

‘IS THE PICTURE WORTH A THOUSAND WORDS?’: THE INTERPERSONAL MEANINGS OF A DIALOGUE IN AN EFL TEXTBOOK

¹Ahmad Sugianto, ¹Ilham Agung Prasetyo, ¹Widy Asti

¹English Education Study Program, Faculty of Language and Literature Education,
Universitas Pendidikan Indonesia, Indonesia
Corresponding Author Email: ahmadsugianto@upi.edu

Article Info	Abstract
Article History Received: December 2021 Revised: February 2022 Published: April 2022	<i>The present study was aimed at scrutinising a multimodal text embedded in a dialogue of an EFL textbook. To this end, a descriptive qualitative study using Systemic Functional Multimodal Discourse Analysis (SF-MDA) focused on interpersonal meanings consulting grammar of visual design and intersemiotic complementarity drawing on systemic functional linguistics were employed to analyse the artifact, a dialogue within a part named 'communication' taken from an EFL textbook for a primary school level. The findings revealed that declarative clause and modalisation (epistemic modality) of probability was found to be the most common system used in the verbal text. Meanwhile, high modality and validity were found in the visual image indicated by the realizations and representations of detailed abstraction and full-color saturation. Finally, the study draws a conclusion that there is a cohesive interaction to a certain extent between the verbal text and the visual image represented in the multimodal dialogue.</i>
Keywords Dialogue; EFL Textbook Evaluation; Interpersonal Meaning; Multimodal Analysis; Systemic Functional Linguistics	

How to cite: Sugianto, A., Prasetyo, I. A., & Asti, W. (2022). ‘Is the Picture Worth a Thousand Words?’: The Interpersonal Meanings of a Dialogue in an EFL Textbook, *JOLLT Journal of Languages and Language Teaching*, 10(2), pp.187-198. DOI: <https://doi.org/10.33394/jollt.v%vi%i.4510>

INTRODUCTION

Pictures have essential roles and become one of the essential parts of human life. It is on account of the fact that particular messages can be communicated through a picture (Sugianto, 2021; Sugianto, Prasetyo, Aria & Wahjuwibowo, 2022; Sugianto, Andriyani, & Prasetyo, 2021; Sugianto, Denarti, & Prasetyo, 2021; Sugianto & Prastika, 2021). Additionally, Halliday (1990) points out that historically the development of today's writing system (either logogram or alphabet) derives from pictorial representations. Moreover, in recent times, pictures constitute one of the profound and growing areas scrutinized in any discipline (Kress, 2010). Besides, Unsworth (2006) reveals, particularly in the educational context, that due to the development of media, either printed or electronic, the use of various modes, i.e., images and language, is the key issue to literacy education influence the school curricula. Additionally, the use of pictures in students' English learning is proven to be helpful to assist the students in generating ideas in writing (Deviga & Diliyana, 2020). Thereby, based on these notions, going through the use of multimodality in the education domain can be worth scrutinizing.

The education context possibly constitutes one of the fruitful and significant places in which multimodality is utilized. For instance, the use of textbooks through which students learn various cultures (Sugianto & Wirza, 2021) is regarded as one of the main sites and artifacts to be scrutinised (Sugianto, Andriyani, & Prasetyo, 2021). Besides, historically, the use of multimodality embedded in an instructional textbook is noted by Gaudin (2019) to exist initially by the mid 18th century. Additionally, at that time, one of the most inspiring ones, entitled *Spectacle* that was in the form of an encyclopedia containing thematic dialogues

(p. 21). A study conducted by (Jauhara et al., 2021) shows that a dialogue that utilises pictures were found to help build the way the viewers/readers feel and perceive. Meanwhile, on the one hand, it is deemed as a key role in English language teaching and learning; in this case, it constitutes one of the means through which students learn the English lessons (Richards, 2002); on the other hand, Cunningsworth (1995) asserts that it is, on the other hand, its use should not be taken for granted; nevertheless, the best and appropriate materials should be selectively chosen.

Concerning the textbook selection, a thorough evaluation is required so as to make judicious use of the textbook can be generated. In this regard, to evaluate the content of the textbook had to do with multimodality, there are two fundamental frameworks that are commonly used. The first one has to do with analyzing the verbal text, drawing mainly on systemic functional linguistics (SFL). SFL constitutes a framework that sees 'language as a strategic, meaning-making resource' (Eggins, 2004, p. 2). There are two main aspects of SFL associated with using language to negotiate things in social contexts, namely *systemic* and *functional* (Gunawan, 2020). The former refers to the notion with respect to a set of systems or options (e.g. tense, conjunctions, persons, and so on) through which meaning is constructed (Emilia, 2014). The latter refers to the way a language is used (Halliday, 1994). There are three types of linguistic functions in SFL, commonly known as metafunctions encompassing ideational metafunction associated with 'constructing experience', interpersonal metafunction, which refers to the way using a language to build interpersonal rapport in social interactions, and textual metafunction having to do with the way the two previous metafunctions are organised in discourse (Halliday & Matthiessen, 2006).

The present study was focused only on the interpersonal meaning on account of some grounds. To begin with, by delimiting only one metafunction, a richer construal of the meaning leading to deeper implications of the way a message is communicated can be obtained (Hermawan & Sukeyadi, 2020). Also, another rationale is based on the credence that dialogue is closely related to the interpersonal metafunction in which dialogic or conversational discourse expressing the negotiation and exchange meaning emerge. In addition, Martin and Rose (2007) point out that negotiation in conversation can encompass a range of speech roles, such as making statements, asking questions, offering services, and demanding goods. Such speech roles are known in the interpersonal system as Mood (Halliday, 1989). The Mood system in a clause is constructed by two major aspects, namely MOOD and RESIDUE (the use of capitalised MOOD along with RESIDUE is to differentiate the MOOD as the element or constituent of a clause and the Mood as a system (Eggins, 2004)). The former consists of the elements or constituents such as Subject and Finite (divided into two types, '*temporal finite verbal operator*' which typically provides tense and '*finite modal operator*' which typically refers to a modality or the modal operator), whereas the latter may consist of the constituent(s) such as Predicator (a part of verbal element), Complement (represented by a nominal group or adjectival element), and a particular Adjunct (commonly in the form a particular adverbial or prepositional elements) (Eggins, 2004; Halliday & Matthiessen, 2014).

Another system of interpersonal metafunction of the verbal text is modality. It is, as Halliday (1994, p. 88) asserts, considered as the 'intermediate degrees, between the positive and negative pole'. Moreover, to construe modality, there are two terms that should be taken into account, namely modalisation, which refers to the propositions (information) realising the degree of probability and the degree of usuality, and modulation, which refers to proposals (goods and services) realising the degree of obligation and inclination (pp. 88-89). Moreover, the degrees of probability and usuality are divided into three categories realized by modal operators and modal adjuncts, namely *high* (*must, certainly and always*), *median* (*may, probably, and usually*), and *low* (*might, possibly, and sometimes*); additionally, the modal

operators and modal adjuncts of the continuum above have different values if they are realised in the negative polarity, for instance, *high* (could not possibly or certainly ... not; never), *median* (probably ...not or not usually), and *low* (possibly might not or not always) (Eggs, 2004). Meanwhile, the values degrees of obligation and inclination in the positive polarity encompass *high* (realised by *must* or *be required*), *median* (realised by *should* or *be supposed*), and *low* (*can* or *be allowed*); meanwhile, regarding the negative polarity, the values encompass *high* (realised by *cannot*, *be required not*, or *be not allowed*), *median* (realised by *should not*, *be supposed not*, or *be not supposed*), *low* (realised by *need not*, *be allowed not*, or *be not required*) (Halliday, 1994).

Furthermore, to construe the visual meaning of the interpersonal metafunction of the multimodal dialogue, the interactive meaning advocated by Kress and van Leeuwen (2006) becomes the alternative framework commonly used. In this regard, there are three aspects that are required to be taken into accounts, such as contact, social distance, and attitude. These are represented and summarised in Figure 1 below. To begin with, the contact with respect to interactive meanings in images represented in Figure 1 is realised by the presence or absence between the represented participants and interactive participants. In this regard, there are two types of contacts, *demand* and *offer*. On the one hand, the former refers to the portrayal of the represented participant(s) that look(s) at directly to the viewers or readers; hence, the represented participant(s) demand(s) something to the readers/viewers. On the other hand, the latter refers to the portrayal of the represented participant(s) that do(es) not look at directly to the viewers or readers; hence, the represented participant(s) acts as an 'item of information or object of contemplation for the viewers/readers (Kress & van Leeuwen, 2006, pp. 115-116).

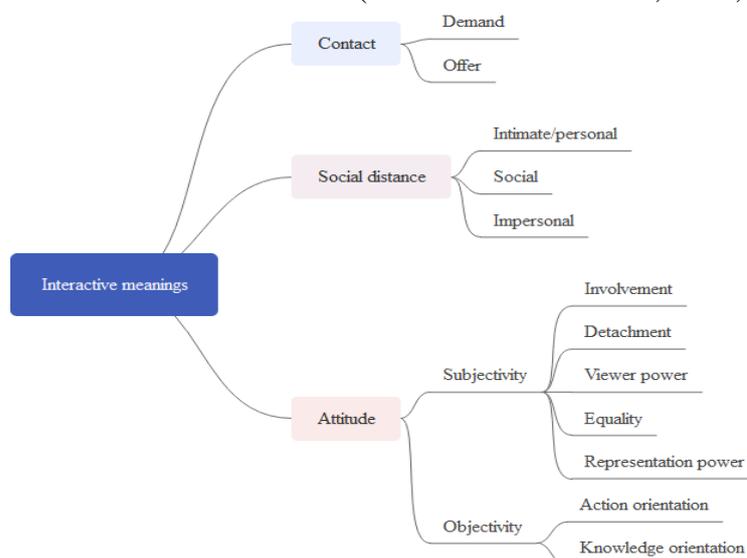


Figure 1. Interactive meanings in images (Kress & van Leeuwen, 2006)

Moreover, the social distance shown in Figure 1 above is realised by the size of the frame and image location. Torres (2015, p. 246) provides the summaries of the explication concerning this social distance deriving from Hall's (1966) theory of social distance and Kress and van Leeuwen's (2006) framing. It is illustrated in Table 6.

Table 1

Criteria of framing and social distance (Torres, 2015, p. 246)

Classification of social distance (Hall, 1966, pp. 110-120)	Vision Range (Kress & van Leeuwen, 2006, pp. 125)	Shot Size (Kress & van Leeuwen, 2006, 124)
Intimate	Face or head only	Very close
Close personal	Head and shoulders	Close
Far personal	Waist up	Medium close

Close social	Whole figure	Medium long
Far social	Whole figure with space around it	Long
Public	Torso, min 4 to 5 people	Very long

In addition to the subjectivity of the interactive meanings in images shown in Figure 1 above, the involvement and power between the represented participants (i.e., the people, things, or places illustrated in the image) and interactive participants (i.e., the producers or viewers of the image) can be depicted through some realisations, horizontally and vertically. The horizontal and vertical realisations are summarised in Table 2 below.

Table 2

Realisations of involvement and power (Kress & van Leeuwen, 2006, pp. 133-148)	
Type of involvement and power	Angle
Involvement, i.e. the represented participants and the interactive participants, get involved and connected with one another.	Frontal
Detachment, i.e. the represented participants and the interactive participants, do not get involved and connected with one another.	Oblique
Viewer power, i.e. the interactive participants, have power over the represented participants.	High
Equality or no difference between the represented participants' power and interactive participants'.	Eye level
Represented participant power, i.e., the represented participants have power over the interactive participants.	Low

Moreover, to construe the modality features encompassing contextualisation, representation and abstraction degree, and texture of the visual mode, a continuum is used. In this case, Royce (1998) proposed a continuum named as 'naturalistic visual continuum' to provide the description of the range of the modality of visual mode, depicted in Figure 2 below.

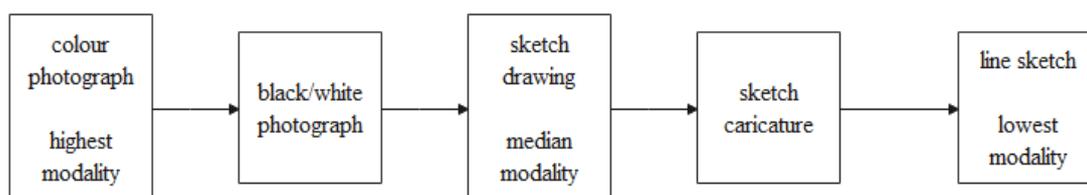


Figure 2. Modality criteria based on the naturalistic visual continuum (Royce, 1998, p. 40)

Meanwhile, in regard to construing the interpersonal intersemiosis between the two modes, the verbal text and visual image of the multimodal dialogue, the analysis is conducted by going through the interaction between the two modes. In this case, Royce (1998) points out that the intersemiotic interactions are constructed by the elements of Mood realised by speech functions such as offer, command, statement, and question and the Modality having to do with degrees of reality, possibility, truthfulness, or necessity.

Apart from its concepts, there are a number of studies concerning the interpersonal meaning associated with other variables. For instance, the scrutinies focusing on verbal texts, such as interpersonal along with ideational metaphors with respect to thesis texts (Ngongo & Benu, 2020), interpersonal meaning with respect to non-native students' thesis abstract (Arifin, 2018), interpersonal meaning in students' personal letters (Nasita, Sugianto, & Thoyyibah, 2020), modality system associated with hortatory exposition texts (Suciati,

Rustandi, & Sugiarto, 2021), interpersonal meaning with respect to workplace material of textbooks (Cheng, Lam, & Kong, 2019), interpersonal meaning in pre-intermediate textbooks for teenagers (Boccia, 2021), and interpersonal meaning in association with EFL textbook for primary and secondary levels (Chen, 2009). Meanwhile, the scrutinies have to do with multimodal texts, for example, interpersonal meaning with respect to English textbooks for junior high school (Dewi, Rukmini, & Saleh, 2020; Jauhara, Emilia, & Lukmana, 2021). Based on the previous studies above, the present study has some distinctions from them in terms of the artefact used and the materials delimited. Thus, the present study attempted to fill the gap left by the previous studies above; in particular, it has to do with the investigation of interpersonal meaning in a multimodal dialogue of an EFL textbook for a primary school level.

RESEARCH METHOD

Research Design

Descriptive qualitative research utilising the SF-MDA approach was employed. This was deemed as a suitable approach to be utilised for, as O'Halloran and Fei (2014) assert, it can be used to analyse several phenomena in regard to multimodal texts ranging from two dimensions, either printed or digital text, to three dimensions such as museum texts or gestures, and also that which involves time-based texts encompassing music, film, television and so on (Knox, 2013). Besides, the SF-MDA covers objectives constitute another ground for using it, which in this regard, namely to understand the meanings systems encompassing image and verbal texts and comprehend social functions owned by a text (Jewitt, Bezemer, & O'Halloran, 2016). In the present study, the meanings between the two modes were focussed on interpersonal meanings. Additionally, several frameworks were used to construe the interpersonal meanings of the visual and verbal text and the interrelationships between the two modes. In this case, they comprised the interpersonal metafunction of systemic functional linguistic (Halliday, 1994; Halliday & Matthiessen, 2004, 2014), Kress and van Leeuwen (2006, 2021) interactive meanings of the grammar of visual design, and Royce's (1998, 2007) interpersonal intersemiotic complementarity.

Analysis Unit

A multimodal text in the form of dialogue accompanied by an image taken from an EFL textbook entitled *Super Minds: Student's Book 6* (Puchta, Gergross, & Lewis-Jones, 2017) constitutes the artefact that was scrutinized. The textbook was selected on account of the fact that it was used by some primary schools in some cities in Indonesia; besides, it was published by a reputable publisher with credible and reliable authors, even one of them had teaching experience in Indonesia (cf. *Authors*, 2021b); also, it encompassed the artefacts in the forms of dialogues covering not only verbal texts but also images constitutes another ground for the textbook selection to be scrutinised. Additionally, so as to cope with the copyright issue, the artifact was recolored.

Research Procedure and Data Analysis

To analyse the artefact, as Hermawan and Sukyadi (2020) point out, each mode, the visual mode and verbal mode, was analysed separately before the intersemiotic analysis between the two modes were conducted. In detail, firstly, concerning the analysis of interpersonal meanings of images, the analysis was started by identifying the mood, which was conducted by going through the presence/absence of the address/gaze, followed by examining the involvement realised by the horizontal perspective and power realised by vertical perspectives, which then was investigating the social distance realised by the way the shot was taken (size of frame); next after the mood investigation associated with visual image above was conducted, another inspection had to with modality comprising identifying the

contextualisation realised by presence/absence of background, followed with identifying the representation/abstraction degree, texture, illumination, and colour saturation. Secondly, regarding the analysis of interpersonal meanings of the verbal text, the analysis was conducted by identifying and determining the forms of clauses (declarative or imperative; demand or offer); another analysis in regard to interpersonal meanings of the verbal text is going through the modality (deontic modality encompassing obligation and inclination or epistemic modality encompassing probability and usuality) followed by the investigation of the use of pronouns to greet the readers; Thirdly, the interpersonal intersemiotic analysis was started by comparing the analysed items of the visual meanings and verbal text meanings, which then was followed with the interpretation of the comparison divided into two main fashions, i.e., in terms of the mood (encompassing identification of offers, commands, statements, and questions) and modality (encompassing reinforcement of address, attitudinal congruence, and attitudinal dissonance) (Hermawan & Sukyadi, 2020, pp. 57-61; Royce, 1998, p. 36, 2007, p. 69).

RESEARCH FINDINGS AND DISCUSSION

The present study was aimed at investigating the interpersonal meaning in a multimodal dialogue of an EFL textbook for a primary school level. The multimodal text in the form of a dialogue is represented in Figure 3. The dialogue was accompanied by a photo depicting two students named Charlie and Olivia. They were discussing a particular topic, i.e. 'joining sport club'. To construe this multimodal text, the inspection was conducted by going through the visual meanings followed by verbal meanings and intersemiotic complementarity, respectively.



1 Look at the photo and answer the questions.
Where are Olivia and Charlie and what are they talking about?
Who do you think is more interested in sport? Why?

2 Read and listen to the dialogue to check your ideas.

Charlie Hi!
Olivia Hi, Charlie! I think we should join a sports club this year.
Charlie Why would I want to do that?
Olivia Well, to get some exercise and to get fit.
Charlie Hm, I'm not so sure.
Olivia Come on. Don't be silly, it'll be fun.
Charlie Well, which sports club do you want to join?
Olivia We could try the athletics club and learn how to do the hurdles and high jump and things like that.
Charlie That's not such a good idea. I hate running and I can't jump.

Olivia OK, what about the rowing club then? Do you fancy that?
Charlie Sorry, but I can't swim at all and I'm afraid of water.
Olivia Tennis? Are you interested in tennis?
Charlie I don't think so. I tried it once and I was rubbish.
Olivia Really? Hang on, I know!
Charlie What?
Olivia We could go to the gymnastics club. I love gymnastics.
Charlie Me? Doing gymnastics? Are you serious?
Olivia Oh, I give up. I'll do it by myself.
Charlie Now that is a good idea.
Olivia Let's go!

Figure 3. Excerpt of the multimodal dialogue in the 'communication' material (Puchta et al., 2017)

Visual Meanings of the Multimodal Dialogue

Concerning the visual meanings associated with the multimodal modal text, some aspects, as represented in Table 8 previously, are scrutinised comprising address, involvement and power, social distance, and modality. To begin with, in terms of the address, the image was considered to be 'offer', as Kress & van Leeuwen (2006, p. 119) explicates, i.e., the represented participants, namely the male student named Charlie and female student named Olivia, were deemed as a piece of information through which the producer offered. In this regard, it was also indicated by the absence of the gaze given by the represented participants to the readers/viewers or the producers. They, the represented participants, did not look at directly the viewers/readers. Thereby, there was supposed no engagement between the

represented participants and the viewers (p. 120). Also, the viewers/readers were not required to react for the represented participants; instead, they were supposed to get only information offered, or in this case, Royce (2007, p. 89) argues that the viewers/readers can only 'agree/disagree, acknowledge, or contradict' to the information offered. Based on this finding in regard to the address aspect, the illustration given can be associated with a certain genre, which in this regard, as informed by Kress and van Leeuwen (2006), the film, television drama, and scientific illustration are the most frequently genre associated with the 'offer' image. Thereby, it can be indicated that the illustrative image fits the common genre of an 'offer' image for indeed the excerpt related to material about communication shown in a dialogue which may be similar to one of the genres of 'offer', i.e. the (television) drama.

Furthermore, the involvement and power levels and relations of the visual mode were indicated by the horizontal and vertical angle the image represents. In this regard, the represented participants were shown in oblique angle, meaning that there was no engagement between the represented participants and the viewers. As there was no engagement found, this was in line with the address aspect previously in which the represented participants were only offering information to the viewers/readers. In this regard, they seemed to offer the information of a certain material concerning the dialogue about a specific topic, namely joining a sports club. Moreover, the represented participants were depicted at eye level, meaning that there was an equal power between the represented participants and the viewers/readers. Moreover, the social distance that the visual mode provides also indicates the engagement between the represented participants and viewers/readers. In this case, the picture was taken in a medium close shot, i.e. the represented participants were depicted in the waist-up vision range. This shot size and the vision range indicate that the represented participants were taken in the far personal social distance, meaning that the relationship between the represented participants and the viewers can be considered to have personal relationship but not close or even intimate (Kress & van Leeuwen, 2006, 2021).

In terms of modality, the visual mode was construed through several elements. To begin with, concerning the contextualisation, the represented participants were accompanied by a particular background and details, i.e. a wall magazine in which a number of pamphlets or flyers related to sports clubs were stuck on it. Such a background provides the viewers/readers with some cues to grasp the setting in which the dialogue between the represented participants was taking place and gives clues about the topic of the dialogue being discussed. Moreover, the visual image depicts two represented participants in quite a detail and adequate illumination and fully saturated colour. For example, in this regard, based on their physical appearance, the reader/viewer might guess that they are approximately primary school students (a male and a female student), around nine to twelve years old. These can be indicated by the way they are illustrated. The male student with short black hair and the female student with long blonde hair is depicted to wear a dark blue pullover with a light blue collar. Both of them wear a schoolbag, which in this case, the male student is wearing a black-yellow-and-light-green school bag with some balls patterns, and the female student is depicted to wear a school bag with triangles pattern coloured in blue, purple, black, and white. The detailed appearance and illustration, added with the aforementioned background above, indicates that the visual image has high modality and validity (Kress & van Leeuwen, 2006, 2021). Thus, it can be considered that the visual image can be regarded to meet the naturalistic criteria. To these findings, the naturalistic image criteria found in the visual image is considered to be useful for students to benefit from them. It can assist them in construing the meaning of the context in which the represented participants emerge (Jauhara, Emilia, & Lukmana, 2021).

Verbal Meanings of the Multimodal Dialogue

Regarding the interpersonal meaning of the verbal text in the dialogue, two representations were employed, namely mood and modality. To begin with, concerning the mood inspection, Table 4 below shows the summary of the mood inspection from the dialogue.

Table 4
Summary of the Mood Analysis of the Dialogue

Speaker	(Independent) Clause								Total
	Indicative *)					Imperative			
	INT		DEC		Exc.	Jussive		Suggestive	
	Wh-	Yes/No	Non-Exc.			Unmarked	Marked		
		Pos	Neg						
Charlie	2 (14.3%)	2 (14.3%)	5 (35.7%)	5 (35.7%)	-	-	-	-	4
Olivia	-	2 (16.6%)	7 (58.3%)	-	1 (8.3%)	1 (8.3%)	-	1 (8.3%)	2

*) INT: Interrogative; DEC: Declarative; Exc.: Exclamative

In terms of modality, it was found that the verbal text employs two main types of modalities, namely the epistemic modality and deontic modality. Nevertheless, the epistemic modality was found to outnumber the deontic modality in the text. In this regard, there are two deontic modalities with the modal operators *should* and *'ll* (the contraction of *will*) encountered. The modal operator *should* was found in the clause *I think we should join a sports club this year*. The modal operator in the clause refers to obligation, which in this case, conveyed by Olivia to invite Charlie to join the sports club. Additionally, the modal operator *should* is considered to be in median continuum (Halliday & Matthiessen, 2014). Moreover, the median value of the modal operator is emphasised and realised explicitly by the Mood Adjunct *I think*, indicating Olivia's judgement (Egins, 2004), which in this case it has to do with her idea concerning the obligation of joining the sports club. Meanwhile, another deontic modality with the modal operator *'ll* was encountered in the clause *I'll do it by myself*. Similarly, the modal operator *'ll* (*will*) is included in the median continuum, but it is regarded as an inclination (p. 697). These uses of deontic modality within the medium continuum are considered to have the subjective implicit orientation on account of the fact that the modality represented by the modal operators are embedded or realised in the main propositions (Thompson, 2014). Moreover, the use of the modal operators included within the median continuum turn in the information given in the clause debatable (Hermawan & Sukyadi, 2020). In regard to the context of the dialogue, it can be interpreted that the interactant who conveys the information, which in this case is Olivia, is not trying to forcefully or highly 'persuade' or recommend Charlie to join the sports club; hence, these result in an open discussion feasible between Charlie and Olivia.

Furthermore, the epistemic modality found in the text encompasses the uses of modal operators *would*, *'ll* (the contraction of *will*), *could*, and *can't*. All these modal operators indicate probability. In this regard, two out of six modal operators are found within the high-value continuum, namely, the modal operator *can't* (Halliday & Matthiessen, 2014). The modal operator *can't* was found in two clauses expressed by Charlie, namely *'I hate running and I can't jump'* and *'I can't swim at all and I'm afraid of water'*. These negative modal operators in the clause indicate, with respect to the dialogue, that Charlie is quite sure that he is not able to jump and swim; despite the high degree of probability, Halliday (1994) asserts that even as the interactant uses the modal operator with a high degree of certainty, there is a doubt value in it. Additionally, the remaining four modal operators having to do with the epistemic modality comprise *would*, *'ll*, and *could*. These modal operators indicate probability within the median continuum. Thereby, the information in the clause unit using these modal

operators constitute information debatable and result in some open discussions between the interactants. The analysis of the modality of the verbal text is summarised in Table 5 below.

Table 5

Types, Orientation, Value and Polarity of the Modality Used

Speaker	Statement	Type	Orientation	Value	Polarity
Olivia	I think we should join a sports club this year.	Modulation (deontic): obligation	Subjective: explicit	Median	Positive
Charlie	Why would I want to do that?	Modalisation (epistemic): probability	Subjective: implicit	Median	Positive
Olivia	It' ll be fun.	Modalisation (epistemic): probability	Subjective: implicit	Median	Positive
Olivia	We could try the athletics club	Modalisation (epistemic): probability	Subjective: implicit	Median	Positive
Charlie	I hate running and I can't jump.	Modalisation (epistemic): probability	Subjective: implicit	High	Negative Transferred
Charlie	I can't swim at all and I'm afraid of water.	Modalisation (epistemic): probability	Subjective: implicit	High	Negative Transferred
Charlie	We could go to the gymnastics club.	Modalisation (epistemic): probability	Subjective: implicit	Median	Positive
Olivia	I' ll do it by myself.	Modulation (deontic): inclination	Subjective: implicit	Median	Positive

Interpersonal Intersemiotic Complementarity of the Multimodal Dialogue

Based on the visual meaning and verbal meaning aforementioned, some interactions within a certain degree between the modes were found. To begin with, in terms of the mood aspect, the visual image is considered to interact cohesively with the verbal text. In this case, the visual image realised by the represented participants are illustrated by the producers or authors in oblique fashions meaning that they are regarded as a unit of contemplation; such a finding is in line with the verbal text in the dialogue that is not addressed the viewers/readers directly, i.e. commonly indicated by the use of the second personal pronoun *you*. Albeit some *you-s* were found in the dialogue, they refer to the interactants in the dialogue, i.e. either Charlie or Olivia. In addition, the use of the visual image as an object to be contemplated is emphasised by the instruction outside the dialogue box mentioning *Read and listen to the dialogue to check your ideas*. The function of the visual image as an object of information, hence readers/viewers are necessarily required to evaluate, interacts cohesively with the presence of modality markers such as *would*, *'ll (will)*, *could*, *could*, and *can't* indicating probability, the modality marker *should* showing obligation, and the modality marker *'ll (will)* indicating inclination. Thereby the uses of different modality markers are required to be taken into account for each has different values and functions.

The other thing that is used to indicate whether the visual image acts as an object or item of information is the evaluative words appear in the dialogue (Jauhara et al., 2021), for instance, *afraid* (affect: -security), *rubbish* (judgement: -capacity), and *good idea* (appreciation: +reaction). Additionally, these evaluative words interact cohesively with the facial expressions of the represented participants of the visual image. In this case, they are indicated, for example, by the *emoter* [borrowing Martin and White's (2005) term] Charlie

experiencing the emotions aforementioned is not into sports clubs as he has flat facial expressions. Charlie's facial expression is different from that which is shown by Olivia, which in this regard, she is illustrated with a mouth open indicating smiling and enthusiasm (Chen, 2009), and in the verbal text, she is the one who initiates and asks Charlie to join one of the sports clubs at school. In addition, the use of declarative mood and some modal operators ranging from median to high is in line with the modality and validity shown by the visual image with the full-colour saturations indicating the naturalistic portrayal (Hermawan & Sukyadi, 2020; Kress & van Leeuwen, 2006, 2021).

CONCLUSION

Based on the findings aforementioned, the present study reaches some conclusions. Concerning the verbal meaning, the declarative mood included in the indicative clause was found to be the most frequently used in the verbal text. Additionally, the epistemic modality (modalisation) referring to probability ranging from median to high value was found to outnumber deontic modality (modulation). Next, in terms of the visual meaning, the visual image is illustrated with high modality and validity realised by the detailed abstraction and naturalistic portrayal indicated by the full-colour saturation. Thereby, to some extent, there is a cohesive interaction between the verbal mode and visual mode represented in the textbook. This cohesive interaction is necessarily required and essential for it assists the readers, which in this case, students to understand the meaning of the dialogue. Furthermore, the present study only focused on and confined the discussion about the interpersonal meaning of the multimodal text, some explorations of other metafunctions or meanings are considered to be worth scrutinising. Moreover, it also suggests that teachers or instructors are required to do a textbook selection and evaluation carefully, and the textbook providing the meaningful visual image along with the verbal text are necessarily and inevitably required to be taken into account. Also, scaffolding or systematic assistance from students' significant others in using multimodal texts are required so as to achieve adequate literacy capability.

REFERENCES

- Arifin, A. (2018). How non-native writers realize their interpersonal meaning? *Lingua Cultura*, 12(2), 155–161. <https://doi.org/10.21512/lc.v12i2.3729>
- Authors. (2021). <https://www.cambridge.org/gb/cambridgeenglish/authors/peter-lewis-jones>
- Boccia, C. (2021). Teaching and learning interpersonal meanings in EFL in the school years. *System*, 101, 102571. <https://doi.org/10.1016/j.system.2021.102571>
- Chen, Y. (2009). *Interpersonal meaning in textbooks for teaching English as a foreign language in China : A multimodal approach*. University of Sydney, Australia.
- Cheng, W., Lam, P. W. Y., & Kong, K. C. C. (2019). Learning English through workplace communication: Linguistic devices for interpersonal meaning in textbooks in Hong Kong. *English for Specific Purposes*, 55, 28–39. <https://doi.org/10.1016/j.esp.2019.03.004>
- Cunningsworth, A. (1995). *Choosing your coursebook*. Macmillan Education.
- Deviga, L., & Diliyana, Y. F. (2020). Using picture series in teaching writing skill for the first semester students of medical record program in Stikes Bhakti Husada Mulia Madiun. *English Teaching Journal: A Journal of English Literature, Linguistics, and Education*, 8(2), 75–81. <https://doi.org/10.25273/etj.v8i2.7827>
- Dewi, A. K., Rukmini, D., & Saleh, M. (2020). The interpersonal meaning of verbal text and visual image relation in English textbook for junior high school grade viii. *English Education Journal*, 10(1), 110–114.
- Eggs, S. (2004). *An introduction to systemic functional linguistics* (2nd ed.). Continuum.
- Emilia, E. (2014). *Introducing functional grammar*. Pustaka Jaya.

- Gaudin, J. (2019). A history of the multimodal classroom from antiquity to the nineteenth century. In H. de S. Joyce & S. Feez (Eds.), *Multimodality across classrooms: Learning about and through different modalities* (pp. 7–29). Routledge.
- Gunawan, W. (2020). *Systemic functional linguistics (SFL) and teacher education in EFL context*. UPI Press.
- Hall, E. T. (1966). *The hidden dimension*. Doubleday. [https://doi.org/10.1016/0003-6870\(70\)90049-9](https://doi.org/10.1016/0003-6870(70)90049-9)
- Halliday, M. A. K. (1989). Register variation. In M. A. K. Halliday & R. Hasan (Eds.), *Language, context, and text: Aspects of language in social semiotic perspective* (2nd ed., pp. 29–43). Oxford University Press.
- Halliday, M. A. K. (1990). *Spoken and written language*. Oxford University Press.
- Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). Edward Arnold Ltd.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd ed.). Hodder Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2006). *Construing experience through meaning: A language-based approach to cognition*. Continuum.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.
- Hermawan, B., & Sukyadi, D. (2020). *Analisis multimodal pada buku teks Sains*. UPI Press.
- Jauhara, D., Emilia, E., & Lukmana, I. (2021). Re-contextualising 'Greeting': A multimodal analysis in an EFL textbook. *Indonesian Journal of Functional Linguistics*, 1(1), 15–24. <https://doi.org/https://doi.org/10.17509/ijfsl.v1i1.32621>
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge. <https://doi.org/10.4324/9781315638027>
- Knox, J. S. (2013). Multimodality and systemic functional analysis. In *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd. <https://doi.org/10.1002/9781405198431.wbeal0836>
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.
- Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.). Routledge.
- Martin, J. R., & Rose, D. (2007). *Working with discourse: Meaning beyond the clause*. Continuum.
- Martin, James R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.
- Nasita, D., Sugianto, B. R., & Thooyibah, L. (2020). The realization of interpersonal meaning on male and female students' personal letter. *JALL (Journal of Applied Linguistics and Literacy)*, 4(1), 57–76.
- Ngongo, M., & Benu, N. (2020). Interpersonal and ideational metaphors in the writing of thesis texts of undergraduate students of English study program: A systemic functional linguistic approach. *RETORIKA: Jurnal Ilmu Bahasa*, 6(2), 113–120. <https://doi.org/10.22225/jr.6.2.2320.113-120>
- O'Halloran, K., & Fei, V. L. (2014). Systemic functional multimodal discourse analysis. In S. Norris & C. D. Maier (Eds.), *Interactions, images, and text: A reader in multimodality* (pp. 137–154). Walter de Gruyter, Inc.
- Puchta, H., Gerngross, G., & Lewis-Jones, P. (2017). *Super minds: Student's book* (Special ed). Cambridge University Press.

- Richards, J. C. (2002). *Curriculum development in language teaching* (2nd ed.). Cambridge University Press.
- Royce, T. D. (1998). Synergy on the Page: Exploring intersemiotic complementarity in page-based multimodal text. In *JASFL Occasional Papers* (Vol. 1, Issue 1, pp. 25–49).
- Royce, T. D. (2007). Intersemiotic complementarity: A framework for multimodal discourse analysis. *New Directions in the Analysis of Multimodal Discourse*, 63–109. <https://doi.org/10.4324/9780203357774>
- Suciati, H., Rustandi, A., & Sugiarto, B. R. (2021). Modality realised on hortatory exposition texts used in senior high grade xi textbook. *JEEP (Journal of English Education Program)*, 8(1), 35–42.
- Sugianto, A. (2021). 'Can we see it?': Contextualizing ' deforestation' from an English-medium science textbook for a primary school level. *J-Lalite: Journal of English Studies*, 2(2), 86–102. <https://doi.org/10.20884/1.jes.2021.2.2.5072>
- Sugianto, A., Andriyani, D., & Prasetyo, I. A. (2021). The visual-verbal text interrelation: Lessons from the ideational meanings of a phonics material in a primary level EFL textbook. *EnJourMe (English Journal of Merdeka) : Culture, Language, and Teaching of English*, 6(1).
- Sugianto, A., Denarti, R., & Agung, I. (2021). Uncovering the anti-Islamic sentiment in The New Yorker cover issued on July 21 , 2008: A semiotic analysis. *International Journal of English Linguistics, Literature, and Education (IJELLE)*, 3(1), 44–54. <https://doi.org/doi :10.32585/ijelle.v3i1.1450>
- Sugianto, A., Prasetyo, I. A., Aria, D., & Wahjuwibowo, I. S. (2022). Demystifying the hegemony of the English language: Scrutiny of 'Gak Bisa Bahasa Inggris!' advertisement within a semiotics lens. *Elsya: Journal of English Language Studies*, 4(2). <https://doi.org/10.31849/elsya.v4i2.7582>
- Sugianto, A., & Prastika, M. A. (2021). 'Are they merely pictures?': Delineating the images represented in acrostic poems of a primary school level EFL textbook. *IJECA (International Journal of Education and Curriculum Application)*, 4(3), 273–282. <https://doi.org/10.31764/ijeca.v4i3.5834>
- Sugianto, A., & Wirza, Y. (2021). Cultural contents of an EFL textbook: How is the potential for students ' intercultural communicative competence development during the COVID-19 outbreak? *Proceedings of the Thirteenth Conference on Applied Linguistics (CONAPLIN 2020)*, 546, 1–6.
- Thompson, G. (2014). *Introducing functional grammar* (3rd ed.). Routledge.
- Torres, G. (2015). 'Reading' World Link: A visual social semiotic analysis of an EFL textbook. *International Journal of English Language Education*, 3(1), 239–253. <https://doi.org/10.5296/ijele.v3i1.7200>
- Unsworth, L. (2006). Multiliteracies and a metalanguage of image/text relations: Implications for teaching English as a first or additional language in the 21st century. *Tales out of School: Identity and English Language Teaching. Special Edition of TESOL in Context, Series S(1)*, 147–162.