

Applying Multimodal Discourse Analysis to Learning Objects' User Interface

George Vorvilas
Thanassis Karalis
Konstantinos Ravanis
University of Patras, Greece

Abstract

This article presents a framework of semiotic analysis that could be used for interpreting learning objects. Many learning objects are multimodal representations that aim at servicing specific educational objectives. Consequently, an urgent need arises to know what kind of meanings these representations produce and what kind of pedagogic relationships are shaped between students and them. Taking a concrete learning object as an example, we deploy a sample of multimodal discourse analysis in order to elucidate these issues. Finally, we conclude with a few thoughts about the possibility of elaborating such a framework in relation to the effective design and the implementation of learning objects.

Keywords: *Learning objects; Multimodal discourse analysis; Multimodal representations; Ideational metafunction; Interpersonal metafunction; Textual metafunction.*

Theoretical Framework

The main idea for designing and implementing artefacts called “learning objects” is the creation of digital educative items that can be reused in different digital educational contexts (Churchill, 2006; Polsanyi, 2003; Wiley, 2002). These digital entities can be accessible from anyone in the World Wide Web (www) through repositories in which they are looked after and are recognizable through metadata which describe their attributes or their context of use (e.g. title, format, learning resource type etc.). From a technical perspective, the two fundamental advantages of using learning objects in e-learning are supposed to be:

- The reduction of learning resources’ cost production, since the re-use of the same resources is allowed, thus avoiding the repeated and costly accumulation of educational material for each training circumstance.
- The re-use of learning resources that can satisfy various needs of the teachers and students at several educational situations.

However, the pedagogic approach of using learning objects has been criticized for its perseverance in a behavioral as well as mechanistic transmission of knowledge and also for its supposed context-independent and value-neutral learning content (Butson, 2003; Friesen, 2004; Lim, 2007; Parrish, 2004). Moreover, it has been pointed out that great emphasis is given in the technological standardization of learning objects and the concomitant economical

benefits from their use, while few studies examine the social-historical, pedagogic, educational and anthropological dimensions of this use (Friesen & Cressman, 2007). Recently, a more promising approach has associated learning objects with learning design strategies, in an effort to achieve their educational contextualization (Lockyer, Bennett, Agostinho, & Harper, 2009).

These new forms of knowledge representation require further understanding and skills based on visual literacy so much for teachers as well as for students. An effective pedagogic use of learning objects should take into account the different meanings produced while reading a text, viewing static or dynamic images, hearing sound extracts, as well as the specific dynamics which emerge through their combination. From this multiple perspective, significant concepts such as multiliteracy and multimodality have emerged in the last few years. Multiliteracy concerns the perceptual abilities and skills required for the intelligibility of the variety which characterizes several semiotic systems. Multimodality concerns the modes through which such a variety expresses itself and permits the meaningful organization of semiotic systems (Baldry, 2000; Cope and Kalatzis, 2000; Jewitt and Kress, 2003; Kress & van Leeuwen, 2001).

Even though particular emphasis is placed on the creation of effective user interfaces for learning objects (Black et al, 2007; Notargiacomo et al, 2007; Simbulan, 2007), through the combination of instructional design theories, principles of multimedia design and design for e-learning in general (Clark and Mayer, 2008), there does not as yet exist a systematic examination of the production and the educational use of learning objects with the lens of social semiotics, and the conclusions the Social Semiotics draws from the aspects of multiliteracy and multimodality which characterize modern educational environments, digital or non digital (Baldry, 2000; Dimopoulos, Koulaidis, & Sklaveniti, 2003; Unsworth, 2006; Unsworth et al, 2005).

Social semiotics studies the practices in which people are involved in order to create and communicate meanings to each other in various social environments. Being rooted in Systemic Functional Grammar (Halliday & Matthiessen, 2004), Social Semiotics considers that linguistic as well as non-linguistic semiotic systems are organized and described through three fundamental metafunctions: (a) the ideational metafunction which describe the way in which various semiotic resources are represented and interconnected with each other; (b) the interpersonal metafunction which describes the relations developed between the addresser of the resources and the addressee; (c) the textual metafunction which describes the different ways by which semiotic resources produce cohesive multimodal texts and meanings.

The basic hermeneutic tool of social semiotics is multimodal discourse analysis which examines the above metafunctions in various socio-cultural fields such as the interpretation of images (Kress & van Leeuwen, 2006), advertisement in television (Thibault, 2000), documentaries (Iedema, 2001), speech, music, and sound (van Leeuwen, 1999) movement and gestures (Martinec, 2004), scientific discourse (Levine & Scollon, 2004), art and architecture (O'Toole, 1994), World Wide Web (Djonov, 2007; Lemke, 2002), and literacy practices (Unsworth, 2006). Multimodal discourse analysis examines how multiple semiotic resources in the aforementioned discursive fields are combined in order to create particular kinds of meanings.

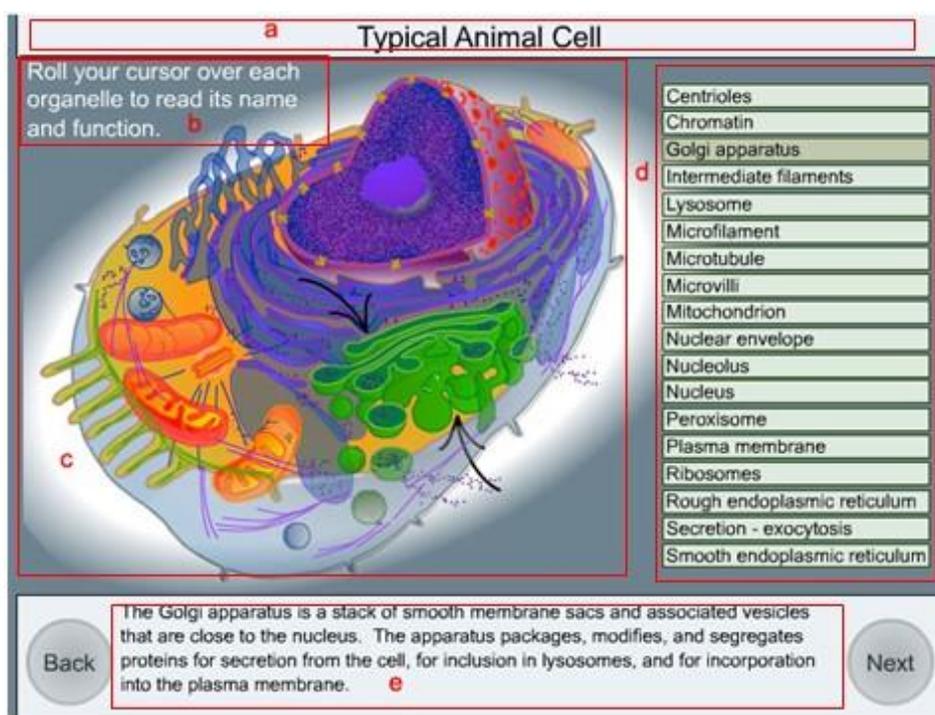
The contribution of social semiotics in the field of learning objects could initially concern itself with two fundamental questions:

- 1) In which manner multimodal representations are organized on the two-dimensional surface of screen in cohesive meaningful wholes?
- 2) What kind of pedagogic relations are created between students and learning objects due to the cohesive meaningful wholes contained in learning objects?

In this article we hope to contribute towards an outline of a conceptual framework for creating and using learning objects, in which an appropriate semiotic meta-language (aimed on how the several multimodal digital resources are organized) could be fabricated. Such a meta-language could offer on the one hand a set of sound and pedagogically appropriate choices in designing learning objects that could be useful for writers-designers, and on the other hand it would equip teachers with a vocabulary for choosing effective and reliable learning objects. Due to the obvious restrictions of space we will limit our analysis in elaborating the application of concepts that have been applied in the field of still images and written text and we will leave aside the analysis of sound representations and filmic text.

The Organization of Meanings on Learning Objects' User Interface

The aforementioned theoretical background imposes the adoption of concrete directions for the semiotic analysis of a learning object. In the next sections we will try to draw the results of such a multimodal analysis through the elaboration of a simple learning object called: 'Typical Animal Cell' (pic.1).



Picture 1. Second screen of the LO: 'A typical animal cell' (source: <http://www.wisc-online.com>)

This object consists of four screens. The first screen reports the learning objective of this learning object which is the recognition and identification of the organelles of an animal cell as well as their operations. In the second screen, which is the one we analyze at great length

(pic.1), the user that moves the cursor over the animated visual depicting the cell and its parts, can see the name of each organelle in the corresponding labels to the right side of the screen (whose color changes slightly becoming dark grey) as well as information about its operation, in the alternating text to the bottom part of the screen. After clicking on the 'next' button, the user is guided to the third screen in which she should do a memorization exercise: Here the animated visual represents the organelles separately, almost distincted from their environment (the cell) and the student is asked to identify them by using a suitable name from the alphabetically arranged column of terms which is displayed always to the right side of screen. The fourth screen informs the student that the activity was completed and that he/she can visit again the initial screen if he/she wishes, the names of learning object's authors are also reported. For a better description we have deliberately separated the second screen in five blocks with the names a, b, c, d and e respectively. The meanings which are developed in terms of the three metafunctions which characterize the semiotic systems are analyzed below.

Ideational Metafunction

At the ideational metafunction level we are interested in knowing the types and functions of various visual elements according to which elements (humans, animals, places, symbols, etc) and their properties are represented. Dimopoulos et al. (2003) distinguish three types of images: realistic, conventional, and hybrid. Realistic images depict their elements through representations which approximate human optical perception like photographs or drawings. Conventional images adopt a peculiar, abstract pictorial symbolism familiar to the scientific field in which they belong such as maps, diagrams, charts etc. Hybrid images combine realistic and conventional elements like the animated image of block c whose represented element is an animal cell.

According to their functions, images can be separated in four main categories involving narrative, classificational, analytical, and symbolic representations (Kress & van Leeuwen, 2006). In narrative representations, the represented elements relate to each other through depictions of change, process and action (often changes and actions that relate such elements are represented through real or imaginary vectors). In classificational representations many elements are arranged on the same surface in overt or covert relations of hierarchy or equality so as to exhibit the common features that group them or the class in which they belong. In analytical representations the relations that characterize represented elements have a whole-part structure, where a main element is analyzed at the attributes which constitute it. Finally, in symbolic representations a main element acquires its identity and meaning through other symbolic elements which carry its attributes. The symbolic and non-literal value carried by these elements is validated from image's cultural context.

In our example, the animated image in block c with the contribution of block d, constitutes an analytical representation characterized by whole-part relations: A main element (the animal cell) is represented, as well as its attributes (the organelles) whose identification has been left to the user.

Apart from the types and functions of the various optical representations, ideational meanings result also from image-text interactions. This meanings are described through the interdependency relations as well as the logico-semantic relations that are developed between image and text (Kong, 2006; Martinec & Salway, 2005; Unsworth, 2007). Interdependency relations are distinguished in relations of parataxis and hypotaxis. When image and text are of

unequal status, then image subordinates to text or vice versa, in this case we have a hypotactic relation. When image and text are of equal status, then they complement or diverge each other, in this case we have a paratactic relation.

In our example, we can see a case of paratactic relation between the text in block a and the animated image in block c: the text "Typical Animal Cell" restates in words what the entire animated image represents, functioning in a complementary way towards the constitution of a greater syntagm. Hypotactic relations exist between the animated image and the labels of block d as well as between the image and the text of block e: the animal cell's image is more general, furthermore, the term 'Gogli apparatus' (block d) as well as its description and function (block e) correspond to a concrete part of this image (whose identification is facilitated through the use of indexical vectors), not to its whole. Thus, text subordinates to image.

Logico-semantic relations are expressed through the subsystems of expansion and projection. Expansion shows how the meaning of a text or image expands through the three categories of elaboration, extension and enhancement. In elaboration an element (image or text) expands the meaning of another one without providing new information about it, by describing it, clarifying it, restating it or specifying it. In extension an element expands the meaning of another one by adding new information, giving an exception to it or offering an alternative. In enhancement an element expands the meaning of another one by enriching it with new information through circumstantial features of time, place, purpose, cause, condition, manner, motivation etc.

In projection the meaning of an element appears through another element either as idea or locution. When the second element of relation (image or text) represents thoughts, projection is mental. When it represents speech, projection is verbal. Characteristic cases of verbal and mental projections constitute the balloons and clouds we found in comics, which express the speech and thoughts of the represented persons or animals respectively.

Several logico-semantic relations exist between text and image in our example. Thus, an elaboration relation can be detected between the sentence: "The Gogli apparatus is a stack of smooth membrane sacs and associated vesicles that are close to the nucleus." in block e, and the intensely colored part of the image to which it refers. Here, the text restates in words what this concrete visual element depicts, without adding new information. On the contrary, the relation between the sentence: "The apparatus packages, modifies, and segregates proteins for secretion from the cell, for inclusion in lysosomes, and for incorporation into the plasma membrane." The above mentioned visual element, is an extension. Here, text adds new information that is not traced in the depicted item. Finally, the relation between blocks b and c is enhancement: The text in block b enhances the animated image by prompting the student to roll the cursor over each one of the organelles that are depicted in it (motivation), in order to obtain the required information (thus a purpose is implied).

Interpersonal Metafunction

At the interpersonal metafunction level, we are interested in examining the ways in which represented elements interact with students. Kress and van Leeuwen (2006) distinguish three parameters that are dealt with interpersonal meanings: contact, social distance, and attitude.

Contact with the viewer is accomplished through the function of images as 'image acts' which request her attention (demand-images) or offer to her visual information (offer-images). In our example, block c creates contact with the student functioning as offer-image which provides information on the structure of the animal cell. On the contrary, language uses more communicative acts for accomplishing contact during social interactions, by offering goods and services through offers, demanding goods and services through commands, giving information through statements, demanding information through questions (Halliday & Matthiessen, 2004). The verbiage that accompanies the image of our example creates contact with the user by giving information through blocks a, d, e and by demanding services through block b.

Social distance that images create can be categorized as personal, social, and impersonal. At the level of visual display it is expressed by the 'frame size' of shots: A close-up expresses an intimate relation between the viewer and the image (e.g. a picture depicting a person's head and face only), a medium shot expresses social distance between them (e.g. the same person portrayed from the waist up), while a long shot expresses an impersonal relation (e.g. torsos of several people with space around them). Block c in picture 1 creates a social distance with the student, the cell is depicted as if it is relatively near the spectator (medium shot) so that she can observe it carefully.

Social distance that language creates can also be personal, social, and public. Personal style in language is presented through a sparing syntactic structure with many idioms and a dependence on the intimate context of situation in which this language is developed; briefly it can be equated with Bernstein's 'limited code'. Social style corresponds to the daily level of professional and social interactions and requires a standard syntax along with a precise vocabulary due to the social distance between addresser and addressee. Public style requires a clarified and precise adaptation of the message in a context of situation in which an impersonal distance between transmitter and receiver is imposed. It corresponds to Bernstein 'elaborated code' (Macken-Horarik, 2004). In block b of our example, verbiage is in personal style, while in block e it is rather in public style, hence it strengthens the distance between addresser and addressee.

Attitude refers to the relations of power and involvement that are developed between viewers and images/texts. At the level of images, attitude is expressed according to Kress and van Leeuwen (2006) through the horizontal and vertical angle of shots. More concretely, viewers' involvement in the image is expressed through the horizontal frontal angle while his detachment is expressed through the horizontal oblique angle. The power relations between the spectator and the image are expressed through three representational choices: High vertical angle which expresses viewer's power over the image, low vertical angle which expresses image's power over the viewer, eye level angle which expresses equality between them.

Dimopoulos et al. (2003) have interpreted the pedagogic relationships which are created between visual representations and students by relating the aforementioned framework to Bernstein's notion of 'framing'. It refers to the degree of control that teachers and students exert upon the educational process (Bernstein, 1996). When framing is strong, control is exerted from the addresser's side (e.g. teacher, educational material, learning object). When framing is weak, control belongs to the addressee (student). The pedagogic relationships are considered as relations of power and involvement. In power relationships strong framing results from low angle representations through which images are imposed on the student. In moderate framing student and image have an equal relationship through the presence of eye

level angle representations, while in weak framing student imposes over the image through the presence of high angle representations. In involvement relationships, framing is strong when there is an oblique angle and a medium or distant shot (thus student's involvement with the image content is minimum), it is moderate when there is an oblique angle and a close shot or frontal angle and distant shot (thus a moderate involvement with the image content is prompted), and finally it is weak when there is a frontal angle and a close or medium shot (student can involve with the image content in a considerable degree).

The oblique angle and the close shot of the animal cell's representation rather increases student's involvement with the image as well as the control from his/her side (high angle), allowing him/her a considerable degree of freedom in observing the represented elements. The relationships of power are characterized here by weak framing, while the relationships of involvement are characterized by moderate framing.

Dimopoulos et al. (2005) have also connected the concept of framing with particular lexicogrammatical choices that generally shape the pedagogic relationships of power and involvement. More specifically, in power relationships framing is strong when the addresser (teacher or educational material) takes control of the learning process through imperatives. A moderate framing exists when the addressee is supported with some options in answering interrogatives that have been posed by the addresser, while a weak framing operates through the presence of declaratives which denote a less clear authority of the learning content. In involvement relations, framing is strong when the text denotes in an explicit manner the conditions of student's involvement through the use of second singular person (you), it is moderate when the text presents a less clear picture of involvement's conditions through the use of first and second plural person (we, you), while it is weak when the text rather focuses on the learning content itself than on the communicating agents, through the use of the third singular or plural person (he, she, it, they).

In our example of picture 1 we can detect power relationships, particularly in blocks a and e, where texts consist of declarative sentences implying weak framing, and in block b where an imperative sentence exists, implying a strong framing. A weak framing of involvement is also provided by blocks d and e which are focusing on the learning content itself rather than on an interaction with the student. Finally, block b explicitly denotes student's involvement with the image, thus a strong framing is provided.

Textual Metafunction

At the textual metafunction level we are interested in the way images and texts are spatially arranged, highlighted, divided etc. to create larger visual compositions of ideational and interpersonal meanings. Kress and van Leeuwen (2006) distinguish three interrelated variables that contribute to this direction:

1) *Informational value*. Visual elements, depending on the structure of the visual composition, have specific informational value. Thus, in Given/New structure the initial information is placed on the left side of the two-dimensional surface (printed text or screen), through an image or a text, while the new information is placed on the right side, through an image or a text. In Centre/Margin structure a prominent element is placed in the center of the surface giving meanings to the secondary elements placed around it. In Ideal/Real structure a visual element with abstract and idealized aesthetic value is placed on the top side of the surface, while

lexicogrammatical elements that lend factual information in the whole composition are placed on the bottom side of the surface.

In our example, informational value is produced through Given/New and Ideal/Real structures. A Given/New structure can be discerned between blocks c and d. Block c carries the given pictorial information while block d constitutes the new information by establishing a relational identifying process between each name (identifier) and its pictorial correlate (identified). An Ideal/Real structure results from the array of elements between blocks c, d and e. Here, the animated image with the contribution of block d, functions as an idealized element with great aesthetic value, while the descriptions of organelles' functions in block e constitute the realistic and factual information which is given to the student.

2) *Framing*. The term here refers to the visual compositions' layout, not to Bernstein's notion. More concretely, it refers to the way in which various elements of the visual composition are connected to each other or are disconnected, through frame devices, creating cohesive meanings. Some of these devices are: (1) Segregation: The represented elements remain separated in different parts and in different order; (2) Separation: The elements are separated from each other via an empty space which, although keeping them in distance, connotes their potential resemblance; (3) Integration: Image and text occupy the same space, thus their natural connection is implied; (4) Overlap: Where image and text frames are partially mixed, thus causing proximity; (5) Rhyme: Elements in different frames are connected to each other via qualities of color, shape, posture etc, implying sameness; (6) Contrast: The qualities of various elements like color, posture, size, etc. are accentuated so that the difference between these elements can be emphasized better (van Leeuwen, 2005).

Certain framing devices contribute to the creation of textual meaning in the second screen of our learning object. For example, a segregation relation characterizes blocks a, d and e, implying their different function in the whole visual composition, but there is rhyme between them through the filling of their frames with the same grey color, implying their sameness due to the informative content they convey (terms and descriptions). The labels that constitute block d are connected to each other with a separation device: They are separated by an empty space which keeps them in distance but allows them to maintain a relative resemblance, implying that all these discrete elements have something in common, that is, they are subordinated under a more general concept. We can also discern an integration device between blocks a and e; all of them possess the same background space implying their noematic connection.

3) *Saliency*. Some elements of the visual composition have been designed as the most salient in order to catch the viewer's attention. This is accomplished through the use of big size, intense and rich color, tone (more brightness), focus (eliminating background), foregrounding or overlapping of visual elements (Machin, 2007). For example, the salient element of the visual composition in picture 1 is the animated image, which catches the viewer's attention through its big size and its intense and reach colors.

Conclusion

In this article we briefly presented and applied in an exemplary fashion a multimodal discourse analysis framework on a learning object. We tried to answer the two fundamental questions posed in our introduction: What kind of meanings do learning objects create as multimodal

representations of knowledge and what kind of pedagogic relationships these meanings develop between learning objects and students?

Answering the first question we saw that three categories of meanings are produced: Ideational meanings through the way learning object's elements are represented, textual meanings through the way distribution of informational value and emphasis among the textual and visual elements of the learning object are organized, and interpersonal meanings through the way verbal and visual resources construct the nature of relationships among addresser/addressee. Answering the second question we found that the above mentioned interpersonal meanings produce particular pedagogic relations of power and involvement. Also, the notion of framing which determines the degree of pedagogic control among addresser/addressee has been demonstrated.

A conceptual framework for the design and use of learning objects for e-learning should take into consideration the above mentioned types of meanings and pedagogic relationships. From this perspective, the socio-cultural as well as educational context in which learning objects will be used is a matter of great importance (Karalis, Sotiropoulos, & Kampeza, 2007; Sotiropoulos, 2003). Instead of a technological focus on learning object's creation, such a framework could be addressed to the semantic/pragmatic dimensions of the relations among their structural elements (video, audio, image, text) and these contexts.

A first step towards this direction could be facilitated by the notion of framing. Beyond the description of pedagogic relations of power and involvement, the notion also refers to the rules structuring the learning procedures in a learning context (content selection, sequencing, assessment etc). Thus, strong framing indicates that the addresser explicitly regulates the content, sequencing, pacing and assessment that constitute the learning context. This could characterize for example a drill and practice application that teaches a procedure. Here student's options should be deliberately restricted; the control should belong to the procedural, step by step organization of several content elements. The assessment should also, in this case, be strongly framed through questions in the form of true/false, multiple choices, fill in the gap etc.

On the contrary, relatively weak framing indicates that the addressee has increased and apparent control in respect to sequence, pacing and assessment. This could characterize, for example, a problem solving or a case study application that students are prompted to engage in. In such a context, a large part of the educational control can be given to students whose options are reinforced. Assessment could also be based on open questions, without predetermined answers, whose aim is to promote learners' critical thought rather than guided response.

Therefore, a modest constructivist approach for designing and using learning objects could be adopted, which takes into consideration so much the necessity of guidance during the learning process (Kirschner, Sweller, & Clark, 2006; Mayer, 2004; Ravanis, 1996; Ravanis, 1999) as well as the importance of scaffolding and collaborative construction of knowledge (Cindy, Ravit, & Clark, 2007; Matthews, 2005; Ravanis, 2005). In the proposed conceptual framework, the affordances of learning objects as artefacts aimed to serve particular educational objectives through the organization of several semiotic resources would be a matter of further research.

References

- Baldry, A. (Ed.) (2000). *multimodality and multimediality in the distance learning age*. Campobasso, Italy: Palladino Editore.
- Baldry, A. & Thibault, P. (2006). *Multimodal transcription and text analysis*. London: Equinox.
- Bernstein, B. (1996). *Pedagogy, symbolic control and identity: Theory, research, critique*. London: Taylor and Francis.
- Black, B., Heatwole, H., & Meeks, H. (2007). Using multimedia in interactive learning objects to meet emerging academic challenges. In A. Koohang & K. Harman (Ed.), *Learning objects: Theory, praxis, issues and trends* (pp.209-257). California: Informing Science Press.
- Butson, R. (2003). Colloquium. Learning objects: Weapons of mass instruction. *British Journal of Educational Technology*, 34(5), 667-669.
- Churchill, D. (2006). Towards a useful classification of learning objects. *Educational Technology Research and Development*, 55(5), 479-797.
- Cindy, E.H.M., Ravit, G.D., & Clark, A.C. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99-107.
- Clark, R.C. & Mayer, R.E. (2008). *E-learning and the science of instruction*. San Francisco: Jossey-Bass Pfeiffer.
- Cope, B. & Kalatzis, M. (Eds.) (2000). *Multiliteracies: Literacy learning and the design of social futures*. Melbourne: McMillan.
- Dimopoulos, K., Koulaïdis, V., & Sklaveniti, S. (2003). Towards an analysis of visual images in school science textbooks and press articles about science and technology. *Research in Science Education*, 33, 189-216.
- Dimopoulos, K., Koulaïdis, V., & Sklaveniti, S. (2005). Towards a framework of socio-linguistic analysis of science textbooks : The Greek case. *Research in Science Education*, 35, 173-195.
- Djonov, E. (2007). Website hierarchy and the interaction between content organization, webpage and navigation design: A systemic functional hypermedia discourse analysis perspective. *Information Design Journal*, 15(2), 144-162.
- Friesen, N. (2004). Three objections to learning objects. In R. McGreal (Ed.), *Online education using learning objects* (pp.59-70), London: Routledge.
- Friesen, N. & Cressman, D. (2007). The politics of e-learning standarization. In A. Koohang & K. Harman (Ed.), *Learning objects: Theory, praxis, issues and trends* (pp.507-525). California: Informing Science Press.
- Halliday, M. A. K. & Matthiessen, C. (2004). *An introduction to functional grammar* (3rd edition). London: Arnold.
- Iedema, R. (2001). Analyzing film and television: A social semiotic account of hospital: An unhealthy business. In Jewitt, C. & T. van Leeuwen (Eds.). *Handbook of visual analysis* (pp.183-204). London: Sage.
- IEEE Learning Technology Standards Committee. (2003). WG12: *Learning object metadata*. IEEE LTSC WG 12. Retrieved 5 September 2008 from <http://ltsc.ieee.org/wg12/>

- Jewitt, C. & Kress, G. (Eds.). (2003). *Multimodal literacy*. New York: Peter Lang.
- Karalis, T. Sotiropoulos, L. & Kampeza, M. (2007). 'La contribution de l'éducation tout au long de la vie et de l'anthropologie dans la préparation professionnelle des enseignants : Réflexions théoriques', *Skholê*, 1, 149-155.
- Kirschner, P., Sweller, J. & Clark, R.E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75-86.
- Kong, K. (2006). A taxonomy of the discourse relations between words and visual. *Information Design Journal*, 14(3), 207-230.
- Kress, G. & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. London: Routledge.
- Kress, G. & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.
- Lemke, J. (2002). Travels in hypermodality. *Visual Communication*, 1(3), 299-325.
- Levine, P. & Scollon, R. (2004). *Discourse and technology: Multimodal discourse analysis*. Washington, DC: Georgetown University Press.
- Lim, G. (2007). Instructional design and pedagogical considerations for the ins-and-outs of learning objects. In A., Koohang & K., Harman (Eds.), *Learning objects and instructional design* (pp. 1-38). California: Informing Science Press.
- Lockyer, L., Bennett, S., Agostinho, S., & Harper, B. (Eds.). (2009). *Handbook of research on learning design and learning objects: Issues, applications, and technologies*. New York: Information Science Reference.
- Machin, D. (2007). *Introduction to multimodal analysis*. London: Hodder Arnold.
- Macken-Horarik, M. (2004). Interacting with the multimodal text: Reflections on image and verbiage in ArtExpress. *Visual Communication*, 3(1), 5-26.
- Martinec, R. (2004). Gestures that co-occur with speech as a systematic resource: The realization of experiential meaning in indexes. *Social Semiotics*, 14(2), 193-213.
- Martinec, R. & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication*, 4(3), 339-374.
- Matthews, R. (2005). Vygotsky's philosophy: Constructivism and its criticism examined. *International Educational Journal*, 6(3), 386-399.
- Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59(1), 14-19.
- Notargiacomo, M. P., Frango, S. I., Omar, N., Dotto, S. M. (2007). Structure of storyboard for interactive learning objects development. In A. Koohang & K. Harman (Ed.), *Learning objects and instructional design* (pp.253-279). California: Informing Science Press.
- O'Toole, M. (1994). *The language of displayed art*. London: Leicester University Press.
- Parrish, P.E. (2004). The trouble with learning objects. *Educational Technology Research and Development*, 52(1), 49-67.

- Polsanyi, P. R., (2003). Use and abuse of reusable learning objects. *Journal of Digital Information*, 3(4). Retrieved 14 January 2009 from <http://jodi.tamu.edu/Articles/v03/i04/Polsani/>
- Ravanis, K. (1996). Stratégies d'interventions didactiques pour l'initiation des enfants de l'école maternelle en sciences physiques. *Spirale*, 17, 161-176.
- Ravanis, K. (1999). Représentations des élèves de l'école maternelle: Le concept de lumière', *International Journal of Early Childhood*, 31(1), 48-53.
- Ravanis, K. (2005). Les sciences physiques à l'école maternelle: Eléments théoriques d'un cadre sociocognitif pour la construction des connaissances et/ou le développements des activités didactiques. *International Review of Education*, 51(2-3), 201-218.
- Simbulan M.S., (2007). Learning objects' user interface. In A. Koohang & K. Harman. (Ed.), *Learning objects: Theory, praxis, issues and trends* (pp.259-336), California: Informing Science Press.
- Sotiropoulos, L. (2003). La recherche anthropologique en éducation : Quelques adaptations de la méthode, *Spirale*, 31, 85-90.
- Thibault, P. (2000). The multimodal transcription of a television advertisement: Theory and practice. In A. Baldry (Ed.), *Multimodality and multimediality in the distance learning age* (pp.331-385), Campobasso, Italy: Palladino Editore.
- Unsworth, L. (2006). *E-literature for children: Enhancing digital literacy learning*. London and New York: Routledge/Falmer.
- Unsworth, L. (2007). Image/text relations and intersemiosis: Towards multimodal text description for multiliteracies education. In B. Leila, B. & T. Berber-Sardinha (Eds.), *Proceedings of the 33rd International Systemic Functional Congress* (pp.1165-1205), PUCSP, São Paulo, Brazil.
- Unsworth, L., Thomas, A., Simpson, A., & Asha, J. (2005). *Children's literature and computer based teaching*. London: McGraw-Hill/ Open University Press.
- van Leeuwen, T. (1999). *Speech, music, sound*. London: Macmillan.
- van Leeuwen, T. (2005). *Introducing social semiotics*. London & New York: Routledge.
- Wiley, D. (2002). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. Wiley, D. (Ed.), *The instructional use of learning objects*. Bloomington, IN: Agency for Instructional Technology and Association for Educational Communications and Technology. Retrieved 14 March 2009 from <http://reusability.org/read/>

Correspondence: George Vorvilas, Department of Educational Sciences and Early Childhood Education University of Patras, Greece.
