

Multimodal Literacy: What it Means for English Teachers in Indonesia and the Challenges it Brings¹

Budi Hermawan², Universitas Pendidikan Indonesia, West Java, Indonesia

Abstract

The official launch of the latest curriculum, *Kurikulum Merdeka* [Freedom Curriculum], has marked a significant shift in the concept of literacy, especially in English teaching and learning in Indonesia. This curriculum also serves as recognition that the literacy we are aiming for now in English teaching and learning is no longer focused on the literacy that merely highlights verbal language as the only mode to make meaning in the classroom, but it now aims for the literacy that recognizes the use of multimodes to make meanings. This article discusses multimodal literacy in the context of English curriculum in Indonesia. It describes the shift in the type of literacy required in the curriculum to change from verbal literacy to multimodal literacy, which is carried out through the teaching and learning of multimodal texts; the text genre serving as the learning material presented in an English textbook under the latest *Kurikulum Merdeka*. This article defines what multimodal literacy means for teachers and students. It identifies likely theoretical and practical challenges teachers can face in teaching multimodal texts in order to develop multimodal literacy in the classroom. An example of multimodal activities is provided for teachers. The article then highlights the need to provide the teachers with the metalanguage of multimodality as it is an essential part of developing multimodal literacy.

Resumen

El lanzamiento oficial del plan de estudios más reciente, *Kurikulum Merdeka* (Currículo de la Libertad), ha marcado un cambio significativo en el concepto de alfabetización, especialmente en la enseñanza y el aprendizaje del inglés en Indonesia. Este plan de estudios reconoce que la alfabetización a la que se aspira actualmente en la enseñanza del inglés ya no se centra únicamente en el lenguaje verbal como el único modo de construir significado en el aula, sino que ahora apunta a una alfabetización que reconoce el uso de múltiples modos para generar sentido. Este artículo aborda la alfabetización multimodal en el contexto del plan de estudios de inglés en Indonesia. Describe la transición en el tipo de alfabetización requerida —pasando de la alfabetización verbal a la multimodal—, la cual se lleva a cabo mediante la enseñanza y el aprendizaje de textos multimodales; estos textos actúan como material didáctico en los libros de texto de inglés bajo el marco del *Kurikulum Merdeka*. El artículo define qué implica la alfabetización multimodal para docentes y estudiantes, e identifica los posibles desafíos teóricos y prácticos que los profesores pueden enfrentar al enseñar textos multimodales con el fin de desarrollar dicha alfabetización en el aula. Asimismo, se ofrece a los docentes un ejemplo de actividades multimodales. Finalmente, el artículo destaca la necesidad de dotar a los docentes de la metalengua de la multimodalidad, ya que constituye un elemento esencial para el desarrollo de la alfabetización multimodal.

Introduction

Many people consider that verbal (oral) language is their main representational tool despite the fact that in communicating we now never employ only one tool to create meaning since communication now increasingly includes many different kinds of tools. They use oral language to communicate their ideas and seem to believe that it is the only tool available for them to communicate. However, there are now a growing number of people with increasing awareness that other 'languages', including non-verbal/oral language can be used along with oral language or with other non-verbal languages to communicate their ideas. These 'languages' or tools of communication are termed as modes in multimodal studies. Indeed, communication is becoming more and more multimodal (Kress & van Leeuwen, 2012; Christie, 2005, as cited in Bezerra, 2012).

Multimodal as a term was initially used to refer to the fact that when we communicate, we often do not use one single mode, that is verbal language only, but a combination of the visual, sound and other modes in addition to the oral language. Hence, we communicate multimodally (Kress & van Leeuwen, 1996). Being 'multimodal' then becomes 'multimodality' the noun of being multimodal, which is defined and understood as "the combination of different semiotic modes — for example, language and music — in a communicative artifact or event" (Van Leeuwen, 2005, p. 28); "representation, communication and interaction as something more than language" (Jewitt, 2014, p. 1); and "a recognition of the fact that people use multiple means of meaning making" (Bezemer & Jewitt, 2018, p. 282). As more research from different traditions investigating multimodality as an important phenomenon in communication has been done, the concept of multimodality

¹ Received: 17 September, 2023. Accepted: 10 November, 2023. Published: 17 August, 2026.

² linquistikmakna@gmail.com, 0000-0002-7192-6439

has varied widely, both across and within the different disciplines and research traditions (Bezemer & Jewitt, 2010).

As the use of different modes at the same time is one of the important features of today's communicative acts, our success of encoding our messages and decoding the messages communicated to us, then, to a large extent depends on our access and mastery of the modes. This is also true for the communication that takes place in classrooms which is largely related to the learning materials that are delivered by the teachers to the students, often guided by the textbooks used. Textbooks used in the classroom are now increasingly multimodal with more images serving the semiotic function previously carried out by writing alone (Bezemer & Kress, 2016). The success of the student's communicating and understanding the learning materials presented in the textbook depends on their access and understanding of the modes themselves. As the textbooks used often reflect the 'types' of literacy aimed to be mastered by the students, multimodal textbooks are aimed at teaching and developing literacy that recognizes the use of different modes for students to use for communicating both in the classroom and in their future workplaces. In other words, multimodal textbooks aim at developing multimodal literacy in the students.

Multimodal literacy as a term has gained more and more popularity. With increasing multimodal investigation done in the field of education, multimodal literacy is a term which is highly discussed by educators, researchers, and text analysts. Many have tried to elaborate on it, however, as many other important terms, its meaning often shifts from what the term initially was intended to mean. This article attempts to define what multimodal literacy is in the context of the English language curriculum in Indonesia. It explains what multimodal literacy means for teachers and students and elaborates the implication of multimodal literacy on the teaching and learning of English and the challenges faced by the teachers. Additionally, this article provides an example of activities aimed at developing multimodal literacy for teachers to get some idea of applications in their class.

Defining multimodal literacy

Multimodal literacy as a term derives from social semiotics, and points to the study of language that combines different modes of meaning at the same time (Mills & Unsworth, 2017). Multimodal literacy originates from social semiotics developed by Hodge and Kress (1988). Later, it was also introduced by van Leeuwen (2005). Multimodal literacy focuses on studying the media and modes of communication that people use including that the choice of modes used to communicate messages are motivated by the interest of their users.

The term mode was adopted from Halliday (1978) by Kress and van Leeuwen (2001, 2006) who extended its use to label representational tools other than verbal language. Halliday had so much influence on Kress and van Leeuwen that many of the theories Halliday developed were adopted and integrated into their works. In addition to modes, Kress also adopted Halliday's metafunction theory which holds that (verbal) language is used to transmit users' experiences, as a tool to enact and maintain interpersonal relations, and that it should be organized into a coherent text to achieve those two purposes. They then extended the metafunction theory into other representational tools people use to form meaning. Kress, van Leeuwen, and multimodal text analysts affirm that modes are socially and culturally shaped resources for meaning-making (Kress, 2009; 2010), and as verbal language also possess those three strands of meanings.

Other representational tools to make meaning have been relied on more and more by people to communicate their meanings along with the verbal as the main representational tool. Kane and Pear (2016) reported that there were 3.8 trillion photos taken in all human history until mid-2011, and 1 trillion photos were taken in 2015 alone. These photos, they reported, were not only used in different social media which originated as chiefly visuals but also were increasingly used on online platforms that had not been designed to be chiefly visual (*Facebook, LinkedIn, and X (once Twitter)*). The increasing use of the visuals has been also observed in English as a foreign language (EFL) textbooks. Bezemer and Kress (2009) reported that textbooks used in England had become increasingly visual, a trend that continues to be documented in teaching materials even though verbal texts retain a dominant role in their structure (Zhou & Liao, 2026). There has not been any report whether the same phenomenon is also true with the English textbooks used in Indonesia.

The increasing use of visuals to communicate ideas whether used alone or accompanying a verbal text has likely become one of the significant features of literacy in this century. Jewitt (2009) refers to the phenomenon of the increasing number and frequency of communications using verbal and visual in particular and using different modes at the same time in general as 'multimodal turn'. As the world now is becoming

increasingly multimodal in nature, it is an inevitable consequence then that people should have the knowledge and skills to be able to contribute and function well in such a world. In other words, to be considered literate in today's world, it is necessary to be able to communicate using different modes (Kress, 2003; Lim, 2018). This literacy in essence is constituted by the ability to use and combine different semiotic modes in ways that are appropriate to a culture and the knowledge of what can be done with the different semiotics modes and how they can be organized into a coherent multimodal text (Van Leeuwen, 2017). Simply stated, being able to make meaning using different modes to achieve the communication goal one sets serves as the core of multimodal literacy.

From the perspective of literacy as a concept and a theory, multimodal literacy is closely linked to multiliteracies as proposed by The New London Group (1996). Cope and Kalantzis (2009) stated that "what was happening to meaning making and representation in the worlds of work, citizenship and personal life" (p. 166) has prompted them and those affiliated to the NLG to reconsider their approach to literacy teaching and learning and they have offered literacy pedagogy, which they called multiliteracies (New London Group 1996; Yi, 2014). The New London Group (1996) said that their pedagogy of multiliteracies was a complement to traditional literacy which they believed emphasized only language and that their pedagogy of multiliteracies focused on modes of representation apart from only language (Cope and Kalantzis, 2000).

Multiliteracies is closely related to multimodality which in this context is referred to the belief that meanings of a text are the result of the orchestration and the working together of different modes used and the practice of making meaning using different modes, hence, multimodal literacy. Yi (2014) asserted that both multiliteracies and multimodality are often used together without distinguishing the two terms from each other when literacy researchers address the multiplicity of communication modes and linguistic/cultural diversity. The closeness between these two terms has been well-stated by Roswell and Walsh (2011), proponents of multimodality, which in their view is the field that takes account how individuals make meaning with different kinds of modes, illuminates how they make meaning, and through multiliteracies as a possible pedagogy, provides tools for making meaning using those modes.

The NLG and those who share this idea hold that we need to distance ourselves from the literacy tradition that emphasized only verbal language to the literacy tradition that acknowledges the use of other 'languages' or so-called modes other than verbal language, that is multiliteracies. This 'new literacy' is the literacy that the NLG believes should be taught and developed, including, of course, in the classrooms for and by the students.

English language curriculum in Indonesia

The curriculum in Indonesia has experienced many, sometimes rapid changes, often regretfully depending on who is in charge of the Office of the Ministry of Education. English language curriculum in specific, has undergone many changes designed to improve learning within contexts, both in the dynamic and unique contexts which are full of challenges, both internal and external (Madya, 2007). It was initially positioned as the first foreign language that should be taught in schools. Its position was later shifted into a compulsory subject taught in secondary schools thanks to the issuance of Presidential Decree No. 28/1990 (Zein et al., 2020).

Along with the change in the curriculum, the change in the focus of English teaching in Indonesia was very obvious. Madya (2007) observed that the teaching of English in Indonesia moved from structure as emphasized in the 1975 curriculum to language use, as emphasized in the 1984 curriculum, and then to the mastery of communicative skills to be acquired through the implementation of the so-called meaning-based approach in the 1995 curriculum. In 2006, the government issued what was called a school-based/level curriculum. Based on the curriculum, schools were given the right to develop their own curriculum with the guidance of national standards determined by the government. Leaning on the curriculum, the teaching of English was directed towards achieving sets of standardized competencies students needed to have, determined by teachers individually with the reference to national competency standards. The competency-based standards were still used as the basis for English teaching in the 2013 curriculum which was designed to replace the 2006 curriculum. Specifically, based on this curriculum English was taught to help students to acquire communicative competence in three types of discourse: interpersonal, transactional, and functional using the text genre taught in each respective level.

Without a doubt the courses have always been motivated towards enabling students to master and develop a set of knowledge and skills under the umbrella of literacy. However, the type of literacy aimed at has

almost always been 'verbal literacy'; literacy that is based on the belief that the verbal mode is the main mode students use to make meaning. However, verbal literacy ignores the other modes available for students to make meaning. In the English textbook this is reflected in how the text genre presented, questions that do not touch the co-deployed modes, and exercises that students need to complete and which do not allow them to either use other modes than the verbal language to make responses nor require them to elaborate on other modes present in the textbook. This is understandable for two reasons. First, verbal mode has always been considered the main representational tool. Second, the awareness of other modes apart from verbal language that can be used to make meaning has grown only recently.

With the pandemic, education in general and teaching English in particular were significantly impacted. To face this daunting challenge, *Kurikulum Merdeka* was designed and released by the Ministry of Education. Based on the curriculum, the teaching and learning should be project-based to develop soft skills and characters that are relevant for the *Pancasila*-based profile of learners, focused on essential materials to allow an in-depth learning for such basic competences as literacy and numeracy, and flexible for teachers to implement differentiated learning (<https://kurikulum.kemendikbud.go.id>).

As an important part of implementing the curriculum the government provides textbooks which often serve as the main learning resources. These learning resources deliver learning contents and activities which are purposely designed not only to be engaging for students, but also to be able to arm them with the knowledge and skills required today. For the purposes of providing quality, cheap and well-distributed around-the-country books, the government, through the Minister of Law and Human Rights based on Law No. 3, 2017, prepares the textbooks for the students and accompanying guidebooks for the teachers. The National Center for Curriculum and Books as the 'operational body' tasked with the provision of textbooks, then commissions writers to design and prepare the textbooks.

The English textbooks based on *Kurikulum Merdeka* for grades 10 to 12 are designed based on the belief that the teaching and learning of English should be directed toward enabling the students to learn different types of text and equip them with the knowledge and skills to produce the texts. The different types of texts students need to learn should no longer be 'traditional' texts in the sense that they only rely on the use of verbal language to communicate their meanings, and that their meaning is perceived and understood to be communicated by the verbal elements of the texts. Students, however, should be able to use texts with different semiotic resources of verbal, visual, audio and the possible combination of these three and other possible resources whose meanings are perceived and understood to be the result of the blending of all the modes used.

This is because first, the texts to which students are exposed, the texts they consume and the texts they produce are increasingly becoming multimodal in the sense that the texts use the combination of different semiotic resources with the visual mode taking more significant role. Hence, the making of text nowadays is concerned with working in different modes and media which include pictures, writing, and video both in the print-based resources and the digital versions (Jewitt, 2012).

Second, authors of the textbooks must believe that literacy cannot be interpreted as a linguistic accomplishment (Jewitt, 2008). Literacy now "extends beyond the conventional view of literacy as printed and written texts" (p. 241). It should include meaning-making practices using digital technologies (e.g., video games, weblogs, mobile texts, etc.) (Sang, 2017) as these have transformed the way in which texts are produced, distributed, and consumed (Kress, 2005). It is fair to say, then, that English language learning under *Kurikulum Merdeka*, shifts the focus of the teaching and learning of English in Indonesia from verbal literacy to multimodal literacy.

Challenges for Teachers

The multimodal nature of the textbook in particular and the multimodal nature of the learning environment in general poses challenges to both the teachers and students such as how to work effectively with the technology, with understanding of the affordances of new texts and resources, as well as knowing the metalanguage with which to talk about them with the students (O'Mara, 2006; Ryan et al., 2010). While technology plays an important role, multimodal texts are not always technologically constructed. Multimodal texts can be paper-based, live, or digital. Paper-based multimodal texts are picture books and of course textbooks. Live multimodal texts include dance, and oral story telling while digital multimodal texts include among others film and slide shows. In addition to the above-mentioned challenges, teachers are also faced with the difficulties of designing learning activities that facilitate students to learn multimodal literacy. In

fact, the challenges faced by the teachers are both theoretical and practical. Some of the challenges are elaborate below.

The metalanguage challenge including the essential terms related to multimodal texts is theoretical as teachers need to understand and master the theories that serve as the ground for the metalanguage. Meanwhile, the practical challenge is concerned with how the teachers carry out the actual teaching learning process.

Theoretical challenge

Textbooks are complete multimodal texts that most teachers have not been trained to use well (Álvarez Valencia, 2016). So, to be able to fruitfully deliver the materials from multimodal textbooks teachers must have good grasp of the language of the textbooks. As a multimodal English textbook uses different languages or modes to present the lesson, teachers need to be equipped with the knowledge of what modes are and the relation between the verbal text and other modes present in the textbook to construct and present its learning messages. To meet this need teachers should comprehend Systemic Functional Linguistics (SFL) as developed by Halliday (1994), Halliday and Matthiessen (2004) and the visual metalanguage as offered and inspired by the works of Kress and Van Leeuwen (2006), which was then further developed by Callow (2013), and Painter et al. (2013).

Developed by Halliday, Systemic Functional Linguistics (SFL) serves as the first essential theory because many of its ideas have been adopted into multimodal studies including into, of course, the concept of multimodal text. Meanwhile, visual metalanguage is essentially required in order for the teachers to be able to help students to comprehend and compose visual meaning in multimodal texts. To be specific in SFL, teachers need to comprehend the clause complex relations from SFL as it has been adopted to explicate the intersemiotic relations of different semiotic resources co-deployed in multimodal texts (see Hermawan, 2019; Liu & O'Halloran, 2009; Martinec & Salway, 2005; Roehrich, 2016; Unsworth, 2006a, b). Intersemiotic relations, the logico-semantic relations between different modes co-deployed in texts, is one of the features of multimodal texts that students need to discuss to unravel the message of the text as a whole. As with the visual metalanguage, teachers need to comprehend such visual semiotic resources for expressing and developing ideas in images as symbols, vector and gaze, size, and colour because these semiotics resources are what teachers can teach students to identify and analyze in multimodal texts to facilitate their obtention of a comprehensive message of the visual images both when appearing alone and when co-deployed with verbal text.

Related to these theories is the operational terms teachers need to comprehend. There are a great number of 'operational terms' whose meanings and referents teachers need to know to avoid misunderstanding when elaborating and discussing a multimodal text with their students. Misunderstanding the terms can result in inappropriate elaboration of the multimodal text discussed with the students. There are at least four core concepts which are common across multimodal research (Bezemer, 2012): mode, semiotic resource, modal affordance, and inter-semiotic relations. The present article will elaborate on only two of them because of the limitation of the scope of the article and because the two terms are closely related as well often create confusion. The first term is 'mode' and the second is 'semiotic resources'. The rest of the terms and other related terms can be learned by the teachers by accessing such sites dedicated to list and explicate terms used in multimodality and multimodal studies as, for example, *Glossary of multimodal terms* (<https://multimodalityglossary.wordpress.com/%20mode-2>) and a highly recommended site in which teachers can learn about teaching and learning multimodality and multimodal text, Vic.Gov.Au (<https://www.schools.vic.gov.au/teaching-materials-and-methods>)

It is essential for teachers to recognize and understand what a mode is as the recognition of different modes in a text and the use of the different modes to make meaning separate multimodal literacy from verbal literacy and multimodal text from monomodal texts. As a term, 'mode' is rapidly growing in its definitions as different multimodal texts' analysts and theorists offer and infuse their perspectives into enriching and expanding its definition. To have a better grasp of what mode is we can turn into social semiotics as one of the three main approaches in multimodal study. According to Mills and Unsworth (2017) social semiotics defines mode as "the socially and culturally shaped resources or organized semiotic structures for making meaning" (p. 5). Kress who is one of the important figures in social semiotics defines mode differently on different occasions with different co-authors. With Gualberto and Kress (2019) defines mode as "the material resources for the production and materialization of signs" (p. 2). Kress and van Leeuwen (2001) defined

modes as “resources which allow the simultaneous realisation of discourses and types of (inter) action” (p. 21). Kress (2010) defined mode as “a socially shaped and culturally given semiotic resource for making meaning” (p. 79). Despite the fact that it is probably one of the most central term, mode is elusive and as lamented by Forceville (2006) “...it is at this stage impossible to give either a satisfactory definition of ‘mode’ or compile an exhaustive list of modes” (p. 382).

The second essential term teachers need to know is semiotic resources. As a term, it originated in the work of Halliday (Lim, 2004; Van Leeuwen, 2005). Van Leeuwen (2005) proposes that semiotic resources are the actions, materials and artefacts we use for communicative purposes, whether produced physiologically or technologically. They have a meaning potential based on their past use. Leaning on opinions of different experts, Jewitt (2009) stated that for Kress and van Leeuwen, semiotics resources can be considered as the connection between representational resources and what people do with them; they are a system of meaning that people have at their disposal (O’Halloran, 2010; O’Toole, 1994). On a different occasion, Kress (2010) defined semiotic resource as resources that are used to organize understanding of the world and to be meaningful in communication with others, or to us (see also Danielsson and Selander, 2021).

Looking at the definitions, the difference between ‘modes’ and ‘semiotics resources’ may not be completely clear. They seem similar to each other even if they don’t refer to the same thing. To this relation, multimodal theorists may be classified into two: those who hold that the two terms refer to two different things (Lim, 2004; O’Halloran, 2010) and others who use the term interchangeably (Kress, 2010; Kress & van Leeuwen, 2002; van Leeuwen, 2005). For Lim (2004) a semiotic resource must have both an expression plane and a content plane; the system for expressing the ideas and the operational set of grammar. Also, colors are not considered a semiotic resource as they are only the system in the expression plane of, for example, a traffic light, which is the semiotic resource. On the contrary, colors for Kress & Van Leeuwen (2002) are a semiotic mode or resource as they can express ideas and also have an operational set of ‘grammar’. Teachers can of course subscribe to either opinion. What is important at the first stage for teachers is to be aware that there are representational tools other than verbal language they must recognize and help their students not only recognize this but also enable them to identify their function in a text, its relation to verbal language, as well as to facilitate them to access and use them to make meaning. They can advance their understanding on the two terms later.

Of the many different definitions of mode offered, the one that seems to be the same in the function it has, is to make meaning. So, to use another definition we can say that a mode is something other than verbal language through which we make meaning. The question, then, is how that ‘something’ is qualified as a mode; how that ‘something’ can be considered as an ‘acceptable’ mode people can use to make meaning. According to Kress (2010) for something to be qualified as a mode first, it must be used in a community as a representation tool to make meaning. Second, the members of the community must believe that it is a mode to make meaning. Other experts (i.e., Kress & Van Leeuwen, 2006; Unsworth, 2006 a, b) add metafunction to the testing questions; for something to be qualified as a mode it should also have the capability to be used to construct ideational, interpersonal, and textual meanings at the same time. Armed with good understanding of what mode is teachers can, then, with students identify what modes are co-deployed in the reading texts used to carry out the learning materials. Teachers can design questions to check reading comprehension that will make students aware of the different modes used in the text and how they collaborate in the text to communicate the whole meaning of the text.

Practical challenge

The second challenge teachers will have to deal with is concerned with how to teach the skills and knowledge of multimodal literacy in the classroom. In other words, practical challenges relate to how teachers design and carry out the teaching and learning, including the assessment in the class. The author’s observations made during tutoring teachers nationally through the program designed by the Office of Ministry of Education to facilitate teachers to get their teachers’ certification confirms that many teachers were not able to design learning steps that were systematically designed for their class. Many of the teachers as their documents showed could not breakdown the learning objectives that should have served as their guide into learning steps that were supportive to the objectives. They must integrate multimodal literacy into the learning activities teachers design which is certainly another tough challenge teachers have to deal with.

Multimodal literacy recognizes modes so that students can select and use them to make meanings. For this reason, teachers need to employ teaching instruction which recognizes and allows this in their classroom.

Teachers need to be able to design classroom activities that facilitate and allow students to be able to select semiotic resources they can access and use to represent their learning experiences in the classroom. Embedded in the teaching's instruction is the assessment teachers need to design which is found to be an ongoing challenge as it, among others, involves deciding which elements or components should be considered to evaluate in multimodal projects students made in the class (Yi & Angay-Crowder, 2016). This teaching practice is a multimodal teaching method which is an important part of multimodal pedagogy; curriculum, pedagogy and assessment practices which focus on mode as a defining feature of communication in learning environments (Stein, 2008; Stein & Newfield, 2006).

It is a great idea to design multimodal literacy classroom activities based on the students' own unique class situation and each student's characteristics. Teachers can adopt and adapt the multiliteracies pedagogy framework proposed by Cope and Kalantzis (2000) which posits that literacy practise should be:

Situated practise

This refers to the skill and knowledge teachers develop in the classroom which should be based and related to what their students need. Literacy should be grounded in the environment of the students who learn it since the skills and knowledge should be directed at facilitating students to learn what skills and knowledge become useful in their environment at the moment of learning. In practice, teachers base and relate their teaching instruction to the needs and knowledge that students have. In practice, this can be done by, for example, allowing students to bring texts of their choice to be discussed together in the classroom thereby expanding their knowledge of the text.

Overt instruction

This feature leads to explicitly teaching the metalanguage students need to know to understand the text. Teachers can label, name, and create a 'grammar' of the text they discuss. While this activity sounds creative, teachers themselves should master the metalanguage necessary to unearth the text that goes beyond simply talking about the parts of speech used in the text.

Critical framing

It is understood here as critically understanding the texts discussed in the classroom. The texts in point are dismantled among other in terms of their purposes, who the target audience is and what values the text tries to disseminate. In other words, a critical framing activity encourages students to question and examine the text beyond simply what is explicitly stated in the texts.

Transformed activity

This is an "immersion in meaningful practices within a community of learners who are capable of playing multiple and different roles based on their background and experiences" (The New London Group, 1996, p. 85). It is the knowledge application moment. In practice, what teachers do is to facilitate students to apply the knowledge they have gained into a learning artefact. Teachers and students create a particular text from what they have learned.

Multimodal Literacy in Reading Session: An Idea for Teachers

To provide teachers with the idea of how multimodal literacy activity could be carried out in their classroom and how the knowledge of the metalanguage could help both the teachers and the students understand of the text from their textbook more critically, the following is the example of paper-based multimodal literacy activity from reading session taken from Hermawan et al. (2022a, b) which elaborated the steps teachers can apply in their classroom. The following example of reading activity is based on the steps.

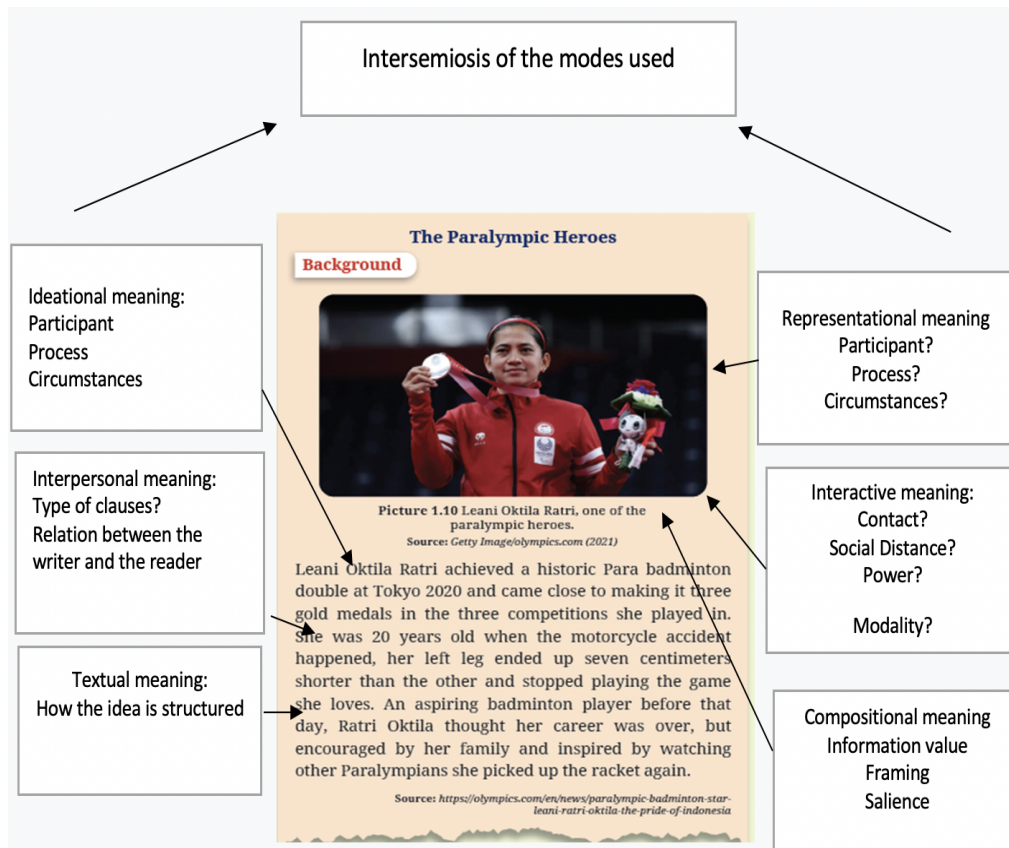


Figure 1: Intersemiosis of the modes used

Figure 1 shows the basic important points teachers can teach students to be critical readers. Teachers can start by critically discussing the text, bringing students' attention to the verbal mode in the text. The verbal text has three layers of meaning; ideational, interpersonal, and textual, or metafunction (Halliday and Matthiessen, 2014). The ideational meaning concerns what the text represents and what the text is about. This can be discovered by the teacher and students analyzing the process and the circumstances of the clauses constituting the text. A participant of a clause is that which becomes the subject or object of a clause, a process of a clause is the verb of the clause, and the circumstances of a clause are those similar to adverbs in a clause. In other words, the teachers and students discuss the transitivity system of the clauses. Transitivity system in Systematic Functional Linguistics is a system that construes the world of experience into a manageable set of process types (Halliday, 1994). The type of process identified will determine the social role of the participant either as an actor, sensor, sayers, existent, or a token. The type of process dominantly used in the text unravels how the text writer perceives and interprets the world.

The second layer of meaning of the text is concerned with the relation between the writer of the text and the teacher and students as readers. The interpersonal meaning of a text deals with clauses as exchanges (Halliday and Matthiessen, 2014). It represents "the idea that language can be used as a means of communicating information" (Halliday, 1975, as cited in Andersen, 2017). It is "the meaning which expresses a speaker's attitudes and judgements" (Gerot & Wignell, 1995, p.13). The interpersonal meaning of clauses can be explained by mood system. As a start, the teacher and students can identify the type of mood, one part of a mood system, used in the clauses; are they declarative, imperative, or interrogative? They can then look into the modality used in the clause. The modality used shows the text writer's attitude towards the world. The mood type and modality used represent the interpersonal relation between the writer of the text and the readers.

The third layer of the text, the textual meaning, includes how the information in a text is structured (Halliday and Matthiessen, 2014). Elaborating the point, Forey and Sampson (2017) said, "the textual metafunction is realized by the choices a speaker or a writer makes in combining the ideas and reality that he or she wishes to express(ideational) along with the relationship that he or she hopes to project and develop

(interpersonal)” (p. 131). The teacher and students identify what the point of departure of the information in each of the clauses is and how it influences the information communicated.

Continuing the discussion, teachers can move to the image and then point out that like the linguistic or the verbal text, the image in the reading can also possess three levels of meaning. Kress and Van Leeuwen (2006) employ the terms representational, interactional, and compositional meaning to label the meanings possessed by visual images. Representational meaning is similar to the ideational meaning of a verbal text; it is about who the represented participant (RP) in the image is, the structure that the images use, narrative or conceptual, and the circumstances through which the RP is represented. In the reading passage in Figure 1, the RP is an athlete as indicated by the sport jacket she is wearing and the medal that is showing. The image is using a narrative structure which is indicated by the vector formed by the athlete’s hand.

The second level of the image’s meaning is interactive. Images, like verbal clauses, can represent a particular social relation between the producer of the image, the viewer, and the object represented despite being imaginary (Kress & Van Leeuwen, 2006). The interactive meaning of images is contact, social distance, power, and modality. Contact is concerned with the existence or non-existence of eye contact or gaze directed frontally toward the viewers (the Interactive Participants, IP) or whether the gaze is directed in oblique which is toward other directions not towards the IPs. The gaze directed frontally to the IPs by the RP creates a demand. The RP demands the IPs be involved with what the RP is doing. The non-existence of eye contact, or a gaze that is directed in another direction from the the IPs creates an offer or an invitation in which the RP is represented as the object of viewers’ contemplation. Social distance is realized with the types of shot taken to represent RP, close, medium, or long shot. Power is shown by whether the RP is represented as looking down to the viewers, looking up to the viewers or whether it is represented at the the viewers’ eye-level.

Looking at the interactive meaning construed in the image; there is no eye contact made between the RP and the viewers. The RP is directing her gaze to another direction. She is represented using a medium close shot showing the RP’s body from head to the middle of her body and is at the eye-level with the viewers. In terms of the modality, defined as the degree to which things represented “as though they are real, as though they actually exist in this way, or as though they do not-as though they are imaginings, fantasies, caricatures, etc.” (Kress & Van Leeuwen 2006, p. 156), the picture uses full color, creating high modality.

The absence of a frontal gaze from the RP in the image, the viewers interpreted the athlete, who is the RP of the image, as an offer to the viewers. The viewers were invited to join in on the celebration caused by the athlete winning the gold medal in the badminton match. The medium close shot taken represents the RP as someone who is quite close to the viewers. This is accentuated by the high modality the picture employs, making the picture realistic, as something that viewers can see in real life. This high modality strengthens the idea that the RP is someone who is quite close to the viewers, as someone the viewers will normally meet on the streets.

The third level is composition. The RP is positioned in the middle of the frame while the background is blurred making the RP salient, which has more informational value than what might be in the background. The frame of the image is constructed by the color, separating it from other elements on the page. However, the similar color and font of the caption used as those of the verbal reading text logico-semantically restate what is depicted in the picture and link the image to the verbal reading text. In the image the RP is the most important element. In relation to the verbal text, the image is positioned in the middle top, creating an ideal-real compositional layout. The athlete is what readers of the text are inspired to be, a possibly positively dream of becoming while the verbal reading text is the real; a story that viewers may have heard before.

In sum, the basic reading activity applies the important metalanguage teachers need to comprehend and explicitly teach to the students to allow them to critically discuss and talk the paper-based multimodal text they are required to learn.

Conclusion

While it is always highly challenging to learn new things that actually benefit their classroom, teachers should make organized actions to prepare themselves to deal with multimodal text to facilitate their students’ learning and developing multimodal literacy; the skills and knowledge that is believed to be of great import in the future. Part of the efforts should be centered on understanding and mastering the metalanguage of multimodal literacy. Failing to understand and to have a good grasp of metalanguage will

result in teachers who are not able to make students learn appropriately, critically, and thoroughly from the texts they have to study. This can lead to students not willing to participate and later, to become contributing members of the society.

References

- Álvarez Valencia, J. A. (2016). Meaning making and communication in the multimodal age: Ideas for language teachers. *Colombian Applied Linguistics Journal* 18(1), 98-115. <http://dx.doi.org/10.14483/calj.v18n1.8403>
- Andersen, T. H. (2017). Interpersonal meaning and the clause. In T. Bartlett., & G. O'Grady (Eds.), *The Routledge handbook of systemic functional linguistics* (pp. 115-130). Routledge.
- Bezemer, J. (2012, 16 February). *What is multimodality?* Institute of Education. <https://mode.ioe.ac.uk/2012/02/16/what-is-multimodality>
- Bezemer, J., & Jewitt, C. (2010). Multimodal analysis: Key issues. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 180-197). Continuum.
- Bezemer, J., & Jewitt, C. (2018). Multimodality: A guide for linguists. In L. Litosseliti (Ed.), *Research methods in linguistics* (2nd ed., pp. 281-304). Continuum.
- Bezemer, J., & Kress, G. (2008). Writing in multimodal text. A social semiotic account of designs for learning. *Written Communication*, 25(2), 166-195. <https://doi.org/10.1177/074108830731317>
- Bezemer, J., & Kress, G. (2009). Visualizing English: A social semiotic history of a school subject. *Visual Communication*, 8(3), 247-262. <https://doi.org/10.1177/1470357209106467>
- Bezemer, J., & Kress, G. (2016). The textbook in a changing multimodal landscape. In N.-M. Klug & H. Stockl (Eds.), *Handbuch sprache im multimodalen kontext* (pp. 476-498). De Gruyter
- Bezerra, F. (2012). Multimodality in the EFL classroom. *Brazilian English Language Teaching Journal*, 2(2), 167-177. <https://revistaseletronicas.pucrs.br/belt/article/view/9671>
- Callow, J. (2013). *The shape of text to come: How image and text work*. Primary English Teaching Association (PETAA).
- Cope, B., & Kalantzis, M. (2000). Multiliteracies: The beginnings of a new idea. In B. Cope & M. Kalantzis (Eds.), *Multiliteracies: Literacy learning and the design of social futures*. Routledge.
- Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New literacies, new learning. *Pedagogies*, 4(3), 164-195. <https://doi.org/10.1080/15544800903076044>
- Danielsson, K., & Selander, S. (2021). *Multimodal texts in disciplinary education. A comprehensive framework*. Springer.
- Forceville, C. (2006). Non-verbal and multimodal metaphor in a cognitivist framework: Agendas for research. In G. Kristiansen, M. Achard, R. Dirven, & F. J. Ruiz de Mendoza Ibáñez. (Eds.), *Cognitive linguistics: Current applications and future perspectives* (pp. 379-402). Mouton De Gruyter.
- Forey, G., & Sampson, N. (2017). Textual metafunction and theme. What's 'it' about? In T. Bartlett., & G. O'Grady. (Eds.), *The Routledge handbook of systemic functional linguistics* (pp. 131-145). Routledge.
- Gualberto, C., & Kress, G. (2019). Social semiotics. In R. Hobbs., & P. Mihailidis (Eds.), *International encyclopedia of media literacy*. Wiley-Blackwell.
- Gerot, L., & Wignell, P. (1995). *Making sense of functional grammar*. Antipodean Educational Enterprise.
- Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the development of language*. Edward Arnold.
- Halliday, M. A. K. (1978). *Language as social semiotics*. Edward Arnold.
- Halliday, M. A. K. (1994). *Introduction to functional grammar*. Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *Introduction to functional grammar* (3rd ed.). Hodder Arnold.
- Hermawan, B. (2019). *Analisis intersemiosis teks verbal-citra visual pada buku teks sains*. [Analysis on the intersemiosis of verbal text: Visual images in science textbooks] [Unpublished dissertation], Universitas Indonesia. URI: <https://lontar.ui.ac.id/detail?id=20502547&lokasi=lokal>
- Hermawan, B., Haryanti, D., & Suryaningsih, N. (2002a). *Buku panduan guru bahasa Inggris*. [English textbook: A teachers' guide]. Kementrian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia
- Hermawan, B., Haryanti, D., & Suryaningsih, N. (2022b). *Work in progress. Buku bahasa Inggris SMA/SMK/MA Kelas X*. [Work in progress. English textbook for Senior High Schools in Indonesia]. Kementrian Pendidikan, Kebudayaan, Riset dan Teknologi Republik Indonesia.
- Hodge, R., & Kress, G. (1988). *Social semiotics*. Polity Press.
- Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, 32(1), 241-267. <https://doi.org/10.3102/0091732X07310586>
- Jewitt, C. (2009). An introduction to multimodality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 14-27). Routledge.
- Jewitt, C. (2014). An introduction to multimodality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (2nd ed., pp. 1-17). Routledge.
- Kane, G. C., & Pear, A. (2016, 4 January). *The rise of visual content: Image really is everything to the digital economy. Big Idea. Digital Leadership*. <https://sloanreview.mit.edu/article/the-rise-of-visual-content-online>
- Kress, G. (2003). *Literacy in the new media age*. Routledge.

- Kress, G. (2005). Gains and losses: New forms of texts, knowledge, and learning. *Computers and Composition*, 22(1), 5–22. <https://doi.org/10.1016/j.compcom.2004.12.004>
- Kress, G. (2009). What is mode? In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 54–67). Routledge.
- Kress, G. (2010). *Multimodality. A social semiotics approach to contemporary communication*. Routledge.
- Kress, G., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. Routledge.
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold.
- Kress, G., & Van Leeuwen, T. (2002). Colour as a semiotic mode: Notes for a grammar of colour. *Visual Communication*, 1(3), 343–368. <https://doi.org/10.1177/147035720200100306>
- Kress, G., & van Leeuwen, T. (2006). *Reading images. The grammar of visual design* (2nd ed.). Routledge.
- Lim, F.V. (2004). Problematising 'semiotic resource'. In E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspective on multimodality* (pp. 51–64). John Benjamins Publishing.
- Lim, F.V. (2018). Developing a systemic functional approach to teach multimodal literacy. *Functional Linguistics*, 5(13). <https://doi.org/10.1186/s40554-018-0066-8>
- Liu, Y., & O'Halloran, K. L. (2009). Inter-semiotic texture: Analyzing cohesive devices between language and images. *Social Semiotics*, 19(4), 367–388. <https://doi.org/10.1080/10350330903361059>
- Madya, S. (2007). Curriculum innovations in Indonesia and the strategies to implement them. In Y. H. Choi & B. Spolsky (Eds.), *ELT curriculum innovation and implementation in Asia* (pp. 1–38). Asia TEFL.
- Martinec, R., & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication*, 4(3), 339–374. <https://doi.org/10.1177/1470357205055928>
- Mills, K. A. & Unsworth, L. (2017). Multimodal literacy. In *Oxford research encyclopedia of education*. Oxford University Press.
- O'Halloran, K. L. (2010). Historical changes in the semiotic landscape: From calculation to computation. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis*, (pp. 98–113). Routledge.
- O'Mara, J. (2006). Becoming I(it)erate: Pre-service English teachers and ICTs. *English in Australia*, 41(3), 44–47. <http://hdl.handle.net/10536/DRO/DU:30004006>
- O'Toole, M. (1994). *The language of displayed art*. Leicester University Press.
- Painter, C., & Unsworth, L. (2013). *Reading visual narratives: Image analysis of children's picture books*. Equinox.
- Roehrich, L. (2016). Intersemiosis in science textbooks. In A. Archer & E. Breuer (Eds.), *Multimodality in higher education* (pp. 195–215). Brill.
- Roswell, J., & Walsh, M. (2011). Rethinking literacy education in new times: Multimodality, multiliteracies, & new literacies. *Brock Education Journal*, 21(1), 53–62. <https://doi.org/10.26522/brocked.v21i1.236>
- Ryan, J., Scott, A., & Walsh, M. (2010). Pedagogy in the multimodal classroom: An analysis of the challenges and opportunities for teachers. *Teachers and Teaching*, 16(4), 477–489. <https://doi.org/10.1080/13540601003754871>
- Sang, Y. (2017). Expanded territories of "literacy": New literacies and multiliteracies. *Journal of Education and Practice*, 8(8), 16–19. <https://www.iiste.org/Journals/index.php/JEP/article/view/36029/37023>
- Stein, P. (2008). *Multimodal pedagogies in diverse classroom. Representation, rights and resources*. Routledge.
- Stein, P., & Newfield, D. (2006). Multiliteracies and multimodality in English in education in Africa: Mapping the terrain. *English Studies in Africa*, 49(1), 1–21. <https://doi.org/10.1080/00138390608691341>
- The New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), 60–92. <https://doi.org/10.17763/haer.66.1.17370n67v22j160u>
- The New London Group. (2000). A pedagogy of multiliteracies designing social futures. In B. Cope & M. Kalantzis (Eds.), *Multiliteracies: Literacy learning and the design of social future* (pp. 9–38). Routledge.
- Unsworth, L. (2006a). Towards a metalanguage for multiliteracies education: Describing the meaning-making resources of language-image interaction. *English Teaching: Practice and Critique*, 5(1), 55–76. <https://files.eric.ed.gov/fulltext/EJ843820.pdf>
- Unsworth, L. (2006b). Image/text relations and intersemiosis: Towards multimodal text description for multiliteracies education. *Proceedings of the 33rd International Systemic Functional Congress, São Paulo, Brazil*, 1165–1205. https://www.pucsp.br/isfc/proceedings/Artigos%20pdf/59pl_unsworth_1
- van Leeuwen, T. (2005). *Introducing social semiotics*. Routledge.
- Van Leeuwen, T. (2017). Multimodal literacy. *Viden Om Literacy*, 21, 4–10. https://www.videnomlaesning.dk/media/2127/21_theo-van-leeuwen.pdf
- Yi, Y. (2014). Possibilities and challenges of multimodal literacy practices in teaching and learning English as an additional language. *Language and Linguistics Compass*, 8(4), 158–169. <https://doi.org/10.1111/Inc3.12076>
- Yi, Y., & Angay-Crowder, T. (2016). Multimodal pedagogies for teacher education in TESOL. *TESOL Quarterly*, 50(4), 988–998. <https://doi.org/10.1002/tesq.326>
- Zein, S., Sukyadi, D., Hamied, F. A., & Lengkanawati, N. S. (2020). English language education in Indonesia: A review of research (2011–2019). *Language Teaching*, 53(4), 491–523. <https://doi.org/10.1017/S0261444820000208>
- Zhou, Y., & Liao, Y. (2026). More than meets the eye: An image-text meaning construction of contemporary senior high school English textbooks in China. *Studies of Applied Linguistics in Asia*, 1(1), 65–77. <https://doi.org/10.53941/sala.2026.100006>