

Original Research

Multimodal discourse in EFL instruction: Implications for language competence, cognitive engagement, and multimodal literacy

Victoria L. Malakhova^{1*}

¹MGIMO University, Moscow, Russia

Multimodal discourse has become increasingly relevant in language education. Although conventional texts continue to dominate English as a Foreign Language (EFL) instruction, interest in multimodal forms of communication has grown in recent years. This study systematises the features and functions of multimodal texts and evaluates their pedagogical relevance for language learning. The research combines theoretical analysis with an empirical investigation involving bachelor's students from non-linguistic majors and English language teachers. Data was collected through a pedagogical experiment, classroom observation, and questionnaire surveys. The study analyses how multimodal texts contribute to the development of students' language and communicative competence and identifies potential difficulties students may encounter when working with multimodal materials. The findings suggest that multimodal discourse may support the development of language proficiency, communicative competence, and learner motivation in EFL instruction. At the same time, certain challenges related to the interpretation of multiple semiotic resources were identified. Based on these results, the study proposes a structured instructional sequence for working with multimodal texts in the language classroom. The results indicate that multimodal discourse represents a potentially valuable resource for foreign language teaching and may contribute to the development of more effective pedagogical practices in diverse educational settings.

Keywords: language education, English as a foreign language, EFL, multimodal discourse, multimodal text, language competences, communicative competences, cognitive engagement

Article history:

Submitted December 20, 2025

Revised March 2, 2026

Accepted March 13, 2026

CRediT Author Statement:

Victoria L. Malakhova: Conceptualisation, Methodology, Validation, Data Collection, Formal Analysis, Investigation, Resources, Writing – Original Draft, Writing – Review & Editing, Visualisation.

Conflict of interest:

The author declared no conflict of interest.

Data availability statement:

The data supporting this study's findings are available from the corresponding author upon reasonable request.

Funding:

No funding was reported for this research.

Doi:

10.22363/2521-442X-2026-10-1-114-125

For citation:

Malakhova, V. L. (2026). Multimodal discourse in EFL instruction: Implications for language competence, cognitive engagement, and multimodal literacy. *Training, Language and Culture*, 10(1), 114–125.



This is an open access article distributed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) (CC BY-NC 4.0), which allows its unrestricted use for non-commercial purposes, subject to attribution. The material can be shared/adapted for non-commercial purposes if you give appropriate credit, provide a link to the license, and indicate if changes were made.

1. INTRODUCTION

The requirements of the times and the rapid expansion of digital technology have brought about significant changes in business and education. Foreign language instruction is no exception in this regard (Akhmadullin, 2023; Carrió-Pastor, 2022; Dmitrenko & Osiyanova et al., 2024), as it is expected to support the development of

professional competencies required of modern specialists in the labour market. Language education therefore faces the task of identifying pedagogical strategies and innovations that can improve language proficiency and communicative competence. These developments help young professionals communicate effectively with people from diverse cultural backgrounds and participate in international business

collaboration (Diem, 2025; Galskova et al., 2025; Osiyanova et al., 2024;) and may also enhance graduates' employability and broaden their access to job opportunities both locally and internationally.

Furthermore, both educators and employers note that changes in students' interpersonal communication styles require new tools to prepare them for future careers (Peredy et al., 2024; Zierock et al., 2025). Contemporary communication practices, strongly influenced by computer-mediated interaction, multicultural environments, and changing value systems (Peredy et al., 2024; Yelenevskaya & Protassova, 2021), are also connected to young people's early integration into the digital world. This influences how they perceive and process information and is associated with changes in behavioural patterns, language norms, and the ways in which ideas are formulated and interpreted. As the boundary between offline and online communication in educational and professional settings becomes increasingly blurred, social platforms have developed into important spaces for sharing news, ideas, and opinions in both personal and professional contexts (Kharkovskaya et al., 2022; Malyuga et al., 2019); for cultural transmission and the circulation of values (Malyuga et al., 2019; Snezhko, 2025); and for developing skills relevant to academic and professional communication (Galskova et al., 2025; Osiyanova et al., 2024).

Processing information in digital environments requires the ability to filter content quickly. In this regard, language economy and visibility become increasingly important (Malyuga et al., 2019; Snezhko, 2025). Younger generations tend to favour visual content; today's students often prefer to look rather than read. In everyday communication, visuals (photos, videos, infographics, memes, emojis, stickers, etc.) are often perceived as clearer and more convenient than lengthy verbal explanations, as they allow the message to be grasped instantly. In many cases, images replace text.

These communication trends among young people cannot be ignored in the learning process. Communication practices modelled in English classes should reflect natural patterns of interpersonal interaction. Language teachers therefore need to identify new ways of creating conditions that support the achievement of learning objectives. Multimodal discourse offers valuable possibilities in this regard, since in many educational settings, including second language classrooms, academic discourse is inherently multimodal. Multimodality refers to the use of diverse semiotic resources (verbal and non-verbal modes such as language, images, sounds, and gestures) within a communicative act to convey information (Carrió-Pastor, 2022; Sukhova et al., 2021). Considering the ways in which contemporary students perceive and process information, the non-verbal modes of multimodal discourse represent a promising tool for language education, as they

contribute to richer academic discourse, increase students' engagement and interest in learning, and can support improved learning outcomes.

Although not a fully innovative pedagogical concept or tool, multimodal discourse has attracted increasing attention from English language teachers. It complements formal methods of English instruction by introducing visually engaging forms of communication. Academic multimodal discourse is typically represented by extended texts combining several semiotic resources and described as multimodal, hybrid, polycode, or creolised texts. Their didactic potential can support the development of students' language and communication skills.

Multimodality has become an increasingly important concept in contemporary communication and language education. As multimodal texts are widely encountered in everyday life, students are expected to interpret and produce messages that combine linguistic and visual elements. This development raises questions about how multimodal forms of communication can be effectively incorporated into language education and what role they may play in supporting students' communicative development. These considerations provide the basis for the present study, which examines the pedagogical potential of multimodal discourse in EFL instruction.

Hence, the study pursues two main objectives. First, it aims to describe the linguistic characteristics of multimodal discourse by analysing how different semiotic resources contribute to meaning making and which multimodal techniques are used to achieve communicative goals. Second, it seeks to specify the methodological functions of multimodal discourse and to evaluate its effectiveness in EFL instruction. To achieve these aims, the study addresses the following tasks: (i) to conceptualise multimodal discourse as a system of interacting semiotic resources; (ii) to analyse meaning-making processes arising from the integration of linguistic and visual elements; (iii) to examine the pragmatic functions of multimodal discourse in academic communication; (iv) to assess the communicative competences that multimodal texts can support in EFL instruction; and (v) to analyse the pedagogical functions of multimodal texts and the factors influencing students' and teachers' preferences for their use in academic activities.

To examine the role of multimodal discourse in educational communication, the present study employs a pedagogical experiment (observation) and a survey conducted among bachelor's degree students and teachers who use multimodal texts in the learning process. The research intends to ascertain whether text multimodality can contribute to improvements in students' English proficiency, support the development of communicative competences, and enhance learning motivation. At the same time, the study also considers potential limitations associated with the use of multimodal discourse in English language education.

2. METHODOLOGY

2.1. Research design

The study is grounded in communication theory, which treats the communicative act as a multimodal process. In this perspective, communication reflects reality through interconnected verbal and non-verbal sign systems. Multiple channels of information transmission interact with one another, and meaning is distributed across linguistic, visual, auditory, and gestural resources. The study therefore adopts a multimodal perspective to examine how these modes combine in meaning-making processes. In addition, a communicative-pragmatic perspective is applied to analyse language and text as multimodal systems and to examine the semantic and pragmatic effects produced through the interaction of different modes.

The methodological design combines several analytical procedures. Multimodal discourse analysis, multimodal interaction analysis, and content analysis are used to examine multimodal discourse, its linguistic features, its internal organisation, and the processes through which meaning is constructed. This analytical framework enables the decomposition of multimodal texts into their verbal and visual components while also allowing the analysis of how these components function together within a communicative whole.

To investigate the pedagogical role of multimodal discourse in EFL instruction, the study employs: (i) qualitative and quantitative methods for data collection and sampling; (ii) general scientific procedures of analysis and synthesis, description, and generalisation; as well as (iii)

empirical methods that model situations of intercultural and business communication in English language classes through the application of multimodal discourse.

2.2. Pedagogical procedure design

To examine the proposed hypothesis concerning the effectiveness of multimodal discourse in language education, the study employed an observational pedagogical experiment. The experiment was organised as classroom observation focusing on students' engagement in activities involving multimodal discourse and their learning outcomes. This design was selected to observe the use of multimodal materials under natural classroom conditions rather than in a controlled experimental setting. For this reason, the study did not include separate experimental and control groups. Instead, students' activity, motivation, and performance were analysed by comparing their work with multimodal texts and with traditional instructional materials.

The experiment was conducted during the 2024–2025 academic year at MGIMO University. The participants included 65 undergraduate students majoring in Trade and Commerce and Ecology and Nature Management (see Table 1) who were studying English as a foreign language. Their English proficiency levels corresponded to intermediate and upper-intermediate. In addition, 15 instructors participated in the study.

The procedure involved observing students' classroom work with multimodal texts and analysing indicators of engagement and learning outcomes.

Table 1
Student participants by major and year of study

YEAR OF	MAJORS AND NUMBER OF PARTICIPANTS		
	Trade and Commerce	Ecology and Nature Management	Total
First year	9	10	19
Second year	7	8	15
Third year	10	8	18
Fourth year	8	5	13
Total	34	31	65

Additional data for the present study was collected through questionnaire surveys administered to both students and teachers. The survey was completed anonymously via an online platform and was organised in several stages.

The questionnaire included both closed-ended items requiring yes/no responses and open-ended questions inviting explanations and comments.

Stage 1 aimed to evaluate students' familiarity with multimodal texts, their defining characteristics, and students' experience of working with extended texts. At this stage, the questionnaire included the following questions: (i) Do you know what a multimodal text is? (ii) Can you provide a definition or explain the specific characteristics of a multimodal text? (iii) Do you know the differences between a regular (traditional) text and a multimodal text?

- (iv) Have you ever encountered a multimodal text in any academic discipline? If so, was the experience successful?
- (v) Have you encountered multimodal texts in English classes? If yes, what types of information did they present?
- (vi) Do you think multimodal texts have advantages over regular texts? Why or why not?

Stage 2 focused on analysing students' perceptions of the benefits of multimodal discourse and assessing their experience of working with such texts. Students were asked to respond to the following questions: (i) What advantages do you think multimodal texts have over traditional texts? (ii) How can learning through multimodal texts benefit your understanding of academic material? (iii) Do multimodal texts increase your interest in learning English and your engagement in classroom activities? (iv) What skills or abilities do you think multimodal texts can develop more effectively in English classes than traditional texts? Why? (v) What limitations, if any, do multimodal texts have compared with regular texts? (vi) What difficulties do you encounter when working with multimodal texts? (vii) Into which stages do you usually divide your work with a multimodal text? (viii) Are you able to create a multimodal text? If so, what challenges do you usually face?

Stage 3 examined students' preferences regarding the type of text used as a source of information and the content typically presented in multimodal texts. The questionnaire included the following questions: (i) As a source of information, do you prefer paper-based or digital texts? Please explain your choice. (ii) In your learning experience, would you prefer to work with multimodal texts or traditional texts? Why? (iii) What types of information or learning material do you think are best presented through multimodal texts? (iv) Do you generally prefer working with multimodal texts under the teacher's guidance or independently?

The teachers were asked similar questions, though these focused primarily on the didactic and methodological potential of multimodal discourse.

Both paper-based and digital texts were used in the study. The multimodal texts selected varied in the relative prominence of their components, meaning that in some cases visual elements predominated while in others linguistic elements were more central. The content of the multimodal texts covered topics from general English for junior students and business-related themes for senior students.

The texts were selected so that they corresponded with the organisational, motivational, communicative, and content-related objectives of the lessons, ensuring that they supported rather than distracted from the instructional tasks. In addition, the texts were intended to supplement lesson materials with additional informational and value-oriented content.

The texts were used for presenting new material, organising discussion, and checking students' comprehension. At later stages of the experiment, the tasks became more complex and required students to produce their own multimodal texts based on the content of conventional written texts.

2.3. Analytical procedure

The expected outcomes of the use of multimodal discourse in English learning were evaluated in terms of students' learning activity, engagement, and performance. Student performance was assessed according to the following criteria: (i) students' readiness and willingness to respond to the teacher's questions about the text; (ii) participation in teacher-led, pair, or group discussions on the topic or issue under consideration; (iii) appropriate use of active vocabulary and grammatical structures, including the number of grammatical, lexical, and stylistic errors; (iv) the expression of personal opinions, attitudes, or comments; (v) the quality of oral and written texts produced by students; and (vi) the quality of multimodal texts produced by students, including the appropriate use of verbal and visual communicative modes.

The selected methodology and the nature of the pedagogical process allow for a detailed analysis of the role of multimodal discourse in language learning, while also establishing certain limitations for the interpretation of the results. Several methodological limitations should therefore be acknowledged.

First, there were no specific instruments for measuring indicators such as engagement and motivation. As a result, the analysis relied primarily on teacher observations and evaluations, which introduces a degree of subjectivity into the interpretation of the data. Although these observations were informed by theoretical principles of linguistic and pedagogical research and by the results of the classroom experiment, some degree of evaluative judgement remains inevitable.

Second, the quantitative scope and time frame of the study impose certain constraints. The analysis of the use of multimodal discourse in English instruction is sufficient to identify its key characteristics and observed effects within the stated conditions and time period, but it does not claim to provide an exhaustive and all-embracing evaluation of its effectiveness.

Third, the content of instructional materials also presents limitations. Not all academic topics can be effectively addressed through multimodal discourse, and not all types of information can easily be transformed into multimodal texts produced by students.

These limitations do not diminish the value of the findings but rather define the boundaries within which the results should be interpreted and indicate possible directions for future research.

3. THEORETICAL BACKGROUND

3.1. Multimodality and the multimodal nature of communication

The term *multimodality*, introduced by Kress and Van Leeuwen (2001), is now widely used in contemporary research. A substantial body of scholarship examines multimodality across different disciplines, including linguistics (Kress & Van Leeuwen, 2001; Sanchez-Stockhammer, 2012), education (Diem, 2025; Sihombing et al., 2022), sociology and cultural studies (Kress, 2009; Snezhko, 2025), and psychology (Demuth, 2022).

In line with linguistic research on multimodality (Kress, 2009; Carrió-Pastor, 2022; Sihombing et al., 2022), it can be understood as a system of verbal and non-verbal sign resources and the principles governing their interaction within a communicative act. In this perspective, communication is realised through a combination of textual, visual, auditory, and gestural modes that together contribute to the construction of meaning. Because these different modes – including written text, images, video, sound, and bodily expression – are used simultaneously, recipients must engage multiple perceptual channels to interpret the intended message (Kress, 2009; Ledin & Machin, 2020).

Human communication therefore relies on a variety of semiotic resources rather than a single sign system. Although linguistic expression often occupies a central role, communication cannot be reduced to language alone. Instead, meaning is constructed through the interaction of several communicative modes (Carrió-Pastor, 2022; Kress, 2009). Within this perspective, it is customary to distinguish between verbal and non-verbal communicative resources. Verbal resources refer to linguistic codes realised in written or spoken language. Non-verbal resources include visual, auditory, paralinguistic, and extra-linguistic elements.

Visual and auditory modes encompass such elements as images, graphic design, page layout, colour, sound, video, and other audiovisual materials. Paralinguistic features include vocal characteristics such as tempo, timbre, tone, and intonation. Extra-linguistic resources involve bodily expressions such as gestures, posture, facial expressions, and eye contact. The combination of these elements forms a system in which different modes may complement one another or modify the interpretation of other components within the communicative act.

The development of digital technologies has further expanded the role of multimodal communication. The widespread use of video, voice messages, graphic symbols, and other visual elements (such as emoticons and stickers) has contributed to the emergence of new communicative practices. Such forms of multimodal communication are now widely observed not only in everyday interaction but also in educational and professional environments. As a

result, extended multimodal texts and discourse have become increasingly common in contemporary communication.

3.2. Meaning construction in multimodal discourse

Multimodal discourse represents a complex semiotic system in which different symbolic resources interact to construct meaning in communication (Diem, 2025; Jewitt, 2017). Discourse may be viewed as an open, holistic system capable of self-organisation through dynamic transformation, with nonlinearity considered one of its defining properties (Ponomarenko et al., 2020; Van Dijk, 2015). This can also be applied to multimodal discourse as meanings are formed by discursive components (in this case, verbal and non-verbal resources) and interpreted in the process of discourse development, when the discourse switches between different modes. As Baldry and Thibault (2006) maintain, although all resources combined and integrated within the certain discourse are separate modes of meaning-making, they form a synergistic whole which cannot be reduced to or explained in terms of the mere sum of its separate parts. Many linguists also emphasise this in their works. Thus, Kress (2009) suggests that modes do not exist in isolation but always in conjunction with others. Ledin and Machin (2020) term it a ‘composition’ that possesses an overarching wholeness, with the constituent parts interrelated in various ways. Diverse multimodal techniques are not simply juxtaposed but strategically connected with one another effectively conveying the intended meanings. Interplay between the components of multimodal discourse reinforces the communication effect and potentially amplify their influence on the recipient (Bateman, 2014; Sukhova et al., 2021).

Multimodal discourse in language education may also be treated as specialised discourse, given that language instruction in higher education is commonly linked to students’ disciplinary fields. In this respect, discourse reflects general, academic, and professional domains of knowledge and practice (Galskova et al., 2025). Specialised discourse is typically characterised by a specific communicative situation, a clearly defined thematic focus, the use of specialised terminology, and the professional community within which communication occurs. According to Galskova et al. (2025), there is also ‘a range of factors that affect communication, such as the setting, type of event, the psychological atmosphere within the community, relationships among its members, and their professional interests and perspectives’ (Galskova et al., 2025, p. 61).

Facing a real challenge to find a proper balance between conventional and cutting-edge methodological tools, educators search for new ways of text design and text production. The present study suggests that the use of multimodal texts in language education represents a promising direction for pedagogical development in this respect.

3.3. Core features of a multimodal text

A multimodal text is an extended text that combines several semiotic resources. In pedagogical traditions, related terms such as *polycode*, *hybrid*, and *creolised text* are also used (Novospasskaya & Zou, 2021; Platonova et al., 2015; Sanchez-Stockhammer, 2012). In the present study, the term *multimodal text* is preferred, with particular attention given to texts integrating verbal and visual components. The verbal mode represents the linguistic component of the text. The visual mode includes images, photographs, tables, diagrams, page layout, typography, colour, and, in digital multimodal texts, video and film clips.

In general, both components contribute to the construction and interpretation of meaning. However, the relationship between verbal and visual elements may vary. In many cases, the verbal component predominates, and the removal of visual elements does not substantially affect the core informational content of the text. In such cases, visuals primarily serve an illustrative or attention-directing function, drawing the reader's attention and providing contextual framing for the message.

In other cases, the visual component is essential for the interpretation of the text. Removing the image may compromise the integrity of the message and lead to incomplete or incorrect interpretation. In such instances, visual elements may summarise key information, represent relationships between concepts, or extend the information presented verbally. This type of relationship is common in charts, diagrams, and concept maps. In the present study, this latter type of multimodal text predominated.

Building on classifications proposed in previous research (Carrió-Pastor, 2022; Danielsson & Selander, 2021; Unsworth, 2006), the analysis distinguishes several types

of relationships between verbal and visual components: (i) *parallel correlation*, where the meaning conveyed visually corresponds directly to the meaning expressed verbally; (ii) *cross correlation*, where verbal and visual elements partially overlap in their semantic and pragmatic contribution; (iii) *opposite correlation*, where the visual component contradicts the verbal message, often producing humour or prompting critical interpretation; (iv) *interpretive correlation*, where the relationship between the two modes is indirect or associative and requires the reader to infer the connection between them; and (v) *complementary correlation*, where one component supports or elaborates the meaning conveyed by the other.

A multimodal text should therefore be interpreted as a unified whole in which different modes contribute jointly to the construction of meaning.

4. STUDY RESULTS

The results of the study indicate that EFL instruction can benefit from the use of multimodal texts. Their potential lies in their dual nature, as multimodal discourse functions simultaneously as a linguistic and a pedagogical phenomenon. Based on the analysis of the questionnaire data, several core features of multimodal texts were identified from both linguistic and didactic perspectives. These features include: F1 combined modes of information presentation; F2 message focusing; F3 concise presentation of information; F4 attention attraction; F5 achievement of multiple communicative goals; F6 diversity of learning materials; F7 reduction of tension in the learning process; F8 comprehension checking; and F9 entertainment. Table 2 presents the distribution of responses provided by students and teachers regarding these features.

Table 2
Core features of multimodal discourse in language instruction

	LINGUISTIC FEATURES				DIDACTIC FEATURES				
	F1	F2	F3	F4	F5	F6	F7	F8	F9
Participants									
Students (n=65)	46	63	57	60	17	54	62	58	63
Teachers (n=15)	13	10	8	15	12	15	12	15	12
Total score per point	59	73	65	75	29	69	74	73	75
Total percentage per point	73.75	91.25	81.2	93.75	36.25	86.25	92.5	91.25	93.75

Different communicative modes (textual, graphical, and paralinguistic) complement one another and contribute to the construction of meaning. Their combined use encourages students to engage with both explicit and implicit aspects of message formation. From a pedagogical perspective, multimodal discourse may increase learner

engagement and motivation. It can also support classroom interaction and facilitate learning outcomes. In addition, multimodal texts contribute to the development of a range of student abilities, skills, and competences. Table 3 summarises the main abilities and competences that multimodal texts were observed to support in the learning process.

Table 3
Student abilities and competences supported by multimodal texts

LANGUAGE COMPETENCES	COGNITIVE ACTIVITIES	MULTIMODAL LITERACY
(i) activation of active and passive vocabulary and grammar	(i) critical and analytical thinking	(i) comprehension skills
(ii) development of oral and written communication skills	(ii) diverse forms of thought expression	(ii) contextual inference
(iii) expansion of linguistic and professional knowledge	(iii) visual perception	(iii) interpretive strategies
	(iv) imagination	(iv) information summarisation
	(v) creativity	

The experiment revealed observable changes in students' ability to search for relevant information, filter and synthesise it, evaluate evidence, identify examples and arguments, draw conclusions from texts and visual materials, and express and justify their opinions. The results also support the assumption that the non-verbal component of a multimodal text — in this study primarily the visual element — can stimulate reflection and discussion when combined with the linguistic component of the text. Students were more willing to engage in the introductory stage of discussions when a topic or text was presented through multimodal materials.

The observations further suggest that students did not passively receive information but actively constructed knowledge through interaction with the learning materials and with their peers. This process appeared to support collaboration and group interaction, as students participated

more readily in classroom discussions and collaborative activities. Multimodal academic materials therefore contributed to the creation of a more engaging learning environment.

The questionnaire results indicate that most students and teachers expressed positive attitudes toward the use of multimodal discourse in language education. Students reported that multimodal materials help reduce the monotony associated with traditional instructional formats and support their learning process. Many students also noted that multimodal texts were beneficial for both foreign language learning and the study of their academic disciplines. Some respondents emphasised the usefulness of such materials for independent learning and self-directed study. Teachers similarly pointed to the potential of multimodal texts to support the development of a range of student competences.

Table 4
Preferences for multimodal discourse in foreign language instruction

PARTICIPANTS	MULTIMODAL TEXTS USAGE
First-year students (n=19)	83%
Second-year students (n=15)	67%
Third-year students (n=18)	54%
Fourth-year students (n=13)	78%
Teachers (n=15)	75%

The results of the experiment suggest that students developed substantial experience in working with multimodal texts in English classes, which may also be applicable in their future professional activities.

Overall, the findings indicate that multimodal discourse in EFL instruction may: (i) bring classroom communication closer to authentic academic and professional interaction; (ii) support a more individualised learning process; (iii) facilitate multiple forms of monitoring students' progress and learning outcomes; (iv) encourage student

self-assessment and self-regulation of knowledge and skills; (v) increase learner engagement and motivation; (vi) introduce elements of creativity and interest into language learning; and (vii) support the adaptation of teaching practices to the demands of contemporary digital communication.

The results suggest that multimodal discourse represents a valuable didactic resource for language education and may contribute to enriching students' learning experiences.

5. DISCUSSION

Having examined the main characteristics and functions of multimodal texts, the study conceptualises them in terms of three interrelated aspects: linguistic, instructional, and educational. The three aspects are closely interconnected and mutually dependent, forming a coherent structure. Analysis of the model components indicates the active integration of the linguistic features of multimodal discourse into EFL instruction.

In many cases, multimodal discourse serves as a valuable resource for creating a more engaging academic learning environment. The findings suggest that the use of multimodal discourse in English learning may increase student engagement and support more effective learning outcomes. To examine students' interest in the learning process, levels of classroom engagement and learning performance were compared when students worked with traditional texts and with multimodal texts.

Table 5
Students' motivation and performance in English learning through multimodal discourse

PARTICIPANTS	MOTIVATION (%)	PERFORMANCE (%)
First-year students (n=19)	82	72
Second-year students (n=15)	65	80
Third-year students (n=18)	52	75
Fourth-year students (n=13)	78	87
Total	69.25	78.5

The findings indicate that the use of multimodal discourse is associated with increased student engagement and learning activity in English classes. A comparison of motivation levels across different years of study showed that the highest level of motivation was observed among first-year students (35%), while the lowest was recorded among third-year students (10%). Second- and fourth-year students demonstrated comparable levels of motivation (25% and 30%, respectively). When explaining their interest, junior students most frequently referred to the entertaining character of multimodal materials and the reduced cognitive strain involved in processing information. Senior students, in contrast, emphasised the development of multimodal literacy and the ability to present information in a concise and structured form.

Classroom observations also indicated that visual elements may facilitate and strengthen the comprehension of textual messages, a view supported by approximately 75% of the teachers surveyed. In addition, 70% of students and 83% of teachers reported that pictorial components function in combination with textual content to reinforce the intended message. The questionnaire results further showed that most students (approximately 71%) expressed positive attitudes toward the use of multimodal discourse in language education.

These perceptions appear to be related to several functions performed by multimodal texts. First, visual elements illustrate and complement the verbal component (informative and illustrative functions), which may support rapid comprehension of the text and allow readers to

anticipate the topic before engaging with the written content (introductory function). Second, pictorial components attract attention (attention-directing function) and introduce variety into conventional text-based activities (entertainment function). Third, images can convey information implicitly, enabling interpretation of meanings not directly stated the verbal content (descriptive function). Fourth, visual elements may evoke emotional responses and contribute to students' aesthetic perception of the material (expressive and aesthetic functions). Finally, the integration of verbal and visual components creates a richer communicative situation and may encourage collaborative interaction among students (communicative function).

However, some students (approximately 20%) reported several difficulties when working with multimodal texts. These students indicated that they sometimes experienced confusion when information was presented through multiple semiotic resources. Approximately 8% of respondents specifically mentioned inconsistencies between verbal and visual components, the unclear function of certain visual elements, or situations in which graphical elements appeared to hinder rather than support comprehension of the multimodal text, including its verbal component. Some students also noted that visual symbols could distract their attention and make it more difficult to concentrate on the core learning material. Several respondents emphasised that the visual component should be unambiguous and clearly related to the textual content. Others considered the verbal component to be more important than the visual elements.

According to the observations, such difficulties were more frequently reported by students with an intermediate level of English proficiency. One possible explanation is that these students may be more accustomed to conventional instructional formats and may have less experience in interpreting multimodal materials. Limited background knowledge or developing analytical skills may therefore contribute to uncertainty when learners are confronted with multiple semiotic systems simultaneously.

The questionnaire results also indicate that most English teachers (75%) recognise the value of multimodal discourse in language education. This suggests a growing interest among instructors in developing new teaching techniques and in addressing the limitations of traditional single-mode instructional practices. Nevertheless, a minority of teachers (approximately 25%) did not consider multimodal discourse to be significantly beneficial in their teaching practice.

Application of multimodal texts in English language instruction appears to support the development of several learner competences. Observations from the study indicate improvements in the following areas: (i) language proficiency, including greater accuracy and fluency in the target language, as well as communicative competence, understood as the ability to recognise situational requirements and audience expectations and to respond to different rhetorical conventions; (ii) practical language skills related to working with information presented in different genres and formats; (iii) multimodal literacy, including the ability to interpret and integrate information conveyed through different semiotic modes; (iv) cognitive activity, including analytical and critical thinking, imagination, and creativity, together with the development of sociocultural knowledge; (v) the production of both multimodal and conventional texts; and (vi) the expansion of general and professionally relevant knowledge.

Working with extended multimodal texts also appears to broaden students' interpretive experience through exposure to multiple perspectives and representational formats. Such materials encourage students to consider how different communicative modes function together in conveying meaning and how verbal and visual resources complement one another. In this respect, multimodal texts contribute to the development of multimodal literacy, as they encourage students to analyse relationships between visual and verbal elements and to apply these principles in their own text production (for example, reports, PowerPoint presentations, web pages, posters, charts, or concept maps). Students gradually acquire skills related to multimodal text design, including the selection of appropriate semiotic resources in accordance with their communicative potential, their functional role in the text, and the expectations of the intended audience (Danielsson & Selander, 2021).

The development of communicative competence also prepares students to function more effectively in real communicative situations, as it encourages the use of different communicative modes depending on audience and purpose. These abilities may also support students' future professional activity, as multimodal communication is increasingly used in a wide range of professional contexts.

In addition, visual elements often reflect cultural features and contemporary social trends associated with the target language. In this way, they may convey cultural norms and values characteristic of English-speaking communities. Exposure to such materials may contribute to the development of students' intercultural awareness and influence their communicative and social behaviour (Snezhko, 2025; Xiong et al., 2022).

Classroom work with multimodal texts can be organised in various ways. The present study proposes an instructional sequence for working with multimodal texts in English language teaching that consists of several stages. This step-by-step organisation of activities supports the gradual interpretation of meaning during reading and discussion and contributes to both comprehension of the text and the development of students' multimodal literacy.

Stage 1. Preparatory (lead-in) phase. This stage introduces the topic or issue and focuses on the interpretation of the visual component of the text. Two directions of activity are possible at this stage. First, the teacher may introduce the topic through preliminary discussion by asking questions such as: *What do you know about ...? Have you ever encountered ...? Can you explain what ... is?* A picture, image, photograph, or short video may then be presented to support and extend responses. Description and interpretation of the visual element serve as a link between the initial discussion and the subsequent work with the text.

Alternatively, the activity may begin with the presentation of a visual element without revealing the title or slogan of the text. Students may be asked questions such as: *What issue do you think this image illustrates? What aspect of the topic does the image emphasise? Why do you think these visual elements were chosen? What emotions does the image evoke?* After this discussion, the title or accompanying slogan can be introduced so that students can compare their interpretations with the intended message.

From a pedagogical perspective, this stage stimulates critical and analytical thinking and activates vocabulary and grammatical structures learned in previous lessons. It also allows the teacher to identify gaps in students' knowledge that may require further clarification. In addition, the activity engages students' personal experience, background knowledge, and sociocultural awareness, including knowledge related to their future professional fields.

Stage 2. Text comprehension phase. At this stage, students work with the verbal component of the multimodal text. They read it, identify its main idea and message, and

‘Multimodal discourse involves multiple semiotic resources – such as language, images, sounds, and gestures – which jointly contribute to meaning construction and information transmission. This mode of communication differs from more traditional text-based forms of expression. The present study focused primarily on multimodal texts that combine verbal and visual elements.’

discuss the key points. Students also analyse how the verbal and visual elements interact in conveying meaning. This stage reinforces the active vocabulary and grammatical structures associated with the unit or section.

Stage 3. Comparative analysis phase. During this stage, students analyse the information presented in the text by distinguishing between familiar and new facts, identifying relevant and interesting information, and considering how the material may relate to their academic specialisations. This stage contributes to the expansion of students’ knowledge. Students also compare the verbal and visual components of the text. They consider whether both modes convey the same meaning or whether the visual element expands or modifies the information presented verbally. In addition, they analyse the effect created by the visual component (for example, whether it introduces humour or emphasises the seriousness of the issue) and discuss which component – verbal or visual – most effectively supports comprehension of the text’s central idea.

Stage 4. Creative production phase. In the final stage, students present the text they have studied. This can be carried out during the lesson or assigned as independent work. A variety of formats may be used, including oral retelling, written summaries, presentations, concept or mind maps, diagrams, and charts. Activities may also include monologues, dialogues, group discussions, or role-play tasks. This stage develops students’ ability to produce coherent oral and written texts, including multimodal texts. It also strengthens oral and written communication skills and promotes students’ ability to interpret communicative situations from a rhetorical perspective.

Although the proposed instructional sequence is relatively extensive, not all stages are required for every text. Depending on the learning objectives and the content of the material, the teacher may focus on selected stages or even a single stage of the procedure.

6. CONCLUSION

Although conventional texts remain widely used in English language teaching, interest in multimodal discourse has increased in recent years. A growing number of educators view extended texts as a useful resource in language learning. Multimodal discourse may offer several advantages for EFL instruction compared with conventional

text-based materials. At the same time, the use of multimodal resources may present certain practical challenges for teachers and learners.

Multimodal discourse involves multiple semiotic resources – such as language, images, sounds, and gestures – which jointly contribute to meaning construction and information transmission. This mode of communication differs from more traditional text-based forms of expression. The present study focused primarily on multimodal texts that combine verbal and visual elements.

The study examined multimodal discourse from the perspectives of linguistics, communication theory, and foreign language pedagogy. It described the characteristics of multimodal discourse and multimodal texts, identified their potential advantages in teaching English to students of non-linguistic majors, and systematised the key student competences associated with the use of multimodal materials. The findings suggest that multimodal texts may contribute to the development of language and communication competence, multimodal literacy, and certain cognitive abilities. They may also support the expansion of students’ general and professional knowledge and encourage learner motivation and autonomy.

One possible explanation for these effects is that multimodal texts often approximate real-life communication and professional environments. They frequently incorporate authentic materials and reflect contemporary social, cultural, and professional practices. In addition, distributing information across different semiotic modes may facilitate comprehension of complex material. The study also proposed an instructional sequence for working with multimodal texts in the EFL classroom.

The theoretical contribution of the study lies in the systematisation of the features and functions of multimodal texts, which contributes to ongoing research on multimodality in linguistics. The practical relevance of the study concerns the potential application of multimodal texts in foreign language teaching, as the findings may inform the use of multimodal techniques in EFL instruction.

Contemporary communication – both offline and online – combines multiple semiotic resources, including textual, visual, and audio elements. As a result, extended texts integrating several modes within a single communicative structure are becoming increasingly common. This development poses challenges both for traditional discourse analysis and for language education. In linguistic analysis, exclusive attention to verbal language may overlook significant aspects of meaning construction. In language teaching, these changes encourage the search for effective pedagogical strategies that account for multimodal forms of communication. For this reason, multimodal discourse represents a relevant area of interest for language education, particularly in the context of expanding digital and online learning environments.

References

- Baldry, P., & Thibault, P. J. (2006). *Multimodal transcription and text analysis*. Equinox.
- Bateman, J. (2014). *Text and image: A critical introduction to the visual/verbal divide*. Routledge. <https://dx.doi.org/10.4324/9781315773971>
- Carrió-Pastor, M. L. (2022). Teaching multimodal metadiscourse in academic English as a foreign language. *Porta Linguarum*, 4, 155–172. <https://doi.org/10.30827/porta-lin.vi.21423>
- Danielsson, K., & Selander, S. (2021). *Multimodal texts in disciplinary education*. A comprehensive framework. Springer. <https://doi.org/10.1007/978-3-030-63960-0>
- Demuth, C. (2022). Multimodal interaction analysis in cultural psychology research. In M. Watzlawik & S. Salden (Ed.), *Courageous methods in cultural psychology: Theory and history in the human and social sciences* (pp. 131–155). Springer. https://dx.doi.org/10.1007/978-3-030-93535-1_8
- Diem, P. N. (2025). Application of the multimodal discourse analysis theory to the teaching listening and speaking skill. *South Asian Research Journal of Arts, Language and Literature*, 7(2), 34–40. <https://doi.org/10.36346/sarjall.2025.v07i02.001>
- Dmitrenko, T. A., & Akhmadullin, A. A. (2023). Current trends in teaching a foreign language in the digital era. *Training, Language and Culture*, 7(4), 16–23. <https://doi.org/10.22363/2521-442X-2023-7-4-16-23>
- Galskova, N. D., Komochkina, E. A., & Poliakova, N. V. (2025). The axiological framework of professional foreign language communication. *Training, Language and Culture*, 9(4), 56–72. <https://dx.doi.org/10.22363/2521-442X-2025-9-4-56-72>
- Jewitt, C. (2017). Multimodal discourses across the curriculum. In S. Thorne & S. May (Ed.), *Language, education and technology: Encyclopedia of language and education* (pp. 31–43). Springer. https://dx.doi.org/10.1007/978-3-319-02237-6_4
- Kharkovskaya, A. A., Aleksandrova, O. V., & Ponomarenko, E. V. (2022, April 22–23). The routes of intercultural professional communication in the XXI century perspective. In *Proceedings of the 10th International Research Conference Topical Issues of Linguistics and Teaching Methods in Business and Professional Communication* (pp. 254–257). RUDN University Press.
- Kress, G. (2009). *Multimodality: A Social semiotic approach to contemporary communication*. Routledge. <https://doi.org/10.4324/9780203970034>
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold Press.
- Ledin, P., & Machin, D. (2020). *Introduction to multimodal analysis* (2nd ed.). Bloomsbury Academic. <https://dx.doi.org/10.5040/9781350069176>
- Malyuga, E., Maksimova, D., & Ivanova M. (2019). Cognitive and discursive features of speech etiquette in corporate communication. *International Journal of English Linguistics*, 9(3), 310–318. <https://dx.doi.org/10.5539/ijel.v9n3p310>
- Novospasskaya, N. V., & Zou, H. (2021). The formation of polycode text theory. *RUDN Journal of Language Studies, Semiotics and Semantics*, 12(2), 501–513. <https://doi.org/10.22363/2313-2299-2021-12-2-501-513>
- Osiyanova, O. M., Burikova, S. A., & Radyuk, A. V. (2024). Technology for improving students' communicative skills in the university digital educational environment. *Training, Language and Culture*, 8(4), 24–35. <https://doi.org/10.22363/2521-442X-2024-8-4-24-35>
- Peredy, Z., Vigh, L., Qiongyu, W., & Muiyi, Jiang. (2024). Analysing Generation Z communication attitudes, values and norms. *Acta Periodica*, 30, 4–19. <https://doi.org/10.47273/AP.2024.30.4-19>
- Platonova, Yu., Tarasova, E., & Golubinskaya, A. (2015). Creolized text as a form of modern educational discourse. *Procedia – Social and Behavioral Sciences*, 214, 788–796. <https://doi.org/10.1016/j.sbspro.2015.11.719>
- Ponomarenko, E. V., Magirovskaya, O. V., & Orlova, S. N. (2020). Introduction: Professional discourse in the focus of functional linguistics. In E. Malyuga (Ed.), *Functional approach to professional discourse exploration in linguistics* (pp. 1–20). Springer. https://dx.doi.org/10.107/978-981-32-9103-4_1
- Sanchez-Stockhammer, C. (2012). Hybridization in language. In P. Stockhammer (Ed.), *Conceptualizing cultural hybridization* (pp. 133–157). Springer. https://dx.doi.org/10.107/978-3-642-21846-0_9
- Sihombing, P. S. R., Herman, H., & Saputra, N. (2022). How to teach English conversation? An implementation of a multimodal discourse analysis through images. *English Review: Journal of English Education*, 10(2), 431–438. <https://doi.org/10.25134/erjee.v10i2.6244>
- Snezhko, G. E. (2025). Hashtags as pragmatic devices of cultural reproduction in digital discourse. *Training, Language and Culture*, 9(2), 104–113. <https://doi.org/10.22363/2521-442X-2025-9-2-104-113>
- Sukhova, N. V., Dubrovskaya, T., & Lobina, Y. A. (2021). Correction to: Multimodality, digitalization and cognitivity in communication and pedagogy. In N. V. Sukhova, T. Dubrovskaya, & Y. A. Lobina (Ed.), *Multimodality, digitalization and cognitivity in communication and pedagogy* (pp. C1). Springer. https://dx.doi.org/10.1007/978-3-030-84071-6_14
- Unsworth, L. (2006, May 10–12). Image/text relations and intersemiosis: Towards multimodal text description for multiliteracies education. In *Proceedings of the 33rd International Systemic Functional Congress 2006* (pp. 1165–1205). PUCSP.
- Van Dijk, T. A. (2015). Critical discourse analysis. In D. Tannen, H. Hamilton, & D. Schiffrin (Ed.), *Handbook of discourse analysis* (pp. 466–485). Wiley Blackwell. <https://doi.org/10.1002/9781118584194.ch22>
- Xiong, T., Feng, D., & Hu, G. (2022). Researching cultural knowledge and values in English language teaching textbooks: Representation, multimodality, and stakeholders. In T. Xiong, D. Feng, & G. Hu (Ed.), *Cultural knowledge and values in english language teaching materials* (pp. 1–18). Springer.

Yelenevskaya, M., & Protassova, E. (2021). Teaching languages in multicultural surroundings: New tendencies. *Russian Journal of Linguistics*, 25(2), 546–568. <https://doi.org/10.22363/2687-0088-2021-25-2-546-568>

Zierock, B., Schulze, J., & Angar, S. (2025). From Gen Z to Generation Alpha: Navigating the evolution of communication in a digital age. *SCIREA Journal of Education*, 10(2), 69–80.

ABOUT THE AUTHOR

Victoria L. Malakhova

DSc in Linguistics, Associate Professor

English Language Department No. 5, Institute of International Trade and Sustainable Development

MGIMO University, Moscow, Russia

Postal address: 76B Prospekt Vernadskogo, 119454 Moscow, Russia

Email: v.l.malakhova@inno.mgimo.ru

ORCID: <https://orcid.org/0000-0002-0261-5493>