

The Multimodal Approach in Language Teaching



Kieran Donaghy

Executive Summary

The Multimodal Approach in English Language Teaching is an innovative pedagogical framework designed to reflect the realities of communication in our digitally and visually mediated world. As learners increasingly engage with multimodal texts such as short films, social media posts, podcasts, memes and video essays, language education must evolve to integrate not only the traditional skills of listening, speaking, reading and writing, but also the equally essential skills of *viewing* and *representing*.

Grounded in research from multimodal literacy, cognitive science and sociocultural theory, the approach draws on Vygotsky's social constructivism and Mayer's Cognitive Theory of Multimedia Learning to support deeper, more meaningful learning. Through a structured five-stage process—Expectation, Exploration, Evaluation, Enhancement and Expression—learners are guided to critically interpret multimodal texts and creatively produce their own, reinforcing vocabulary, developing linguistic competence and fostering empathy.

The approach also promotes inclusion and accessibility, offering learners multiple pathways to understanding and self-expression. Featured in all Film English lesson plans, the Multimodal Approach makes language learning more authentic and engaging—empowering learners not only to communicate effectively, but to connect emotionally, think critically and express their identities in an interconnected world.

The Multimodal Approach in English Language Teaching

Introduction

In today's media-rich environment, the vast majority of texts students engage with daily—such as YouTube videos, Instagram reels, TikTok clips, streaming series, video games and blog posts—are not only **digital** but also **multimodal**. These texts combine language with image, sound, movement and spatial design to convey meaning, often requiring learners to interpret and respond across multiple modes simultaneously. As digital technologies evolve, the pervasiveness and complexity of multimodal communication grow, reshaping both how we communicate and what it means to be literate.

Contemporary communication extends far beyond speech and writing; it is a layered, dynamic process that includes visual, aural, gestural and spatial elements. As such, **literacy** can no longer be understood solely as a linguistic skill. To participate fully in today's social, academic and professional contexts, learners need more than the ability to read and write alphabetic text—they need to **view, interpret, evaluate and represent** meaning through a range of semiotic resources. In this context, traditional notions of literacy are insufficient and language education must evolve to reflect the **multimodal** nature of communication in the 21st century.

A mode

In very simple terms, a mode is a means of communicating.

There are five modes of communication: linguistic, aural, visual, gestural and spatial, as illustrated in Figure 1.

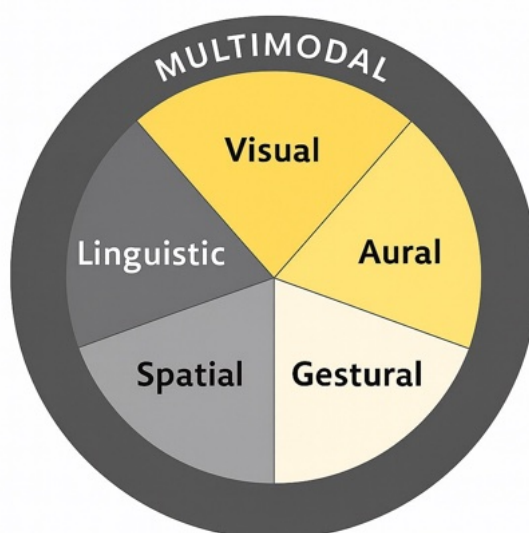


Figure 1

Linguistic mode

The **linguistic mode** refers to the use of spoken or written language to convey meaning. It includes vocabulary, grammar, syntax, word choice, and the structure of sentences and texts. The linguistic mode forms the foundation of reading, writing, speaking and listening activities in language learning.

Visual mode

The **visual mode** relates to the meaning communicated through what is seen. This includes images, illustrations, symbols, signs, diagrams, graphs, charts and videos. It also encompasses elements of visual design such as colour, layout, font style, and size, which influence how information is perceived and understood.

Aural mode

The **aural mode** involves meaning conveyed through sound. This includes spoken language, music, ambient noise, silence and sound effects. Meaning in the aural mode is shaped by features such as tone, pitch, volume, rhythm and pace, all of which affect emotional and interpretive responses to auditory input.

Gestural mode

The **gestural mode** refers to communication through physical movement. This includes facial expressions, gestures, body language, eye contact, posture and interaction between individuals. Gestural communication often complements or enhances spoken language by adding emotional nuance and emphasis.

Spatial mode

The **spatial mode** concerns how meaning is created through the arrangement of space. It includes layout, organisation and proximity in both physical and digital environments—for example, the arrangement of elements on a page or screen, or the distance between people in an interaction. Spatial design influences how information is navigated, prioritised and interpreted.

A multimodal text

A **multimodal text** is a text that combines two or more modes of communication—such as linguistic, visual, aural, gestural and spatial—to convey meaning. These texts draw on multiple semiotic resources to enhance understanding, engagement and expression.

Multimodal texts can be **paper-based** or **digital** in nature:

- **Paper-based multimodal texts** include newspapers, picture books, textbooks, graphic novels, comics, posters, flyers, brochures, magazine articles and instructional diagrams. These texts typically combine **linguistic** (written language), **visual** (still images) and **spatial** (layout and arrangement) modes to communicate meaning.
- **Digital multimodal texts** include YouTube videos, TED Talks, social media posts, vlogs, slide presentations, video games, interactive eBooks, video calls (e.g., Zoom or Teams) and websites. These texts often incorporate **linguistic** (written and spoken language), **visual** (still and moving images), **aural** (sound, music, silence), **gestural** (body language, facial expressions) and **spatial** (digital layout and positioning) modes.

As communication becomes increasingly digital and media-rich, learners need to develop the ability to interpret and create multimodal texts fluently. Understanding how these modes interact supports deeper comprehension, more effective communication and the development of critical media literacy skills.

Multimodality

Multimodality refers to the use of two or more modes of communication—linguistic, visual, aural, gestural and spatial—to construct meaning within a single text. It recognises that meaning is not conveyed by language alone but is co-constructed through the interaction of various semiotic resources. Multimodality highlights how each mode contributes uniquely to a message and how the combined effect of these modes enhances communication.

Figure 2 illustrates how multimodality involves the use of two or more modes of communication within a single text.

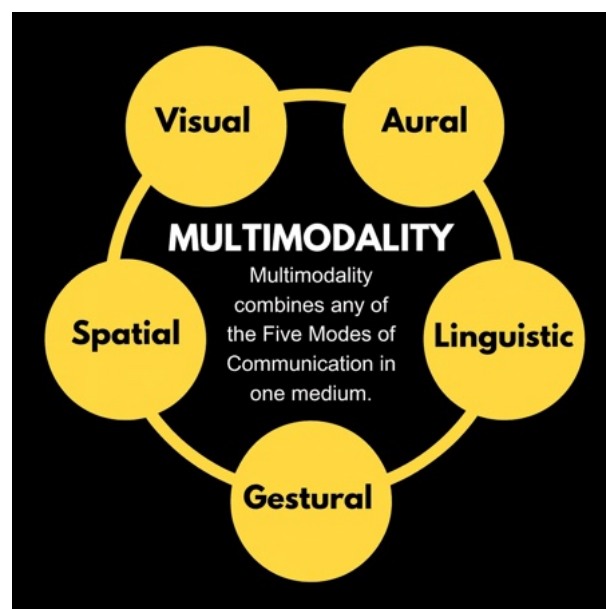


Figure 2

As communication becomes increasingly digital, multimodality has become central to how we interact, learn and share information. Today's texts—whether social media posts, video essays or infographics—are rarely monomodal. The growing focus on multimodal texts in education reflects this shift, emphasising the need for learners to understand how meaning is made across multiple modes in order to become effective, critical communicators in a media-rich world.

Multimodal literacy

Multimodal literacy refers to the ability to interpret, evaluate and create texts that combine multiple modes of communication—such as linguistic, visual, aural, gestural and spatial elements. It brings together the interrelated practices of **viewing** (comprehending and

analysing multimodal texts) and **representing** (producing multimodal texts), positioning them alongside the traditional language skills of listening, speaking, reading and writing. As contemporary communication increasingly takes the form of digital, image-rich and media-driven texts, education systems worldwide have begun to adapt their curricula accordingly. For example, Singapore's 2020 English Language Syllabus integrates viewing and representing explicitly, recognising the need for learners to make meaning across modes in diverse multimodal contexts. Multimodal literacy equips learners with the analytical and creative skills necessary to navigate this evolving communication landscape.

Viewing is the active, interpretive process of engaging with multimodal texts. It requires learners to notice and decode the various modes that contribute to a text's meaning—such as images, gesture, sound, and spatial layout—while responding both emotionally and critically. Viewing goes beyond surface comprehension to involve analysis of visual and auditory elements, evaluation of the creator's intent and reflection on personal or cultural implications. As a receptive skill, viewing complements and extends traditional reading and listening practices in a multimodal environment.

Representing refers to learners' ability to communicate ideas, feelings and information through the creation of multimodal texts. These may include print-based compositions such as posters, comic strips or collages, as well as digital formats like infographics, video essays, memes and slideshow presentations. Closely linked with the productive skills of writing and speaking, representing encourages learners to synthesise knowledge, use design principles, and develop voice and perspective across modes. Multimodal composition fosters critical digital literacy, creative expression, and personal agency—essential components of a well-rounded language education in the 21st century.

Why multimodal literacy is increasingly important in language teaching

As communication today increasingly takes place through multimodal texts—combining language with images, sound, gesture and spatial design—language education must evolve to reflect this reality. While language remains central to English language teaching, developing **multimodal communicative competence** is now essential. Learners must be able to analyse how visual and other non-linguistic modes interact with written and spoken language to construct meaning.

Incorporating multimodal literacy into the classroom bridges the gap between traditional language instruction and learners' real-world experiences. Outside the classroom, students are regularly exposed to multimodal texts such as YouTube videos, social media posts, blogs, websites and memes. By providing opportunities to view and analyse these texts critically in class, teachers help students navigate the visual and cultural codes embedded in everyday media. Learners become more discerning consumers of information and more effective communicators.

Research shows that working with multimodal texts can significantly improve learners' **reading and listening comprehension, vocabulary acquisition and retention, motivation**

and critical thinking. In addition to interpreting these texts, students also benefit from composing their own multimodal texts—vlogs, infographics, memes, digital stories—which allows them to apply language in personally meaningful and creative ways. This dual process of viewing and representing reinforces language learning and enhances communicative competence.

Importantly, multimodal literacy also promotes **accessibility and inclusion**. Learners with reading difficulties, sensory impairments or diverse learning needs often struggle with traditional print-based instruction. Multimodal texts provide alternative pathways to meaning, allowing information to be accessed and expressed through visuals, sound, movement and space. This not only supports learners with special educational needs but benefits all students by offering multiple modes for engagement and expression. By embedding multimodal materials and encouraging students to compose multimodal texts, teachers create **more equitable and inclusive classrooms**. Multimodal learning reflects how people communicate in real life—and, crucially, supports the development of diverse literacies needed for participation in contemporary society.

In summary, integrating multimodal literacy into language teaching fosters deeper engagement, enhances learning outcomes and equips students with the skills to navigate and contribute meaningfully to today's complex communicative landscape.

The Multimodal Approach

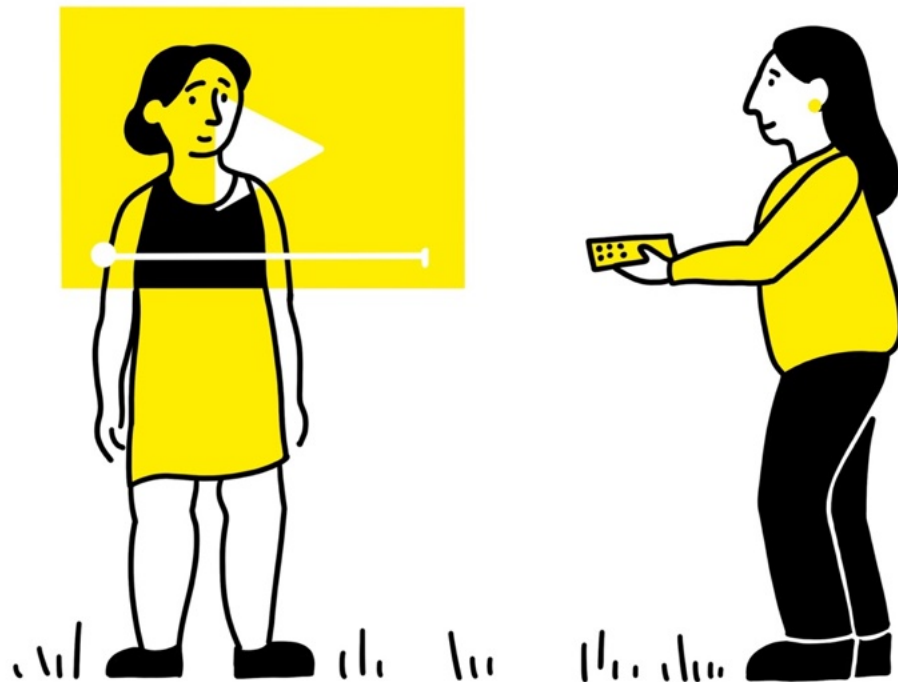


When integrating **multimodal literacy** into classroom practice, it is important that the skills of **viewing** and **representing** do not feel radically unfamiliar or overwhelming to teachers. These skills should not be perceived as an add-on but rather as natural **extensions of existing literacy instruction**. **Viewing**, for example, can be taught alongside **listening** and **reading**, while **representing** can be incorporated into **writing** instruction. By framing these multimodal skills as connected to **traditional literacy processes**, teachers can approach them with greater **confidence** and **coherence**. Both teachers and learners must understand that the processes of **viewing** and **representing** are just as essential to **meaning-making** as **listening**, **reading** and **writing**.

The **Multimodal Approach** offers an **innovative pedagogical framework** that introduces new literacies—specifically the skills of **viewing** and **representing**—while simultaneously integrating and adapting key elements of **traditional literacy instruction**. It supports teachers in bridging the familiar with the new, making multimodal teaching both **accessible** and **pedagogically sound**. Most teachers are already using multimodal texts in their classes but need structured and robust frameworks to help them integrate multimodal teaching more effectively.

At its core, the **Multimodal Approach** is grounded in the recognition that **communication today is no longer exclusively linguistic**. Instead, it involves the dynamic interplay of

multiple semiotic modes—language, image, sound, gesture and spatial design. This is especially relevant in today’s **media-rich environment**, where learners engage daily with multimodal texts such as **YouTube videos, films, infographics, podcasts, social media posts** and **interactive apps**. The approach reflects this reality, preparing learners not only to **interpret** such texts **critically** and **meaningfully** but also to **produce multimodal texts** that communicate effectively across diverse contexts.



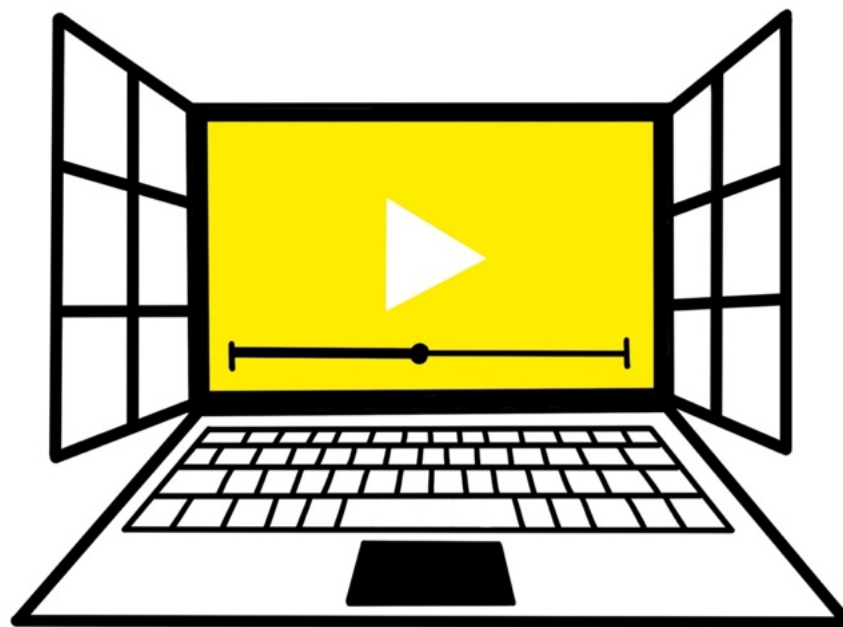
The **Film English** lesson plans, developed using the Multimodal Approach, exemplify this pedagogy in action. Each lesson is built around an authentic **short video** or **film** and structured to develop learners’ skills—**listening, viewing, speaking, reading, writing** and **representing**—in an **integrated, meaningful sequence**. Rather than isolating skills, the lessons model how **language functions within a broader system of meaning-making**, where **visuals, sound, gesture, facial expression, layout** and **tone** all contribute to **communication**. This approach makes language learning more **authentic, cognitively engaging** and **emotionally resonant** for students.

The Multimodal Approach is underpinned by a **robust theoretical foundation**, drawing on insights from **cognitive science, sociocultural theory** and **multimodal literacy research**. Central to this is **Richard Mayer’s Cognitive Theory of Multimedia Learning**, which asserts that learners understand and retain information more effectively when it is presented through **both verbal and visual channels**. This **dual-channel processing** enhances **comprehension, retention** and **engagement**—core principles embedded in the design of multimodal lessons.

The Multimodal Approach also aligns with **Vygotsky’s sociocultural theory**, which views learning as a **socially situated** process, mediated by **tools, language** and **interaction**. **Multimodal texts**—such as **short films, memes, infographics** and **video blogs**—act as

cultural artefacts, offering students opportunities to **negotiate meaning** in socially and emotionally resonant ways.

The Multimodal Approach further draws on **Stephen Krashen's Affective Filter Hypothesis**, which suggests that **emotional engagement lowers anxiety** and increases **receptivity to input**. The approach deliberately uses **emotionally powerful, visually expressive texts** to create meaningful, personal connections. Themes such as **emotional intelligence, character, values, empathy, personal development, identity, relationships, global issues and social issues**, encourage learners to relate to characters and content on an **affective level**, which supports deeper **processing of language and meaning**.



In addition to emotional engagement, the approach addresses **cognitive load**, incorporating principles from **Cognitive Load Theory** (Sweller, 1988). Providing learners with a **glossary of key vocabulary** before viewing helps reduce **extraneous cognitive load**, allowing students to focus on **comprehension** and **multimodal analysis**. This **scaffolding** not only facilitates learning but also contributes to lowering the **affective filter** by increasing learner **confidence** and reducing **anxiety**.

Another defining feature of the Multimodal Approach is its explicit emphasis on **fostering empathy**. Unlike traditional literacy practices, multimodal texts integrate **facial expressions, tone of voice, body language, visual symbolism, music and silence**—elements that convey **emotional subtext** and **cultural nuance**. These give learners opportunities to **interpret emotional states, consider diverse perspectives** and build **emotional literacy**. The result is a language learning experience that cultivates **intercultural awareness, perspective-taking** and **emotional intelligence**.

The design of the Multimodal Approach is informed by the pioneering work of **Gunther Kress**, whose scholarship laid the theoretical foundation for understanding **multimodality** as

central to contemporary communication and literacy. Kress argued that meaning is made through the **interaction of multiple modes**—not just through language, but also through image, layout, gesture, sound and spatial arrangement—and that these modes are chosen and orchestrated according to the **social, cultural and communicative needs** of the moment. His concept of **multimodal design** reframed literacy as a dynamic, socially situated act of meaning-making across modes, profoundly influencing how educators understand and teach communication in the 21st century.

Building on this foundation, the Multimodal Approach is also shaped by the work of **Victor Lim** and **Lynda Tan-Chia**, whose work on **multimodal literacy pedagogy** underscores the importance of **structured, developmental learning processes**. While the Multimodal Approach adapts and extends their framework, it retains their focus on **progressive meaning-making** across multiple modes. Their influence is evident in the **sequencing of learning processes** within the Multimodal Approach framework—from **initial interpretation**, through **critical analysis**, to **creative multimodal production**—all scaffolded to develop both **critical** and **empathetic engagement** with texts.

Pedagogically, the Multimodal Approach supports **communicative competence, vocabulary acquisition and retention, critical thinking** and **emotional intelligence**. Learners engage in **pre-viewing vocabulary scaffolding, multimodal analysis, and meaningful communicative tasks**—including **discussion, reflection, writing** and **roleplay**—that integrate new language in **contextualised, emotionally resonant ways**. This holistic process leads to **deeper engagement, stronger motivation** and more **authentic learning outcomes**.

The Five Multimodal Learning Processes

The Multimodal Approach framework incorporates five interconnected learning processes that develop learners' skills in viewing and support their ability in representing meaning across modes.

Figure 3 illustrates the five learning processes of Expectation, Exploration, Evaluation, Enhancement and Expression, and these are considered in more detail below.

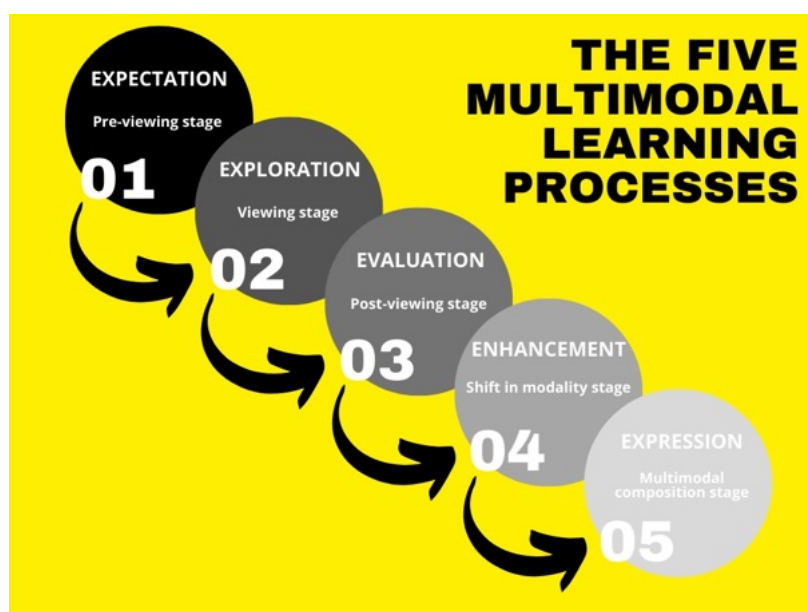


Figure 3

1. Expectation (*Pre-viewing stage*)

The first process, **Expectation**, focuses on activating learners' **prior knowledge and schema**. Learners prepare for viewing by reflecting on personal experiences related to the video's theme, engaging in discussion questions and interpreting related quotations. They also speculate on the content based on the video's title, narrator background or context. These activities establish purpose, spark curiosity and prompt prediction—critical components of schema activation.

Additionally, learners are introduced to a **video glossary** of key vocabulary. This reduces cognitive load and lowers the affective filter, enabling more confident comprehension and later productive use of the target language.

2. Exploration (*Viewing stage*)

In the **Exploration** stage, the focus shifts to **cognitive understanding**. Learners view the video multiple times, each with a clear purpose. They answer comprehension questions, compare responses with peers and build multimodal awareness. The teacher facilitates this process by guiding learners to identify, interpret and support their responses using evidence from different modes (e.g. spoken language, written language, gesture, facial expression, tone, sound).



3. Evaluation (*Post-viewing stage*)

The third process, **Evaluation**, encourages learners to adopt a **critical perspective** on the multimodal text. They move beyond comprehension to analyse how meaning is constructed through various modes. Tasks in this stage include responding to post-viewing multimodal analysis questions, reflecting on what they've learned, engaging in critical thinking discussions, and participating in roleplay scenarios related to the video's topic. Learners actively use target vocabulary in these productive tasks, applying language meaningfully while questioning representations and viewpoints in the video.

4. Enhancement (*Shift in modality stage*)

The **Enhancement** process involves a shift in modality—learners revisit the topic through a **thematically related reading text**. This enhances learning by:

- Reintroducing vocabulary in a new modality.
- Offering additional perspectives.
- Strengthening reading strategies (prediction, skimming, scanning, inference).
- Providing reflective space for more introverted learners.

From a multimodal literacy standpoint, this shift from **audiovisual to written** mode deepens understanding by allowing learners to engage with the same theme through a different semiotic system. This strengthens meaning-making and cross-modal integration.

5. Expression (*Multimodal composition stage*)

The final process, **Expression**, centres on **multimodal text creation**. Here, learners become active producers of meaning. They are encouraged to respond to what they've viewed and read by **creating their own multimodal texts** using language, image, sound, gesture and layout. Tasks include producing vlogs, infographics, comic strips, memes, or social media posts. Typically completed as homework, these creative projects empower learners to express their ideas, identities and interpretations in authentic, engaging formats—consolidating their multimodal competence while fostering voice and autonomy.

How the Multimodal Approach Develops Linguistic Competence through the Five Learning Processes

The Multimodal Approach supports the development of linguistic competence by engaging learners in meaningful interaction with language at all stages of the learning cycle. In the **Expectation stage**, learners are introduced to the theme of the video and engage in pre-viewing discussion tasks using guided prompts. These activities activate prior knowledge and provide opportunities for spoken interaction, laying the foundation for vocabulary use and discourse awareness. Learners are also exposed to new language in the form of a glossary, priming them for receptive understanding and future production.

During the **Exploration stage**, learners watch a short film or video and respond to structured viewing tasks that focus on both literal and inferential comprehension. These tasks develop listening sub-skills—such as identifying main ideas and understanding tone and intention. Crucially, learners observe how linguistic elements operate in context alongside visual, aural and gestural cues, enhancing their ability to interpret authentic spoken language with natural speed, intonation and variation in register.

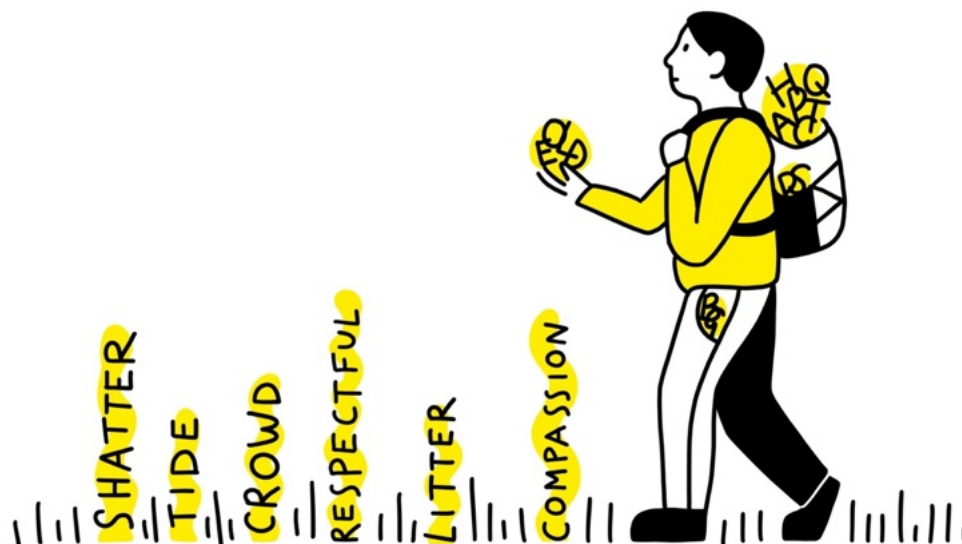
The **Evaluation stage** strengthens linguistic competence by encouraging learners to articulate interpretations, opinions and arguments about the video. In post-viewing tasks such as multimodal analysis of the video, reflective writing, critical thinking discussions and roleplays, learners use the target language productively in extended speech or writing. These tasks develop grammatical accuracy, fluency, coherence and the pragmatic use of language in socially and culturally appropriate ways.

In the **Enhancement stage**, students read a thematically related article that revisits key vocabulary and structures in a new genre. This supports syntactic development and lexical expansion by exposing learners to complex sentence structures and cohesive devices. It also builds awareness of how language varies across modes and purposes, helping learners consolidate their receptive linguistic competence in reading while reinforcing cross-modal transfer of language knowledge.

Finally, in the **Expression stage**, learners create their own multimodal texts—such as digital presentations, storyboards, blog posts or video recordings—using target language in personalised and purposeful contexts. These tasks integrate reading, writing, speaking and listening in an authentic communicative cycle. Learners focus not only on linguistic form but also on how language interacts with other modes to convey meaning clearly and persuasively.

Through this carefully scaffolded process, the Multimodal Approach cultivates all components of linguistic competence—phonological, lexical, grammatical, semantic and pragmatic—by embedding language learning in emotionally engaging, cognitively challenging and socially meaningful tasks.

How the Multimodal Approach Develops Vocabulary Acquisition and Retention through the Five Learning Processes



The Multimodal Approach promotes robust vocabulary learning by embedding key vocabulary across all five stages of its learning cycle, ensuring repeated, meaningful exposure to new language in varied and multimodal contexts. In the **Expectation stage**, learners are introduced to a glossary of essential words and phrases drawn directly from the video they will view. This pre-teaching reduces cognitive load and primes learners for receptive understanding, allowing them to recognise and anticipate vocabulary use in context. By activating prior knowledge and discussing thematic concepts using these new terms, learners begin to make initial connections between word, meaning and context.

In the **Exploration stage**, learners encounter target vocabulary through spoken dialogue and captions and receive other rich, multimodal input— gestures, facial expressions, sound effects and visuals. This layered input strengthens word recognition and deepens semantic understanding. Through guided comprehension tasks, learners are encouraged to notice how language operates across modes, reinforcing both form and meaning. The **Evaluation stage** prompts students to actively retrieve and use new vocabulary during critical discussions, written reflections and roleplays. This productive use, especially when linked to emotionally engaging themes, solidifies retention through deeper cognitive processing.

Vocabulary learning continues in the **Enhancement stage**, where students engage with a thematically linked written text. Here, key words reappear in a new modality, offering a second encounter in a different discourse context. This reinforces lexical patterns, expands understanding, and introduces collocations or synonyms. Finally, in the **Expression stage**, learners are tasked with creating their own multimodal compositions—vlogs, short animations, slideshows, infographics, comic strips, social media posts, memes—using newly acquired vocabulary to communicate original ideas. This stage offers opportunities for ownership, personalisation and creative use of language, which are powerful drivers of long-term retention.

By recycling vocabulary across multiple stages, tasks and modalities, the Multimodal Approach ensures learners encounter, process and produce new language in dynamic, memorable and meaningful ways.

How the Multimodal Approach Fosters Empathy through the Five Learning Processes



Empathy—the ability to understand and share the feelings of others—is fundamental to meaningful communication, social cohesion and intercultural competence. The Multimodal Approach fosters empathy by guiding learners through five interconnected learning processes that develop not only language proficiency but also emotional intelligence and human connection.

In the **Expectation** stage, learners are invited to share personal experiences and reflect on their identities in relation to the video's theme. By listening to their teacher's and peers' stories and revealing something about themselves, learners begin to foster mutual understanding and emotional openness. Discussing thematic questions and quotations, and making predictions encourages them to anticipate diverse perspectives and prepare to engage with the emotional and conceptual world of the video. The videos learners will explore in the next stage often address powerful, and socially and emotionally relevant themes such as emotional intelligence, character, values, empathy, personal development, identity, relationships, global issues and social issues. These topics prompt learners to confront and reconsider their assumptions and biases. Engaging in pre-viewing discussions about these issues not only supports language development but also nurtures global citizenship competencies—including respect for difference, ethical awareness and emotional

sensitivity. This stage thus lays the foundation for empathy through introspection, connection and perspective-taking.

During the **Exploration** stage, as learners view the video multiple times, they attend closely to multimodal elements—facial expressions, tone of voice, gesture, background music, silence and visual framing. These rich, affective cues support learners in perceiving the emotions and intentions of characters, making implicit meaning more accessible. This close, repeated observation fosters emotional resonance and encourages learners to infer feelings beyond what is said—deepening their emotional literacy.

In the **Evaluation** stage, learners begin to interpret the multimodal text through a critical and personal lens. They question the values presented or imagine what they would do in similar situations. Roleplays and reflective writing based on the video allow learners to inhabit the perspectives of others. These activities explicitly ask learners to adopt alternative viewpoints and articulate them in the target language, bridging cognitive reasoning with emotional engagement—essential for developing empathetic understanding.

The **Enhancement** stage, which involves reading a thematically linked article, offers learners a chance to revisit the topic in a new modality. Reading deepens reflection, especially for quieter or more introspective learners who may process emotion more privately. Articles often provide additional voices, cultural perspectives or background knowledge that expand learners' worldview. This shift in mode reinforces the topic while encouraging perspective-taking and empathy through language-rich analysis.

Finally, in the **Expression** stage, learners become creators of meaning through multimodal text production—vlogs, infographics, comic strips, or social media posts. These tasks allow learners to express their views, tell stories and reflect on what they've learned emotionally and cognitively. When learners speak through the voice of a character or advocate for a social issue explored in the video, they take empathetic understanding into the realm of creation, positioning themselves as emotionally aware communicators.

At every stage, the Multimodal Approach integrates visual, aural, gestural, linguistic and spatial meaning-making to support both linguistic and empathetic growth. It cultivates a classroom culture where learners are not only language users but also listeners, interpreters and storytellers who value emotional and cultural connection. In this way, empathy is not an add-on—it is an outcome embedded in the very design of multimodal language learning.

Conclusion

The **Multimodal Approach** is an innovative pedagogical framework that reflects the realities of communication in a visually and digitally mediated world. It integrates traditional language skills—listening, speaking, reading and writing—with the newer, equally essential skills of **viewing and representing**, enabling learners to engage critically, creatively and empathetically with a wide range of multimodal texts.

Grounded in **multimodal pedagogy**, this approach recognises that meaning is made not only through language but also through the interplay of visual, aural, gestural and spatial modes. By actively working across these modes, learners become more proficient, flexible and expressive communicators.

The Multimodal Approach draws on several key educational theories. **Vygotsky's social constructivism** underpins the view that learning is a socially mediated process shaped by interaction, dialogue, and tools—including digital and visual media. **Mayer's Cognitive Theory of Multimedia Learning** supports the integration of visual and verbal channels, showing that learners retain and understand information more deeply when multiple sensory pathways are engaged. These theoretical foundations justify the use of multimodal texts such as short films, vlogs, infographics, and podcasts, which offer rich, layered input that can support both language development and emotional engagement.

Central to the Multimodal Approach is the cultivation of **empathy**. By engaging learners in themes related to emotional intelligence, character, values, empathy, personal development, identity, relationships, global issues and social issues, the approach provides opportunities for students to encounter diverse perspectives and deepen their emotional and ethical understanding. Multimodal texts—especially short videos and films—are uniquely positioned to foster empathy, as they offer visual and auditory cues (such as facial expression, body language, music and silence) that help learners connect with characters on an affective level.

The Multimodal Approach embeds a sequence of carefully structured learning processes: **Expectation, Exploration, Evaluation, Enhancement** and **Expression**. These interrelated stages guide learners through activating prior knowledge, understanding multimodal input, developing critical perspectives, reinforcing content through alternative modes and finally creating their own multimodal texts. This process nurtures both linguistic competence and emotional intelligence by inviting learners to respond reflectively, communicate authentically and create meaning purposefully.

In essence, the Multimodal Approach enables students not only to **learn English** but also equips them with the skills to interpret and produce the kinds of complex, multimodal texts they encounter every day, while also developing empathy, intercultural awareness and the confidence to express their identities and ideas in meaningful ways.

References

- Byram, M. (1997). *Teaching and Assessing Intercultural Communicative Competence*. Clevedon: Multilingual Matters.
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press.
- Lim, F. V., & Tan-Chia, L. (2022). *Designing Learning for Multimodal Literacy*. Taylor & Francis.
- Kress, G. (2003). *Literacy in the new media age*. Routledge.
- Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). New York: Cambridge University Press.
- Ministry of Education (MOE), Singapore. (2020). *English language teaching and learning syllabus 2020*. Ministry of Education.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

About the Author



Kieran Donaghy is the founder and creative force behind Film English. A leading expert in the use of film and video in English language teaching, his work bridges the fields of visual literacy, multimodality and language pedagogy and has shaped the way educators use multimodal texts in the classroom. He is the author of books for students and teachers of English as a foreign language. His publications include *Film in Action* (Delta Publishing), *Writing Activities for Film* (ELT Teacher2Writer) and *Video, The Image in ELT* (ELT Council), *Language Hub* (Macmillan) and *Multimodality in ELT* (Oxford University Press). He trains teachers in Barcelona and online at his specialist teacher development institute, The School for Training.

Find out more about Kieran at: <http://kierandonaghy.com/>

You can contact Kieran at: kieran@theschoolfortraining.com

About Film English



Film English is a groundbreaking, award-winning website dedicated to providing high-quality teaching resources based on short videos and feature-length films. Designed for English language teachers seeking fresh, contemporary materials, the site offers a wide collection of Video Lesson Plans, Extensive Viewing Guides, and Film Worksheets—all created to harness the educational potential of authentic, multimodal content.

Find out more about Film English at: <https://film-english.com/>

About Film English Club



Membership of Film English Club allows unlimited access to all the Video Lesson Plans, Extensive Viewing Guides and Film Worksheets from our vast library.

Find out more about Film English Club at: <https://film-english.com/product-category/memberships/>

About The School for Training



The School for Training is a centre of excellence in teacher development offering innovative courses to language teachers in the centre of Barcelona and online. The school offers innovative courses on important areas of language pedagogy not normally covered on CELTA or DELTA courses such as how to use short films and videos, the role of empathy in language education, how to introduce social justice issues in the language classroom, and the fifth skill of Viewing in the language curriculum. Find out more about The School for Training at: <https://theschoolfortraining.com/>



Materials that matter for teachers who care

FILM ENGLISH

