

## Understanding the roles of images and intermodal relationships for optimized use of visual and verbal resources in Vietnam's textbooks for lower secondary levels

### Memahami peran gambar dan hubungan intermodal untuk optimalisasi penggunaan sumber visual dan verbal dalam buku teks bahasa Inggris untuk siswa sekolah menengah di Vietnam

Yusnita Febrianti <sup>a</sup> \*  , Thao Vu <sup>b</sup>  

<sup>a</sup> Universitas Negeri Malang, Indonesia

<sup>b</sup>Department of Foreign Languages, Haiduong College, Vietnam

Submitted: September 12, 2022; Accepted: February 22, 2023; Published: February 27, 2023

#### KEYWORDS

images,  
 language,  
 English text-  
 books,  
 language-image  
 relationship/  
 interaction

#### ABSTRACT

The study investigates the roles of images and intermodal relationships of both language and images in two English textbooks used in Vietnamese lower secondary schools, namely Tieng Anh and Solutions textbooks to gain insights on the ways to optimise the use of both language and images as resources in teaching and learning activities. Kress and Van Leeuwen's (1996, 2006) framework on the grammar of visual design is used to analyse the images, looking at the types of image representations and the compositional meanings. The language-image relationships in the books, on the other hand, are analysed using the framework of Intermodal Identification (Unsworth & Cleirigh, 2014). Intermodal identification is built on the notion that language complements the meaning of the image and vice versa. Language identifies image by glossing the image participants which are not encoded in the language elements. Image identifies language in three aspects: intensive to visualise quality such as shape, colour, or texture, possessive to visualise additional participants which are not explicitly addressed in the language and circumstantial to visualise the elements of locations in the language. While image in textbooks has always been considered essential as a source of teaching and learning materials as well as helping students to learn, this study suggests that the role of image is augmented when juxtaposed with the accompanying language. The study implies the need for further investigation, for example in the classroom action research on how language and image resources can be utilised in teaching and learning activities. Also, the result of the study may be replicated to analyse language-image relationship in different samples of textbooks.

#### KATA KUNCI

bahasa,  
 gambar,  
 buku teks  
 bahasa Inggris,  
 hubungan/  
 interaksi  
 bahasa-  
 gambar.

#### ABSTRAK

Penelitian ini menyelidiki peran gambar dan hubungan antar moda dari kedua bahasa dan gambar dalam dua buku teks bahasa Inggris yang digunakan di sekolah menengah pertama Vietnam, yaitu buku teks Tieng Anh dan Solutions untuk mendapatkan wawasan tentang cara mengoptimalkan penggunaan bahasa dan gambar sebagai sumber daya di kegiatan belajar mengajar. Kerangka kerja dari Kress dan Van Leeuwen (1996, 2006) tentang tata bahasa desain visual digunakan untuk menganalisis gambar, melihat jenis representasi gambar dan makna komposisi. Kemudian, hubungan bahasa-gambar dalam buku dianalisis menggunakan kerangka Identifikasi Intermodal (Unsworth & Cleirigh, 2014). Identifikasi intermodal dibangun di atas gagasan bahwa bahasa melengkapi makna gambar dan sebaliknya. Bahasa mengidentifikasi gambar dengan mengilapkan peserta gambar yang tidak dikodekan dalam elemen bahasa. Gambar mengidentifikasi bahasa dalam tiga aspek: intensif untuk memvisualisasikan kualitas seperti bentuk, warna, atau tekstur, posesif untuk memvisualisasikan peserta tambahan yang tidak secara eksplisit dibahas dalam bahasa dan sirkumstansial untuk memvisualisasikan elemen lokasi dalam bahasa. Sementara gambar

\* Corresponding author: [yusnita.febrianti.fs@um.ac.id](mailto:yusnita.febrianti.fs@um.ac.id)

---

dalam buku teks selalu dianggap penting sebagai sumber bahan ajar dan pembelajaran serta membantu siswa untuk belajar, penelitian ini menunjukkan bahwa peran gambar bertambah ketika disandingkan dengan bahasa yang menyertainya. Penelitian ini menunjukkan pentingnya penelitian lanjutan, misalnya dalam bentuk penelitian tindakan kelas untuk mengetahui bagaimana bahasa dan gambar dapat dimanfaatkan dalam kegiatan belajar mengajar. Juga, hasil penelitian ini dapat direplika untuk menganalisis hubungan bahasa-gambar dalam sampel buku teks yang berbeda.

---

**How to cite this article:**

Febrianti, Y. & Vu, T. (2023). Understanding the roles of images and intermodal relationships for optimized use of visual and verbal resources in Vietnam's textbooks for lower secondary levels. *Bahasa dan Seni: Jurnal Bahasa, Sastra, Seni, dan Pengajarannya*, 51(1), 54–72. <https://dx.doi.org/10.17977/um015v51i12023p54>

---

## Introduction

Language is a fundamental resource of making meaning (Mickan, 2006); however, in very basis we encounter a large number of texts with the incorporation of different semiotic modes, such as image. Research on meaning arising from the increasing use of multiple semiotic modes in a range of written, printed and electronic texts has become an exciting area of investigation. A number of studies on the theory and practice of analysing the meaning potential conveyed by images and their interaction with language have been conducted (Kress & Van Leeuwen, 1996, 2006; Lemke, 1998; O'Halloran, 2008; Royce, 1998). Although the image and language semiotic systems utilise meaning-making resources in their own particular ways, they not only share the semantic load between each other, but also multiply the overall meaning of the text (Painter et al., 2013; Royce, 2002). This multimodal reality has recently become a significant focus of literacy education research and strongly influenced school curricula (Kress, 2000; Unsworth, 2001, 2006, 2014). Unsworth and Chan (2009) argued that it is now no longer adequate to consider that literacy development simply involves the possessing of language information alone, and students need to be equipped with the ability to negotiate meanings across semiotic modes. In the domain of English language teaching and learning, Unsworth and Ngo (2015) indicated that the integration of image and language reflects the multimodal nature of a wide range of authentic texts in English; therefore, for language learners in contexts of learning English as a first, second or foreign language, learning to construct meaning from the interaction of visual and verbal language should be seen as a crucial aspect of language competence.

The inclusion of multimodal resources that involve image and language semiotic modes in school textbooks has also attracted scholarly attention for a long time. Levin and Mayer (1993) explained the significance of images in the classroom as they can make texts more concentrated, concrete, coherent and comprehensible; and thus, facilitating learning. Peeck (1993) advocated the role of images in assisting the learning process, including increasing motivation, focusing attention, clarifying text contents and facilitating dual-coding processing. Carney and Levin (2002) reviewed a number of studies on the roles of pictorial illustrations in students' learning and concluded that carefully constructed images in combination with texts generally enhanced learner performance on a variety of text-dependent cognitive outcomes. Given the role of images which have been long supported to aid learning, in the tradition of language textbook development, more serious attempts have been devoted to make textbooks a richer resource for classroom teaching and learning by including attractively illustrated texts with images. However, images in textbooks for teaching English as a foreign language

(EFL) are far less investigated compared to textbooks for other school subjects. Studies on the multimodal features of EFL textbooks include Hsu and Yang (2007) who explored the effect of image and language integration on reading comprehension; Chen (2010) who identified and analysed types of multimodal resources that enabled dialogic engagement in textbooks in China, and Unsworth and Ngo (2014) who examined a number of images in EFL textbooks in Vietnam that informed changing textbook design to better support learning. Other studies on EFL textbook evaluation by Mukundan and Nimehchisalem (2012), Mukundan and Kalajahi (2013), Chegeni, et al. (2016) used evaluation checklists to evaluate textbooks. Although the visual is included as a category of evaluation in the checklist, there have been lack of a detailed and principled analysis of the visuals and its relation to the verbal texts as a significant aspect of textbook considerations.

The current study investigates the roles of images and intermodal relationships of both language and images in two English textbooks used in Vietnamese lower secondary schools, namely Tieng Anh and Solutions textbooks to gain insights on the ways to optimise the use of both language and images as resources in teaching and learning activities. One of the rationales in the selection of the sample textbooks is the unique implementation, which is the combination of locally produced and imported textbooks for the English teaching at the secondary schools in Vietnam. The initial hypothesis for our argument is built based on the observation that the meaning potentials in language-image interaction multiplies when they are juxtaposed on the same page. Therefore, they are potential to provide resources for teaching and learning activities. The analyses are facilitated using the principles of grammar in visual designs by Kress and Van Leeuwen (1996, 2006) and Intermodal Identification by Unsworth and Cleirigh (2014).

### *Theoretical frameworks*

The main analytical tool used in the study is based on the work of Kress and Van Leeuwen (1996, 2006) on analyzing the grammar of visual designs. They offer a framework to understand the structure of meanings in different types of texts that comprise mostly images or a mix of image and language resources. Their concept of visual grammar was inspired by Systemic Functional Linguistics (SFL), reflecting the realization of a simultaneous threefold purpose in every text: the representation of the outside world, the relationship of the communicating participants, and the pertinent organization in the communication method (Painter et al., 2013, pp. 6-7). As a result, the framework for visual grammar is laid up in such a way that it unfolds across the three metafunctions of representational, interpersonal, and compositional meanings.

Representational meanings refer to how images represent reality. This metafunction is realised in both narrative and non-narrative types of images. The narrative image involves the use of various elements in the images to tell an action and reaction by means of a vector, a directional line between the actors in the images to connect the participants and goals. Another type of narrative image involves the use of speech or thought bubbles to represent verbal and mental events respectively. Non-narrative images are also known as conceptual images, depicting the properties in the images either by classificatory or analytical relationship between elements in the images.

The interpersonal or interactive metafunction in images establishes two kinds of relationships: relationship between the participants within the images and relationship between the actors in the image and the readers or audience of the text. These relationships are realised by gaze. Gaze between participants within the image puts the readers as an observer of the images. When an image of an actor is made with a gaze directed towards an audience, a pseudo interpersonal relationship is created between the image and the viewer. Another important dimension of this metafunction is the establishment of social distance. Social distance is created by the sizes of image (either framed or not) hence determines how viewers perceive the images. Images may be displayed in close up, medium shot and long shot to indicate closeness or remoteness with the audience. Finally, the horizontal and vertical perspective of images relates to attitudinal aspects. Horizontal perspective hints involvement by showing images ‘in front of’ the audience and detachment when the images are shown obliquely. From the vertical point of view, images create some sense of power. Images with low angle look down on the audience, a vertical angle creates a sense of relative equality between image and audience, and high-angled images position audience in a more powerful place than the actors in the image.

The arrangement of images on a medium of delivery is termed compositional meanings. Composition or layout of images on a page facilitates the organization of both the representational and interpersonal metafunctions. One of the main aspect in compositional meaning is realised by the placement of elements on a page to express information value, consisting of left-right, top-bottom and centre-margin to indicate given and new information respectively. An element of composition such as framing indicates the connection or disconnection of images to the rest of the page contents. Strong framing, for example bold borders, disconnects an image from the other elements on the page. Weak framing, on the other hand, allows an image to blend with the other elements in the page. Other aspects of composition such as colour and size determines the prominence and salience of images on a page.

The significance of the intermodal relationship between language and image has been explicated in Unsworth and Cleirigh (2014). They argue that language and image have natural synergistic relations that play significant role in constructing meanings (p. 177). They outline a framework called Intermodal Identification. Developed from the grammatical concept of relational grammatical identification from SFL, the Intermodal Identification recognizes two main relationships. Basing on the departure point of the identifier, the framework views the relations between visual and verbal from two perspectives. From the perspective of the verbal elements, words can be used to decode or gloss images with suitable lexis in association with the shared meaning. From the point of view of the visual, images can be used to visualize the verbal in three ways. First, intensive identification is a relation in which image visualises unverbaised qualities such as shape, colour, texture of the identified participants. Secondly, in intermodal possessive relationship, images visualise additional participants that are not explicitly mentioned in the verbal elements. Finally, should there be elements of locations that are not mentioned in the verbal resources, the circumstantial identification recognizes the visualisation of this aspects in the images.

*The new English textbooks for Vietnamese students*

Again, the rationale of the sample textbook selection is based upon Vietnam's unique implementation of the combined locally produced and imported English textbooks for the secondary school students. Tieng Anh, a new series of English textbooks, was created as a part of Project 2020, a nationwide effort to promote foreign languages. The series was created with the overall aim "to renovate thoroughly the tasks of teaching and learning foreign languages within the national education system." It is expected that school graduates "will be able to use a foreign language confidently in their daily life, study and work in a multicultural and multilingual environment, making foreign languages a comparative advantage of development for Vietnamese people" (Prime Minister of Vietnam, 2008). This project is regarded as the most ambitious language endeavor in Vietnam's educational history, with a budget of 9,378 billion VND (426 million USD) (Le, 2015, Bui & Nguyen, 2016). The project shows a significant commitment by the government to improving English language proficiency in the country's educational system, which is important given the low per-capita GDP and a lot of unresolved educational difficulties. The project would help an estimated 22 million English language learners in Vietnam, according to estimates from 2013 (MOET 2014).

Teams of textbook authors from the Ministry of Education and Training (MOET) and Pearson, an American textbook publisher, worked together to create the Tieng Anh series. With "richly illustrated, cross-curricular, and theme-based units, concentrating on delivering student motivation, memorable lessons, and a pleasurable learning experience," it adheres to a methodical, cyclical, and topic-based syllabus (Hoang et al., 2015a, p. iii). It was ensured that communicative language teaching (CLT) and participatory learning are supported in the classroom is one of the guiding concepts for creating the textbooks (Hoang et al., 2015a). This curriculum seeks to replace old pedagogies that treated texts and settings separately from grammar and lexis by encouraging students to become more proficient communicators through a variety of communicative exercises.

The new Tieng Anh series for the lower secondary level consists of Tieng Anh 6, Tieng Anh 7, Tieng Anh 8, and Tieng Anh 9, with Volume 1 for the first semester and Volume 2 for the second semester in each book. At the lower secondary level, 420 lessons lasting 45 minutes each are allotted for English language instruction. Each book has six components built around four major themes of Our Communities, Our Heritage, Our World, and Visions of the Future. Each unit begins with a dialogue, which incorporates topic-related phonological and lexical items, specific functions and notions, and the grammatical structures which express them.

Prior to the teaching of the four macro skills, the language parts, functions, and concepts are practiced and then expanded in the following sections (reading, speaking, listening, and writing). The final project of the unit should provide students with opportunity to apply the language and abilities they have acquired to carry out communicative activities either independently or in groups (Hoang et al., 2015b). The pilot curriculum and texts have been used in several lower secondary schools since 2012 with great success. However, there haven't been any empirical studies investigating how these new textbook series are actually used in the classroom.

## Method

Data for this investigation is drawn from two English textbooks used in the lower secondary schools in Vietnam: Tieng Anh 6 (henceforth TA6) (Hoang et al., 2015b), a newly piloted textbook within the new communicative-based curriculum initiative, and Solutions Pre-intermediate (henceforth SP) (Falla & Davies, 2013). While TA6 is used as a mandated textbook at schools, SP is sometimes used by teachers as a supplementary material. SP is one of a five-level English course series for teenagers published by Oxford University Press. The book is claimed to employ a supportive approach to teaching speaking and writing, provide achievable activities and consistent practice. This is a commercialized textbook aiming for aimed at any learners of English as second/foreign language.

In the present study, the selection of these two books as the source of data is not for the purpose of textbook comparison; rather, the study aims to collect rich samples of images from both books in order to provide comprehensible analysis and discussion drawing on the meaning potential conveyed by images and their relations with texts. In accordance to the aim of the study, Unit 6 from TA6 (Sports and Games) and Unit 2 from SP (Winning and losing) are selected to exemplify the analysis and discussion. These units are developed within a similar topic of sports. The selection of the units as the data for analysis is based on the following rationales. First, the unit contains visual resources produced from a combination of animated drawings and photographs which gives a balanced perspective in analyzing the images. Secondly, the unit is seen to have reasonably balanced comparison between image and language resources at every page. Finally, the topic of the units is considered a neutral topic. This type of topic would be commonly found in any English textbooks in an Asian context, which in turn will serve as an acceptable rationale if the results of the study are to be generalized to assess English textbooks in different contexts of use.

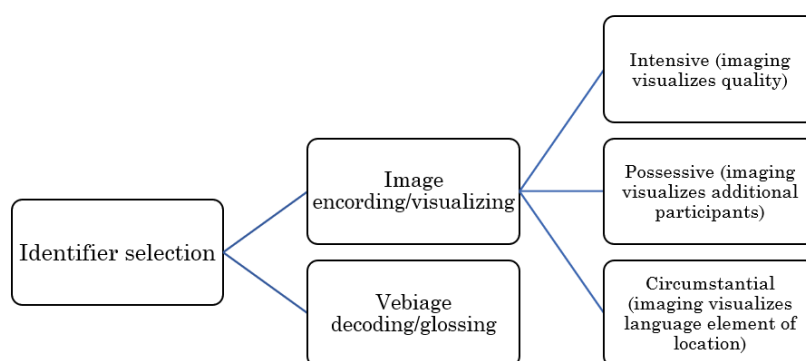


Figure 1. Intermodal Identification framework (Unsworth & Cleirigh, 2014)

The main analytical methodology to analyze the images is Kress and Van Leeuwen's (1996) Grammar of Visual Design, looking at two analytical aspects of an image namely the compositional and representational meanings. Each page of the units is assessed based on the textual compositions by looking at the organization of information value and framing. The information value refers to the positioning of image resources and language resources on the page that will determine the salience of information on each page. Framing of any image resources on the page will determine its relations to the

rest of the page content. Detailed accounts of the image resources will be analyzed by looking at the narrative and/or conceptual representations i.e., how the elements of images construct meanings through the vectors or absence of vectors. Analysis of the Intermodal Relationship (Unsworth & Cleirigh, 2014) between the image and language in the textbooks is conducted based on the system in the Figure 1.

## Results and discussion

Both selected units in TA6 and SP are laid out in a similar number of 10 pages. A general observation of the units shows that the proportion of images' occupation ranges from 25% to 50% of the page space. The images included in the units are analyzed and discussed based on their roles in serving as an individual resource facilitating the learning activities, and their interaction with the verbal elements in the task instructions and the content of the verbal texts. The analysis results show that while some images in the units are potentially able to fulfil the instructional functions and multiply the meaning potential of the text to aid learning, a few drawbacks are found, mostly due to the irrelevant relationship between the selected visual and the written wordings.

### *Images for topic elicitation*

The opening image in TA6, bearing the function of introducing the topic of the whole unit, occupies the whole of page 16. Presented in framed boxes, the verbal elements are overlaid on the foreground of the image occupying a half section on the page on the right side. This creates visibility of image on the left-hand side of the page. In this arrangement of information value, following the common western left-to-right reading path (Kress & Van Leeuwen, 1996, 2006, p. 177), the image becomes the departure point of 'seeing' before 'reading' the verbal content.

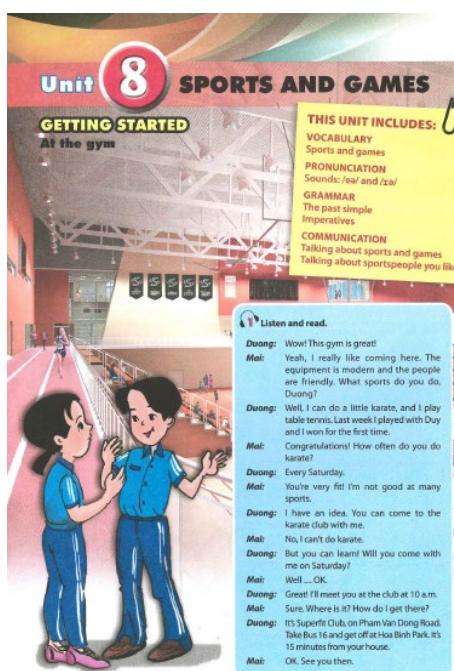


Figure 2. Opening of the unit in TA6

The overall image in this opening page features a combination of photographic image of an indoor sporting hall in the background and animated drawings of the narrative

representation of the two characters in the foreground. The language-image cohesion in this opening page can be seen from two relationships: between the title of the unit and the overall image and between the dialog and the overall image. The title of the unit, Sports and Games, intermodal circumstantial relation can be seen in the visualisation of a location related to the lexis sports and games i.e. the image of the interior part of a sporting hall and people doing activities in it. Intermodal intensive identification is seen in the depiction of the two characters shown wearing their sport uniforms in bright colours.

While a general look at the image seems to suggest an appropriate introduction of the topic, a closer examination reveals some inconsistencies. Firstly, the image of a photographic character running towards the animated characters creates discrepancy in the whole coherence of the image. The depiction of image of a running track and an athlete running on the track are seen to be situated in the second storey of the sporting hall, which seems to be a made-up situation that may not be relatable to an actual gym. The standing position of the two animated characters appears to be unclear to the viewer as it is not observable whether they are standing on the running track or on an unidentified position on the second floor of the gym.

Secondly, a focused analysis from the intermodal perspective between the two animated characters and the written dialog uncovers some mismatches. The depiction of the animated characters can be seen as the visualisations as formulated in the intermodal possessive identification system in which the two secondary school students involved in the dialog named Duong and Mai are shown. The vectors created by the gaze towards each other and hand movements suggest that they are talking to each other. Despite the salience of the image, there seems to be a limited relationship between the image and the language content. The only connection may be only lies in the lexis *This* in the first line, namely *This gym is great* referring to the image of the gym depicted in the image. The next line showing *The equipment is modern and the people are friendly* is not supported with any visualization in the image. This limited relationship would mean that the overall image is no more than a mere decoration. The close examination shows that having an oversize image on a page for the topic introduction does not guarantee a useful relevance to all details involved on the page.

SP, on the other hand, opens the unit by using the top-bottom composition. Instead of using one oversize image which occupies a considerable space, SP puts a series of image of sporting characters at the top half of the page. The use of frames wraps each character, distinguishing it from each other; but the close adjacency creates a cluster of image, showing that these images are bound by a shared topic. Each image within the frame displays a sport person with sufficient visualization of the sport character depicted in the outfit, equipment and background setting of the images. While there is a total of four different images, they relate to a single verbal element to visualize the question: *Can you name these sportspeople and their sports?*



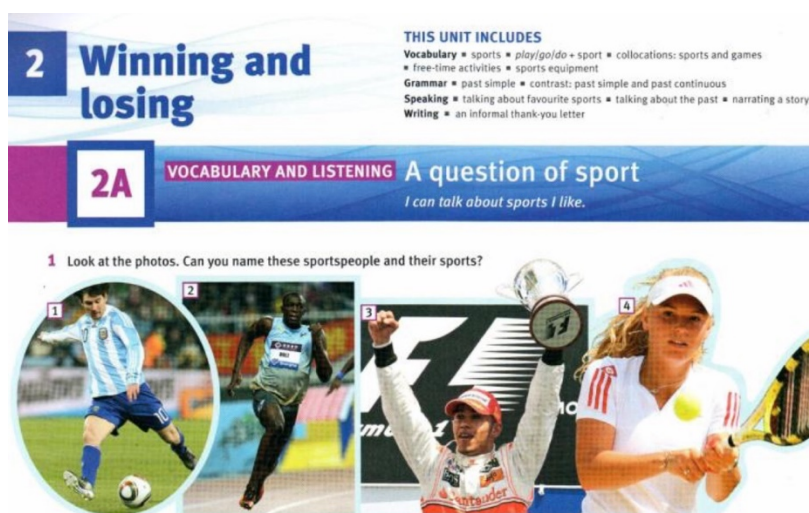


Figure 3. Opening of the unit in SP

Minimum number of words to relate to the images means minimum risk of irrelevance between the image and the wordings. However, minimal use of visual and verbal resources in SP opening section is still quite provocative for stimulating interest in the introduced topic. Using authentic images such as these also stimulates students' background knowledge of the topic to support the learning.

The variants of images at the opening of the unit, as observed from TA6 and SP, are considerable in terms of size and content. Either being an individual oversized image, such as in TA6, or a cluster of small images occupying considerable space on the page (SP), opening images play an important role in leading the lesson into the main topic that becomes the core idea that develops into every task within the unit. Images in this opening section, therefore, provide an important foundation for the students to build background knowledge before moving on to the other tasks in the unit.

### *Images for vocabulary building*

In the case of ESL/EFL teaching and learning, vocabulary building activities are considered important in providing students with a lexical input to progress through the subsequent tasks. Vocabulary building, therefore, is often found at the beginning sections in textbooks right after the opening. Among the many types of vocabulary building activities, using images glossed with suitable verbal elements has been observed as commonly found activities in English textbooks. Unsworth and Cleirigh (2014) describe this relationship in their intermodal system as a relationship where a language element identifies an image participant (p. 184). In both TA6 and SP, vocabulary building activities are presented using a set of both conceptual and narrative images related to the topic of sports i.e. by displaying the equipment of sport to relate to the types of sports in question.


In TA6, the exercise of vocabulary building is given in two tasks. In the first task, a list of Vocabulary is given (Exercise 2) before the images are presented. Narrative images for question numbers 1, 2, 4, 5, and 6 show participants in the settings with outfits and equipment relevant to various kinds of sporting activities as provided in the glossary. A meticulous look at image for question number 3, however, reveals an irrelevance. The participants in this particular image are depicted to be running, but their

outfits do not reflect the set up for a sporting activity. This selection implies a careless association between the provided glossaries of running with just any available image of people running instead of connecting with the set topic of sport. The second task for the vocabulary building utilises conceptual images in which the equipment of sports is individually framed to prepare for the next task of relating the equipment with the kind of sports. The images in this task mostly presented in photographic style, with the exception of the image for number 2 which is an animated drawing.

**2 Listen and repeat these words and phrases.**

1. boxing	2. fishing	3. aerobics
4. chess	5. table tennis	6. karate
7. cycling	8. swimming	9. volleyball
10. tennis	11. skiing	12. running

**3 Using the words in 2, name these sports and games.**



1. \_\_\_\_\_

2. \_\_\_\_\_

3. \_\_\_\_\_

4. \_\_\_\_\_

5. \_\_\_\_\_

6. \_\_\_\_\_

Figure 4. Vocabulary task 1 in TA6

**2 Now write the words under the pictures.**



1. \_\_\_\_\_

2. \_\_\_\_\_

3. \_\_\_\_\_

4. \_\_\_\_\_

5. \_\_\_\_\_

6. \_\_\_\_\_

7. \_\_\_\_\_

8. \_\_\_\_\_

Figure 5. Vocabulary task 2 in TA6

*The inconsistent style is acceptable as the images do not deviate from the intended purpose of being glossed with the correct wordings.*

A similar activity is found in SP, which requires matching the wordings with the given image. A listening activity is followed in order to check the answer and help with the pronunciation. While both TA6 and SP use similar types of image selection and presentation for the vocabulary building, SP seems to employ a slightly better strategy in the presentation. Featuring animated drawings, SP uses consistent style in the images, albeit using both conceptual and narrative images, making use different shapes in various colours as the background of each image to represent a particular kind of sport. The sizing of the images is also cleverly managed; hence yielding 18 representative images occupying about 20% of the page space. This is in contrast to the image for the second task in TA6 which displays 8 images occupying almost 50% of the page space. As a result, effective use of the page space is achieved with a better strategy for space exploitation in SP. While TA6 has to provide vocabulary building exercises in two tasks, it yields the same number of vocabulary that can be achieved in one task in SP.

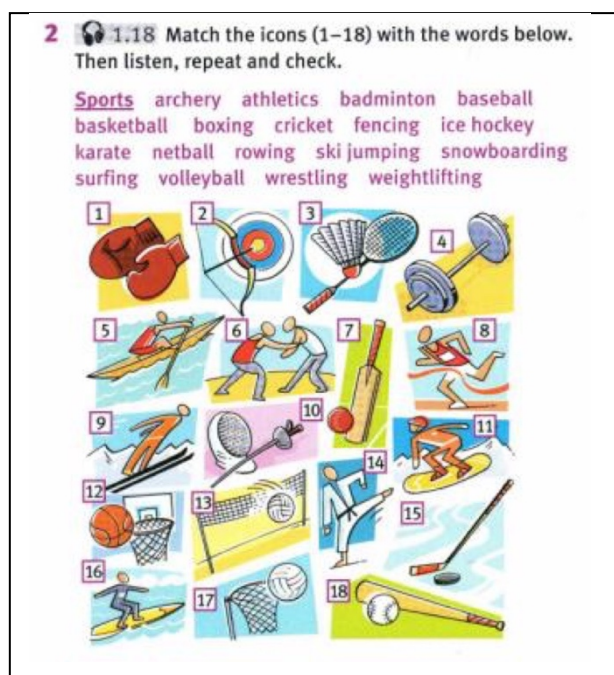


Figure 6. Vocabulary task in SP

It is important to note that in the exercises for vocabulary building in both TA6 and SP, the role of framing is essential. In TA6, for example, each image is segregated by means of a clear framing of thin blue line. This type of framing confines the image item within the box. In SP, although no clear lines are used for framing, each image item is separated by means of colourful shapes that simultaneously become a background for each image. White space between images is also used for segregating one image from the other, these framing strategies lead to exact representation of a single wordings to gloss each of the image.

#### *Images for reading comprehension*

In the data, images that accompany a reading passage seem to be ‘decorating’ the written verbal elements. However, Intermodal identification system recognizes that language-image relationship is naturally mutual, including in various reading texts where images correspond to main text segments (Unsworth & Cleirigh, 2014, p.184). A close examination of images surrounding a reading passage can be seen from the composition, types of image and how the images relate to the wordings.

In TA6, for example, the reading text is presented in a frame of a green box. The main image of the reading passage is positioned on the top left and bottom right corner, allowing the written verbals to flow from top to bottom. The reading path, therefore, flows from the look at the image of the character of the image, Pele, on to the reading text from the top to bottom, ending with the look at another image of Pele at the bottom right corner. The significance of the selected image to accompany the text can be detailed by looking at the narrative representation of the images. The image of Pele is shown to raise his fist up in the air, with his gaze directed towards the viewer. As another character in the image looks away, there is no interaction between Pele and the other character. Instead, a ‘pseudo’ interaction is created between Pele and the readers of the book. This is how the image interacts to the students, facilitating a sense of

engagement with the set of the meaning resources in the reading text. As the image at the bottom right corner does not interact with the audience, i.e. the gaze looks away from the audience, the image serves as an image to strengthen the character being described in the reading text.

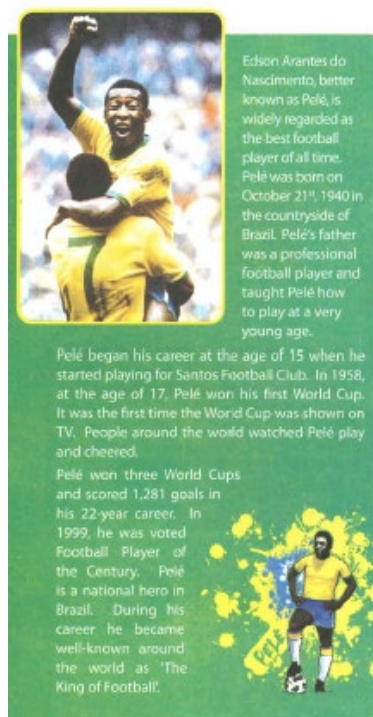


Figure 7. Reading text in TA6

The visualisation of the character described in the story, Edson Arantes or Pele, strengthens the meaning that is intended to be understood by the students through the reading text. Human character is obviously understood better through visual characterisation than a mere verbiage description. Other qualities of this character is visualised in the choice of image of the character Pele wearing a football uniform, which gives much better idea when related to the verbiage the best football player of all time in the first paragraph of the reading text. The colour yellow attributed in his uniform serves as an extra information for the student about his nationality mentioned in the reading text, Brazil, and yellow being the colour of Brazil's national soccer team. Since the reading text discusses about Pele's achievement in his sporting career, the verbiage is expressed in the lexis such as best football player, won, people cheered, scored, and well-known. All these positive qualities are summarised and transferred into the image by depicting Pele in with a cheerful facial expression (top image) and in a proud, open shoulder standing pose (bottom image).

In much the same way, SP uses similar strategy of image-language combination in its reading text. Overall, the image cleverly blends with the background of the reading text, fusing the wordings and the image elements in the reading text layout. The main image is seen at the bottom right corner of the layout, depicting a character surfing. The use of indirect gaze from the character's image detached the interaction between the image and the audience. Audience, or readers, are left as observers whose understanding of the image relies on the understanding of the reading text. One main

immediate relation can be drawn between the title of the text, surfing superstar, and the main image. This relation essentially introduces the content of the reading text. A browse on the first line of the text informs the reader of the name of the character in the image, Bethany Hamilton. A few other wordings from the reading text such as the sky was clear, the sun was shining, wave and board are visualised in the image, with the first two lexes are especially helped with the set of lighting used in the image. The rest of the lexes in the text are carried through to the understanding of the text and how these are set out to be as background knowledge to go through the next task

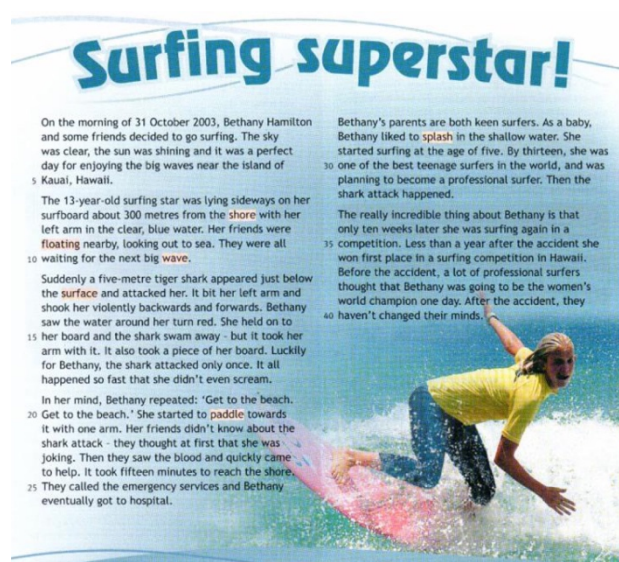


Figure 8. Reading text in SP

While providing the image for the climax of the story in the reading text: It bit her left arm and shook her violently backwards and forwards may be too graphic for a school textbook, the aftermath is presented in a more positive depiction of the character on the next page.

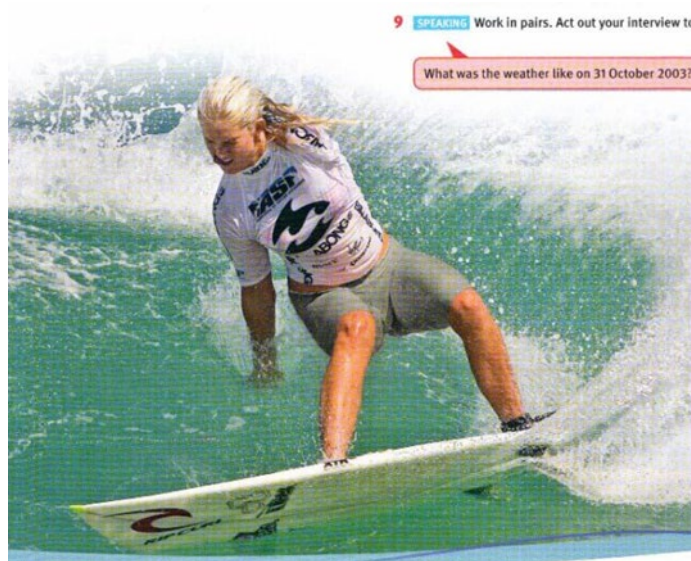


Figure 9. Second visual in reading text in SP

Shown with only one arm, the image visualizes the story in the third paragraph of the passage. Because the reading text says: Suddenly, a five-metre tiger shark appeared

just below the surface and attacked her, it bit her left arm and shook her violently backwards and forwards, this could have been a graphic image when included as a visual in the school textbook. Instead, this incident is visualised using the image as shown in Figure 8 to emphasise the opposite idea to emphasise the positive sense of meaning i.e. despite losing one of her limbs, the character is still an active surfer. This image, hence, will imprint more positive idea to the readers who are students.

No clear drawbacks have been observed from the use of the images accompanying the reading texts. On the contrary, the images juxtaposed with the reading texts augment the delivery of the overall meanings. One notable difference between the set of visual and verbal in both TA6 and SP is seen in terms of framing strategies. TA6 utilises a clear framing in yellow lines to wrap the image of Pele, giving it a sense of detachment with the content of the reading passage. The unity, however, is reinforced in the contrastive colours of green and yellow. SP, on the other hand, opts to blend the visual and the verbal of the reading text with the absence of framing. The image and the background setting of the image have essentially become the space which is occupied by the verbal elements. This strategy creates a strong coherence between the image and language, making the visual and the verbal inseparable resources to understand the whole meaning in the passage of Surfing Superstar!

#### *Some ineffective language-image arrangement*

One of the challenges of using images in textbooks is to make sure that every selection of image has significance in fulfilling the instructional purposes. Otherwise, the images will not function optimally and will seem more like decoration to fill the blank space on a page. Consider, for example, the following spread found on pages 20-21 of TA6.



Figure 10. A spread of pp. 20-21 in TA6

Moving from left to right, following the normal western reading path, the spread can be seen as a unit of meaning which is arranged in one composition. The images on the

left part of the layout (exercise 4) are useful in prompting students with the required terms for the exercises. Also, the image in the right part of the spread (page 21) is useful in providing context for exercise 1 in the communication section. The image of the vertical depiction of stadium positioned in the middle of the layout, however, is irrelevant to the rest of the page. The framing of this particular image is detached from the adjacent images. It is clearly not relevant to exercise 5 as the instruction says: Tell your friends what to do and what not to do at the gym, and it does not relate to any exercises in the next page as another image has been supplied at the top half of page 21. Therefore, the function of providing this image lacks relevance with the rest of the elements on the spread and is redundant.

The image supplied for the exercise of Sports Quiz in the Communication section would have sufficed for the tasks provided. This image relates to the given questions for the Sports Quiz. The phrase football match in question number 1 is very well visualised in a picture depicting a match going on in a stadium, complete with the image of spectators seen from a distance in the tiered seating of the stadium. The perspective of the image of the cheering spectators in the foreground gives a sense of involvement with the readers. The lexis Olympic games, which is repeated in question numbers 3, 4 and 6, is visualised by showing the logo of the Olympic Rings. The positioning of the logo in a salient location. i.e. the top right corner in balanced size to the overall image, guarantees an ease of vision by students. The task in this section is cunningly followed by the similar idea of asking students to visualise a sports personality using the clues given in verbiage. While this task is quite challenging i.e. asking students to draw a sports person, it can be modified into collecting or cutting and pasting images from any available source and/or being expanded into a presentation. The task can be further expanded into a speaking task by using students' own ideas into a Pecha Kucha describing a sports person such as exemplified by Rahmaningtyas et al., (2017, pp. 305 – 323).

Another example of flawed images in TA6 unit is found on page 23. The images' size seems to outweigh the listed 5 questions related to the image. Together, the set of questions and the supplied images occupy almost 50% of the whole page.

3 Listen to the passages again. Fill in each blank to complete the sentences.

- Hai practises at the \_\_\_\_\_ three times a week.
- Trung can \_\_\_\_\_ the guitar.
- Alice likes \_\_\_\_\_ ice skating.
- \_\_\_\_\_ is at Rosemarrick Lower Secondary School.
- Trung \_\_\_\_\_ swimming on hot days.

The worksheet includes a large yellow box with horizontal lines for writing answers. Below the text are two photographs: one showing a swimming pool with people swimming, and another showing a group of people in a gym setting.

Figure 11. Task on page 23 in TA6

Secondly, visual-verbal relations between the image and the question shows a slight irrelevance in meanings. Question number 3, Alice likes \_\_\_\_\_ ice skating, is visualised in the group of young girls doing skating. Due to the size of the shot, the setting of ice skating is not very clear. Also, the similar colour between the skating shoes and the ice skating ground is not clearly defined, hence requiring extra attention when reading the image. The image of the swimming pool is related to question number 5: Trung \_\_\_\_\_ swimming on hot days. While there is no other clear context to the required sentence completion, it seems to suggest a casual swimming activity. The questionable part of the image, then, lies in the fact that the image displays a swimming pool with ropes and floats, suggesting a context of swimming competition. Again, this seems like a careless selection of image that may distract from meanings arising from the wordings and visual resources.

In contrast, an effective use of the spread is in SP page 16-17. A general observation of the spread recognises a balanced use of image and wordings. The sizes of the image vary, following the wordings that relates to each image. The image on page 16, for example, is supplied to visualise the main information in the reading passage: The competitors have to swim 110 meters through cold, dark, muddy water – without using their arms! A careful blend of the green colour unites the verbal and the visual to reinforce the meanings. Other visual images on the next page also fulfil their function to reinforce the meanings. Considering the space provided for each task, the sizing seems appropriate. Not only does each image framed in clear dividing lines either using colour separation or clear lines, but also the framing of each task conveys the separation of one task to another. In this way, each image can only be related to its assigned task, minimising risk of misinterpreting the visual and verbal resources in the task.

**2C CULTURE Village sports**  
I can understand information about a sporting event.

**1 VOCABULARY** Look at the photo. How would you describe this sport?  
a. noisy   b. cold   c. slow   d. tiring

**2 LISTENING** Read the text. Is bog snorkelling only popular with British people?

**3 VOCABULARY** Match the highlighted words in the text with the definitions below.

**4 READ THE TEXT AGAIN.** Answer the questions.

**5 LISTENING** Listen to the radio programme. Do you believe that this event can be exciting? Why? Why not?

**6 MULTIPLE CHOICE** Listen again and choose the correct answers.

**7 GRAMMAR** Work in pairs. Cover the text and describe the photo in exercise 1.

**8 GRAMMAR** Work in groups. Find information about an unusual sporting event in your own country or abroad. Make notes about:

**GRAMMAR BUILDER 2D, PAGE 104**

**2D GRAMMAR Past simple and continuous**  
I can tell a short story using past tenses.

**1 READ THE TEXT.** What is unusual about this team photo?

**2 LOOK AT THE PAST CONTINUOUS FORMS** in the text. Then complete the table with the correct form of the verb be.

**3 READ THE TEXT AGAIN.** Find one example of the uses 1–3 in the text in exercise 1.

**4 COMPLETE THE TEXT.** Use the past simple or past continuous form of the verbs in brackets. Then listen and check.

**5 LOOK AT THE CARTOON STORY** and answer the questions. Use the words in brackets to help you.

**6 WORK IN PAIRS.** Look at the cartoon story again for one more minute. Close your books and tell the story to your partner. Use the past simple and past continuous.

**UNIT 2 • Winning and losing**

Figure 12. Spread on pp. 16-17 in SP

All in all, the sample spreads from both TA6 and SP show that effective use of image on a textbook page is not solely determined by the size of the image. In other words, bigger size of image does not always mean better use of image. Rather, the relevance of images with the corresponding tasks matters.



## Conclusions

In line with Mickan's (2017) observation that the field of language teaching has been developed with a tradition of treating grammar and vocabulary as basic building blocks which are separated from authentic texts that consequently devoid of meaning and disregarding the essential relationship of text to context (p. 20), this study finds the segregation of tasks for different language skills in the data. While the structure of tasks is not the main in this study, the relation between image and language throughout the observed units in the books revolves around the role of images and their relations to the language elements in fulfilling instructional purposes in separate tasks. Three main functions are outlined. First, to effectively aid in eliciting the general topic of a unit, images may utilise sizing to offer salience in order that students can relate the main verbal and visual at the opening of a unit. While Widiastuti (2017) has found that some textbooks lack cohesiveness between topic and section, this study specifically addresses the topic of the units and their relations to the supplied image. It is found that while the topic of the units plays an important role in setting up students' readiness in learning the aspects of language throughout the unit, the image supplied in the opening helps this process provided that the image is relevant to the wordings. Second, as vocabulary building is an important task for students, especially in ESL contexts, presentation of images to aid vocabulary building should consider clarity by utilising effective framing strategy that will result in easily distinguishable images to refer to the necessary vocabulary. Finally, while the analysis has shown that language and image resources combine to convey the overall meanings of the passage, the reading path is determined by the strategic positioning of both resources. On a final note, the study observes ineffective language-image relation due to the layout of visual and verbal resources either on a page or a spread owing to the lack of relations between the images and the other adjacent elements. Although the discussion in the paper is limited to only one unit of two separate books, it is possible to replicate the method to analyse the language-image relationship in the remainder of the book. Future exploration of textbooks remains open, for example, by expanding the number of samples or by examining the usefulness of images and language-image relationships of the units in the whole book.

## References

- Bui, T. T. N., & Nguyen, H. (2016). Standardizing English for educational and socio-economic betterment – A critical analysis of English language policy reforms in Vietnam. In R. Kirkpatrick (Ed.), *English language education policy in Asia* (pp. 363-388). Springer.
- Carney, R., & Levin, N. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14(1), 5-26.
- Chegeni, N., Kamali, B., Noroozi, A., & Chegeni, N. (2016). General English textbook evaluation: A closer look at "Four Corners". *Theory and Practice in Language Studies*, 6(12), 2325.
- Chen, Y. (2010). Exploring dialogic engagement with readers in multimodal EFL textbooks in China. *Visual Communication*, 9(4), 485-506.
- Falla, T., & Davies, P. (2013). *Solutions: Pre-intermediate student's book* (2 ed.). Oxford University Press.
- Hoang, V. V., Nguyen, T. C., Le, K. D., Phan, C. N., Vu, M. T., Luong, Q. T., & Kaye, D. (2015a). *Tieng anh 6-9 sach hoc sinh*. Vietnam Education Publishing House.

- Hoang, V. V., Nguyen, T. C., Le, K. D., Phan, C. N., Vu, M. T., Luong, Q. T., Kaye, D. (2015b). *Tiếng anh 6 sách học sinh Tập 2*. Vietnam Education Publishing House.
- Hsu, P., & Yang, W. (2007). Print and image integration of science texts and reading comprehension: A systemic functional linguistics perspective. *International Journal of Science and Mathematics Education*, 5(4), 639-659.
- Kress, G. (2000). Multimodality: Challenges to thinking about language. *TESOL Quarterly*, 34(3), 337-340.
- Kress, G. R., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design (1st ed.)*. Routledge.
- Kress, G. R., & Van Leeuwen, T. (2006). *Reading Images: The grammar of visual design (2nd ed.)*. Routledge.
- Le, V. C. (2015). English language innovation for the Vietnamese secondary school: The Project 2020. In B. Spolsky & K. Sung (Eds.), *Secondary school English in Asia: From policy to practice*. Routledge.
- Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J.R. Martin and R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science* (pp.87–113). Routledge.
- Levin, J. R., & Mayer, R. E. (1993). Understanding illustrations in text. *Learning from textbooks: Theory and practice*, 95-113.
- Mickan, P. (2006). Socialisation, social practices and teaching. In P. Mickan, I. Petrescu, & J. Timoney (Eds.), *Social practices, pedagogy and language use: Studies in socialisation* (pp. 7-23). Lythrum Press.
- Mickan, P. (2017) Text-based research and teaching from a social semiotic perspective: Transformative research and pedagogy. In Mickan, P., & Lopez, E., (Eds) *Text-based research and Teaching: A social semiotic perspective on language in use*. Palgrave Macmillan.
- MOET. (2014). *The six-level language proficiency framework for Vietnam*. Hanoi, Vietnam.
- Mukundan, J. & Kalajahi, S. A. R. (2013). Evaluation of Malaysian English language teaching textbooks. *International Journal of Education and Literacy Studies*, 1(1), 38-46.
- Mukundan, J., & Nimehchisalem, V. (2012). Evaluative criteria of an English language *textbook evaluation checklist*. *Journal of Language Teaching and Research*, 3(6), 1128-1134.
- O'Halloran, K. (2008). Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Visual Communication*, 7(4), 443-475.
- Painter, C., Martin, J. R., & Unsworth, L. (2013). *Reading visual narratives: Image analysis of children's picture books*. Equinox Publishing Ltd.
- Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction*, 3(3), 227-238.
- Prime Minister Nguyen Tan Dung. (2008). *Decision 1400: Teaching and learning foreign language education in the national education system in the period 2008-2020*. Hanoi, Vietnam.
- Rahmaningtyas, H., Febrianti, Y., & Inayati, N., (2017) Using student-selected texts in speaking classes. In Mickan, P., & Lopez, E., (Eds.), *Text-based research and teaching: A social semiotic perspective on language in use* (pp. 305-324). Palgrave Macmillan.
- Royce, T. (1998). Synergy on the page: Exploring intersemiotic complementarity in page-based multimodal text. *JASFL Occasional Papers*, 1(1): 25-49.
- Royce, T. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy', *TESOL Quarterly*, 36(2): 191–205.
- Unsworth, L. (2001). *Teaching multiliteracies across the curriculum: Changing contexts of text and image in classroom practice*. Open University Press.
- 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. *TESOL in Context*, 16 (Special edition), 147-162.

- Unsworth, L. (2014). Multimodal reading comprehension: Curriculum expectations and large-scale literacy testing practices. *Pedagogies: An International Journal*, 9, 26-44.
- Unsworth, L., & Chan, E. (2009). Bridging multimodal literacies and national assessment programs in literacy. *Australian Journal of Language and Literacy*, 32(3), 245-257.
- Unsworth & Cleirigh, C. (2014) Multimodality and reading: The construction of meaning through image-text interaction. In Jewitt, C., & Jewitt, C., (Eds.), *The Routledge handbook of multimodal analysis*. Routledge.
- Unsworth, L., & Ngo, T. (2015). Vai trò của hình ảnh trong sách giáo khoa dạy tiếng Anh ở Việt Nam [The role of visual images in English textbooks in Vietnam]. *The Journal of Language and Life*, 2015(1), 93-99.
- Widiastuti, R., (2017) A Socio-semiotic perspective on a unit of work in an Indonesian textbook. In Mickan, P., & Lopez, E., (Eds), *Text-based research and teaching: A social semiotic perspective on language in use* (pp. 179-192), Palgrave Macmillan.