



CAMBRIDGE  
UNIVERSITY PRESS

# Visual literacy in English language teaching

Part of the Cambridge Papers in ELT series

August 2016

## CONTENTS

- 
- |   |  |
|---|--|
| 2 | The rise of the visual: Multimodal ensembles |
|---|--|
- 
- |   |                          |
|---|--------------------------|
| 4 | Defining visual literacy |
|---|--------------------------|
- 
- |   |   |
|---|---|
| 5 | Visual literacy and language learning:<br>A new role for the visual |
|---|---|
- 
- |   |  |
|---|--|
| 8 | Visual literacy and ELT materials design |
|---|--|
- 
- |   |  |
|---|--|
| 9 | Visual literacy in action: A framework |
|---|--|
- 
- |    |  |
|----|--|
| 11 | Implications for materials and<br>the language classroom |
|----|--|
- 
- |    |              |
|----|--------------|
| 12 | Bibliography |
|----|--------------|

# The rise of the visual:

---

# Multimodal ensembles

---

As language teachers, it is obvious to us that we should focus on the written word in our classes and in the teaching materials we use. However, these materials have too often ignored the visual aspect, treating it as something merely decorative, without educational value. More recently, there has been a shift in our understanding of the importance of the visual. Nowadays, it is increasingly common for English teachers, when assessing the pedagogical worth of some new material, to evaluate not only the presentation of grammar and vocabulary or the amount and type of skills work, but the visual aspects as well.

This shift towards the visual is, of course, reflected in many fields. After all, the printed word on the page was once our dominant medium of communication, but it is quickly being replaced by a combination of text and still or moving images viewed on increasingly smaller screens. As far back as 2003, Gunther Kress stated: 'The former constellation of *medium of book and mode of writing* is giving way, and in many domains has already given way, to the new constellation of *medium of screen and mode of image*' (Kress 2003:9, italics in original). This change of delivery mode has made communication via image often easier and more practical than via text. The digital age has brought us instant messaging services (e.g., WhatsApp), applications (Skype), social media sites (Facebook) or video-sharing platforms (YouTube), all of which contribute to this extraordinary rise in visual communication.

Likewise, our classrooms are becoming transformed. The increase in tablets equipped with user-friendly apps means that learners can easily carry out tasks involving images, which would have been very time-consuming and complicated in the past. For example, there is a huge number of apps or tools which can help learners in digital storytelling and creating presentations, comics, sketches, mind-maps, word clouds or their own short videos.<sup>1</sup>

This visual turn<sup>2</sup> does not mean the end of the written word — far from it. A fascinating advance is, in fact, the way in which still or moving images and design features combine with written text to create **multimodal ensembles**: '[entities which utilize] a variety of cultural and semiotic resources to articulate, render, represent, and communicate an array of concepts and information' (Serafini 2014:13).<sup>3</sup> Many of the digital creations that learners can produce using apps like those mentioned above are multimodal ensembles, because they consist of different media that contribute to creating multi-layered content that often appears simultaneously on screen.

We are surrounded by such ensembles in online environments. A [recent feature on the BBC News website](#)<sup>4</sup> exemplifies the features of these online multimodal ensembles. The page consists of video showing a 30-minute journey, speeded up to last only 3 minutes, along the extent of London Underground's Victoria Line. The video focuses on the route above ground, but in the corner you can also see synchronized footage from the driver's perspective below ground. As you watch the video, you hear voices in several of the different languages of the people who live along the route. (For example, the journey begins in Stockwell, where there is a large Portuguese and Brazilian community.) Importantly, the page is interactive in that you can jump to any station along the way and/or listen to the language of your choice as both are indicated in a menu bar that also shows the elapsed time.

---

1 Apps include Glogster, Voicethread (digital storytelling); Movenote, Prezi (presentations); Wordle (word clouds); Total Recall, Mindmeister (mind maps); and Skitch (sketches).

2 'Visual turn' refers to a shift in emphasis from verbal towards visual representation.

3 This term will be used throughout this paper to replace the more commonly heard term 'multimodal text', which perhaps overly emphasizes the importance of the written word.

4 'London's Underground languages' (BBC 2016). Available at: [www.bbc.co.uk/news/resources/idt-1e4fbc9-5bd6-4e14-adf5-bf3ab1ba79bb](http://www.bbc.co.uk/news/resources/idt-1e4fbc9-5bd6-4e14-adf5-bf3ab1ba79bb)



You can also scroll down to read supplementary texts, see images of the speakers, study maps showing the population of these communities and watch complementary videos giving background information about the people who live in these different neighbourhoods. The whole experience immerses you in the diverse cultures living in these parts of London.

Frank Rose, author of *The Art of Immersion* (a study of how digital technology is changing how we tell stories), has pointed out (Rose 2011:3) that what is emerging here is a new **multimodal narrative form**, which simultaneously combines text, audio and video 'in a way that's non-linear, that's participatory and often game-like, and that's designed above all to be immersive.'

A multimedia experience such as the London Underground one described above would have been unthinkable just a few years ago, whereas now we are accustomed to 'reading' this kind of rich visual information, effortlessly switching from one medium to the other. But how much of it do we really take in?

In order to understand such a multimedia experience and interact with it, we may need to have some prior knowledge of its content, but we also need to be **visually literate** to interpret the information and respond to it appropriately. This could be difficult, especially at first, as the delivery may be fast, the content multimodal and delivered simultaneously, and the concept or message not entirely clear at first.

**But what exactly do we mean by being 'visually literate'?**

# Defining visual literacy

The whole notion of literacy has been reassessed in these changing times. Traditionally, it was about acquiring a set of cognitive skills largely linked to print media. However, academics such as James Paul Gee (1996) have argued that literacy is actually a broader competence. All texts, regardless of the media through which they are delivered, are social artefacts that appear in particular contexts and practices, and thus need to be interpreted critically and creatively. Being literate today is about interrogating messages, however they are communicated.

This has led to the notion of **multiliteracies**,<sup>5</sup> an important distinction because it suggests that there are different types of literacy and that these literacies overlap and are evolving, especially in this digital age. Indeed, a recent book on digital literacies (Dudeney et al. 2014) identifies no fewer than sixteen distinct literacies that may be required today, from those based around information such as 'filtering literacy' to ones associated with digital creation such as 'remix literacy'.

Stephen Apkon, in his book *The Age of the Image*, defined literacy as 'the ability to express oneself in an effective way through the text of the moment, the prevailing mode of expression in a particular society. ... To be literate, in other words, is to be conversant in the dominant expressive language and form of the age.' (Apkon 2013:13). If this is the case, then it is 'visual literacy' on which we should be focusing today. This term has been defined in a number of ways since it was coined in the 1960s<sup>6</sup> and, although more recent definitions (e.g., Serafini 2014) vary, they have generally expanded to accommodate social as well as cognitive practices.

The consensus seems to be that we need 'a variety of lenses to interpret and analyze [images'] meaning potentials' (Serafini 2014:23). What is appealing about this definition is its connection with critical thinking, in its acknowledgement of the multiple meanings that can be generated by exposure to such diverse media — the 'variety of lenses' referring to the different layers of meaning that can be unravelled from looking in multiple ways. However, what perhaps needs adding here is the issue of **creative production and selection** which other definitions take into account, acknowledging the fact that today's world is as much about curation as creation. We interpret, select and create images to convey a wide range of meanings.

There is a strong connection here with today's participatory culture. Nowadays, it is customary for virtually anybody with an internet connection and a digital camera to set up their own blog or online video channel and upload their own content as a response to others. This establishes a crucial paradigm shift — today's literacies are about encouraging the audience not just to be passive consumers but active contributors of their own digital experience. Likewise, this shift is occurring in an education context, with the focus moving away from the display of an individual's knowledge or skills to that of interaction with others and collaborative learning. The learner thus feels able to introduce their own content into the classroom and not just receive what has been handed out to them.

5 The New London Group (Cazden et al.1996) first coined this phrase.

6 See Debes (1968).



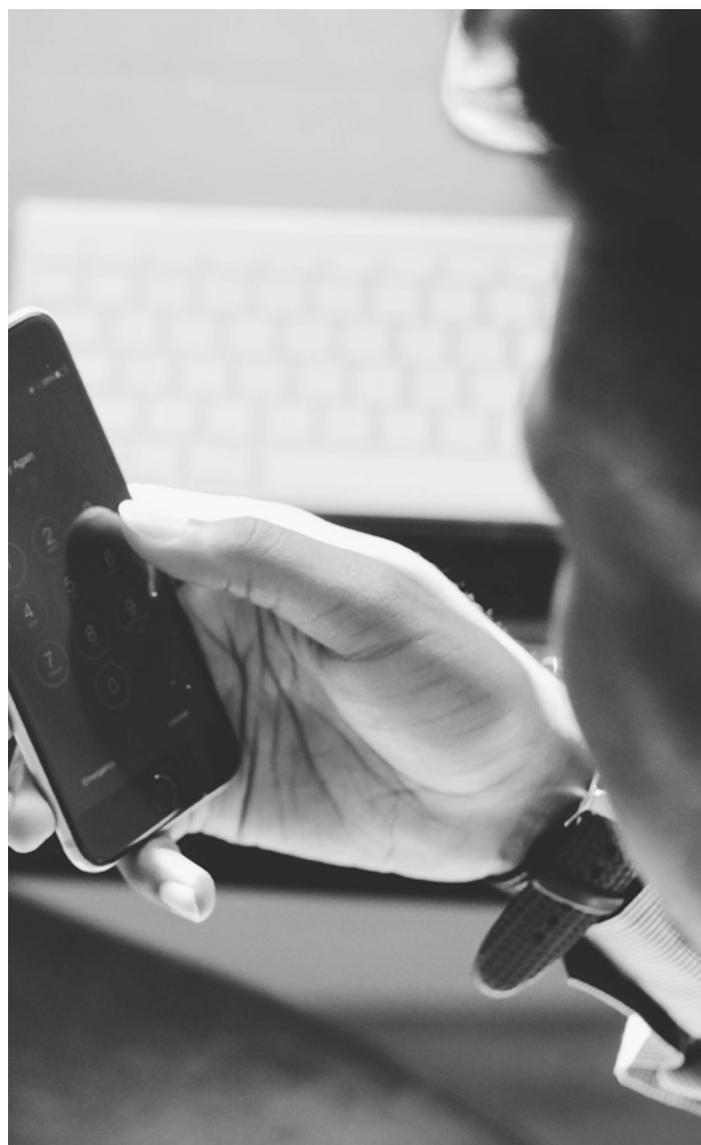
Other new video genres also exploit this relationship between image and text, such as infographic films, animated lectures, mash-ups and response films. For example, in the case of visual poetry, cognitively challenging texts can be made more accessible by incorporating a visual component, as can be seen in the animated version of the Billy Collins (1999) poem 'Forgetfulness'.<sup>7</sup>



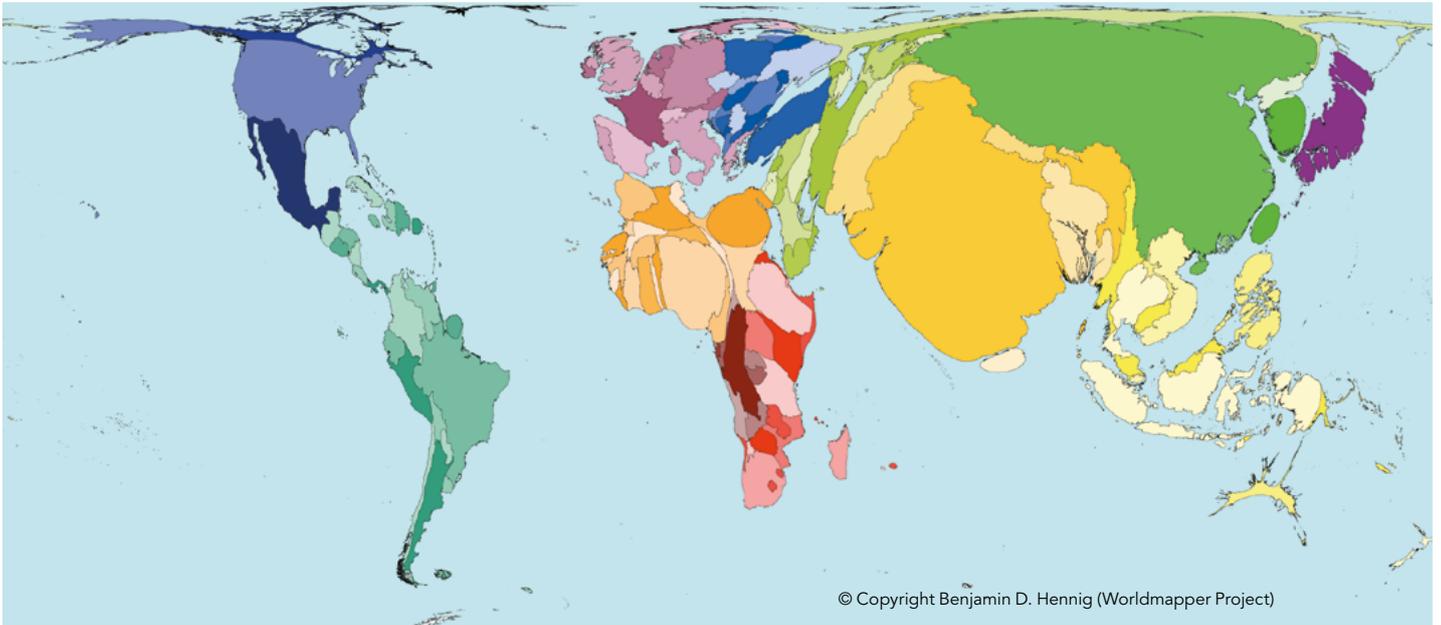
Another important point here is that if students are accessing these types of multimodal ensembles, surely we should attempt to create similar pedagogical engagement when designing lesson content? However, until recently, images were traditionally seen as either decorative or secondary to text. A classic example of this is a 'spot the difference' activity, in which students in pairs have two similar-looking images. They simply have to identify the differences between them and the task is complete. This is typical of an activity that has straightforward answers and in which the student can contribute few (if any) ideas of their own. This is how images have traditionally been exploited in exams as well. In these cases, the image is simply a visual aid to the practice of a very limited range of structures or lexical items. Students merely describe what they are expected to describe and cannot interpret or make the images their own.

It is not difficult, however, to shift the parameters and create tasks in which the image has a more prominent role and in which learners are engaged on a more cognitively challenging level. This is possible with all language learners, even those with a low level of proficiency.

One way to do this is to make the familiar unfamiliar, i.e., take a familiar image and subvert it in some way to create curiosity. For example, if a teacher wanted to present the vocabulary of countries and nationalities or languages, the logical way to do this is to use the image of a map as a visual aid. However, students have seen the classic political map of the world so many times that it may have little impact for them. Consider how many different types of maps now exist online, from interactive or personalizable ones to cartograms which show statistical data in map form.



<sup>7</sup> The full animated poem can be viewed at <https://vimeo.com/38356>. A collection of similar videos are featured on the YouTube channel which accompanies the book *Language Learning with Digital Video* (Goldstein and Driver 2014). Available at: [https://www.youtube.com/playlist?list=PL-oYKB0D9-E3eunEmqxr\\_42m8Pcrojb89](https://www.youtube.com/playlist?list=PL-oYKB0D9-E3eunEmqxr_42m8Pcrojb89)



For example, by choosing a cartogram<sup>8</sup> such as the one above, which shows the world's population density in map form, students can still identify the countries (and thus the vocabulary is covered), but they may need to look more carefully and reflect before they can work out what the map represents.

Questions that could be posed to students might include the following:

- *What does the map represent? How do you know?*
- *Which countries are most distorted in size?*
- *What can you learn from it?*
- *Who made it? What could it be used for?*

An alternative, more game-like, task could be undertaken with the image below.<sup>9</sup>



Questions that could be asked of the image might include:

- *Which countries are inside the circle and which outside?*
- *What is the map trying to contrast? Is it effective? Why/Why not?*
- *Think of an appropriate title for the map.*

These tasks show how an element of visual literacy is required in order to interpret the images successfully. In a sense, asking such questions requires learners to dig a little deeper, to look 'beyond the frame' of the image, in the same way that 'reading between the lines' is required in order to unpack written texts.

<sup>8</sup> A cartogram is a map in which some thematic variable, such as travel time or population, is substituted for land area or distance.

<sup>9</sup> This map shows an area of the world which has more people living within the green-shaded area than those living outside of it.

# Visual literacy and ELT materials design

The visual turn of recent years has led to important changes in the design of ELT materials. Previously, language coursebooks were filled with illustrations<sup>10</sup> but did not take the image seriously, either from a methodological or an aesthetic point of view. However, in an effort to resemble something other than language coursebooks and grant them a magazine-like layout, the size and role of images has increased and illustrations have virtually disappeared. At the same time, tie-ins with media organizations such as National Geographic, TED or Discovery Education have contributed to the prominence of visual material in ELT materials, especially video, which had never been previously integrated into course design.

However, it is still the case that there is a bias towards images which come from large digital image archive banks, rather than photo-journalism based archives. The problem with these images is that they are often used in advertising and can present an aspirational culture, full of attractive people enjoying comfortable lifestyles. These role models may alienate learners who cannot identify with such a reality.

Changes in materials design have been accompanied by a growing interest in all things visual in the teaching profession. In the last few years, the main ELT publishers have produced an increasing number of methodological handbooks for teachers on images and video.<sup>11</sup> At the same time, any number of websites and blogs have been set up with the aim of exploiting images and particularly video for the English language classroom.

Some have proven very successful, for example [film-english.com](http://film-english.com)<sup>12</sup> and [lessonstream.org](http://lessonstream.org)<sup>13</sup>, which largely exploit clips from Vimeo and YouTube respectively.

Meanwhile, from the academic point of view, The Image Conference<sup>14</sup> (an annual English language teachers' event devoted to images, video, film and gaming) was started in 2012 and has been held subsequently in Spain, Brazil and Germany with international speakers sharing their expertise in the field. In the 2014 Image Conference, the organizers issued what they called 'A Visual Manifesto' — a three-minute video clip in which critical reflection of the images in our daily lives is promoted as a methodological advance, as is image and video creation on the part of the learners, calling on skills acquired in their daily use of image and video. The authors explain their rationale (Donaghy et al. 2014) on the accompanying website:

'In the film, we emphasize the need to pause, reflect and think about how to be more deliberate instead of just reacting to today's media overload. ... We felt that by encouraging actual image making and filmmaking in the classroom setting, students will be given even more of an opportunity to practice the skills they are acquiring from their own personal usage of image-making devices and learn that through collaboration, they can actually create something greater.'

10 For example, drawings and cartoons (as opposed to the full range of image possibilities).

11 For example, Cambridge University Press published *Working with Images* (Goldstein 2008) and *Language Learning with Digital Video* (Goldstein et al. 2014). See also Donaghy (2015) and Keddie (2009, 2014).

12 Available at: <http://film-english.com>

13 Available at: <http://www.lessonstream.org>

14 Information available at: <http://theimageconference.org/about>

# Visual literacy in action:

---

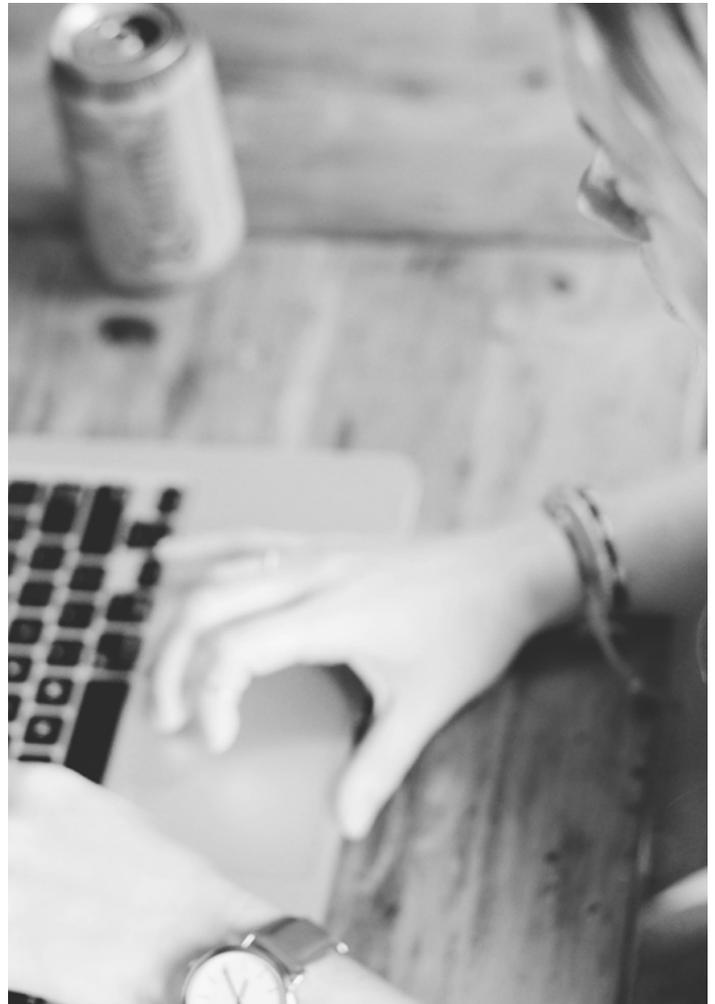
## A framework

---

This brings us to the particular skills and abilities that a visually literate learner can claim to possess. This is perhaps best understood by appreciating how we view an image in different ways.

Callow (2005) established an interesting and accessible paradigm for ways of viewing an image and the checklist of questions below is based on this model. His three dimensions of viewing are **affective**, **compositional** and **critical** (Callow 2005:13). The affective focuses on the individual's sensual and immediate response; the compositional includes the identification of semiotic, structural and contextual elements; and the critical foregrounds sociocultural considerations and highlights the image's message, requiring us often to look beyond an image's literal meaning. Callow quotes Kress and van Leeuwen's assertion (1996:12) here: 'All images, even apparently neutral ones, are entirely in the realm of ideology'.

This paradigm is echoed in the model proposed more recently by Serafini (2014:43), who defines the dimensions of viewing as **perceptual**, **structural** and **ideological**.<sup>15</sup> Based on these models, we can establish certain questions<sup>16</sup> which could be posed to learners when they are confronted by any still or moving image. This is helpful as visual literacy is a concept which many of us would agree should be incorporated into our classroom practice but which, at times, can be hard to systematize.



<sup>15</sup> Serafini's paradigm was highly influenced by Kress and van Leeuwen's own scheme of visual grammar which is laid out in *Reading Images* (1996).

<sup>16</sup> Some of these questions have been selected from a list in *Language Learning with Digital Video* (Goldstein et al. 2014:11–13).

## Questions and tasks to help learners analyze images

### Three dimensions of viewing

AFFECTIVE/ PERCEPTUAL	COMPOSITIONAL/ STRUCTURAL	CRITICAL/ IDEOLOGICAL
How does the image make you feel?	What elements can you see in the foreground/background, etc.?	What message does the image transmit?
Why does it make you feel this way?	What text accompanies the image, if any (a caption, a title, etc.)? What does it add to the image?	Who created it? For what purpose and in what context?
What other images come to mind when you see it?	How is the image framed or composed?	In what forms of media will the image be seen?
What personal relevance does it have for you, if any?	What do you think lies beyond the frame?	Who is the intended audience for the image?
What does the image remind you of?	From what angle or point of view has the image been taken?	In what context did you view the image? The original context or another one? What is the difference?
Do you identify with or relate to the image in any way? If so, how?	Which parts of the image are centrally focused?	In how many different ways could the image be interpreted?
Do you think the image is positive or negative? Or do you feel indifferent towards it? Why?	What has been altered, omitted from or included in the image?	Are any of the images stereotypical, idealized, non-representative or anachronistic?

To complete this list, we need to consider the active viewer as well and engage the students' creative or curative responses to the image. Possible activities which call on this skill could be the following:

- Embed the image in a different context or genre.
- Rewrite the accompanying text or script as if it were a different genre. What changes take place?
- Sequence images in a different order to create your own narrative.
- Design or visualize your own video sequence for a song or poem, etc.
- Narrate a live video sequence.
- Identify the subtext of a particular image or video sequence and present this to others.
- Research and find similar images or videos online.
- Create a collage or mash-up of different images or videos related to what you have seen in class (e.g., create a mosaic of different images which say something about yourself).
- Create a storyboard trailer for a movie and then compare it with the original.

# Implications for materials and the language classroom

Exploiting visual literacy can have a number of implications both for materials and the classroom. Firstly, with regard to materials, focusing on the image for its visual stimulus and moving away from comprehension-based approaches (which rely on text) will enable students to engage with images in a number of different ways. The more visually-oriented learners will benefit from this approach and may be able to contribute more in class when previously they had been silent. This also means that the same piece of visual material can be used for classes of different levels of language proficiency. Free from having to understand the language input, the materials producer can grade the tasks rather than the input itself.<sup>17</sup>

However, although the focus on the visual is important, to do this exclusively and ignore the textual can be a mistake. A number of websites feature activities in which students simply describe what they have seen in a high-impact silent video sequence, but there is very little else in the way of exploitation and little thought for how to extend this to generate a creative response in the learner, although creating their own monologue or accompanying text is clearly an option. As explained above, rather than seeing video and images as separate entities, a focus on multimodal ensembles would seem a more sensible approach to take. Likewise, seeing image and video as being springboards for debate and discussion, as well as presentations and project work, is a way of integrating them into other tasks done in and outside a language class.

With regard to classrooms, clearly, the increased use of video will change how and where learning is carried out. It is perfectly possible now for students to engage with materials and interact with peers in online platforms on many different devices.

The role of the visual is intrinsic to this shifting concept of classroom spaces. In most blended learning<sup>18</sup> programmes, the visual material accessed at home via video in their virtual study space acts as a way to engage learners and reflect on ideas and concepts outside of class. They then have the chance to share this with peers and teachers in the physical classroom space, contributing more to the input and, to a certain extent, changing the role of teacher and learner. However, this visual material cannot simply be dished out to students in a haphazard fashion with the reasoning that 'kids today understand images'. The materials need to have accompanying questions and/or notes which guide the learner and help structure their responses.

It may well be that learners today are far more visually literate than the previous generation because they have so much greater access to visual material. However, these learners may not be fully aware of this, because they have not needed to call on this literacy in an academic context but rather in contexts of entertainment. It is therefore up to those who mediate the learning process, such as language teachers and materials writers, to access motivating material and to ask the right kinds of questions about it in order to engage learners and tap into this literacy. Exploiting visual literacy has the potential to trigger any number of enlightened responses and interpretations. Language learning today has to take into consideration that the messages we receive on a daily basis do not just arrive to us via the written word but in multimodal ensembles that need analyzing and unpacking. Designing materials to embrace these ensembles and all things visual will be one of the most stimulating challenges we face in the future.

17 An approach commonly known as 'grade the task, not the text'.

18 'Blended learning' refers to any course which appropriately combines both synchronous interactive study (usually face-to-face) and asynchronous individual study (usually online).

# Bibliography

- Apkon, S. (2013). *The age of the image: Redefining literacy in a world of screens*. New York: Farrar, Strauss and Giroux.
- Bamford, A. (2003). The visual literacy white paper. [online] Available at: <http://www.images.adobe.com/content/dam/Adobe/en/education/pdfs/visual-literacy-wp.pdf> [Last accessed 10/06/2016]
- BBC (2016). London's Underground languages. [online] Available at: [www.bbc.co.uk/news/resources/idt-1e4fbc9-5bd6-4e14-adf5-bf3ab1ba79bb](http://www.bbc.co.uk/news/resources/idt-1e4fbc9-5bd6-4e14-adf5-bf3ab1ba79bb) [Last accessed 10/06/2016]
- Callow, J. (2005). Literacy and the visual: Broadening our vision. *English teaching: Practice and critique*, 4(1), pp. 6—19.
- Cazden, C., Cope, B., Fairclough, N., Gee, J., Kalantzis, M., Kress, G., Luke, A., Luke, C., Michaels, S. and Nakata, M. (The New London Group) (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), pp. 60—92. Also available at: [http://newarcproject.pbworks.com/f/Pedagogy%2Bof%2BMultiliteracies\\_New%2BLondon%2BGroup.pdf](http://newarcproject.pbworks.com/f/Pedagogy%2Bof%2BMultiliteracies_New%2BLondon%2BGroup.pdf) [Last accessed 10/06/2016]
- Collins, B. (1999). *Forgetfulness. Questions about Angels*. Pittsburgh: University of Pittsburgh Press.
- Debes, J. L. (1968). Some foundations for visual literacy. *Audiovisual Instruction*, 13(9), pp. 961—964.
- Donaghy, K. and Whitcher, A. (2014). A visual manifesto for language teaching. [online] Available at: <http://visualmanifesto.com> [Last accessed 10/06/2016]
- Donaghy, K. (2015). *Film in action: Teaching language using moving images*. Peaslake, Surrey: Delta Publishing.
- Dudeney, G., Hockly, N. and Pegrum, M. (2014). *Digital literacies*. London: Routledge.
- Film English. [online] Available at: [www.film-english.com](http://www.film-english.com) [Last accessed 10/06/2016]
- Gee, J. P. (1996). *Social linguistics and literacies: Ideology in discourses*. London: Taylor and Francis.
- Goldstein, B. (2008). *Working with images*. Cambridge: Cambridge University Press.
- Goldstein, B. and Driver, P. (2014). *Language learning with digital video*. Cambridge: Cambridge University Press.
- Keddie, J. (2009). *Images*. Oxford: Oxford University Press.
- Keddie, J. (2014). *Bringing online video into the classroom*. Oxford: Oxford University Press.
- Kress, G. (2003). *Literacy in the new media age*. London: Routledge.
- Kress, G. and van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London: Routledge.
- Lesson Stream. [online] Available at: <http://www.lessonstream.org> [Last accessed 13/06/2016]
- Rose, F. (2011). *The art of immersion: How the digital generation is remaking Hollywood, Madison Avenue, and the way we tell stories*. New York: Norton.
- Serafini, F. (2014). *Reading the visual: An introduction to teaching multimodal literacy*. New York: Teachers College Press.
- The Image Conference. [online] Available at: <http://theimageconference.org> [Last accessed 10/06/2016]

---

**Ben Goldstein** has taught English for over 25 years in the UK, Spain and Hong Kong. He currently teaches on The New School's online MATESOL program (New York). He is co-author of the secondary series *Eyes Open* and *Uncover*, and of the adult coursebook *English Unlimited Advanced*. He has also published the teachers' methodology handbooks *Working with Images* and *Language Learning with Digital Video* (all the above titles published by Cambridge University Press). His main interests lie in visual literacy, intercultural and identity issues and English as an International Language.

[www.bengoldstein.es](http://www.bengoldstein.es)

---

To cite this paper:

Goldstein, B. (2016). *Visual literacy in English language teaching: Part of the Cambridge Papers in ELT series*. [pdf] Cambridge: Cambridge University Press.

Available at [cambridge.org/betterlearning](http://cambridge.org/betterlearning)