



# ‘Are They Merely Pictures?’: Delineating The Images Represented in Acrostic Poems of a Primary School Level EFL Textbook

Ahmad Sugianto<sup>1</sup>, Marisa Azka Prastika<sup>2</sup>

<sup>1,2</sup>Universitas Pendidikan Indonesia

[ahmadsugianto@upi.edu](mailto:ahmadsugianto@upi.edu)<sup>1</sup>, [marisaazkap@upi.edu](mailto:marisaazkap@upi.edu)<sup>2</sup>

## ABSTRACT

### Keywords:

Acrostic Poem;  
EFL Textbook;  
Grammar of Visual  
Design;  
Image-Text Relation;  
Logico-Semantics;  
Multimodal Analysis.

A picture is considered to have a significant role in facilitating the students' English learning. Thus, understanding the picture represented in an EFL textbook considered one of the sources through which the students learn the English materials is crucial. Nevertheless, to the best of the writers' knowledge, few studies investigating the interrelations between visual and verbal text focusing on the learning material of acrostic poems were found. Thereby, this study aimed to scrutinise the interrelations of the visual-verbal text concerning acrostic poems. To that end, a qualitative research method using Systemic Functional Multimodal Discourse Analysis (SFMDA) was employed to investigate the analysis units, i.e., the acrostic poems included the images of a primary level EFL textbook. Besides, analysis units were investigated based on the relative status and logico-semantic relations and grammar of visual design deriving from systemic functional linguistics. The findings revealed that the visual image could be construed through ideational, interpersonal, and textual/compositional meanings. Besides, there are some interrelations in a certain extent and fashions between the pictures and the verbal text, indicated by the relative status and logico-semantic comprising independent and complimentary equal status with exposition, exemplification, and extension. In summary, the trinocular meanings, i.e. ideational, interpersonal, and compositional meanings, along with the interactions between the visual image and verbal texts, the pictures are considered to have significant roles in assisting the readers/viewers in understanding the poems due to such an interaction built by the images and verbal texts exist.



### Article History:

Received: 10-11-2021  
Revised : 23-11-2021  
Accepted: 28-11-2021  
Online : 06-12-2021



This is an open access article under the **CC-BY-SA** license



<https://doi.org/10.31764/ijecca.v4i3.5834>

## A. INTRODUCTION

Construing pictures in association with language has been a prominent study at the present time. It is due to the development of technology, any information is provided not only through a single-mode, namely through merely a verbal text, but the other semiotic mode such as pictures are required to be taken into account. The availability of software, as Curwood (2012) asserts, make ease of other semiotic modes to be embedded into texts. Moreover, pictures have an essential role, that is, the message of a particular message can be conveyed through it (Sugianto, Denarti, & Prasetyo, 2021), as well as realise some words or depict the communication in real life situation that enable to draw students' attention (Torres, 2015). Thereby, scrutinising the way images or pictures embedded in the verbal text and vice versa can be considered to be a

growing interest (Sugianto, Andriyani, & Prasetyo, 2021), which is particularly a part of social semiotics study (O'Halloran & Smith, 2011) stemming from the notion that views language as an entity dependent upon the socio-cultural context (Halliday & Matthiessen, 2014).

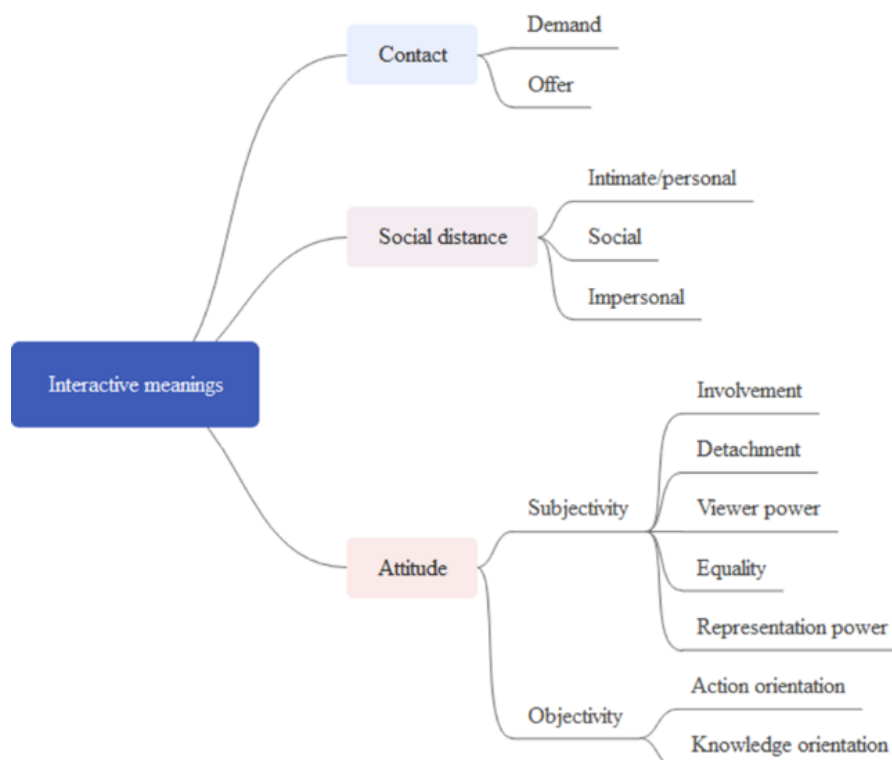
In addition to the socio-cultural context, pictures in EFL teaching and learning have also been considered to be essential, particularly in those that are embedded in an EFL textbook. It is due to the fact an EFL textbook is deemed to have a key role through which students learn certain cultures (Sugianto & Wirza, 2021) and obtain authentic exposures (Marzban & Zokaeieh, 2017). Moreover, to bear in mind, pictures included in an EFL textbook are not merely accessories (Hermawan & Sukyadi, 2020), but these are generated rigorously by collaborations between the art designers and textbook authors (Bezemer & Kress, 2010; Sukyadi, Hermawan, & Rakhmawati, 2018). Regardless, its rigour in terms of textbook production, an EFL textbook without a doubt is required to be used and chosen carefully and selectively because there is a possibility that the textbook may contain any deficiency that may not suit with the students' needs and context (Ayu & Indrawati, 2018; Kusuma & Apriyanto, 2018; Richards, 2001). Thus, evaluating an EFL textbook, in particular associated with the pictures represented in it, is considered to be worth doing. It is not only because of understanding pictures along with verbal text is deemed to be a growing area scrutinised but the deficiency of the textbook can also be spot-lighted.

Numerous frameworks as a tool for analysing multimodal texts have been developed and advocated by scholars and experts. In this case, the present study made use of some particular frameworks, namely Kress & van Leeuwen's (2021) representations of visual meanings and Martinec and Salway's (2005) relative status and logico-semantics relations. These were employed to prove the assumptions that these frameworks can facilitate the scrutiny of the analysis units of the present study.

To begin with, in terms of visual meanings, some representations are advocated by Kress and van Leeuwen (2021). In this regard, they point out the representations of visual images encompass narrative and conceptual representations comprising several processes: 1) action processes in which the participants at least consist of the actor and the goal; the action processes can be divided into several types of actions, i.e., unidirectional transactional action (in which the vector indicated by an arrow connect the participants encompassing an actor and a goal), bidirectional transactional action (in which the vector indicated by a double-headed arrow connects the participants or it can also be in the form in which there are two vectors indicated by two arrows with different directions connect the participants/interactors), and non-transactional action (in which there is only one vector from one participant without pointing to the other participant); 2) reactional processes (in which the vector indicated by an arrow commonly from the participant named as a reactor to point a phenomenon); 3) speech process and mental processes commonly found in comic strips in which the participants comprise senser if associated with the thought ballon and a speaker in a dialogue ballon; 4) conversion processes in which the the participant can have two types of roles simultaneously, namely as the actor and the goal; 5) geometrical symbolism in which only a vector emerges without any participants enclosed; 6) circumstances comprising setting, means, and accompaniment, 7) analytical processes associated with the 'part-whole' structure' involving the participants named carrier (the whole) and possessive attributes 'the part'; 8) symbolic processes involving the participants comprising the carrier and the symbolic attribute (pp. 58-107).

Moreover, in terms of the interactive meanings, there are three components through which the visual images can be represented. They encompass *contact* which can be realised through

the presence (demand) or absence of the gaze from the represented participants to the viewers or the readers (offer); *social distance* which can be indicated by the range of the shots of the image taken (comprising intimate personal indicated by close shot, social indicated by a medium shot, and impersonal indicated by a long shot); and *attitude* which can be divided into two strands, i.e. subjectivity associated with the horizontal angle, i.e. whether the image is taken in a frontal angle indicating involvement or an oblique angle indicating detachment, and another strand having to do with subjectivity relates to the vertical angle, i.e. whether the image is taken from a high angle indicating viewer power, eye-level angle indicating equality, and low angle indicating represented participant power; another representation of attitude is objectivity, i.e. either an image focuses on action representation or knowledge representation (pp. 115-144). All these components are shown in detail in Figure 1 below.



**Figure 1.** Interactive Meanings Realised in Images (p. 143)

In addition, in terms of textual or compositional meanings, there are three essential elements through which an image can be represented. They comprise information value, salience, and framing. To begin with, the information value has to do with 'zone' or placement of the image, i.e. represented left and right, top and bottom, centre and margin; the additional representation has to do with salience associated with the degree of the elements of the image realised by its placement (foreground or background), size, colour, and sharpness; and another realisation of textual or compositional meaning has to do with presence or absence of 'framing devices' represents connection and signification of the elements through 'dividing lines' or 'frame lines' (pp. 181-182).

Furthermore, to delineate the meaning of pictures, the inspection can be associated with the verbal text accompanied. In this regard, Martinec and Salway (2005) advocate a system through which the relations between the image and verbal text can be scrutinised, i.e. the relative status and logico-semantics relations. This system was developed based on Barthe's (1977) basic

theory of image-text relations encompassing three types of possibilities, i.e. *anchorage* (i.e. the verbal text explains the visual image), *illustration* (i.e. the visual image explains the verbal text), and *relay* (i.e. the visual image and verbal text are deemed to be complementary each other). Moreover, the system, particularly having to do with the logico-semantics stems from the logico-semantics system of clause complex in functional grammar encompassing the projection and expansion, and more particularly having to do with the logical metafunction (Halliday & Matthiessen, 2014). Regardless of the similar system to logico-semantics of functional grammar, the system of logico-semantics of the image-text was considered to add another additional strand, which in this regard, has to do with the relative status. The relative status is based on the assumption that the relations of the text can be equal or unequal. The former was divided into two types, independent and complementary, whereas the latter associates whether a verbal text is subordinate to an image and vice versa (Martinec & Salway, 2005) as Figure 2 below.

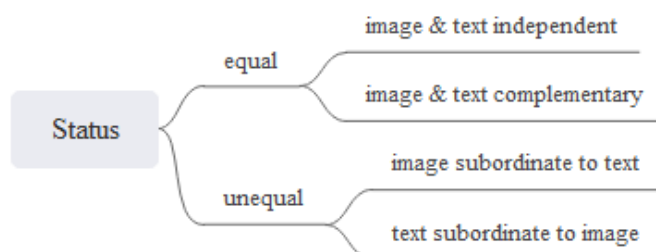


Figure 2. Relative Status Of Image-Text (p. 349)

Logico semantics relations of image-text as the Figure 3 below.

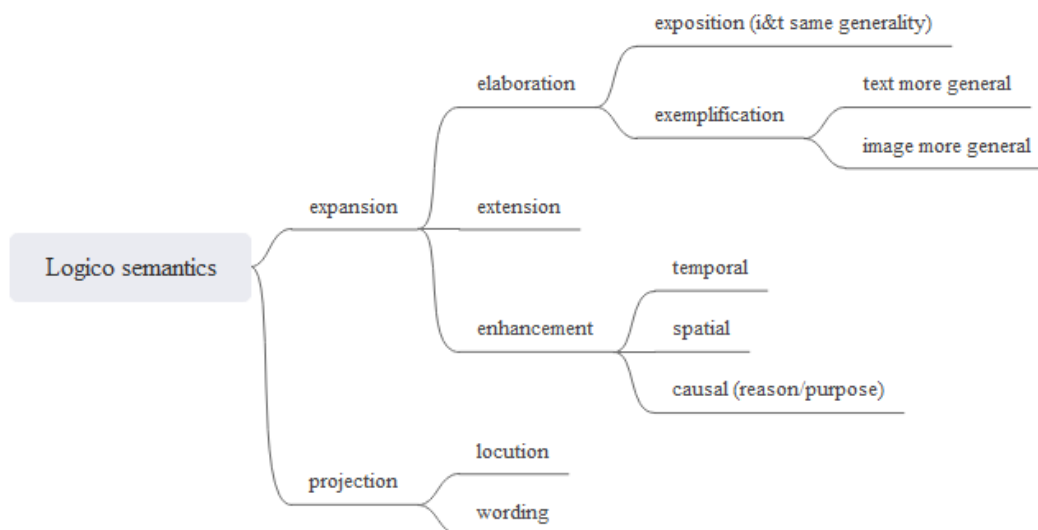


Figure 3. Logico Semantics Relations of Image-Text (p. 358)

Moreover, regarding the logico-semantics relations represented in Figure 3, there are two main strands, expansion and projection. To begin with, expansion can be divided into three encompassing elaboration which is divided based on the aspect of generality; in this regard, if both the image and verbal text have the same generality, the relation is considered exposition, whereas if one of them is more general than the other, relations are called as exemplification. Meanwhile, in terms of extension associated with additional information that goes beyond the things represented. In addition, enhancement can be divided into three types, namely

enhancement by temporal or time, by circumstance or place, and by reason/purpose. Lastly, the projection is similar to functional grammar (Halliday & Matthiessen, 2014), which consists of two categories, namely locution having to do with the use of exact quotation realised through verbal processes and speech bubbles, and idea associated with the reported speech related to the use of mental process realised by thought bubbles (Martinec & Salway, 2005).

Furthermore, the present study tried to shed some light on multimodality in regard to acrostic poems included in an EFL textbook. There are some grounds underlying the scrutiny on this particular subject. To begin with, an acrostic poem is considered to have its own place in students' learning. In this regard, it is one of the topics learned by students at school (Buckingham & Litzler, 2020). Moreover, an acrostic poem is proven to be one of the effective ways through which students learn creative thinking that is required in learning writing skill (Khoii & Amin, 2016). Additionally, Dougherty (2011) reveals that acrostic poems as a part of creative writing can promote students' vocabulary acquisition. In addition, acrostic poem as a part of creative writing has been regarded as an essential subject of EFL students' learning. It was due to the fact that creative writing constitutes an integral strand needed as one of the core skills of learning and innovation named 4Cs required in the twenty-first century (BatteleforKids, 2019). Regardless of the many inspections with respect to acrostic poems conducted by researchers, to the best of the writers' knowledge, few studies concerning the multimodal analysis associated with acrostic poems were found. Thereby, based on the rationale above, the present study was conducted to investigate the pictures and the verbal text relations in an EFL textbook with regard to acrostic poems.

## **B. METHODS**

Qualitative research by using systemic functional multimodal analysis (SF-MDA) was employed. It was used, as Jewitt, Bezemer, and O'Halloran (2016) assert, to go through the meanings systems with respect to social functions. Additionally, the meanings systems comprise the visual image and verbal text systems. Moreover, the analysis units or artefacts investigated were taken from a primary school level EFL textbook (Puchta, Gerngross, & Lewis-Jones, 2017). In particular, they were acrostic poems. The acrostic poems were selected because they were considered to be multimodal texts in which there are visual semiotic mode and verbal semiotic mode used.

To analyse the artefacts, Kress and van Leeuwen's (2021) grammar of visual design and Martinec and Salway's (2005) relative status and logico-semantics relations. In this case, the procedures encompass two steps, i.e., firstly, analysing each semiotic mode emerging in the artefacts, which in this regard comprise visual images and verbal texts; secondly, following the analyses of each semiotic mode, the analyses concerning the relations of the visual mode and verbal text mode were investigated. Moreover, to ensure the trustworthiness of the findings, some techniques were employed. In this regard, describing in a 'thick' fashion, building 'referential adequacy' by consulting some relevant resources, and peer debriefing by communicating the results to the researchers' colleagues (Guba, 1981). Also, concerning the ethical issues, particularly having to do with the copyright issue of the artefacts scrutinised, the colourful multimodal texts were changed into black and white colours (Jewitt et al., 2016) reported in this article.

### C. RESULT AND DISCUSSION

This section uncovers the findings with respect to the meanings of the artefacts, two acrostic poems entitled FRIENDS and DINOSAURS based on narrative representations, interactive representations, compositional meanings, relative status, and logico-semantic relations. To begin with, in terms of the narrative representations, the first artefact entitled FRIENDS (henceforth FRIENDS), there are six children in FRIENDS. They are the represented participants. In this case, the first two participants, a boy with spectacles brown hair, dark blue jumper, green trousers, and brown shoes, carrying a light blue bag on his back seem to have a conversation with a girl with a yellow/blonde ponytail, wearing a red jumper, black trousers, brown boots, and carrying a brown sling bag. These two participants placed in the foreground seem to play as the actors and goals simultaneously or sequentially who are talking interactively, and thus they are considered to be a bidirectional action (Kress & van Leeuwen, 2021). Additionally, the symbolic attributive processes also appear in the first two participants, in which the two children are as the Carrier and their physical appearances are considered as the Symbolic Attribute.

Similarly, there were two children near to the first two children in the foreground. These two children is a boy and a girl. The boy wears a light green T-shirt with an aqua colour and blue shoes holding a basketball showing to the girl with afro hair and wearing blue and yellow sweater with red trousers and white shoes. The images of these two children generate some narrative processes. First, the transactional process is generated by the vector indicated by the boy seeing the curly girl; in this regard, the girl acts as the goal and the boy as the actor. Furthermore, the girl also can act as the actor generating a vector indicated by her glance at the ball which acts as the goal. Moreover, the transactional process created by the image of the boy holding a basketball can shift to be a phenomenon and the girl as the reactor; hence, there is a reactional process.

Additionally, an analytical process and a symbolic attribute emerge in this image, in which the boy and the girl are the Carriers and the (symbolic) possessive attribute process comprises their physical appearances (skin colour, hair, and clothes). Moreover, there are two other children in the background playing a swing. The symbolic attribute constitutes the process found in these two children. They become the Carriers and the symbolic possessive attribute refers to their physical appearances, i.e. the clothes they wear and skin colours as well as their hair. Moreover, another artefact entitled DINOSAURS (henceforth DINOSAURS) has only one represented participant with two types of processes, namely non-transactional action process realised by the represented participant, i.e., Tyrannosaurus Rex (T-Rex) as Actor with an imaginary arrow made through the way it roars; another process is symbolic Suggestive process (pp. 105-106), in which T-Rex is the Carrier with some symbolic attributes such as its sharp teeth, tail, and other physical appearance.

The interpersonal meanings are realised by several aspects, namely contact, social distance, attitude, and modality. To begin with, concerning the contact aspect, some of the represented participants, i.e. the two first two children in the foreground, are oblique, i.e. their gazes do not directly to the readers/viewers, meaning that instead of trying to engage with the viewers/readers they 'offer' information; thus they become the 'objects of contemplation' (p. 120). Meanwhile, the children in the background seem to be illustrated to have a frontal gaze, that is, the children in the foreground look at directly the readers/viewers, meaning that the children 'demand' the readers/viewers to have an engagement with them (p. 118).

The other aspect is the social distance, in which in this case, the first two children in the picture are depicted in long shot indicated by the whole figure representation, meaning that they are regarded to have a far social distance; meanwhile, the other last two children in the background are depicted in public distance due to their long shot depiction, meaning that the relationship between the represented participants and the viewers/readers are like strangers one another (Hall, 1966; Kress & van Leeuwen, 2021). Moreover, on the one hand, in terms of the aspect subjectivity of attitude, the first two children presented in the foreground are depicted in a way that some detachment emerge due to their oblique gaze; hence there seems no interaction built between them and the interactive participants; additionally, due to their depiction at eye-level, there is equality in terms of power between the children in the foreground and the interactive participants (Kress & van Leeuwen, 2021). In addition, concerning the aspect of objectivity of the attitude, the represented participants tend to be either knowledge orientation or action orientation since these two pictures illustrate some processes that the children are doing and the condition which refers to the main topic of the poem FRIENDS.

Furthermore, on the other hand, the represented participants, namely two children in the background, in regard to the aspect of subjectivity of the attitude, due to their frontal gaze they are considered to have an engagement to the viewers/readers, but it tends to be viewer power in which interactive participants are deemed to have power over the represented participants. Additionally, in terms of modality, due to the full-colour saturation and adequate detailed background, yet they are non-naturalistic pictures, i.e., cartoon, it can be considered that have medium modality. Meanwhile, in regard to the artefact DINOSAURS, due to the represented participant's frontal angle looking at directly the viewers/readers, it demands the readers/viewers to have an 'imaginary' emotional bond with it (p. 117). Moreover, because of the low angle, the picture is considered to have represented participant power or authority over the readers/viewers (pp. 138-139). Besides, due to its long shot, the social distance of this picture is included as far social distance (Hall, 1966). Due to its cartoon, non-naturalistic picture, without many colour saturations, the modality is likely to be low to medium. Additionally, the variety of colours employed in the textbook is possibly advantageous for students for it might attract the students' attention to the topic learned for it is created aesthetically (Sa'aleek, 2018).

Furthermore, in terms of textual/compositional meaning, FRIENDS is realised by salience, information value, and framing. In this regard, in terms of salience, the first two children in the foreground are the most salient elements in the composition. It is on account of the fact that they have the largest represented participants depicted in the composition as well as have the most bright colour or light showing by the represented, the female girl with a yellow ponytail carrying a brown sling bag; besides, the red jumper makes it as the more salient represented participant in comparison with the other represented participants (Kress & van Leeuwen, 2002).

Meanwhile, the other two children playing a swing constitute the least salient elements in the composition due to their placement in the background and their size or the way they are illustrated, i.e. the smallest feature and they are depicted rather blurred. Thus, it can also be considered that, in association with the other aspect of compositional meaning, which is information value, the New (shown by the represented participants on the right side) is more salient than the Given (the represented participants in the left side). Also, not only the picture is based on the Given-New structure, but also this picture, along with the verbal text, can be considered to have the Ideal-Real structure, in which the verbal text constitutes the Ideal and the represented participants showing the detailed representations of the acrostic poem constitute

the Real (Kress & van Leeuwen, 2021). Thus, it can be considered that the compositional meaning of FRIENDS utilised combined the Given-New structure and the Ideal-Real structure. Also, due to its name, acrostic poem has a typical feature in which it can be read vertically, the initial letters of each phrase of the verbal text have formed a word FRIENDS, of which colours are made in contrast with the background; thus, the verbal text is deemed to be more salient than the background depicting a light blue sky. In addition to the elements of compositional meaning above, another element, that is, framing. To begin with, the frame lines in a white colour seems to segregate the picture along with the verbal text with the other space of the textbook, including the instruction 'Read these poems, which are called an acrostic. What is an acrostic?'. Also, it seems one of the represented participants, namely a girl with a yellow and blue sweater and brown trousers and white shoes with afro hair, can be a 'framing device' separating the two represented participants playing a swing in the background. Meanwhile, for the artefact DINOSAURS, the information value tends to be a Given-New structure in terms of the verbal text along with the visual image. In this regard, the verbal text is the Given due to its position on the left side and in contrast, due to the fact that it is placed on the right side, the represented participant, T-Rex, is the Given in this compositional structure. In terms of Saliency, the represented participant is considered to be more salient in comparison to the verbal text since its size (even its tail leaves out the frame) and some orange arrows. Similar to FRIENDS, DINOSAURS makes use of white frame line to segregate the picture with space outside of it as well as the instruction of the text.

Moreover, another inspection of FRIENDS has to do with the relative status between the pictures and the verbal text as well as the logico-semantics. In this regard, some of the verbal texts, for example, 'Fun and laughter', 'In our playground', 'Enjoy', 'Not alone' are considered to have equal status in which both the image and text are complementary (Martinec & Salway, 2005). These phrases are related to the whole image, including the represented participants. For instance, the phrases 'Fun and laughter' and 'Enjoy' are related cohesively by synonymy to all the represented participants either in the foreground and background indicated by their smiles; additionally, 'In our playground' is related to the setting in which the represented participants are playing, indicated by the surrounding objects, such as a slide and a swing; besides, the phrases 'Not alone' and 'FRIENDS' are depicted by all the represented participants showing that they, for instance, have a friend to play with or talk to. Nevertheless, the remaining phrases, namely 'Running around', 'Dreams' is cohesively related by hyponymy to the represented participants indicated by the represented participants who are children commonly having dreams or passions, for instance, one of the represented participants in the picture to be a basketball player associated with his physical appearance and accessories, wearing shorts, basketball, and sports shoes; meanwhile, the phrase 'Sunshine' is cohesively related by hyponymy by the setting, i.e. clear weather indicated by the light blue sky.

Hence the relative status is likely to comprise more than one, namely image and text-independent realised by 'Dreams' and 'Sunshine' with the pictures and another relative status is image and text complementary realised by 'Fun and laughter', 'In our playground', 'Enjoy', 'Not alone', 'FRIENDS' and the picture. Meanwhile, in terms of logico-semantics the interrelation between the verbal texts and the images comprise elaboration, particularly exposition in which image and verbal text have the same generality and hence they interact with each other, i.e. one mode restates the other mode and vice versa; the exposition is realised by verbal text 'Fun and laughter', 'Enjoy', 'Not alone', 'FRIENDS' and pictures. The other image-text relation is exemplification in which the text is more general than the image realised by 'In our playground'



with pictures such as the swing and slide and 'Sunshine' with the clear weather indicated by the light blue sky. Another image-text relation is extension in which one mode adds something new to the other mode realised by 'Dream' to the represented participants, for instance, the children, and the passion to be a professional basketball player as realised by one of the represented participants as explained aforementioned above. Meanwhile, the same as FRIENDS, the second artefact, DINOSAURS, has some relative statuses comprising equal status, namely complementary, which is cohesively associated with synonymy by the text 'Dinosaurs were huge', 'Strong', 'And dangerous' Really Scary' associated with the way the represented participant is depicted, namely with its tail leaving out the frame, body, and sharp teeth; additionally, its logico-semantics is considered to be exposition in which the image and text have the same generality (p. 358). Additionally, the other phrases 'In a time far away', 'No people on Eath', 'One hundred million years ago', and 'Under water and on land' can be considered to be realised by hyponymy to the represented participant in which the text and image are independent, and the logico-semantics is deemed to be an extension in which the verbal text provides something new to the image (Hermawan & Sukyadi, 2020).

#### D. CONCLUSION AND SUGGESTIONS

Based on the findings and discussion above, some conclusions are drawn. To begin with, the pictures can be considered interrelated to a certain extent and fashion with the verbal text accompanying them, indicated by the relative status and logico-semantics comprising, independent and complimentary equal status with exposition, exemplification, and extension. Moreover, the trinocular meanings, i.e. ideational, interpersonal, and compositional meanings, along with the interactions between the visual image and verbal texts, the pictures are considered to have significant roles in assisting the readers/viewers in understanding the poems due to such an interaction built by the images and verbal texts exist. This study suggests that teachers are required to make use of the meaningful visual image for it has an essential role to assist the students in receiving the materials learned as well as such pictures are considered to be used as means for fostering critical thinking towards the subject matters since the pictures act as observable objects. Other subject matters might be an area that is worth investigating for future studies with the emphasis on other modes besides visual images, such as the use of gestures along with verbal texts.

#### REFERENCES

- Ayu, M., & Indrawati, R. (2018). EFL textbook evaluation: The analysis of tasks presented in English textbook. *Teknosastik*, 16(1), 21–25. <https://doi.org/10.33365/ts.v16i1.87>
- Barthes, R. (1977). *Image, music, text*. Fontana Press.
- BattleforKids. (2019). Framework for 21st century learning definitions. *Partnership for 21st Century Learning*, 9.
- Bezemer, J., & Kress, G. (2010). Changing text: A social semiotic analysis of textbooks. *Designs for Learning*, 3(1–2), 10–29. <https://doi.org/10.16993/dfl.26>
- Buckingham, L. R., & Litzler, M. F. (2020). Literacy development in EFL textbooks in Madrid's primary schools. *Revista Nebrija de Lingüística Aplicada*, 27, 129–144. <https://doi.org/10.26378/rnlael>
- Curwood, J. S. (2012). Cultural shifts, multimodal representations, and assessment practices: A case study. *E-Learning and Digital Media*, 9(2), 232–244. <https://doi.org/10.2304/elea.2012.9.2.232>
- Dougherty, P. (2011). *Creative writing and vocabulary acquisition: A study of Emirati college*

- students. 57–69. <http://marifa.hct.ac.ae/files/2011/07/Creative-Writing-and-Vocabulary-Acquisition-A-Study-of-Emirati-College-Students.pdf>
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology*, 29(2), 75–91.
- 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., & 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.
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge. <https://doi.org/10.4324/9781315638027>
- Khoii, R., & Amin, E. (2016). Which helps to become a better EFL writer: Creative tasks or communicative tasks? *Literacy Information and Computer Education Journal*, 6(2), 2295–2302. <https://doi.org/10.20533/licej.2040.2589.2016.0303>
- Kress, G., & van Leeuwen, T. (2002). Colour as a semiotic mode: Notes for a grammar of colour. *Visual Communication*, 1(3), 343–368.
- Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.). Routledge.
- Kusuma, H. A., & Apriyanto, S. (2018). Strategy on developing English learning material for specific purposes. *IJECA (International Journal of Education and Curriculum Application)*, 1(3), 39–48. <https://doi.org/10.31764/ijeca.v1i3.2144>
- Martinec, R., & Salway, A. (2005). A system for image–text relations in new (and old) media. *Visual Communication*, 4(3), 337–371. <https://doi.org/10.1177/1470357205055928>
- Marzban, A., & Zokaeieh, S. (2017). Local evaluation of an EFL textbook: 'Mosaic 1: Listening and speaking.' *Theory and Practice in Language Studies*, 7(10), 874–882. <https://doi.org/10.17507/tpls.0710.08>
- O'Halloran, K. L., & Smith, B. A. (2011). Multimodal studies. In K. L. O'Halloran & B. A. Smith (Eds.), *Multimodal studies: Exploring issues and domains* (pp. 1–13). Routledge.
- Puchta, H., Gerngross, G., & Lewis-Jones, P. (2017). *Super minds: Student's book* (Special ed). Cambridge University Press.
- Richards, J. C. (2001). *The role of textbooks in a language program*.
- Sa'aleek, R. A. (2018). A content analysis of the colors in the EFL textbook of Action Pack. *British Journal of Education, Learning and Development Psychology*, 1(1), 57–65.
- 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., & 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.
- Sukyadi, D., Hermawan, B., & Rakhmawati, S. S. (2018). *Multimodalitas dalam pembelajaran Bahasa Inggris*. UPI Press.
- 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>