IMAGE-TEXT RELATION INTERPRETATION:
TEACHERS' VISUAL-VERBAL COMPETENCE IN TEACHING TEXTS

Yustika Nur Fajriah*, Fuad Abdul Hamied, Wawan Gunawan
Universitas Pendidikan Indonesia, Indonesia
*email: yustikanurfajri@upi.edu

Abstract: Multimodal literacy instruction is such a new shift of the literacy in which the construction of knowledge is led to be more socially and contextually bounded. Due to the urgency that teachers have to own multimodal competencies, this study aimed to investigate EFL teachers' competence in interpreting visual-verbal relations to teach multimodal texts. To collect the data, an online test through the Google form platform was distributed. As many as 43 responses were collected from junior and senior high school teachers in one of the cities in Indonesia. A semi-structured interview was also conducted with six purposive participants. The data in this research were then analysed based on Royce's criteria of image-text relation. The analysis found that the teachers only partially possessed multimodal competencies. It means that that they had used images to help them teach the texts but had insufficient knowledge on how to utilize the images as meaning-making sources. Then, based on the finding, it is suggested that the teachers should improve their competences in interpreting multimodal meanings in texts, so images are used not only for making learning materials interesting but also for making more meanings from the texts.

Keywords: meaning-making, multimodal texts, visual images.

HUBUNGAN GAMBAR DAN TEKS DITAFSIRKAN:
KOMPETENSI VISUAL-VERBAL GURU DALAM MENGAJAR TEKS

Abstrak: Pembelajaran teks yang bersifat multimodal merupakan perubahan arah dari pembelajaran literasi di mana pengetahuan dikembangkan secara sosial kontekstual. Penelitian ini bertujuan untuk menganalisis kemampuan guru bahasa Inggris dalam menginterpretasikan hubungan visual dan verbal saat mengajarkan teks yang bersifat multimodal. Untuk itu, tes daring dalam bentuk *Google form* disebar. Sebanyak 43 guru bahasa Inggris di tingkat SMP dan SMA di salah satu kota di Indonesia. Selain itu, wawancara juga dilakukan kepada enam orang guru yang ditetapkan sebagai partisipan inti pada penelitian ini. Data yang sudah diperoleh dianalisa dengan mengacu pada teori Royce terkait hubungan gambar dan teks. Hasil analisis menunjukkan bahwa guru belum sepenuhnya memahami modalitas pada teks. Mereka telah menggunakan gambar dalam pembelajaran, namun kemampuan mereka dalam mengeksplorasi makna suatu gambar dan kaitannya dengan teks masih rendah. Temuan ini merefleksikan bahwa guru seharusnya dapat mengembangkan kemampuannya dalam menginterpretasikan makna multimodalitas pada suatu teks, sehingga unsur gambar pada teks tidak hanya ditujukan untuk aspek daya tarik siswa terhadap membaca, namun lebih pada menggali konstruksi makna pada teks tersebut.

Kata Kunci: konstruksi makna, teks multimodal, gambar visual.

INTRODUCTION

Multimodal literacy instruction has become the trend in this worldwide education. This is such a new shift of the literacy in which the construction of knowledge is led to be more socially and contextually bounded (McKee & Heydon, 2015). Furthermore, the advancement of multimodal literacy instruction has been obviously extended not only for language instruction (Jacobs, 2006; Shanahan, 2013; Wang & Zhan, 2010), but also for other knowledge disciplines such as science class (Murcia, 2014; Tang & Moje, 2010), arts teaching, business education (Tomsett & Trott, 2014). Put differently, the interest in integrating multimodality in literacy instruction is growing higher and sharper.

Lined up with this development, the learning materials and media in the classroom have been also multimodally and technologically

oriented. This is harmony with the statement that students have to be equipped with skills to understand multimodal texts; such as making judgement, recognizing perspective, and clarifying values in the texts. As an example, today the technology supports to develop multimodal character of a text through producing and utilizing visual texts (Lim-Fei & Yin, 2017). Furthermore, multimodal teaching and media are being discussed by numerous researchers; such as interactive whiteboard (Murcia, 2014; Twiner, Coffin, Littleton, & Whitelock, 2010), picturebooks (Hermawan & Sukyadi, 2017; Wu, 2014), gadget apps and software (Lim-Fei, O'Halloran, Tan, & E, 2015; Vungthong, Djonov, & Torr, 2017), and Digital Story Telling (Brailas, 2017; Çıralı & Usluel, 2015; Robin, 2006; Wang & Zhan, 2010). Seeing this, multimodal communication skill is highly needed by TESOL Professionals, specifically teachers (Royce, 2002).

Nevertheless, the importance of multimodal communication skills has not been recognized thoroughly by teachers; especially EFL teachers in the ASEAN context. This is harmony with Hundley & Holdbrook (2013) that due to the established interpretation of literacy and the view of language which is always linguistically-centred, many teachers do not easily welcome the essence of multimodality in their teaching. Peculiarly, many of them keep handling classes with the use of a single resource; named as flat literacy (Jaksic, 2017); such as, only using the power of teachers' verbal explanations in the classroom setting. Even, little is known about the manner to use visual modes; specifically, images as the EFL teaching resource to make meanings (Vungthong, Djonov, & Torr, 2017). Specifically, images are often claimed as the harmonizing ornament in a text teaching. This is also strengthened by some related research findings, one of them in Philippines, that Junior High School teachers have not been endowed by sufficient competence to employ visual modes; images, postures, and others in customizing the teaching and learning process (Gabinete, 2017). In short, many EFL teachers are not conversant to apply the multimodal approach to doing any instruction in the classroom. The report of these studies is also strengthened by other scholars' endeavours showing that many teachers have not been well-prepared and trained to integrate multimodality to the teaching process (Ajayi, 2011; Howell, Butler, & Reinking, 2017; Sewell & Denton, 2015).

Based on the previous studies above, there is still an intriguing area to discuss more; one of them is emphasizing on how images and texts are contributing to meaning-making process. Hence, this study is aimed at investigating EFL teachers' initial competence in understanding the relation of images and texts in constructing particular meanings. Specifically, this study focused on exploring multimodal competence among teachers in part of Indonesian regions in which the teachers' multimodal competences vary.

METHODS

This study used qualitative descriptive method since this focuses on seeking deep understanding on teachers' multimodal skill in viewing image-text relation in teaching texts. Forty-six teachers served as the participants coming from Junior and Senior High Schools in a part of Indonesia's region. As the profile, most of them are categorized as experienced teachers who have been teaching for more than six years. They participated in this study through the electronic data collection; that is, online test through Google form. The questions are set to uncover the teachers' understanding of the concept of image-text relation in making meanings. Peculiarly, six focus participants were chosen to follow semi-structured interview. They were selected based on their different perceptions in teaching texts completed by images. A half of them believed that images must be explained or discussed in the beginning stage, while, the others think that the text (written form) should be clarified in the first place since images are not so important in conveying the meanings.

Generally, the data in this study were analysed inductively (Thomas, Nelson, & Silverman, 2005), consisting of categorizing, identifying, analysing and interpreting. In regards to the text analysis, Royce's classification of image-text relations is used in this study. The theory of Systemic Functional Linguistics (SFL) from Halliday & Matthiessen (2004) and Kress & van Leewen (2006) is also a reference utilized to strengthen the data analysis; focusing on experiential metafunction analysis of the image (participants, process, and circumstances).

FINDINGS AND DISCUSSIONS

This section elaborates the data findings. Specifically, the findings are divided into two major sections; teachers' understanding of the notion of visual images in teaching texts, as well as their understanding of visual literacy, particularly about text-image relation.

Findings

Teachers' Understanding of the Notion of Visual Images in Teaching Texts

This part provides the data about teachers' understanding of the importance of images in the text teaching, the types of images frequently used in the classroom, and the teachers' knowledge in giving multimodal -text instruction.

All forty-six teachers perceived that images role prominent in teaching students' texts. This means that the teachers, who have been teaching for these six years, are little bit aware of the visual literacy. This is also followed up by some interviewed participants claiming that image is important in teaching a certain text since it can help students understand the text more easily. It is also stated by one of the respondents that *'kalau tidak ada gambar, maka kita sebagai guru akan kesulitan menghubungkan makna teks dengan anak karena kan tidak semua anak mempunyai background knowledge terhadap topik yang sedang dibicarakan'*, meaning that only single mode (written text) is not enough for guiding students understand the materials.

Furthermore, all respondents have apply images in teaching texts to their students. It is seen from the Figure 1 about types of images usually used and selected during their teaching of genres.
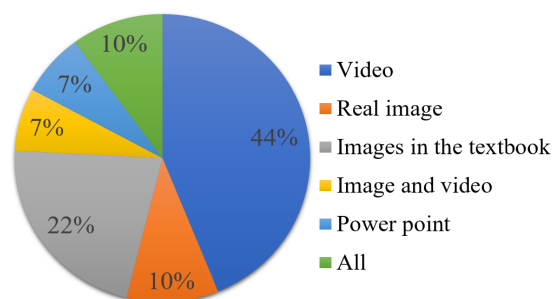

Figure 1. The Source of Taking Images in Teaching Texts

Regarding the types of images, most of teachers usually used video, pictures in the textbook, as well as browsed images from the internet. Three respondents said that video is more 'live' that other image forms since it combines any elements such as sound, moving images, texts, and others; therefore, it contributes to sharping the conveyed meanings. Not only that, they also stated that *'gambar yang ada pada buku teks juga lebih paktis dipakai, karena aktifitas tentang pengajaran teks nya juga sudah lengkap disana'* (the pictures in the textbook are more practical to use since they are completed by the stages to teach the text). This means that the aspect of practicality became the main reason of choosing the images provided in the textbook. The data also show that few of them have combined using pictures in the printed form with the video; meaning that the teachers had tried to use more that one media to present some images in teaching genres.

However, practically, all respondents had no certain aims, strategies, as well as procedures in exploring images while teaching texts. Particularly, three of them said that they discussed the image first (asking about what happens in the images) then students are asked to read the written text. This mechanism is different from the others, telling that students are guided to read the whole written texts, then asked to clarify their understanding by seeing the provided images. Peculiarly, when there were asked about the approach they used in teaching texts, most of them answered that they had no certain approach or method to teach texts and images. Most of them stated that they asked their students to read the text and discuss the images, after that students were asked to answer some questions related to the content of the text, then make the texts by themselves. As an example, a respondent answered, *'tidak ada pendekatan yang pasti,namun intinya saya mengenalkan teks dan gambar, anak anak membaca dan berdiskusi, menjawab pertanyaan seputar teks, dan membuat teks serupa masing masing'* (there are no specific approaches yet I usually use the stages; such as, I introduce the text and the images, I ask students to read and discuss some related questions, and I instruct them to construct texts).

To conclude, the respondents had utilized images in teaching students' texts, yet they had not been well- prepared in designing the meaningful activities in teaching images and texts as well as in determining the focus of

teaching images in constructing meanings.

Image-Text Relation in Meaning-Making

In this sub-section, there are three points to discuss; the function of images in teaching students' texts, the sequence of teaching (image-text or vice versa), and the dependence-relation between those two modes.

First, most of the respondents believe that images contribute to adding specific information of the texts, connecting students' prior knowledge as well as attracting audiences' interest. In addition, some respondents also claimed that image could functions to explain a grading. Figure 2 shows the data telling the function of images as perceived by the respondents.

From the data in Figure 2, it is also clear to say that most of respondents still considered images more as decorative tools rather than as meaning sources. As an example, a respondent said, *'gambar itu bisa membuat anak jadi lebih semangat belajar dikelas, berbeda dengan teks saja karena itu monoton'* (different form a single text which seems monotonous, images could make students more enthusiasm to learn the texts). Another respondent said, *'gambar itu menarik anak karena bentuk nya warnanya, sehingga anak akan lebih termotivasi saat membaca teks'* (images are interesting for students because of any colors and forms that make them more motivated to learn the texts). This reflects that no respondents knew that images could possibly have more contribution in delivering meanings than a written text.

Second, in the teaching practice, most of the respondents teach students the verbal texts first followed by the picture. It is in line with their general statements that mostly the verbal texts have completely conveyed the meanings of the story, and the picture is only the ornamental element in delivering ideas. Some respondents answered, *'teks dulu diberikan baru nanti mendiskusikan gambar'* (I give the text first then followed by the images). The data also reflect that the teachers had not quite understood the meaning construction in an image; such as, image could possibly have more meanings than texts, or meanings in images cannot be conveyed by the verbal modes.

In addition, the teachers were given the sample picture (Figure 3), yet none of the respondents explained the picture comprehensively. They merely said that the picture only strengthens the meanings who have been fully conveyed in the verbal text; that is, *Malin Kundang who hurted his mother*. As an instance, a teacher answered, *'Gambar ini menunjukan seorang anak yang bernama Malin Kundang sedang memarahi ibunya'* (this image shows a man who is angry to his mother). Another one has similar answer,*' Ini menggambarkan seorang anak yang durhaka pada ibunya'* (this depicts a disobedient man to his mother). Another similar response is also written, *'Ini tentang tragedy dimana seorang ibu disakiti oleh anaknya yang bernama Malin Kundang'* (this is about the tragedy when a mother is hurted by her son, Malin Kundang). Furthermore, there is no any explanation of this image comprehensively written by the respondents; such as, telling the setting (circumstance), the role of actors (participants), the possible actions (processes). It means that they did not know the possible meanings which could be explored in an image as well as the text-image relation conceptually.
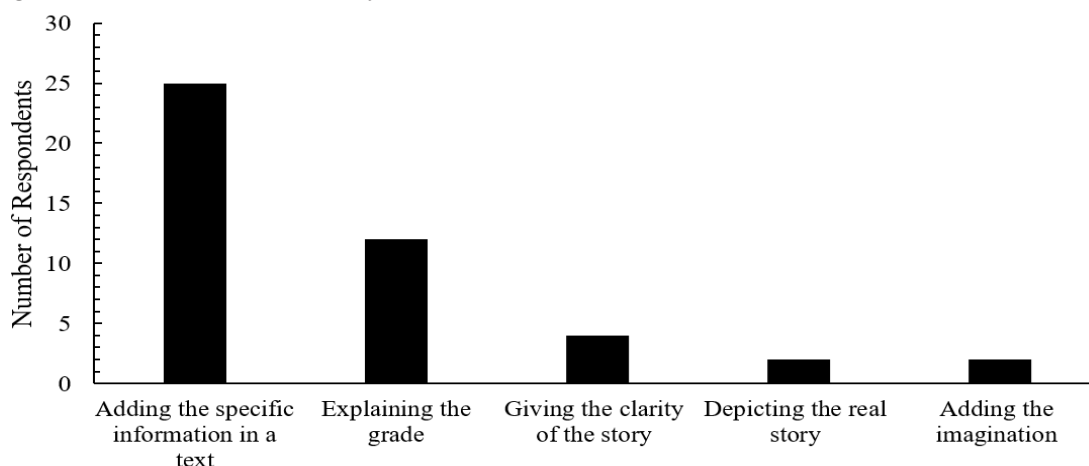

Figure 2. Function of Images in Teaching Texts

**Figure 3. The Sample Image to be Analysed
by the Respondents**
(https://alchetron.com/)

Discussions

This part elaborates the discussions of the data findings as well as the implication of the data findings.

Teachers' Understanding of the Notion of Visual Images in Teaching Texts

All respondents claimed that image is an essential part in teaching students' texts. This is in line with Lundy & Sktephens (2015) stating that image could assist students to comprehend the meanings of the texts beyond the literal context. More than that, visual clues in the images contributes to stimulate students' capacity in expressing their thought (Kedra & Zakeviciute, 2019). In short, image is claimed necessary to use in teaching texts. Additionally, the data also showed that teachers mostly used video in teaching texts. In line with this, it is mentioned that video is claimed as one of media consisting various multimodal modes; such as moving image, gesture, speech, movement (Rakhmawati, 2016). In addition to that, the result of Rakhmawati's study (2016) reveals that moving images shown in the video could help students develop their interpreting skills. It means that the meanings in the video could enrich the students' idea to translate a certain message.

Moreover, the teachers' preference to use pictures as provided in the textbook is also responded by researchers, proven by any studies in EFL teaching context focusing on analysing images in the textbook seen from any angle (da Costa & de Barros, 2012; Hermawan & Rahyono, 2019; Liu & Qu, 2014). Their studies implicitly show that teachers now days still gain the textbook as a good and practical alternative to use in the teaching texts. The studies also point out that the meanings of images in the textbooks are potentially explored by teachers if they know how to read them. In response to this, (Lim-Fei & Yin, 2017)) have explained that multimodal approach deals with a consideration to choose the modes (language, images, and others) in fulfilling the aim of a text, the addressee and the context, and the organization of the notions and information. It means that the teachers must be careful in determining modes in the teaching. Further, it is also stated that it is the teachers' work to help students understand the messages in the multimodal texts; such as, making judgment, analysing the perspective, and elucidating the values in the texts (Lim-Fei & Yin, 2017). Therefore, the images are potentially developing teachers and students' critical thinking in seeing a phenomenon.

Not only that, the result told that students were not explicitly guided to mean the images in discussing a particular genre. This finding is consonant with (Kress & van Leeuwen, 2006) that students are not taught how to understand the meanings in the images; indicating that teachers pay attention more on discussing the meanings of the passage. Therefore, it is explicitly stated that pedagogical consideration to visual literacy is still low (Yus, 2006; Serafini, 2012). It implies that many teachers are not aware of the importance of developing students' visual literacy in the classroom setting (Duchak, 2014) since the priority is delivering the content in the texts. Seeing this, a study has been conducted by (Hermawan & Rahyono, 2019) that transformation and transduction should be employed in teaching multimodal texts to students. This study discusses some opportunities of teachers in teaching students' multimodal literacy.

Above all, it is inferred that the teachers need to be equipped with the knowledge of how to develop their students' multimodal skills, particularly their visual literacy. This need is responded by numerous studies which promote some models, approach or techniques in teaching multimodal texts (Danielsson & Selander, 2016).

Image-Text Relation in Meaning-Making

Regarding the concept of image-text relation as mentioned by the respondent, the first function (giving details on meanings) is hand in hand with the theory of meaning-making in the SFL perspective, that unequal-exemplification could possibly occur between text and image (Royce, 2002) in which the position of text and image is not equal in constructing meanings. Furthermore, Kress & van Leeuwen (2006) stated that images could be treated as language, consisting of three meta-functions; ideational, interpersonal, and textual meta-functions (see also Hermawan & Sukyadi, 2017; Wu, 2014). More specifically, Martinec & Salway (2005) extend the theory (Halliday & Matthiessen, 2004; Kress & van Leeuwen, 2006) that text and image could have equal (image & text independence, image & text complementary) and unequal relation (image subordinate the text or vice versa). Additionally, the second function of image, to activate students' prior knowledge of the topic, is parallel to Zimmerman's (Gabinete, 2017) that visualization, resulting from images, could significantly develop students' skill in reading since it engages with the text mentally with the prior knowledge they possess. Furthermore, it has been obviously stated that interpretive process of readers in a text must be preceded by recognizing what they have been familiar with, and images will work for that (Serafini, 2012). Third, images are claimed as the medium to enhance students' interests. This is in conformity with (Duchak, 2014) stating that visual literacy could facilitate teachers to connect students in more exciting way.

Another finding indicates that most of teachers do not understand that image could create more meanings. Bearne (2009) stated that image could help verbal texts to achieve some purposes, such as explaining a complicated process. Further, it is emphasized that due to the limitation of every mode in conveying meanings, images could be an alternative to use in extending meanings in relation to 'cognition' and 'engagement (Bailey & van Harken, 2014). More than that, based on the interview result, the teachers said that they were not equipped with the knowledge and practical ways to teach multimodal texts. Therefore, the result show that they did not teach students explicitly about how to explore conceivable meanings on images. The teacher only centred on teaching meanings of written texts helped by the existence of images.

This is in keeping with the statement from (Lim-Fei & Yin, 2017) that the teachers may have technical challenges in teaching multimodal texts. Put differently, teachers had not been sufficiently knowledgeable to teach images as 'a language' in constructing a substance. Moreover, the answer of the respondents which is only limited to the statement that the picture is about 'Malin Kundang who hurt his mother' is also confirmed by the interview result to the focus participants. They only presumed that images as a complementary tool since the meanings are majority reflected in the written text (passage). This finding is not relevant to the theory (Jewitt, 2008) that the drawing (design) of image and the word relations give many impacts on the meaning-making process. This implies that sometime, images could not be replaceable by other modes; such as texts, since they are constructing major meanings. Therefore, a written text could be claimed not so important when an image has full or complete meanings to be understood by readers. Likewise, the teachers have not been familiar with the concept of image-text relation.

In more details, the participants did not understand the equality relation between image and text (Martinec & Salway, 2005) as well as the specific relation between those two images from (Royce, 2002). In addition to that, they clearly stated the classification of synonymy, antonymy, hyponymy, meronimy, and collocation they know is only in terms of the word relation. As has been mentioned before, what they have already known is that the existence of image is to attract students' attention or recall the background knowledge of a certain topic, rather than to form the meanings.

From multimodal landscape, there are some core aspects which can be discussed from an image. Seen from Systemic Functional Linguistic (SFL) particularly from transitivity analysis, there are three possible points; activities (process), the participants, and setting (circumstances). Seen from the process, the type of process occurring in the picture is a verbal process, shown by the gesture of Malin Kundang who is saying something to his mother while pointing his hand. In addition to this, in another side, the expression of his mother showing her

sadness, completely by her tears, indicate the behavioral process. From participants, it is clearly seen that the man is the actor and the woman is the target (goal). This could be seen from the position of the man who is doing an action, while the women is the object of the action. From circumstance, the picture is provided by the setting in the verbal texts; such as the boat, sea, and wave. The way to read images has also been explained by three analytical perspectives (Serafini, 2012) including perceptual, structural, and ideological analytical processes. Put differently, images are not sufficiently seen from the perspective of attractiveness and readability of readers, yet every element within them are also meaningful; color, size, position, and others.

Above all, it is clear to say that teachers crave to find out the techniques in teaching multimodal texts. This is hand in hand with the statement that in this 21st century, the teachers are suggested to guide students' critical literacy in understanding any types of text towards multimodal reading (Lim-Fei & Yin, 2017); one of the alternatives is Genre- Pedagogy. This type of pedagogy comes from Systemic Approach which guide the teachers to teach the material explicitly (Lim-Fei & Yin, 2017)Peculiarly, in this pedagogy, the teachers are required to teach generic structure of the texts (language, images, and others) as well as some shared multimodal schemes to engage readers (Lim-Fei & Yin, 2017). Therefore, basically, some teaching techniques in Genre Based Approach (Emilia, 2014; Gibbons, 2015) such as scaffolding and explicit teaching are necessary to undertake. More than that, there is a term namely SFMDA (Systemic-Functional Multimodal Discourse Analysis) by Jewitt, Bezemer, and O'Halloran which is developed from Systemic Functional Theory (Lim-Fei & Yin, 2017).

One of the concepts emphasized in this theory is meta-language; referred to the choice of using modes; language, sound, image, and others and the interaction among those modes. By understanding meta-language, students and teachers are helped to determine the choices of using modes by considering the aim and the effects in meaning-making process. In regard to this, some studies have found that technology is needed to alleviate the text production and consumption (Lim-Fei & Yin, 2017). In their study, students are guided to have multimodal literacy from the implementation of systemic approach (developed from Systemic Functional Linguistics), completed by the use of multimodal software analysis for annotation and the analytical tool. The main findings of that study revealed that both teachers and students believed that multimodal software analysis is beneficial for the multimodal literacy. In addition to this, it is safe to say that teachers basically need relevant particular programs, equipping them with the capacity to effectively use semiotic resources during teaching and learning process; such as, in-service trainings (Lim Fei *et al*., 2015), peer-tutoring, and others. More than that, due to the power of teachers who can build social relationship with their pupils, it is suggested to involve the competency of social and emotion for the teaching practice.

Put differently, there are numerous ways of teachers to explore meanings in the images while teaching texts with the help of proper teaching approaches, content selections, and other technological supports; such as any platforms, software analysis, and many more.

CONCLUSIONS

Due to the partiality of teachers' understanding of the image-text relation concept in text teaching which also impacts to their quality of teaching genre to their students, the teachers' multimodal competencies need to be upgraded. It means that the first thing to do before developing students' multimodal literacy, specifically, visual literacy, is that enhancing teachers' skills in reading visual images. Besides some relevant trainings (In-house training, workshop, and others), multimodal software could be an alternative media to use in helping that teachers' multimodal skill. Through this way; teachers will be more sensitive and critical in meaning the images as a meaningful unit rather than just an ornament. Therefore, it leads to implement the multimodal pedagogy in the classroom which also is a key to confront 21st century teaching in this era.

REFERENCES

Ajayi, L. (2011). A multiliteracies pedagogy: Exploring semiotic possibilities of a disney video in a third grade diverse classroom. *The Urban Review, 43*(3), 396–413. https://doi.org/10.1007/s11256-

010-0151-0.

Bailey, N. M., & van Harken, E. M. (2014). Visual images as tools of teacher inquiry. *Journal of Teacher Education*, *65*(3), 241–260. https://doi.org/10.1177/0022487113519130.

Bearne, E. (2009). Multimodality, literacy and texts: Developing a discourse. *Journal of Early Childhood Literacy, 9*(2), 156–187. https://doi.org/10.1177/1468798409105585.

Brailas, A. (2017). Digital storytelling in the classroom : How to tell students to tell a story. International *Journal of Teaching and Case Studies*, *8*(1), 16–28. https://doi.org/10.1504/IJTCS.2017.084407.

Çıralı, H., & Usluel, Y. K. (2015, 6-8 July). *A descriptive review study about digital storytelling in a descriptive review study about digital storytelling in educational context*. Paper presented at the 7th International Conference on Education and New Learning Technologies (EDULEARN15). Barcelona, Spain.

Da Costa, E. P. M., & de Barros, C. G. P. (2012). Multimodal genres in textbooks: Are students being schooled for visual literacy? *Bakhtiniana: Revista de Estudos do Discurso*, *7*(2), 38–56. https://doi.org/10.1590/S2176-45732012000200004.

Danielsson, K., & Selander, S. (2016). Reading multimodal texts for learning – a model for cultivating multimodal literacy. *Designs for Learning*, *8*(1), 25–36. https://doi.org/10.16993/dfl.72.

Duchak, O. (2014). Visual literacy in educational practice. *Czech-Polish Historical and Pedagogical Journal*, *6*(2), 41–48. https://doi.org/10.2478/cphpj-2014-0017.

Emilia, E. (2014). *Introducing functional grammar*. Bandung: Pustaka Jaya.

Gabinete, M. K. L. (2017). Teachers' beliefs and practices in assessing the viewing skill of ESL learners. *Indonesian Journal of Applied Linguistics*, *7*(1), 19–28. https://doi.org/10.17509/ijal.v7i1.6854.

Gibbons, P. (2015). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom*. Portsmouth, NH: Heinemann.

Halliday, M. A. K., & Matthiessen, C. (2004). *An introduction to functional grammar*. New York, NY: Hodder Arnold.

Hermawan, B., & Rahyono, F. X. (2019). Ideational meanings of science and interpersonal position of readers in science textbooks for basic level in Indonesia. *Indonesian Journal of Applied Linguistics*, *9*(1), 38–47. https://doi.org/10.17509/ijal.v9i1.15932.

Hermawan, B., & Sukyadi, D. (2017). Ideational and interpersonal meanings of children narratives in Indonesian picturebooks. *Indonesian Journal of Applied Linguistics*, *7*(2), 404–412. https://doi.org/10.17509/ijal.v7i2.8139.

Howell, E., Butler, T., & Reinking, D. (2017). Integrating multimodal arguments into high school writing instruction. *Journal of Literacy Research*, *49*(2), 181–209. https://doi.org/10.1177/1086296X17700456.

Hundley, M., & Holbrook, T. (2013). Set in stone or set in motion: Multimodal and digital writing with preservice English teachers. Journal of Adolescent & Adult Literacy, 56(6), 500-509. https://doi.org/10.1002/JAAL.171.

Jacobs, G. E. (2006). Developing multimodal academic literacies among college freshmen. *Journal of Media Literacy Education*, *4*(3), 244–255. https://digitalcommons.uri.edu/jmle/vol4/iss3/5/.

Jaksic, V. (2017). *Teachers' perception of the application of gestures in TEFL classroom* (Thesis, University of Zagreb). http://darhiv.ffzg.unizg.hr/id/eprint/9016.

Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, *32*(1), 241–267. https://doi.org/10.3102/0091732X07310586.

Kedra, J., & Zakeviciute, R. (2019). Visual literacy practices in higher education: What, why, and how? Journal of Visual Literacy, 38(1-2), 1-7. https://doi.org/10.1080/1051144X.2019.1580438.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. New York, NY: Routledge.

Lim-Fei, V., & Yin, S. T. K. (2017). Multimodal translational research: Teaching visual texts. In O. Seizov & J. Wildfeuer (Eds.). *New Studies in Multimodality: Conceptual and Methodological Elaborations*. London: Bloomsbury Academic, pp. 175–199. https://doi.org/10.5040/9781350026544.0014.

Lim-Fei, V., O'Halloran, K. L., Tan, S., & E, M. K. L. (2015). Teaching visual texts with the multimodal analysis software. *Educational Technology Research and Development*, 63(6), 915–935. https://doi.org/10.1007/s11423-015-9395-4.

Liu, X., & Qu, D. (2014). Exploring the multimodality of EFL textbooks for Chinese college students: A comparative study. *RELC Journal*, 45(2), 135–150. https://doi.org/10.1177/0033688214533865.

Lundy, A. D., & Stephen, A. E. (2015). Beyond the literal: Teaching visual literacy in the 21st century classroom. Procedia - Social and Behavioral Sciences, 174, 1057-1060. https://doi.org/10.1016/j.sbspro.2015.01.794.

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.

McKee, L. L., & Heydon, R. M. (2015). Orchestrating literacies: Print literacy learning opportunities within multimodal intergenerational ensembles. *Journal of Early Childhood Literacy*, 15(2), 227–255. https://doi.org/10.1177/1468798414533562.

Murcia, K. (2014). Interactive and multimodal pedagogy: A case study of how teachers and students use interactive whiteboard technology in primary science. *Australian Journal of Education*, 58(1), 74–88. https://doi.org/10.1177/0004944113517834.

Rakhmawati, S. S. (2016). The role of moving images in the conference interpreting classroom. *Indonesian Journal of Applied Linguistics*, 5(2), 262–271. https://doi.org/10.17509/ijal.v5i2.1350.

Robin, B. (2006). *The educational uses of digital storytelling*. In C. Crawford, R. Carlsen, K. McFerrin, J. Price, R. Weber & D. Willis (Eds.), Proceedings of *SITE 2006--Society for Information Technology & Teacher Education International Conference*. Orlando, Florida: Association for the Advancement of Computing in Education (AACE), 709-716. https://www.learntechlib.org/primary/p/22129/.

Royce, T. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Quarterly*, 36(2), 191–205. https://doi.org/10.2307/3588330.

Serafini, F. (2012). Expanding the four resources model: Reading visual and multi-modal texts. *Pedagogies*, 7(2), 150–164. https://doi.org/10.1080/1554480X.2012.656347.

Sewell, W. C., & Denton, S. (2015). Multimodal literacies in the secondary English. *The English Journal, 100*(5), 61–65. http://www.jstor.org/stable/23047804.

Shanahan, L. E. (2013). Composing "kid-friendly" multimodal text: When conversations, instruction, and signs come together. *Written Communication*, 30(2), 194–227. https://doi.org/10.1177/0741088313480328.

Tang, K. S., & Moje, E. B. (2010). Relating multimodal representations to the literacies of science. *Research in Science Education*, 40(1), 81–85. https://doi.org/10.1007/s11165-009-9158-5.

Thomas, J. R, Nelson, J. K, & Silverman, S. (2005). *Research methods in physical activity*. Canada: Human Kinetics.

Tomsett, P. M., & Trott, D. C. (2014). Multimodal pedagogy in international business. *International Journal of Research in Social Sciences*, *4*(2), 50–56. http://www.ijsk.org/volume-4-issue-2/.

Twiner, A., Coffin, C., Littleton, K., & Whitelock, D. (2010). Multimodality, orchestration and participation in the context of classroom use of the interactive whiteboard: A discussion. *Technology, Pedagogy and Education*, *19*(2), 211–223. https://doi.org/10.1080/1475939X.2010.491232.

Vungthong, S., Djonov, E., & Torr, J. (2017). Images as a resource for supporting vocabulary learning: A multimodal analysis of Thai EFL tablet apps for primary school children. *TESOL Quarterly*, *51*(1), 32–58. https://doi.org/10.1002/tesq.274.

Wang, S., & Zhan, H. (2010). Enhancing teaching and learning with digital storytelling. *International Journal of Information and Communication Technology Education*, *6*(2), 76–87. https://doi.org/10.4018/jicte.2010040107.

Wu, S. (2014). A multimodal analysis of image-text relations in picture books. *Theory and Practice in Language Studies*, *4*(7), 1415–1420. https://doi.org/10.4304/tpls.4.7.1415-1420.

Yus, F. (2006). Literacy in the new media age (review). *Language, 82*(4), 953-954. https://doi.org/10.1353/lan.2006.0242.