

Nayara Salbego
nayara.salbego@yahoo.com.br

Viviane M. Heberle
viviane.heberle@ufsc.br

Maria Gabriela Soares da Silva Balen
gabi20soares@gmail.com

A visual analysis of English textbooks: Multimodal scaffolded learning

Análise visual de livros didáticos de Inglês: multimodalidade facilitando a aprendizagem

ABSTRACT - This paper investigates how multimodality in English textbooks may scaffold learning through visual texts. It provides a multimodal analysis of images, integrated with verbal texts and proposed language activities, in order to explain how the visual meanings may enhance students' understanding of language and content. Kress and van Leeuwen's visual grammar grounds the image analysis, which is discussed in light of the three metafunctions – representational, interactional and compositional. Findings show that the visual texts are relevant for beginners to understand the content of the activity. The analysis of the images may contribute to scaffold learning in that they are part of the overall meaning, supporting students' understanding of the activity as a whole. Results also point to the importance of working with multimodality in language learning contexts.

Keywords: multimodality, visual analysis, English textbooks, learning English.

RESUMO - Este trabalho investiga como a multimodalidade em livros didáticos de Inglês pode auxiliar na aprendizagem através da análise de imagens. Fez-se um estudo multimodal de imagens, presentes em textos e atividades didáticas, a fim de explorar como os significados visuais podem auxiliar no entendimento dos alunos com relação à linguagem e ao conteúdo proposto. A gramática visual de Kress and van Leeuwen consiste na base teórica do trabalho, de forma que as três metafunções – representacional, interacional e textual – fundamentam o estudo das imagens. Os resultados mostram que os textos visuais são relevantes principalmente para aprendizes iniciantes entenderem o conteúdo da atividade. A análise das imagens pode contribuir para a aprendizagem, porque fazem parte do significado geral e auxiliam os alunos na compreensão da atividade como um todo. Os resultados também apontam para a importância de se trabalhar com multimodalidade em contextos de ensino e aprendizagem de línguas.

Palavras-chave: multimodalidade, análise visual, livros didáticos de Inglês, aprendizagem de Inglês.

Introduction

In 1996, a group of researchers from the United States of America, Australia and The United Kingdom met to discuss their concerns on literacy pedagogy due to new demands from multicultural societies and diverse technological advances. This group, called *The New London Group*, explained at that time:

If it were possible to define generally the mission of education, one could say that its fundamental purpose is to ensure that all students benefit from learning in ways that allow them to participate fully in public, community, and economic life. Literacy pedagogy is expected to play a particularly important role in fulfilling this mission. Pedagogy is a teaching and learn-

ing relationship that creates the potential for building learning conditions leading to full and equitable social participation (New London Group, 1996, p. 60).

In an attempt to account for these demands, The New London Group (1996) proposed what they called “A pedagogy of Multiliteracies”, and their concerns have become a new ground for studies on literacy pedagogy since then. Thus, concerning the importance of understanding the variety of text forms in the world nowadays, The New London Group (1996, p. 61) tells us that:

[...] literacy pedagogy now must account for the burgeoning variety of text forms associated with information and multimedia technologies. This includes understanding and competent

control of representational forms that are becoming increasingly significant in the overall communications environment, such as visual images and their relationship to the written word.

In accordance with The New London Group, Macken-Horarik (2004, p. 24) also states that:

Whatever the subject, students now have to interpret and produce texts which integrate visual and verbal modalities, not to mention even more complex interweaving of sound, image and verbiage in filmic media and other performative modalities.

As teachers of English in Brazil, we view the new demands on literacy pedagogy as fundamental in our classes. In this paper we focus specifically on the synergy between visual and verbal modes, as a way to stimulate beginner students in their learning of foreign/additional languages¹. In this sense, we are also concerned with critical reading, and due to the prevalence of graphical and pictorial artifacts in today's world as part of people's daily lives, reading images critically requires closer attention.

In any learning context, students are usually surrounded by images, especially in textbooks they carry around with them as well as in the websites they access on the Internet. In order to take advantage of these affordances, learners may need some guidance and specific metalanguage to read these multimodal texts. Language teachers may thus play an important role in instructing their students to make sense of and explore the visual and verbal resources in these texts, that is, the "image-text relations" or the "co-articulation of image-verbiage" (Unsworth, 2006, p. 1165, 1201).

In view of these pedagogical interests, this paper, thus, aims to contribute to the discussions regarding image analysis to foster learning, more specifically, learning English as a foreign language². Communication is increasingly multimodal and more studies on the analysis of images are necessary in educational contexts (Christie, 2005; Heberle, 2010; Unsworth, 2001, 2013). For this reason, images provided along texts and textual language activities are analyzed here in order to show how they may scaffold language learning.

Theoretical background

In contemporary society, images are part of our everyday lives and visual literacy has become very important in educational contexts. As explained by Kress and van Leeuwen (1996, p. 2) "visual structures realize meanings as linguistic structures do also, and thereby point to different

interpretations of experience and different forms of social interaction". By proposing a grammar of visual design (GVD), these authors, in fact, state that "visual literacy will begin to be a matter of survival" (Kress and van Leeuwen, 2006, p. 2) and reading images will be essential to understand and interpret the world. By creating inventories of common visual conventions, Kress and van Leeuwen seek to make explicit what is often implicit in images. It is our belief that teachers can foster students' language learning processes by using GVD and teaching its basic principles.

GVD is based on three language metafunctions previously developed by Halliday (1994; Halliday and Matthiessen, 2004), in systemic-functional linguistics (SFL), which combines lexicogrammar, semantics and context. Halliday sees language as social semiotics, that is, as a resource to understand and produce meanings in any social environment, and it can be regarded as an attempt to describe and understand how people produce and interpret meanings in social settings. Kress and van Leeuwen, thus, clarify that:

[J]ust as grammars of language describe how words combine in clauses, sentences and texts, so our visual 'grammar' will describe the way in which depicted people, places and things combine in visual 'statements' of greater or lesser complexity and extension (Kress and van Leeuwen, 1996, p. 1).

Kress and van Leeuwen state that similarly to verbal language, images can be interpreted according to what they represent. Meaning expressed in language through parts of speech and grammatical structures can be expressed in images through color, tone, angle, framing, among other categories, and this affects what and how images communicate meanings to viewers. The authors' descriptive framework for multimodality assigns representational, interactive and compositional meanings to images. Thus, any image (a) represents an aspect of the world – whether in abstract or concrete ways –; (b) plays a part in some interaction with the viewers; and (c) combines visual elements into a coherent whole. The representational metafunction corresponds to the identification of the represented participants, whether animate or inanimate, the processes or the activities being performed, the attributes or the qualities of the participants and, finally, the circumstances in which the action is being developed. When participants are connected by vectors or by eyelines, e.g. as in narrative images, they are represented as doing something to or for one another. These narrative patterns present unfolding actions and events³.

¹ Nowadays, other terms as English as an additional or international language may be used, which can be seen as more adequate in contemporary society. See Jordão (2014) for a discussion of these different terms in relation to English-teaching practices in Brazil. In our paper we use English as a Foreign Language (EFL), but we acknowledge the relevance of the other alternative terms. We also mention ESL (meaning English as a Second Language, not our case in Brazil), when we refer to Royce's study with ESL students as well as TESOL (Teaching of English to Speakers of Other Languages), in relation to Stenglin and Iedema's study.

² This paper is funded by the Brazilian Science and Technology Research Council, CNPq, Process n. 313144/2014-1, with a grant to Viviane M. Heberle.

³ For studies in Brazil concerning GVD, see Almeida (2008), Lima *et al.* (2009) and Nascimento *et al.* (2011), for example.

The interactional metafunction comprises the social relations between represented participants (people or objects depicted in the images), viewers (people who see the images) and also the image producer (the designer, the photographer, etc.). In the verbal mode, writers address their readers by making statements, asking questions, making offers or requiring some kind of action from them. In the visual mode, producers use visual techniques to get their messages across. Among the visual techniques used to analyze interpersonal meaning we can refer to the absence or presence of facial expressions towards the viewers (demands or offers), gestures which make commands, and offers of information or offers of goods and services (Royce, 2007). Interactive relationships are also defined on the basis of perspective and social distance (long, medium or close shots; angles; and framing⁴).

The compositional metafunction corresponds to the study of aspects related to the layout of the page, to the placement of the visual elements, to “the way in which the representational and interactive elements are made to relate to each other, the way they are integrated into a meaningful whole” (Kress and van Leeuwen, 2006, p. 176). Thus, the compositional features involve the study of the visuals concerning the distribution of the information value, visual salience (size and color) and visual framing. The placement of elements to the left (given information) or to the right (new information), the relative size of the figures in the image and the use of framing are all relevant factors of the compositional meaning.

By using GVD as proposed by Kress and van Leeuwen (2006), teachers may contribute to develop students’ Multimodal Communicative Competence (Royce, 2007; Heberle, 2010), which refers to the skills needed to read and interpret not only written language, but also images. In this sense, Royce (2007, p. 377) states that:

[A]lmost every image can be analyzed in terms of what it presents, who it is presenting to, and how it is presenting, and [...] the concept of metafunctions can be suggestive for the language teacher in developing pedagogical resources targeted to help students extract just what the visuals are trying to ‘say’, to relate these messages to the verbal aspect, and then use them to contribute to developing students’ multimodal communicative skills.

Following the same line, this paper suggests that by integrating visual analysis and helping students to learn how to “read” images, teachers of English as a foreign/additional language may help these students interpret what has been said in written texts and language activities. Accordingly, Kress (2000, p. 337) states that:

[I]t is now impossible to make sense of texts, even of their linguistic parts alone, without having a clear idea of what these other features might be contributing to the meaning of a text. In fact, it is now no longer possible to understand language and its uses without understanding the effect of all modes of communication that are copresent in any text.

In this regard, if teachers explore visual features along with the textual content in textbooks and guide learners through understanding their meanings and what they represent, they will be scaffolding students’ learning processes and understanding of activities. By scaffolding, teachers provide support that will assist learners to develop new understandings, concepts, and abilities in learning (Hammond and Gibbons, 2001).

Hammond and Gibbons (2001) draw on Vygotsky’s and Bruner’s ideas to explain the essentiality of scaffolding in the context of education. Hammond (2001) discusses the concept in relation to Halliday’s theorization of language and indicates that scaffolding represents the driving force for language learning. In relation to visual meanings and scaffolding, Herrel and Jordan (2012) use the term visual scaffolding regarding the use of drawings, photographs and other visuals in order to help students to better understand the language used in each lesson.

Accordingly, in relation to scaffolding students’ learning through the use of images, as a way to enhance language development, activities related to visual literacy may contribute to make students aware that “visuals are not to be seen as a separate or add-on strategy, but as a valid tool in EFL teaching and learning”, as explained by Heberle (2010, p. 113). Similarly, following Kress and van Leeuwen’s GVD, Stenglin and Iedema (2001) also emphasize the relevance of developing students’ skills regarding image-text relations and offer pedagogical suggestions for TESOL.

Likewise, according to Royce (2007), the act of interpreting images in light of the three metafunctions may help ESL students interpret the verbal texts which are accompanied by them. Following this line, Royce (2007) analyses a text with images extracted from an ESL textbook approved by the Japanese Education Ministry to be used in high schools in Japan. The author argues that “...this text is in fact a rich source of multimodal meanings which can be approached in terms of multimodal communicative competence” (p. 377).

According to Royce (2007), analyzing the images before reading the text can ease the students’ interpretation of it. “Activities could be organized which involve the students asking questions of the visuals, and then using their answers to assist in their reading development” (p. 379). Some of the questions may be: who or what is seen in the

⁴ Although framing is part of the compositional metafunction (Kress and van Leeuwen, 2006), it is also a resource to understand social distance in the interactional/interactive metafunction. This illustrates the interrelated meanings among the metafunctions.

image; what they are doing; who or what they are doing it with; where and why they are doing it, for instance. Royce adds that these questions are related to the message-focus or the ideational aspects and they provide a good source of information. Heberle (2010, p. 112) also suggests some questions concerning the three metafunctions:

What is the picture about? Who are the participants involved, and what circumstances are represented in the photograph/image? (representational metafunction)
 What is the relationship between the viewer and what is viewed? (interactional metafunction)
 How are the meanings conveyed? How are the representational structures and the interactive/interpersonal resources integrated into a whole? (compositional metafunction)

Regarding the efficiency of interpreting images before the linguistic text, Royce (2007, p. 380) states that:

The students can ease themselves into a reading and get some idea of what to expect in terms of the who, what, where, why, how and with whom in the image. The effect is that expectancies are being set up in the students' minds, and the process of reading the text will then either give them a confirmation of their interpretation of the information (or story), or in rare cases introduce ambiguities, which the class can then explore in more depth through discussion and follow-up written activities.

In these terms, Royce emphasizes that reading images influences the way learners may understand written texts. While image interpretations may be confirmed and reinforced, others may be clarified by the written text. Consequently, images may not only complement and support the reading of texts, but they are part of the overall meaning in the visual-verbal synergy. Bearing that in mind, this study aims at showing how textual information is depicted in images that accompany language activities and, most importantly, how the analysis of images may contribute to scaffold language learning.

Considering the relevance of visual literacy as previously indicated, this paper shows, in the next sections, the analysis of three images extracted from textbooks for EFL learners to suggest the idea that images should be 'read' in order to ease the students' interpretation of the written texts that accompany them.

Method

For the present study, three images from three textbooks were selected and analyzed in the context of visual social semiotics (Kress and van Leeuwen, 1996; 2006), based on our experience and familiarity⁵ with them. The selected books –*Interchange* (Richards, 1999), *New*

Interchange (Richards *et al.*, 2001) and *Cutting Edge Elementary* (Cunningham *et al.*, 2005) – are well known English textbook series in Brazil and used in a number of language schools. The specified proficiency level concerned activities and texts for beginners since this is the phase in which learners may need more support and images might help to reinforce their understanding of the content.

As stated before, the three metafunctions developed by Kress and van Leeuwen (2006) in their GVD have grounded our qualitative analysis in this paper and our study tries to show benefits from interpreting images coupled with texts or language activities. Along with the image analysis, examples from the textual activities are provided. The analyzed images were named as Images 1, 2 and 3. Image 3 is divided into three other images, which will be addressed as Photos 1, 2 and 3. We emphasize that in our analysis we refer to viewers as learners of English observing the images analyzed.

Analysis and discussion

As we have pointed out, understanding what the images are 'saying' is supportive of learning. In this section, thus, images from the three selected textbooks are analyzed and discussed based on GVD (Kress and van Leeuwen, 2006) in order to show how the visual analysis may help beginner EFL students to interpret verbal texts or activities. Firstly, the selected images are shown, followed by the analysis and discussion.

The first analyzed image (Image 1) was extracted from the book *New Interchange* (Richards *et al.*, 2001), a well-known English book for beginner students in Brazil, which contains activities and texts to develop the so-called four skills – reading, listening, writing and speaking – through the use of colorful drawings (e.g. Images 1 and 2 analyzed in this paper). Image 1 consists of a conversation among three people and a drawing representing this scene.

In **Image 1** students can identify the main characters, the way they interact with each other, and the circumstances. The three main characters, interactive participants, are depicted in a specific social context, which seems to be a café or a restaurant. The students can infer this by analyzing the representation of other participants on the background (three minor represented participants on the upper-part of the picture, with two of them sitting down, probably talking to each other and the other one standing up). The main participants are standing up and looking at each other, and their gestures, represented mainly by their arms and facial expressions, show they are actively involved in a conversation.

⁵ We are familiar with the textbooks analyzed in this article since we are EFL teachers and have used them to teach EFL students in Brazil. We are also aware that nowadays there are more recent EFL books developed specifically for Brazilian contexts, as proposed in the national program PNLD (Programa Nacional do Livro Didático) (http://portal.mec.gov.br/index.php?Itemid=668&id=12391option=com_content&view=article, retrieved in January, 2015).

7 CONVERSATION Asking about someone

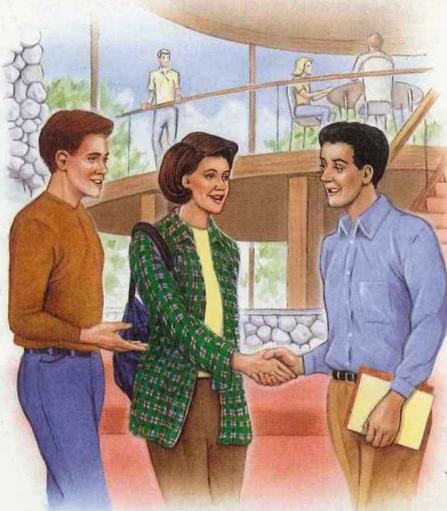
A  Listen and practice.

Sarah: Hi, Tom. How's everything?
 Tom: Not bad. How are you?
 Sarah: Pretty good, thanks.

Tom: Sarah, this is Paulo. He's from Brazil.
 Sarah: Hello, Paulo. Are you on vacation?
 Paulo: No, I'm not. I'm a student here.
 Sarah: Oh, are you studying English?
 Paulo: Well, yes, I am. And engineering, too.
 Sarah: Are you and Tom in the same class?
 Paulo: No, we aren't. But we're on the same volleyball team.

CLASS AUDIO ONLY **B**  Listen to the rest of the conversation.

Where is Sarah from?



5

Image 1. Image from a conversation in *New Interchange*: Please call me Chuck.
 Source: Richards *et al.* (2001, p. 5).

Analyzing this image in terms of a representational structure and considering only the woman and the man on the right side of the picture, it is accurate to state that they represent a transactional reaction – a vector emanates from the man's eyes toward the woman's eyes and vice versa. Their arms also form vectors toward each other, since they are shaking hands, which may indicate that they are being introduced to each other. Another vector emanates from the man's eyes on the left side of the image toward the other man on the right side. Furthermore, the right arm of the man on the left points at the man on the right, forming a vector which suggests he is introducing the man on the right side to the woman. Therefore, his utterance "Sarah, this is Paulo. He's from Brazil" confirms this interpretation.

Analyzing participants' attributes can also help viewers to understand who these participants are and what actions they are involved in. The main participants seem to be young adults (maybe college students) due to the casual clothes they are wearing, and what seems to be a notebook which the participant being introduced to the woman is holding. Therefore, the sentence "I'm a student here" in the dialogue confirms this interpretation.

In terms of interactive meanings, in which the relationship between the viewers and image is taken into consideration, Image 1 represents an offer, since none of the participants depicted is looking at the viewers, demanding an "answer" or reaction from them. On the contrary, "...it 'offers' the represented participants to the viewer as items of information, objects of contemplation,

impersonally, as though they were specimens in a display case" (Kress and van Leeuwen, 2006, p. 119). In relation to Image 1, the viewers are not expected to take part in the conversation, since they are only observing the scene (a narrative, in GVD's term) as if it were a movie.

Concerning the compositional aspects of the image, both the man on the left side and the woman are "given"; they are together, placed on the same oblique angle and interacting with the "new" participant, the man on the right side of the image. Thus, the position of the participants in the image and the act of "shaking hands" can help EFL students to realize that the conversation began between the man and the woman on the left, and the man on the right arrived later, being introduced to the woman. The following extracts from the text exemplify this statement:

"Sarah: Hi, Tom. How's everything?"
 "Tom: Not bad. How are you?"

"Tom: Sarah, This is Paulo. He's from Brazil."
 "Sarah: Hello, Paulo..."

Another compositional aspect in the image is salience, which is represented through the participants' location in the foreground (medium shot) in relation to the participants in the background (long shot), composing the setting and contextualizing the narrative. Besides being represented in a medium shot, which provides them more visibility in the scene, the main participants are portrayed in brighter and more vivid colors than the participants in

the background. The latter, thus, are represented in a long shot which gives them less emphasis with fewer details and lighter colors. Thus, viewers may realize that the focus of the conversation concentrates on the three people in the foreground, and the other participants in the background contribute to suggest where the narrative is taking place.

The visual analysis of Image 1 shows that, in fact, the interpretation of the picture supported by the three metafunctions proposed by Kress and van Leeuwen (2006) may help beginner students to understand the written text. Each clue such as the act of shaking hands, participants' gaze and facial expression, the position of the characters in the picture and salience provide evidence which may contribute to students' understanding of the textual content.

Image 2 was extracted from the text book *Interchange*, which is a previous version of the *New Interchange* (Image 1) and follows the same combination of images (drawings and photographs) and written text. In this image, the represented participants (the man and the woman, drawn as caricatures) are looking at each other and they seem to be arguing, as can be observed by their facial expressions and gestures, mainly represented by their arms. There is a map in the background which can lead the viewers to infer the theme of the discussion, probably related to countries and their location on the map.

Analyzing it through a representational structure, the represented participants represent a transactional reaction, since there is an eyeline vector connecting them. The woman's right arm is pointing upwards, forming a vector. Through this gesture and her confident posture, viewers may infer that she is making a statement: "Actually, Costa Rica isn't in South America. It's in Central America". The man, however, is lifting his shoulders, his arms are pointing at opposite directions, and his facial expression shows that he may be embarrassed or not confident about

something. The following statement makes it clear: "Oh, right. My geography isn't very good".

In terms of interactive meanings, just as in Image 1, this picture represents an offer, since neither of the participants depicted is looking at the viewers. They are, indeed, in a position of "objects of contemplation" (Kress and van Leeuwen, 2006, p. 120). The woman seems to be a little closer to the viewers than the man, but both are in a social distance (medium shot), since their bodies can be seen from their legs upwards. They are also represented in oblique angles, which suggest detachment toward the viewers. Concerning the compositional aspects, the interactive participants are depicted in the foreground and have brighter and more vivid colors in comparison to the map, which is in the background. Similarly to Image 1, the represented participants occupy a foregrounded position in the narrative. In this case, however, contrary to Image 1, the background does not show details in terms of where the narrative takes place, except for the image of a map in the background. The presence of the map allows viewers to infer that the young man and woman are in a classroom, or maybe in an office, but they cannot determine the context in which the conversation is taking place.

Image 3 was extracted from the textbook *the Cutting Edge Elementary* (Cunningham *et al.*, 2005), which presents a variety of activities concerning reading, listening, writing, speaking, grammar analysis and practice, using English in an elementary level. Throughout the book, there are images used to illustrate and support the textual content: every single page from a total of 175 has images, which include drawings and photographs. For this paper, we analyze a multimodal activity which is composed of three integrated photos and a written matching activity in two columns. The discussion again concentrates on how they may support students' understanding of the activity

3 CONVERSATION

1 Listen and practice.

Mark: Where are you from, Laura?
 Laura: Well, my whole family is in the United States now, but we're from Costa Rica originally.
 Mark: Oh, so you're from South America.
 Laura: Actually, Costa Rica isn't in South America. It's in Central America.
 Mark: Oh, right. My geography isn't very good!

2 Listen to the rest of the conversation.

a) Where is Mark from?
 b) Where is his country?

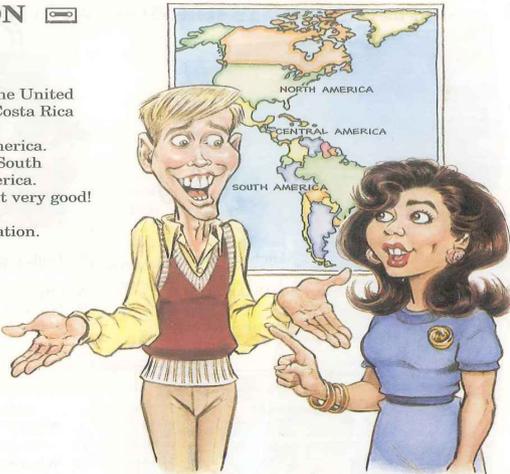


Image 2. Image from a conversation in *Interchange*: Where are you from?
 Source: Richards (1999, p. 15).

1 MD Look at photos 1–3. Match the sentences below to the people.

1 Andrei

2 Toshi and Mariko

3 Marisol

a His name's Andrei. **1**

b Her name's Marisol.

c Their names are Toshi and Mariko.

d They're from Tokyo.

e He's from Russia.

f She's from Valencia in Spain.

g He's twenty-two.

h He's a student at Moscow University.

i He's nineteen.

j She's at the airport.

k She isn't on holiday. She's on business.

l They're tourists.

m She's thirty-five and she's married.

n He isn't married. He's single.

o They aren't married. They're friends.

Image 3. Image from *Cutting Edge*: “Language Focus 2” - Personal information.
Source: Cunningham *et al.* (2005, p. 11).

as a whole. The image containing the three photos and the corresponding language activity is presented below as Image 3, followed by our analysis and discussion. The numbers used in the corner of each image (1, 2 and 3) correspond to the way they are addressed (Photos 1, 2 and 3).

The three different photos in Image 3 (addressed as Photo 1, Photo 2 and Photo 3) correspond to the same language activity in which students are expected to match sentences to the corresponding images. They were not detached from each other in this paper because they form a whole with the activity proposed in the textbook. Reference to the activity proposed is again made to explain why the images contribute to scaffold students' understanding of the textual content provided for them to analyze and match the verbal and visual meanings.

Considering the representational metafunction, Photos 1 and 3 depict action qualities, which classify them as narrative, since there are vectors in these two images, and the represented participants are carrying out activities. In Photo 1, the man's right arm forms a vector showing that he is writing. There is another vector from his eyes, which leads the viewers to see

that the young man is looking at the paper he is working on as well.

Still concerning the representational metafunction, Photo 2 can be seen as conceptual (not narrative), since we can see only the represented participants' faces. As pointed out by Kress and van Leuwen (2006, p. 79), conceptual images ‘represent participants in terms of their more generalized and more or less stable and timeless essence, in terms of class, or structure or meaning’. The participants in Photo 2 are looking straight at the viewers, demanding an action, maybe asking the EFL students to find the appropriate sentences to match their picture. Another important visual meaning concerns these participants' physical characteristics: they have Asian-shaped eyes and straight black hair. Such characteristics allows the students to match these participants' place of origin with sentence ‘d’ – “They're from Tokyo”.

In Photo 3, there is a woman talking on the phone. The background shows some details of the circumstances, such as direction signs (one of them with an icon of an airplane) and very big glass windows, which makes it possible to identify that the woman is sitting in an airport lounge.

In the three photos (Photos 1, 2 and 3), the represented participants are not looking at each other (a non-transactional process), but the pictures contribute to the identification of sentences that may be matched to each of the photos in Image 3. For example, sentence ‘h’ – He’s a student at Moscow University” – can be related to Photo 1, in which the man is writing. The pronoun ‘he’ allows viewers to identify that there is one person in the corresponding photo, and this person should be male. Moreover, this ‘he’ is a student at a university, so some of his attributes in the photo are a book or a notebook and a pencil or a pen. Furthermore, the image portrays a very informal way of studying, sitting on a lawn. Another sentence which can be easily matched because of the visual text is ‘k’: “She isn’t on holiday. She’s on business”. The vectors show that the woman in Photo 3 is talking on the phone, which may be interpreted as doing business on the phone at the airport. Even her tidy clothes (a suit) and accessories (bag and cell phone) are important attributes which show she is a businesswoman, and not a tourist.

Considering the interactive metafunction, angles and levels of modalities are analyzed. As previously stated, Photo 2 is demanding an action from the viewers; so it is possible to classify this image as demand (Kress and van Leeuwen, 2006). Photos 1 and 3 are classified as offer since the represented participants do not look directly at the viewers; they are depicted at eye level, simulating a sense of power balance, that is, neither the viewers nor the represented participants have power over each other. Photo 2, however, is represented from a low angle, which makes the viewers feel that they have less power, since they are looking at the represented participants from a lower angle (Callow, 1999).

The demand on Photo 2 emphasizes viewers’ vulnerability, since someone who is represented as having more power (the represented participants) seems to be expecting something from the viewers who have, in this case, less power. In the textbook, it seems as if Photo 2 is telling students to find the correct corresponding sentences and match them to the images.

Image 3 presents high modality since it uses vibrant and naturalistic colors in the three selected photos of people’s experiences. This fact can influence the way EFL students may interpret these images. Since they are studying a language to be used in real situations in life, high modality may contribute to make them understand the scenes being portrayed. This can also influence the way they learn; it may suggest the feeling that they will use the language in realistic situations as well, connecting learning to the real world. Moreover, the thin framing in all three photos (but especially between Photos 2 and 3) suggests a possible integration between the represented participants, the learners of English and the different situations they may face.

The types of shot (long, medium or close) also tell a lot about how students may use the images to interpret

the textual part (that is, the lexicogrammatical choices) of the activity. Photos 1 and 3 are portrayed in a long shot, in which the viewers can see practically the whole body of the participants and identify the type of activities the represented participants are involved in. For example, the man in Photo 1 is probably an undergraduate studying on a lawn at a university campus. The woman in the airport is also portrayed in a long shot which makes it possible to imagine her interacting with someone by her cell phone. Her attributes (clothes, accessories, style) help viewers realize that she is on business, and the compositional part of the activity helps viewers to see the image as a coherent whole. Photo 2, on the other hand, is depicted in a close shot, which emphasizes the demand being portrayed to viewers.

The visual analysis shows that the interpretation of images may scaffold learners in understanding and doing the language activities proposed in the textbooks. The way the represented participants are depicted in all the images, angles, shots, modality, vectors, among other categories analyzed in the light of the visual grammar proposed by Kress and van Leeuwen (2006) may help learners to understand the content being worked with in the textual activities. The visual analysis of the three metafunctions may thus help foster learners’ possibilities of interpreting and understanding what is being studied in the language class.

Final remarks

It is important to teach students why and how image matters, considering that the visual meanings may foster students’ understanding of the content they are studying. In the analysis carried out in this paper, images reiterate the content being dealt with when they are analyzed and read following Kress and van Leeuwen’s (2006) GVD. Therefore, teachers who are instructed to use the ‘tools’ to approach images may scaffold students’ learning processes by directing them to read images, instead of only written texts.

In this sense, Kress and van Leeuwen’s (2006) GVD is fruitful to drive learners’ and teachers’ attention to the visual world around them. As discussed in this paper, the analysis of the representational, interactional and compositional metafunctions, in terms of participants, angles, vectors, salience, for example, suggests that these visual resources may play a very important role in scaffolding students’ understanding of the language content.

Although this paper has discussed a short preliminary study of images in textbooks, it aimed at contributing to the discussions regarding multimodal literacy and multimodal communicative competence (Royce, 2007). It also intended to raise teachers’ awareness regarding the importance of viewing the visual text as a resource to scaffold students’ learning processes. As a suggestion for

future studies, an analysis of how teachers use images to foster students' learning would add to the practicality of theories in the area of Visual Grammar.

References

- ALMEIDA, D.B.L. (ed.). 2008. *Perspectivas em análise visual: do fotojornalismo ao blog*. João Pessoa, Ed. da UFPB, 87 p.
- CALLOW, J. (ed.). 1999. *Image matters: Visual texts in the classroom*. Marrickville, NSW, Primary English Teaching Association, 125 p.
- CHRISTIE, F. 2005. *Language education in the primary years*. Sydney, UNSW Press, 248 p.
- CUNNINGHAM, S.; CHRIS, R.; MOOR, P. 2005. *New Cutting Edge Elementary*. Student's Book. London, Pearson Education, 175 p.
- HALLIDAY, M.A.K. 1994. *An Introduction to Functional Grammar*. 2nd ed., London, Edward Arnold, 434 p.
- HALLIDAY, M.A.K.; MATTHIESSEN, C.M.I.M. 2004. *An introduction to functional grammar*. 3rd ed., London, Arnold, 689 p.
- HAMMOND, J. 2001. Scaffolding and Language. In: J. HAMMOND (ed.), *Scaffolding: Teaching and Learning in Language and Literacy Education*. Newtown, Primary English Teaching Association, p. 15-30.
- HAMMOND, J.; GIBBONS, P. 2001. What is scaffolding? In: J. HAMMOND (ed.), *Scaffolding: Teaching and Learning in Language and Literacy Education*. Newtown, Primary English Teaching Association, p. 1-14.
- HEBERLE, V. 2010. Multimodal literacy for teenage EFL students. *Cadernos de Letras*, 27:101-116. Available at: http://www.letas.uftj.br/anglo_germanicas/cadernos/numeros/122010/textos/cl-30122010viviane.pdf. Accessed on: 08/08/2014.
- HERRELL, A.L.; JORDAN, M. 2012. *50 strategies for teaching English language learners*. 4th ed., Columbus, Pearson, 336 p.
- JORDÃO, C.M. 2014. ILA - ILF - ILE - ILG: quem dá conta? *Revista Brasileira de Linguística Aplicada*, 14(1):13-40.
- KRESS, G. 2000. Multimodality: Challenges to thinking about language. *TESOL Quarterly*, 34(2):337-340. <http://dx.doi.org/10.2307/3587959>
- KRESS, G.; van LEEUWEN, T. 1996. *Reading images: The grammar of visual design*. London/New York, Routledge, 288 p.
- KRESS, G.; van LEEUWEN, T. 2006. *Reading images: The grammar of visual design*. 2nd ed., London, Routledge, 291 p.
- LIMA, C.H.P.; PIMENTA, S.M.O.; AZEVEDO, A.M.T. (eds.). 2009. *Incurções Semióticas: teoria e prática de gramática sistêmico-funcional, multimodalidade, semiótica social e análise crítica do discurso*. Rio de Janeiro, Livre Expressão Ed, 314 p.
- MACKEN-HORARIK, M. 2004. Interacting with the multimodal text: reflections on image and verbiage in ArtExpress. *Visual communication*, 3(1):5-26. <http://dx.doi.org/10.1177/1470357204039596>
- NASCIMENTO, R.G.; BEZERRA, F.A.S.; HEBERLE, V.M. 2011. Multiletramentos: iniciação à análise de imagens. *Linguagem & Ensino*, 14(2):529-552.
- NEW LONDON GROUP. 1996. A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1):60-91. <http://dx.doi.org/10.17763/haer.66.1.17370n67v22j160u>
- RICHARDS, J.C. 1999. *Interchange. English for international communication*. Student's book A. Cambridge, Cambridge University Press, 134 p.
- RICHARDS, J.C.; HULL, J.; PROCTOR, S. 2001. *New interchange. English for international communication*. Student's book. 19th ed., Cambridge, Cambridge University Press, 209 p.
- ROYCE, T.D. 2007. Multimodal Communicative Competence in Second Language Contexts. In: T.D. ROYCE; W.L. BOWCHER (eds.), *New Directions in the Analysis of Multimodal Discourse*. Mahwah, Lawrence Erlbaum, p. 361-403.
- STENGLIN, M.; IEDEMA, R. 2001. How to analyse visual images: A guide for TESOL teachers. In: A. BURNS; C. COFFIN (eds.), *Analysing English in a global context. - A Reader*. London/New York, Routledge/Open University/Macquarie University, p. 194-208.
- UNSWORTH, L. 2001. *Teaching multiliteracies across the curriculum: Changing contexts of text and image in classroom practice*. Buckingham, Open University, 306 p.
- UNSWORTH, L. 2006. Image/text relations and intersemiosis: towards multimodal text description for multiliteracies education. In: International Systemic Functional Congress, 33, São Paulo, 2006. *Proceedings...* São Paulo, PUC-SP, 1:1165-1205.
- UNSWORTH, L. 2013. Re-configuring image-language relations and interpretive possibilities in picture books as animated movies: A site for developing multimodal literacy pedagogy. In: V. HEBERLE; F. VELOSO (eds.), *Ilha do Desterro*, 64:15-48.

Submetido: 09/12/2014
Aceito: 30/04/2015

Nayara Salbego

Universidade Federal de Santa Catarina
Programa de Pós-Graduação em Inglês
Campus Reitor João David Ferreira Lima, Trindade
8804-586, Florianópolis, SC, Brasil

Viviane M. Heberle

Universidade Federal de Santa Catarina
Programa de Pós-Graduação em Inglês
Centro de Comunicação e Expressão
Campus Reitor João David Ferreira Lima, Trindade
8804-586, Florianópolis, SC, Brasil

Maria Gabriela Soares da Silva Balen

Universidade Federal de Santa Catarina
Programa de Pós-Graduação em Inglês
Campus Reitor João David Ferreira Lima, Trindade
8804-586, Florianópolis, SC, Brasil