenes: L2 students' cognitive processes of multimodal composing and traditional writing, J. Sec Lang. Writ. 59 (2023) 100958, https://doi.org/10.1016/j.jslw.2022.100958.

[41] M. Kessler, Technology-mediated writing: exploring incoming graduate students' L2 writing strategies with activity theory, Comput. Compos. 55 (2020) 102542, https://doi.org/10.1016/j.compcom.2020.102542.

[42] J.-E. Peng, Understanding willingness to communicate as embedded in classroom multimodal affordances: evidence from interdisciplinary perspectives, Ling. Educ. 51 (2019) 59–68, https://doi.org/10.1016/j.linged.2019.04.006.

[43] J.-E. Peng, The roles of multimodal pedagogic effects and classroom environment in willingness to communicate in English, System 82 (2019) 161–173, https://doi.org/10.1016/j.system.2019.04.006.

[44] A. Cheung, Verbal and on-screen peer interactions of EFL learners during multimodal collaborative writing: a multiple case-study, J. Sec Lang. Writ. 58 (2022) 100931, https://doi.org/10.1016/j.jslw.2022.100931.

[45] S.-Y. Ahn, Visualizing the interplay of fear and desire in English learner's imagined identities and communities, Pushing the edge in narrative inquiry 102 (2021) 102598, https://doi.org/10.1016/j.system.2021.102598.

[46] Y. Park, S. Park, Eliciting student participation in synchronous online L2 lessons: the use of oral and written DIUs, Ling. Educ. 71 (2022) 101085, https://doi.org/10.1016/j.linged.2022.101085.

[47] M.-H. Hsieh, H.-H. Chuang, D. Albanese, Investigating student agency and affordances during online virtual exchange projects in an ELF context from an ecological CALL perspective, System 109 (2022) 102888, https://doi.org/10.1016/j.system.2022.102888.

[48] K.W.H. Tai, Managing classroom misbehaviours in the Hong Kong English Medium Instruction secondary classrooms: a translanguaging perspective, System 113 (2023) 102959, https://doi.org/10.1016/j.system.2022.102959.

[49] Y. Kim, S. Kang, Y. Nam, S. Skalicky, Peer interaction, writing proficiency, and the quality of collaborative digital multimodal composing task: Comparing guided and unguided planning, System 106 (2022) 102722, https://doi.org/10.1016/j.system.2022.102722.

[50] A. Rothoni, The interplay of global forms of pop culture and media in teenagers' 'interest-driven' everyday literacy practices with English in Greece, Ling. Educ. 38 (2017) 92–103, https://doi.org/10.1016/j.linged.2017.03.001.

[51] J. Kang, Transfer of knowledge across genres and media: investigating L2 learners' multiple composing practices, J. Engl. Acad. Purp. 56 (2022) 101096, https://doi.org/10.1016/j.jeap.2022.101096.

[52] O.G. Stewart, "Using digital media in the classroom as writing platforms for multimodal authoring, publishing, and reflecting," Comput Compos 67 (2023) 102764, https://doi.org/10.1016/j.compcom.2023.102764.

[53] Y. Kim, S. Kang, Writing to make meaning through collaborative multimodal composing among Korean EFL learners: writing processes, writing quality and student perception, Comput. Compos. 58 (2020) 102609, https://doi.org/10.1016/j.compcom.2020.102609.

[54] E. Tragant Mestres, A. Pellicer-Sánchez, Young EFL learners' processing of multimodal input: examining learners' eye movements, System 80 (2019) 212–223, https://doi.org/10.1016/j.system.2018.12.002.

[55] M.G. Busà, Teaching learners to communicate effectively in the L2: Integrating body language in the students' syllabus," Lingue e Linguaggi, 2015, pp. 83–98.

[56] M. Tecedor, Digital storytelling: changing learners' attitudes and self-efficacy beliefs, Appl. Linguist. (2023) amad002.

[57] Y. Deng, D. Liu, D. Feng, Students' perceptions of peer review for assessing digital multimodal composing: the case of a discipline-specific English course, Assess Eval. High Educ. (2023) 1–14.

[58] K. Takada, K.H. Murakami, A.K. Yamazaki, T. Yamanakac, A preliminary study of VR English training material for personnel in the tourism industry, Procedia Comput. Sci. 207 (2022) 3721–3729, https://doi.org/10.1016/j.procs.2022.09.432.

[59] J.-H. Zhao, Z.-W. Chen, Q.-F. Yang, I do and I understand: a virtual reality-supported collaborative design-assessing activity for EFL students, System 121 (2024) 103213, https://doi.org/10.1016/j.system.2023.103213.

[60] M. Barreda-Ángeles, S. Horneber, T. Hartmann, Easily applicable social virtual reality and social presence in online higher education during the covid-19 pandemic: a qualitative study, Comput. Educ.: X Reality 2 (2023) 100024, https://doi.org/10.1016/j.cexr.2023.100024.

[61] L. Gao, B. Wan, G. Liu, G. Xie, J. Huang, G. Meng, Investigating the effectiveness of virtual reality for culture learning, Int. J. Hum. Comput. Interact. 37 (18) (2021) 1771–1781, https://doi.org/10.1080/10447318.2021.1913858.

[62] Z. Taçgın, The perceived effectiveness regarding Immersive Virtual Reality learning environments changes by the prior knowledge of learners, Educ. Inf. Technol. 25 (4) (2020) 2791–2809, https://doi.org/10.1007/s10639-019-10088-0.

[63] S. Xia, Transcending science in scientific communication: multimodal strategies to incorporate humanistic perspectives in TED talks on biology, Engl. Specif. Purp. 71 (2023) 60–77, https://doi.org/10.1016/j.esp.2023.02.004.