

PONTOS DE POSSÍVEL DIÁLOGO ENTRE APRENDIZADO MULTIMODAL
E ENSINO-APRENDIZADO DE LÍNGUAS ESTRANGEIRAS

ENGAGING MULTIMODAL LEARNING AND SECOND/FOREIGN
LANGUAGE EDUCATION IN DIALOGUE¹

MIGUEL ANGEL FARIAS*
KATICA OBILINOVIC**
ROXANA ORREGO***

RESUMO: Este artigo é uma resenha bibliográfica derivada de um projeto de pesquisa focando na multimodalidade e seu possível impacto na aprendizagem de línguas estrangeiras, com especial atenção aos desenhos de textos (KRESS & VAN LEEUWEN, 2001) de acordo com o informado pelos modelos de aprendizado multimodal (SCHNOTZ, 2005; MAYER, 2001). Em primeiro lugar, o artigo resgata estes modelos e os relaciona ao campo da aquisição de um segundo idioma. Em continuação, descreve-se alguns pontos de possível encontro e diálogos entre estes dois campos de pesquisa. Finalmente, apresenta alguns estudos que se realizaram em torno ao aprendizado multimodal de línguas como exemplos deste diálogo.

Palavras-chave: aprendizado multimodal, modelos sócio-semióticos, ensino-aprendizado multimodal de línguas estrangeiras

ABSTRACT: This article is a partial review of the literature that supports an investigation on the possible impact of multimodality on foreign language learning, with particular emphasis on text design (KRESS & VAN LEEUWEN, 2001) and models of multimedia learning (SCHNOTZ, 2005; MAYER, 2001). To start, the article describes these models and establishes their relation to the field of second language acquisition. To continue, we discuss some meeting points that may justify our interest in making multimodal learning and foreign language education engage in dialog. The article concludes with a brief account of some studies dealing with multimodal foreign language learning that illustrate such dialogue.

Keywords: Multimodal learning, sociosemiotic models, multimodal second/foreign language teaching-learning

1. INTRODUCTION

Multimodality is not an alien concept in the field of foreign language learning and teaching (FLLT). If we look at the instructional practices used in FLLT, we can realize that they have always, by nature, included a multimodal dimension, although probably not acknowledged as what we understand now as multimodality. Presenting the foreign language, even in its most text-based (or monomodal) form, requires the learner to make some sort of visual or aural intervention, be it by highlighting, creating a diagram or

* USACH, Santiago, Chile. miguel.farias@usach.cl

** USACH, Santiago, Chile. katica.obilnovic@usach.cl

*** USACH, Santiago, Chile. roxana.orrego@usach.cl

¹ The elaboration of this article was possible by support from DICYT USACH, Project 030951FF. A preliminary version was presented at the International Conference Multimodality and Learning, University of London, June 2008.

evoking sounds or a translation. If the language textbook comes with illustrations, the learner, then, would make associations between the images and the linguistic code she is learning. In Fariás (2004) we illustrated with the texts depicted in Figures 1, 2 and 3 dating from 1899, 1951 and 1999, the transition from rather monomodal texts to multimodal texts used to teach English in Chile.

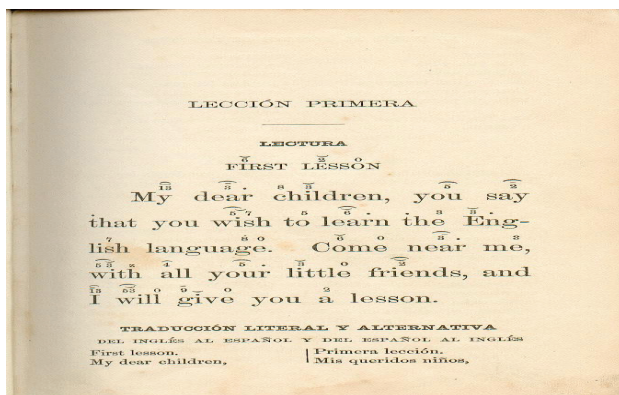


Figure 1 - El inglés al alcance de los niños. Robertson, T.(1899).

In Figure 1 we see a text written in English and, at the bottom of the page, its literal and alternative translation into Spanish. This type of text was used in Chilean schools in the late 19th century. The printed text calls for oral cues, indicated by the transcriptions of some sounds and the primary accent overgliding some words. This textbook does not include any other visual cues.



Figure 2 - Villegas, A. (1951). Little book for children.

Figure 2 shows a text from 1951. Here, we find that text and picture complement each other as the segments of language are iconically referenced to the picture located in the upper section of the page. At least the concrete nouns, with the right pointers, can be clearly referenced: *Fanny* and *bird*; and the same obtains for the action of *feeding*.

8.3 Build your skills

1 Before you start
 Look at the title of the text below, and the pictures. What's this text about?



THE UNITED NATIONS DECLARATION OF HUMAN RIGHTS
 SAYS EVERYONE HAS A RIGHT TO:

- LIVE WITHOUT FEAR OF VIOLENCE
- VOTE IN FREE ELECTIONS
- WORK, GO TO SCHOOL, AND HAVE A DECENT STANDARD OF LIVING
- BE FREE FROM UNLAWFUL ARREST
- HAVE FREEDOM OF THOUGHT AND RELIGION

But ...

- 842 million adults can't read (of these, 598 million are women).
- 1 billion people are without clean safe water.
- 140 million children can't go to school
- 907 million people die before the age of 40
- 158 million children can't get enough food to eat



Sengoara is a 15-year-old boy from Lesotho. He can't go to school because he works: he looks after 700 sheep and 72 cows.

Cows are very important animals in Lesotho – an old African proverb says a cow is 'a god with a wet nose' – and most boys between the ages of seven and fifteen leave school to look after the animals.

Sengoara gets up at 6 o'clock in the morning and goes to sleep at 8 in the evening. He earns \$140 a year – the price of one cow – and it's a difficult job. The boys are lonely, and cold, and hungry. When they leave their jobs, they have problems because they can't read or write. 60% of boys in some areas of Lesotho can't read.



Fifi is 14 and she's from Haiti, in the Caribbean. Her brother and two sisters go to school, but Fifi stays at home and helps her parents. She carries water, and cleans the dishes. Half of the men and women in Haiti can't read, and Fifi is the same.

The family is very poor. Her father buys and sells things, and works 60 hours a week. Her mother works 55 hours a week. The children work 30 hours a week. There's no TV, radio, video recorder, or car. The family of six people earns \$534 a year.

Figure 3 - Elsworth & Rose (1999). Go for Chile.

In terms of visually, Figure 3 displays an extract from an EFL textbook written in English in 1999. As we can see the design has clearly improved in this textbook which was widely used in most Chilean public schools. Thanks to technological advances in printing we now have that color photographs, colors, a distinct layout, and different letter types have been included. At the top of the page on the left side there is an icon representing a tape. This symbol is appealing to the auditory dimension as the same text that the student is reading can also be listened to. With the arrival of new technologies, as the tape recorder, the film and the video, came the necessary support for the language classroom teacher to have a more realistic take of the language being presented, so much that the teaching of French as a foreign language, for example, came to be associated solely with the audiovisual methodology for a long time (PORQUIER & VIVES, 1974).

In the last twenty years or so as we quickly move to a digital culture, one of the first mentions to multimedia in FLLT came from Warschauer (1996) who included multimodality in what he calls the third stage of CALL, the integrative paradigm, which emerged with the introduction of digital technologies, as the CD Rom, in the language classroom. Therefore, the dialog between second language learning, naturalistic, instructed or both, and multimodality has been for long a silent exchange. However, we feel that what has been called the “digital revolution” (TURKLE, 1995; WOLF, 2008) and, particularly the arrival of new hypertextual formats to represent the language, has exacerbated the need to bring the explanatory apparatuses from both sides to unify criteria and share the necessary metalanguage needed for this dialog to happen. And this dialog should also be motivated by the necessary reflection on the uses of technologies in education as cognitive tools (others call them cognitive technology, technologies of the mind, mind tools) that can assist learners to support cognitive processes that would be out of reach otherwise, (HARPER et al., 2000).

In what follows we include a review of the literature on models of multimedia learning and some key concepts in S/FLA that can be used in an integrated model. We then provide an account of some studies including second/foreign language multimodal learning that evidence concrete examples of dialog between both fields and that may serve, in turn, as avenues to explore what has been termed multimodal communicative competence (ROYCE, 2007).

2. MODELS OF MULTIMODAL LEARNING

The literature on models of multimedia/multimodal learning which integrate second language learning/acquisition is an unexplored area of research which needs urgent development. For the purposes of this review we firstly describe two models of multimedia/multimodal learning which, although they make no explicit reference to second language acquisition, can be applied to understand language learning in the context of pedagogical environments that provide multimodally presented linguistic input. One is Mayer’s (2001) multimedia learning and the other one is Schnotz, Bannert & Seufert’s (2002) model of picture and text comprehension which addresses the issue of reading comprehension from multimodal texts including text and pictures. Then, we discuss what we have found to be two good attempts to integrate the two fields: Plass & Jones’ (2005) model of multimedia learning and second language acquisition and Schnotz & Baadte’s (2008) distinction between multimedia domain learning and multimedia second language learning.

2.1. Mayer’s model

Mayer’s model of multimedia learning (2001), as summarized by Farías, Obilinovic & Orrego (2007), includes the discussion about three views of multimedia, two views of multimedia design, two metaphors of multimedia learning, three kinds of multimedia learning outcomes, two kinds of active learning, and seven principles of multimedia design. This author, who belongs to a cognitive multimedia learning tradition, uses an attractive

pedagogical discourse to describe multimedia from three perspectives: as delivery media (combining two or more delivery devices, as a PowerPoint presentation and the lecturer's voice), presentation modes (representations that include words and pictures, as on-screen text and animation) and sensory modalities (visual and auditory senses, as used to process slides and narration, for example).

According to this theory, multimedia presentations have the potential to result in deeper learning and understanding than do presentations that are presented solely in one format. Thus, students who learn with words and pictures learn and remember information better. Mayer's multimedia learning theory is a learner-centered approach based on a constructivist view of learning. This means that "multimedia designs that are consistent with the way the human mind works are more effective in fostering learning than those that are not" (MAYER, 2001, p.10). Consequently, multimedia learning is conceived as knowledge construction. In this regards, Mayer claims that the learner constructs his/her own learning by interacting with the multimedia designs. This is done as follows:

[...] according to the knowledge construction view, the learner's job is to make sense of the presented material; thus the learner is an active sense maker who experiences a multimedia presentation and tries to organize and integrate the presented material into a coherent mental representation. (p. 13)

He also points out that within this conception of learning, the role of teachers is to help learners in that sense-making process. In other words, the teacher is a cognitive guide who provides support to the learner's cognitive processing. To support empirically this multimedia cognitive theory, this author has developed seven multimedia principles which are summarized in Figure 4:

1. Multimedia Principle: Students learn better from words and pictures than from words alone.
2. Spatial Contiguity Principle: Students learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen.
3. Temporal Contiguity Principle: Students learn better when corresponding words and pictures are presented simultaneously rather than successively.
4. Coherence Principle: Students learn better when extraneous words, pictures, and sounds are excluded rather than included.
5. Modality Principle: Students learn better from animation and narration than from animation and on-screen text.
6. Redundancy Principle: Students learn better from animation and narration than from animation, narration, and on-screen text.
7. Individual Differences Principle: Design effects are stronger for low-knowledge learners than for high-knowledge learners and for high-spatial learners rather than for low-spatial learners

Figure 4 - 'Seven research-based principles for the design of multimedia messages', (Mayer 2001, p. 184).

Even though these principles have not been thoroughly tested in FLLT, they represent a challenge to media designers and a call of attention to educators at the moment of preparing and presenting classroom materials. As explained later in this paper, the principle that poses more questions to FLLT is the redundancy principle.

Underlying Mayer's cognitive multimedia model is Paivio's dual-channel theory (DCT) that posits that human beings have separate channels to process visual and auditory

information. DCT is a theory of cognition whose origins can be traced to the cognitive revolution of the 1960s and 1970s. The most basic assumption of this theory (PAIVIO, 1971; SADOSKI & PAIVIO, 2001) when applied, for example, to reading and writing would tell that cognition consists of two separate coding systems of mental representations that are organized hierarchically, one system specialized for language and another one specialized to deal with nonverbal objects and events. According to these authors, mental representations refer to internal forms of information used in memory while coding refers to the ways the external world is captured in those internal forms. In the verbal system the information is processed sequentially, whereas in the nonverbal system information is organized nonsequentially (e.g., spatially). DCT states that there is continuity between perception and memory. From this perspective these authors have claimed that:

External experiences are perceived through the simulation of our various sense modalities, including the visual, auditory, haptic, gustatory, and olfactory sense modalities... [...]. In DCT, all of our mental representations retain some of the original, concrete qualities of the external experiences from which they derive, so that representational structures and processes are modality-specific rather than amodal. This implies that our mental encodings themselves are concrete rather than abstract although they can easily deal with abstract information and concepts such as language symbols, charts or diagrams. (p. 42)

This theory assumes that human beings represent the verbal and nonverbal information by means of two basic units. These are called *logogens* and *imagens*. *Logogens* are the basic representational units in the verbal system known as verbal representations, verbal encodings, mental language and inner speech, while *imagens* are nonverbal representations usually called mental images, or imagery (external and internal). Both *logogens* and *imagens* are concepts introduced to distinguish between the underlying neurological representations and their conscious expression in language and imagery. In other words, these units in DCT are not abstractions without concrete form, they are assumed to have some physical form in neural structures and pathways. Later developments include a Dual Coding Theoretical Model of Reading (SADOSKI & PAIVIO, 2004) which is an hybrid model that considers that mental representations are modality and sensory specific. This model also differs from other models of reading comprehension in the sense that mental representations are not stored in propositions or schemata, as assumed by authors like Van Dijk and Rumelhart. What is more, the authors behind the DCT strongly believe that these processing units are difficult to operationalize and empirically test with human data.

2.2. Schnotz and Bannert's integrated model of text and picture comprehension

Behind Schnotz and Bannert's integrated model of text and picture comprehension there is a very important distinction made by Schnotz (2002) between descriptive and depictive representations. According to Schnotz (2005) descriptions are basic forms of representation. Some common examples of descriptions are texts and mathematical expressions. Descriptions are basically characterized to consist of symbols. Citing Pierce (1931-1958), Schnotz claims that symbols are signs that do not have similarity with their referents. For example, the word *dog* has no resemblance with a real dog; its meaning is based on a convention. On the other hand, we find depictive descriptions that can be

defined as representations that include pictures such as photographs, drawings, paintings, and maps. They are depictive representations because they consist of icons, defined by Schnotz (2005, p. 52) as "...signs that are associated with their referent by similarity or by another structural commonality"; for instance, the picture of a dog has some similarity with the corresponding referent. These descriptive and depictive representations have different uses for different objectives; descriptive ones are more powerful to express abstract concepts. On the other hand, pictures such as photographs, drawings, paintings, maps or small miniatures are depictive representations which are iconic because they resemble their referents more closely. In other words, depictive representations consist of icons and icons are signs that are associated with their referents by similarity or by any other structural commonalities.

Schnotz and Bannert's integrated model of text and picture comprehension also incorporates Baddeley's theory of working memory (1986). This author argues that descriptive and depictive channels are constrained by the limited capacity of working memory. According to Baddeley and Hitch's multiple-component model of working memory (1974, cited in Miyake & Shah, 2004), this memory

[...] comprises a central executive controlling which are specialized for the processing and temporary maintenance of material within a particular domain (i.e., verbally coded information and visual and/or spatial information, respectively).(p. 29)

Schnotz and Bannert's model of text and image comprehension is shown in Figure 5:

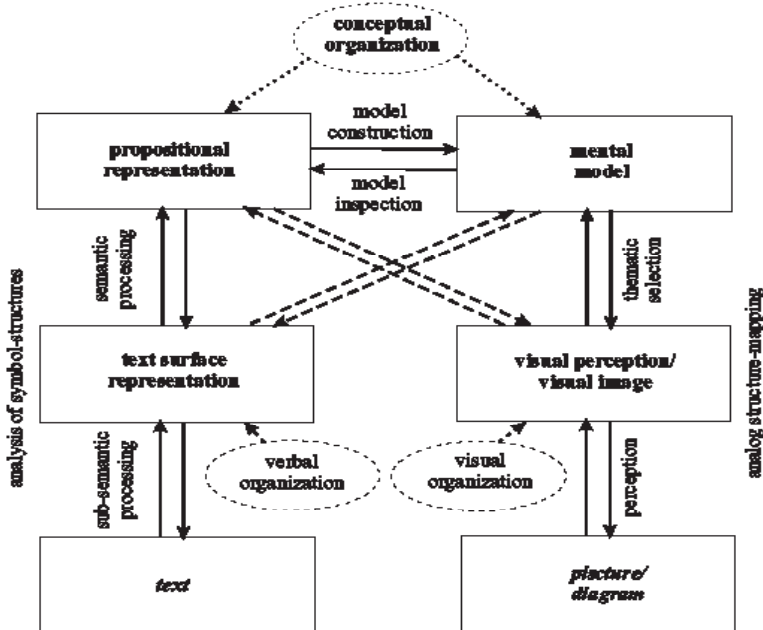


Figure 5 - From "Text and picture comprehension" (SCHNOTZ, BANNERT & SEUFERT, 2002, p. 394).

This model consists of a descriptive side on the left, and a depictive one, shown on the right side of Figure 5. The descriptive column shows the way the textual written information is represented. It includes three levels of representation of the information: an external text, the internal mental representation of the text surface structure and a propositional level representing the text's semantic content. It is important to mention that for these authors the interaction between these descriptive representations is based on symbol processing. On the other hand, the depictive side comprises the external picture, the internal visual perception of the image or picture, and an internal mental model of the content presented in the picture. These levels of representation are based on processes of structure mapping such as analogy relations. The basic assumption to understand how text and picture information are represented and understood is based on the following: regarding textual comprehension, first, the reader of a text constructs a mental representation of the text surface structure, then, a propositional representation is generated and, finally, the reader constructs a mental model of the subject matter described in the text from the text base. During these processes top-down and bottom-up processings are applied.

Along the same lines, when picture comprehension takes place, different processes occur. First, the individual creates a visual mental representation of the picture through perceptual processing, then, he constructs both a mental model through semantic processing and a propositional representation of the subject matter depicted in the picture. Third, the resulting mental representation is the visual perception of the picture in the visuospatial sketchpad of the working memory of the individual. According to Schnotz, Bannert & Seufert (2002) all of these perceptual processes are organized according to Gestalt laws. Finally, the individual constructs a mental model of the depicted subject matter through a schema-driven mapping analogic process. Accordingly, text comprehension and picture comprehension are to be considered as complementary ways of creating mental representations from the textual and visual information. As a conclusion, the authors state that the essential point when comprehending textual and visual information is the fact that propositional representations and mental models are based on different sign systems and different principles of representation that complement one another.

3. INTEGRATED MODELS

The models analyzed so far furnish educators with baseline metalanguage to understand what goes on in the learner's cognitive system as she learns from multimedia. However, this is not enough for the second/foreign language educator since (formal) language learning is different from other types of learning. Such difference has been added in the following two models:

3.1. Schnotz and Baadte

To make the distinction between language learning and other types of learning, Schnotz & Baadte (2008) call the last "domain learning". They mention that domain learning with multimedia takes place when an individual "uses the external representations

as information sources in order to construct in working memory internal (mental) representations of the learning content and store those representations in long-term memory” (p. 27). An example of multimodal domain learning is the use of texts, pictures, graphs, animation and sound as external representations of content such as biology, geography or physics.

Schnotz & Baadte (2008) explain the potentially different relationships that may exist between domain learning and language learning positing that in some cases “to use language as a tool for domain learning, the individual has at first to learn language” (p. 23). Thus, conceptual knowledge of various domains is only possible through the development of linguistic competence. In this case, language learning is certainly prior to the learning of other specific conceptual domains; the individual must be able to master language before s/he can use it as an instrument for learning about other specific domains. However, in early first language acquisition, these authors posit that “domain learning precedes language learning” (p. 23) which is evidenced by the type of communicative activity that most parents and caretakers in general use with their young children. This early communicative interaction is based on those concepts that children are already familiar with.

In the early stages of adult second language learning, which resembles the process children go through with the first language, there certainly is prior conceptual knowledge which in this case is much broader than in early language learning. The individual is engaged in learning the basic platform of the second language and s/he is already familiar with a variety of domains, but needs to add a new code to express the known meanings.

Regardless of the differences between domain learning and (first/second) language learning, when it comes to determining the requirements for meaningful learning to occur, these German researchers emphasize selection of information as a key. They claim that “meaningful learning from text and pictures requires a coordinated set of cognitive processes including selection of information, organization of information, activation of prior knowledge, and active coherent formation by integration of information from different sources” (p. 36).

For meaningful learning to be possible, then, the mind of the learner must first select the information to be processed, organize it activating the already existing knowledge and finally integrate it into his/her cognitive structures. Using the metalanguage that characterizes Second Language Acquisition (SLA), we should say that in order to transform part of the input that the learner is exposed to into intake s/he must make an appropriate selection of the input that will become learned/acquired intake. For this to occur, attention is crucial.

Gass (2006) recognizes that “in the recent history of SLA research, much emphasis has been placed on the concept of attention and the related notion of noticing” (p. 244). We believe that both the appropriate selection of information for cognitive processing and the subsequent meaningful learning mentioned by Schnotz & Baadte (2008) cannot be possible without the learner’s attention and the closely related phenomenon of noticing, whose relevance to learning has reached consensus among researchers in spite of the lack of a uniform definition. As Lewis (2000) states: “[T]here is a broad consensus that language that is not noticed does not become intake, but there is no agreement on the precise meaning of the word “noticed” (p.161).

Mitchell & Myles, on the other hand, (2004) explain the phenomenon through Schmidt's proposal which is a definition we would like to adhere to:

Schmidt is careful to distinguish among different types of attention that learners might pay to language form. He uses the term noticing to refer to the process of bringing some stimulus into focal attention, that is, registering its simple occurrence, whether voluntarily or involuntarily. (p. 164)

Noticing, then, can be interpreted as one specific type of attention. It seems reasonable to assume that this particular type of attention is directed to the formal aspects of language as in the case of a syntactic arrangement, a spelling or the pronunciation of a word.

3.2. Plass and Jones's model

One of the main references in the attempt to bring the two fields together is Plass and Jones (2005) who postulated an integrated model of SLA and multimodal learning. By adapting theories from SLA and Mayer's multimodal learning these authors organize a model in which the verbal and pictorial input is selected by the processes of apperception and noticing to create a verbal text base and a visual image base. Comprehension takes place as words and images are organized in a verbal model and a visual model which become integrated with the participation of the learner's background knowledge.

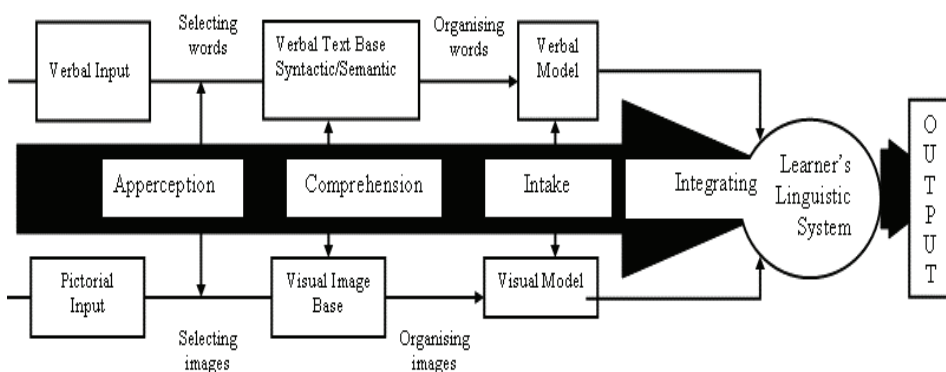


Figure 6 - Integrated model of multimedia learning and SLA, Plass and Jones (2005).

Figure 6 shows the integrated model of multimedia learning and SLA proposed by Plass and Jones (2005). In this model Plass and Jones bring multimodality and second language acquisition together through the concepts of apperception, comprehension, and intake. Apperception constitutes the very first stage in the process. It is defined as the selection of input that learners must make before processing what is presented to them. Thus, if the information they are exposed to is verbal, the selection is mentally represented in a text base. When the information is of a pictorial nature, the learner's mind takes it and places it in a visual image base. After this selection process has taken place, the material is organized into visual mental representations and verbal mental representations.

The authors explain that within interactive processing for comprehension of information, meaningful interaction with the material is important. Only through meaningful interaction can the learner succeed in constructing meaning. Intake is defined as input that has been successfully comprehended and that can be integrated into the learner's linguistic system, as shown in figure 6. The end result of the whole process is output or the learner's production of meanings using his/her linguistic system.

As language educators we share the main concern that triggered the elaboration of this model, synthesized in the following question: "In what way can multimedia support second-language acquisition by providing comprehensible input, facilitating meaningful interaction, and eliciting comprehensible output?" (PLASS & JONES 2005, p. 471).

We find this to be a good theoretical attempt at integrating multimodality and SLA, but it is insufficient in incorporating the complex processes involved in interlanguage and the construction of a bilingual mind. If viewed from a sociosemiotic perspective (KRESS & VAN LEEUWEN, 1996), the multimodal input the foreign language learner is exposed to triggers cognitive responses and processes that are both mental and social in that they are influenced by the individual's background knowledge and, at the same time, conditioned by the social context. For these authors visual representations are socially constructed out of the affordances made available by particular cultures. This tension between social and individual meanings is in tandem with Doughty & Long (2006) when they mention that researchers recognize that SLA takes place in a social context but that it also is "ultimately a matter of change in an individual's mental state" (p. 4).

It is actually from this sociosemiotic view that FLLT can establish a stronger dialog with multimodal theories to the extent that the very setting in which learning takes place, which includes the classroom, the teacher, other learners, the textbook and the teaching materials, can be looked at not as mere linguistic objects but as cultural artifacts whose architecture, designs, affordances and learning potential are not neutrally constructed.

4. SOME SLA CONCEPTS IN THE LIGHT OF MULTIMODALITY

Our main interest as second language educators is to describe how multimodal presentations influence the teaching and learning of English as a foreign language. We are particularly interested in studying reading for comprehension and written production and the focus of our attention is the lexicon, particularly the retention and retrieval of new vocabulary items in a second language. Therefore, in this section we present a synthesis of some of the most relevant findings in studies in S/FLA in these areas and we identify three areas of possible convergence.

Even though there is no consensus on the percentage of utterances that are the result of rule-application, i.e. the generative component of language acquisition, and the percentage of utterances that derive from the memorization of fixed phrases or prefabricated speech, linguists agree that the formulaic component of the second language vocabulary and the lexical component in general are of vital importance in the language learning process and, as a consequence, in the teaching of a target language. Larsen-Freeman (2003, p. 14), for example, reminds us that

..it has become increasingly clear these days, with the use of million-word language corpora, that a great deal of our ability to control language is due to the fact that we have committed to memory thousands of multiword sequences, lexicogrammatical units or formulas that are preassembled (e.g. I see what you mean; Once you have done that, the rest is easy) or partially assembled (e.g. NP + tell + tense + the truth as in “Jo seldom tells the truth”; “I wish you had told me the truth”).

Therefore, the lexical component, including the so-called “canned speech”, is no longer subordinated to structures and has become an important element to consider when teaching a target language. Like the grammatical component, the vocabulary of the second language must be incorporated in the conceptual store of our learners not only for its retention, which is in itself a complex process, but also in terms of its transformation into active knowledge, which the students must be capable of producing (or reproducing) orally and in writing, with the necessary fluency. In other words, like all the other aspects of language, the lexicon is only declarative knowledge until it becomes procedural knowledge, hopefully through the systematic use of effective techniques on the part of the language teacher. Segalowitz (2006), summarizing the Active Control of Thought Theory developed by Andersen, points out that this theory “assumes that skill acquisition involves a transition from a stage characterized by declarative knowledge to one characterized by procedural knowledge.” He goes on to say that:

The transition from declarative knowledge to procedural knowledge through the application of production rules occurs via a process called proceduralization. This involves passing from a cognitive stage where rules are explicit, through an associative phase where rules are applied repeatedly in a consistent manner, to an autonomous stage where the rules are no longer explicit and are executed automatically, implicitly in a fast, coordinated fashion. (p. 394, 395)

We can no longer advise our novice teachers to implement an entire class around mechanistic techniques, although we are aware of the fact that the repeated exposure that learners need to a new item in order to retain it invites some sort of repetitive and mechanical work. Instead, the presentations of the new items can be done through more than one mode, redundantly even, to compensate for the lack of formal repetitive work in the foreign language classroom. By exposing the students to new items presented through narration (the teacher’s oral input), visual images (static pictures or pictures with movement) and on-screen text, we can promote enough noticing which will eventually promote the basic degree of retention needed by our learners prior to storing the new information in their cognitive structures, assuring a certain stability for future automatic use. As we explained earlier, the learner must be able *to select* (which is the term used by SCHNOTZ & BAADTE) the part of the input that may eventually become active intake. Selection is only possible if the learner pays close attention to the input (at least to relevant parts of it). If the attention paid is directed to the form of the input and not only to the comprehension of the meaning conveyed, the phenomenon of noticing should occur. If the input is presented not only through the teacher’s narration (oral input) but also through multimedia, contributing to the redundant presentation of the new information, meaningful learning should be enhanced.

Our interest in conducting research that integrates multimodality and second language acquisition is, then, theoretically supported by the following meeting points where, as we see, SLA converses with multimodality:

- a. The need to expose the learner through more than just one mode to more than just one “encounter” to each new lexical item in order to contribute to its retention. Although “remembering” is only a basic cognitive skill, it is fundamental for the use of higher-order cognitive skills such as application or analysis. Hulstijn (2006, p. 367) emphasizes the need to provide several instances of rehearsal in order to attain long-term retention of vocabulary. We consider that the benefit of narration (spoken text or input), on-screen text and animation (images) is one way of compensating for the low frequency in the exposure received in a foreign language context.
- b. The need to transform declarative knowledge into procedural knowledge. For example, for the language rules to become automatic in use, multimedia may serve, as Schnotz & Baadte (2008, p. 27) claim, “as a vehicle to convey formal rules of the target language such as syntactic, prosody or semantic rules presenting this explicit linguistic knowledge in a visual model with the help of graphics or annotations and in an auditory mode by presenting pronunciation guides or narratives”.
- c. Language learning (defined as retention and active use of lexicon in our case), requires, as was mentioned, a special type of attention called “noticing”, regardless of whether the learning is intentional or incidental. For this, we need techniques that can help us provide enhanced input that can also cater for individual differences. This can be done in multimedia learning environments that can furnish “the situational context in which communication takes place and therefore stimulate the learner to comprehend the semantic content of the messages and contribute to the ongoing (simulated) conversation” (SCHNOTZ & BAADTE 2008, p. 27).

In his discussion of foreign language teaching, De Beaugrande (1997, p. 5) has claimed that “the teaching situation can under no conditions suffice to impart the *entire* (his italics) language”. A possible solution for such problem, he goes on arguing, is to design what he calls an “artificially restricted intersystem” that would function “with maximally powerful rules and options”. In our opinion, multimodal presentations provide the appropriate vehicle for such intersystem, and the various modalities offer the necessary options for the varieties of cognitive styles learners bring and deploy in foreign language learning.

5. MULTIMEDIA SECOND LANGUAGE ACQUISITION RESEARCH

Research that may support the three arguments mentioned include Dubois and Vial (2000, p. 157), reporting on a study on the retention of new (Russian) vocabulary. They explain that “although classical theories in psychology enable the prediction of a student’s performance with verbal or illustrated material...few models have been developed to describe learners when they must process several types of information at the same time...” Dubois and Vial mention other researchers that have studied the potential influence of visual images on the learning of vocabulary and who, for example, have trained a group of learners to imagine scenes as they were being exposed to new vocabulary; yet other researchers, have provided the mental images to the learners. These two studies confirmed

the effectiveness of using a keyword whose meaning helped the students associate each word with its translation. Dubois and Vial also mention other investigators who demonstrated that providing a drawing that integrated keywords and translations was very useful. The objective of Dubois and Vial's research was to analyze the effects of different modes of presenting information on vocabulary learning of Russian as a foreign language. They predicted that their students would show better recall when textual information was presented with visual and auditory information provided that there were semantic and phonetic links between the elements. Their theoretical rationale for this claim was that "when textual, visual, and auditory materials are integrated in this way, the learner may be forced to engage in additional processing that leads to better memorisation". (p. 159). Three important prerequisites for better recall are, a) if the visual elements depict information that is relevant to the particular text used, b) if they illustrate new content that is important to the new task, and c) if they establish a connection with the text.

Among the results obtained, Dubois and Vial explain that auditory information presented together with visual elements fostered more learning than textual information presented with the same image. It is important to point out that these findings are in line with the findings of other authors, as Mayer and Moreno (1998), for example. Dubois & Vial explain the consistency of the results saying that "a presentation where both elements presented (image and text) are only visual results in less learning than if both are visual and auditory (less cognitive load)"(p. 163). This coincides perfectly well with Mayer's modality principle (See Fig. 4 above).

Research in the field of foreign language teaching also includes Plass, Chun, Mayer and Leutner (1998) who conducted a study on English speaking learners in multimedia second language learning environments in which the students were asked to read a text in German presented through a computational program that provided the translation of key concepts, an image of a video clip illustrating the word, or both. The results showed that the students could remember the translations in a better way when they chose both the verbal and the visual annotations and that they understood the story in a better way when they chose their favorite mode of annotation. These two types of learning, vocabulary retention and transfer, represent the main challenges for foreign language learning as they are the foundations on which other competences are built. Another important concept included in this study is that of individual differences since the authors measured how visualizers differ from verbalizers and concluded that those students who used their preferred mode did better in comprehension. Some of the conclusions of this study suggest that "learners should have options for selecting and processing material presented both in visual and verbal modes" (p. 34).

On the other hand, Plass (1998) has done research on the interfaces utilized in the interaction with multimedia in foreign language learning software. The author presents four approaches used for describing the design of user interface (craft, enhanced software engineering, technologist and cognitive) and explains that he has adopted the cognitive model as the most appropriate for the design and evaluation of user interface computer programs for second language acquisition due to the fact that such an approach can incorporate in its design both the user as well as the learning task.

One particular aspect of multimodal theories that has caught our attention as foreign language learners and teachers is the redundancy principle proposed by Mayer (2005). Even though the applications of such principle as explained by Mayer and Moreno (2003) and Mayer and Anderson (1991) have not been to the field of foreign language learning, we were intrigued when we first read about it because it apparently contradicted a long lasting tradition in language teaching: that of presenting the language in writing (text), orally (teacher's voice or tape or other verbal medium), and via animation. As we have explained earlier, Mayer's redundancy principle states that a representation consisting of narration and images is much more effective than one consisting of narration, images, and on-screen text due to the unnecessary use of the same (visual) channel twice, which overloads the learner's capacity (cognitive load). Preliminary results from various pilot studies we have conducted at USACH invite speculations as to several variables that would be affecting outcomes when learners are exposed to redundant versus non-redundant multimodal presentations. One of them has to do with the level of proficiency: we hypothesize that redundant presentations are helpful for adult low-proficiency learners who have not yet attained full automaticity and thus would need the reassurance the written text offers (as subtitles to a movie). Another issue to be considered is the type of language input to be presented, particularly as it concerns visual representations. In Farias, Obilinovic & Orrego (2009) we reported on a pilot study that tested the redundancy principle using idiomatic expressions. The interesting question that came out from this study is to what extent the literal visual representation of an idiomatic expression helps understand its metaphorical or figurative meaning. Results lead to more complex grounds as we noticed that, on one hand, the very nature of the idiomatic expression may cause differences in learning. On the other hand, the iconicity of the visual representation poses a challenge as idiomatic expressions convey a metaphorical meaning that transcends its literal representation.

As concerns the redundancy principle, Chandía et al. (2007) set up a pilot study to investigate the differences in vocabulary retention in secondary students exposed to English as a foreign language presented in two variants: animation, narration and on-screen text, and animation and spoken text, i.e., Mayer's redundancy principle and multimodal principle, respectively. Alternatively, this study hypothesized that students with a predominant visual learning style (imagers) would retain more vocabulary than students with a predominant verbal learning style (verbalizers) after being exposed to multimodal lessons. The results showed that students exposed to animation, narration and on-screen text retained more vocabulary than those exposed to animation and spoken text. These results serve to ratify our assumption that low proficiency learners benefit from the redundant element. In turn, their results point that imagers scored higher than verbalizers in the vocabulary test.

Macis (2007) brought together several interesting topics related to second and foreign language learning and teaching. These are: *enhanced input* and its impact on the learners' "noticing" of new forms, *retention* of high frequency *collocations*, Mayer's *redundancy principle* and its effectiveness or lack of effectiveness when it is applied in the discipline of second language acquisition. Her research questions were: What is the role of redundancy in second language acquisition? Does enhanced input - through the application of Mayer's

redundancy principle - facilitate or hinder the retention of collocations if compared to only two modes, that is narration (the teacher's input) and monomodal text? Her findings show no difference in the effect of the two types of input (enhanced versus non-enhanced input) on noticing and retention of the collocations; however, more than half of the students from the experimental group agreed on the beneficial effects of the images for a better understanding and retention of the collocations.

As for the gestural dimension involved in the learning process, Fariás & Acevedo (2007) investigated the types of gestures used by dyads of Chilean EFL learners solving a semi-guided communicative task. Based on some gesture typologies (MCNEILL, 2005; CASSELL 2007), the researchers concluded that gestures function as constitutive concurrent elements when learners deploy interactional and mediational strategies. Iconic gestures serve as compensatory strategies that learners use in the absence of lexical access in the target language. The authors speculate about the possible correlation between language proficiency and progression in the use of gestures; from deictic to iconic to metaphoric gestures as proficiency increases.

6. FINAL REMARKS

Based on this review, it should be clear that FLLT and multimodality can effectively and productively engage in dialog. As Mayer (2002, p. 69) has asserted, this dialog can be built when there is a "two-way street between cognitive science and instruction" that allows for an expeditious interaction between theory and practice. We need to add that such dialog also requires a critical pedagogy approach to the contexts in which such learning takes place and the media used for instructional purposes. In countries like Chile, the context has traditionally been that of learning a "foreign" language with the purpose of communicating with native speakers. However, authors like Pennycook (1994), Phillipson (1992) and Canagarajah (1999) among others, have called our attention to critical approaches to world Englishes or to English as an international language that question, among other concepts, the notion of native speaker and the traditionally accepted concepts of linguistic standards. Such critical approaches to English language teaching should also integrate questions as to the media in which the language is portrayed, and this should be the aim of critical multimedia literacy incorporated into FLLT. The media has been traditionally the textbook, in most cases containing acritical and bland portrayals of what some have called the three D's of consumerist EFL culture: dinner parties, dieting, dating (WALLACE, 2002). As learners face more and more other types of texts, as hypertexts, and as we embark in pedagogies in which such texts receive a critical treatment by unveiling their very constitution as particular genres, we can also establish communicating vessels between critical language awareness (FARIÁS, 2006) and critical multimedia literacy.

The quick pace of change from print-based to more visually oriented and digitalized presentations of information involves also a quick response from language teachers and educators to take advantage of multimodality to engage learners in meaningful cognitive, social and critical understandings. As the theories and applications that we have reviewed point out, attention to the meaning-making potential of the various designs of multimodal

discourse is an important component of visual literacy that can help language learners to cope more efficiently as they face new modes of information portrayal. The dialog has started.

REFERENCES

- ATKINSON, R. C., & RAUGH, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, 104, p. 126-133.
- BADDELEY, A. D. & HITCH, G. J. (1974). Working memory. In: G. H. Bower (ed.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, v. 8, p. 47-89.
- BADDELEY, A. D. (1986). *Working memory*. New York: Oxford University Press.
- CANAGARAJAH, A.S. (1999). *Resisting linguistic imperialism in English teaching*. Oxford: OUP.
- CASSELL, J. (2007). A framework for gesture generation and interpretation. In: CIPOLLA, R. and PENTLAND, A. (eds). *Computer vision in human-machine interaction*. Cambridge: CUP.
- CHANDÍA, F. et al (2007). *Multimodality and vocabulary learning: evaluating the impact of the modality and redundancy multimedia principles on the lexical retention of EFL secondary students classified as imagers and verbalizers*. Undergraduate thesis, Licenciado en Educación en Inglés, USACH. Santiago.
- DE BEAUGRANDE, R. (1997). *Text, discourse and process. Toward a multidisciplinary science of texts*. Washington: National Academy Press. Available at: <<http://www.beaugrande.com/TDPOpening.htm>>. Accessed on: 14 February 2008.
- DUBOIS, M. & VIAL, I. (2000). Multimedia design: the effects of relating multimodal information. *Journal of Computer Assisted Learning*. v. 16, p. 157-165.
- DOUGHTY J. C. & M. H. LONG (eds.) (2006). *The handbook of second language acquisition*. Malden, Ma.: Blackwell Publishing Ltd.
- ELSWORTH, S. & ROSE, J. (1999). *Go for Chile*. Primer Año de Enseñanza Media. Santiago de Chile: Addison Wesley Longman.
- FARÍAS, M. (2004). *Textos multimodales en tiempos de competencias múltiples*. In: Actas del XIV Congreso Nacional de SONAPLES. Univ. de Los Lagos. Osorno.
- _____. (2006). Critical language awareness in foreign language learning. *Literatura y Lingüística*, n. 16, p. 209-220. Available at: < <http://www.scielo.cl/scielo.php>>. Accessed on: 22 June 2009.
- FARIAS, M. & ACEVEDO, I. (2008). El rol de los gestos en el discurso multimodal de aprendices de inglés como lengua extranjera. In: FARÍAS and OBILINOVIC (eds.) *Aprendizaje Multimodal/Multimodal learning*, Santiago de Chile: Publifahu USACH, p. 97-110.
- FARÍAS, M.; OBILINOVIC, K. & ORREGO, R. (2007). Implications of multimodal learning models for foreign language teaching and learning. *Colombian Applied Linguistics Journal*. n. 19, p. 174-200.
- _____. (2009). Estudio exploratorio de aplicación del principio de redundancia en el aprendizaje de expresiones idiomáticas en lengua extranjera. In: Actas del II Congreso Nacional de la Cátedra UNESCO Lectura y Escritura. Universidad de Los Lagos. Osorno.
- GASS, S. (2006). Input and interaction, In: *The handbook of second language acquisition*. DOUGHTY, C. J. & LONG, M. H., (eds.). Malden, Ma.: Blackwell publishing Ltd, p. 224-255.

- HARPER, B., HEDBERG, J., CORDEROY, B., & WRIGHT R. (2000). Employing cognitive tools within interactive multimedia applications. In: S. LAJOIE. (ed). *Computers as cognitive tools*. Mahwah, NJ: Lawrence Erlbaum, p. 227-245.
- HULSTIJN, J. H. (2006). Incidental and intentional learning. In: *The handbook of second language acquisition*. DOUGHTY, C. J. & LONG, M. H. (eds.). Malden, Ma.: Blackwell Publishing Ltd, p. 349-381.
- KRESS, G. & VAN LEEUWEN, T. (1996). *Reading images. The grammar of visual design*. London: Routledge.
- _____. (2001). *Multimodal discourse: The modes and media of contemporary communications*. London: Edward Arnold.
- LARSEN-FREEMAN, D. (2003). *Teaching Language: From grammar to grammaring*, Canada: Thomson-Heinle.
- LEWIS, M. (ed.) (2000). *Teaching collocation, further developments in the lexical approach*. Hove: Language Teaching Publications.
- MACIS, M. (2007). *The effect of enhanced input through multimedia presentations on performance in the retention of collocations*. Master's thesis for the Program in Linguistics, Universidad de Santiago de Chile, Santiago.
- MAYER, R. & ANDERSON, R. (1991). Animations need narrations: an experiment test of a dual-coding hypothesis. *Journal of Educational Psychology*. v. 83, n. 4, p. 484-490.
- MAYER, R & MORENO, R. (1998). A split attention effect in multimedia learning: evidence for dual processing systems in working memory. *Journal of Educational Psychology*. v. 90, n. 2, p. 312-320.
- MAYER, R. (2001). *Multimedia learning*. New York: Cambridge University Press.
- _____. (2002). Cognitive theory and the design of multimedia instruction: an example of the two-way street between cognition and instruction. In: HALPERN, D. & HAKEL, M. (eds.). *Applying the science of learning to university teaching and beyond*. San Francisco, CA: Jossey-Bass, p. 55-71.
- _____. (ed.) (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.
- MCNEILL, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- MIYAKE, A. & SHAH, P. (2004). (eds.) *Models of working memory. Mechanics of active maintenance and executive control*. New York, NY: CUP.
- MITCHELL, R. & MYLES, F. (2004). *Second language learning theories*. New York: Hodder Arnold.
- PAIVIO, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston. (Reprinted 1979, Hillsdale, NY: Lawrence Erlbaum Associates).
- PENNYCOOK, A. (1994). *The cultural politics of English as an international language*. London: Longman.
- PHILLIPSON, R. (1992). *Linguistic imperialism*. Oxford: OUP.
- PLASS, J. (1998). Design and evaluation of the user interface of foreign language multimedia software: a cognitive approach. *Language Learning & Technology*. v. 2, n. 1, p. 40-53.
- PLASS, J.; CHUN, D.; MAYER, R., & LEUTNER, D. (1998). Supporting visual and verbal learning preferences in a second language multimedia learning environment. *Journal of Educational Psychology*. v. 90, n. 1, p. 25-36.
- PLASS, J. & JONES, L. (2005). Multimedia learning in second language acquisition. In: Mayer, R. (ed.) *The Cambridge handbook of multimedia learning*. New York: CUP. p. 467-488.

- PORQUIER, R. & VIVES, R. (1974). Sur quatre methods audiovisuelles. *Langue Francaise*. v. 24, p. 105-122.
- ROBERTSON, T. (1899). *El inglés al alcance de los niños y de los jóvenes en las escuelas y en la enseñanza particular*. New York: D. Appleton & Co.
- ROYCE, T. (2007). Multimodal communicative competence. In: ROYCE, T. and BOWCHER, W. (eds.) *New directions in the analysis of multimodal discourse*. Mahwah: NJ: Lawrence Erlbaum Associates, p. 361-390.
- SADOSKI, M. & PAIVIO, A. (2001). *Imagery and text*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- SADOSKI, M. & PAIVIO, A. (2004). A dual coding theoretical model of reading. In: RUDELL, R. B. & UNRAU, N. J. (eds.) *Theoretical models and processes of reading*. Newark, NJ: International Reading Association, p. 1329-1362.
- SCHNOTZ, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*. v. 14, n. 1, p. 101-120.
- _____. (2005). An integrated model of text and picture comprehension. In: R. MAYER, (ed.) *The Cambridge handbook of multimedia learning*. New York: CUP, p. 49-69.
- SCHNOTZ, W., BANNERT, M., & SEUFERT, T. (2002). Toward an integrative view of text and picture comprehension: Visualization effects on the construction of mental models. In: OTERO, J. A., LEÓN & GRAESSER, A. (eds.) *The Psychology of science text comprehension*. Mahwah, NJ: Lawrence Erlbaum, p. 385-416.
- SCHNOTZ, W. & BAADTE, C. (2008). Domain learning versus language learning with multimedia. In: FARÍAS and OBILINOVIC (eds.), *Aprendizaje multimodal/Multimodal learning*. Santiago de Chile: Publifahu USACH, p. 21-49.
- SEGALOWITZ, N. (2006). Automaticity and Second Languages. In: DOUGHTY, C. & LONG, M., (eds.). *The handbook of second language acquisition*, Malden, Ma: Blackwell Publishing Ltd., p. 382-408.
- TURKLE, S. H. (1995). *Life on the screen: Identity in the age of the internet*. New York: Simon & Schuster.
- VILLEGAS, A. (1951). *Little book for children*. Santiago de Chile.
- WALLACE, K. (2002). Local literacies and global literacy. In: Block, D. & Cameron, D. (eds.) *Globalization and language teaching*. London: Routledge, p. 101-114.
- WARSCHAUER, M. (1996). *Computer-assisted language learning: An introduction*. In: FOTOS, S. (ed.) *Multimedia language teaching*. Tokyo: Logos International, p. 3-20.
- WOLF, M. (2008). *Proust and the squid. The story and science of the reading brain*. Cambridge, UK: Icon Books.

Recebido: 05/05/2010

Aceito: 11/05/2011

