

Tanulmány

Sheida Marzban & Gyöngyi Fábíán
**Second language learners' reading strategies.
The case of intersemiotic relations**

Abstract

The current paper discusses a pilot study with a focus on L2 learners' multimodal reading habits. More specifically, the study seeks to investigate (1) how second language learners read multimodal texts of different image-text relations, (2) which image-text relation builds a stronger visual or verbal mental representation of the text, and (3) whether age affects multimodal reading processes. In the present study, we focus our attention on students' reading practices of two types of multimodal texts which demonstrate the image-more-general-than-text and text-more-general-than-image relations (Martinec & Salway 2005) of visual and verbal representations of meaning. To do so, 70 B1 level Hungarian English language learners attend the research to read and respond to an online multimodal reading test. The preliminary results reveal that not all intersemiotic relations lead to a stronger mental model of multimodal texts; also, the generality of the mode and the degree of participant's visual involvement with the text affect both the processing and speed of reading and information recall. In addition, as far as the participants' age is concerned, adult second language learners spend longer time on reading multimodal texts and rely on images more frequently to respond to the texts. The present research may develop more comprehensive theoretical accounts of multimodal reading among L2 learners and inform pedagogical practices.

Keywords: multimodality, L2 reading, intersemiotic relations, multimodal literacy

1 Introduction

Reading has long been multimodal including various modes such as font size and letter size to accompanying hyperlinks, audios and images in recent years. The advent of technology in the digital age allows for creating texts of varied formats to present language and has changed “the how and the what of reading” (Abraham & Farías 2017: 59). This shift to multimodal texts as the primary source of communication in the digital age adapts and transforms traditional reading models which consider reading as an individual and often passive activity and incorporates the semiotic approach to examine reading multimodal texts and develop literacy practice (Abraham & Farías 2017).

Second language learners are also actively involved in dealing with multimodal texts (written or spoken texts and images) which are believed to improve reading comprehension and facilitate learning (Pellicer-Sánchez 2022). Despite the extensive use of multimodal texts in second language learning, little research has been done to investigate how L2 readers create meaning from multiple sources of semiotics in learning materials and how these affect the

readers' attentional demands and cognitive processes (Pellicer-Sánchez 2022). Previously, multimodal texts were typically addressed in relation to the specialisation of modes, claiming, for instance, that images describe spatial information and speech or writing is apt for logic reasoning or sequential information (Kress 2003). However, this is a simplification of the multimodal perspective (Unsworth 2007) and it is essential to investigate various resources and the interaction between them and subsequently develop intersemiotic meaning making skills.

Having said that, we believe that teachers and researchers need to study learning and teaching processes from both the *production* (how information in multimodal messages is designed and delivered by the sign-maker) and the *reception* (how multimodal messages are perceived and interpreted by the users) perspectives (Bucher 2011) in the multimodal environment of the contemporary classroom. Reception studies also help the researchers understand how readers exploit different semiotic resources to reach their final understanding and whether all the modes are equally effective and helpful in learning and reading comprehension.

Thus, it is crucial to analyse second language learners' multimodal reading strategies and investigate cognitive processes during their reading and understanding of multimodal discourse. For this reason, the present article attempts to give a better understanding of second language learners' multimodal reading processes related to two types of text-image relations.

1.1 Multimodality

Visual displays form an increasing significant part of our everyday life. Moreover, compared to a few decades ago, more pictures, diagrams and graphs are used for learning and teaching to exert “supportive effects on communication, thinking, and learning” (Schnotz 2002: 102). As a result, it is not possible to think about literacy solely as a linguistic accomplishment anymore and that the time for the co-existence of language, print literacy and learning is over (Kress 2003). Thereby, learners need to comprehend, interact with and produce multimodal texts using various and potentially complex semiotic modes such as images, audios and videos (Jewitt 2008). The form of representation is essential to meaning and learning and shapes what and how is to be learned (Jewitt 2008). Hence, it is essential to better understand how students interact with the multimodal teaching and learning materials in the contemporary multimodal classroom.

Paivio's dual coding theory (1986) acknowledges the benefits of multimodal texts for literacy development and explains how pictures facilitate learning from texts (Clark & Paivio 1991).

According to his theory, two interconnected but still independent systems serve cognition, one of which processes verbal (e.g., text and speech) and the other one encodes visual information (e.g., images and graphics). Meaningful learning involves building connections between visual and verbal representations and their active processing (Mayer & Moreno 2003). Building on the Dual Coding Theory, Mayer (1997) also develops the Cognitive Theory of Multimedia Learning, according to which visual and verbal information is processed and encoded through different cognitive subsystems. Readers choose the relevant words and images, create their prepositional representations and develop the information into visual and verbal mental models established on the text which will be later integrated and mapped onto each other. Based on his multimedia theory, Mayer (2009) compiles a list of principles which defines successful multimedia learning. The *multimedia* principle explains that students learn

better from visual and verbal information integrated than from verbal information alone. The *coherence* principle states that students learn better when irrelevant material is excluded and *redundancy* occurs when the same information is presented in simultaneous multiple forms (Mayer 2001). The Generative Theory of multimedia learning (Mayer 1997) is concerned with meaningful learning which involves cognitive activities of selecting, organizing and integrating verbal and pictorial representations. It explains that the amount of information which can be processed by learners via each channel is limited (limited-capacity assumption). Cognitive Load Theory suggests that “effective instructional material facilitates learning by directing cognitive resources toward activities that are relevant to learning rather than toward preliminaries to learning” (Chandler & Sweller 1991: 293). Thus, learning is enabled when redundant information is removed from multimodal materials and avoids the learners’ additional cognitive load (Kalyuga, Chandler & Sweller 1999).

Despite Paivio and Mayer’s principles, which support the use of pictures alongside text in multimodal texts, a great deal of empirical research has been conducted in the context of domain learning, especially science in L1 (Johnson & Mayer 2012). Jakobsen (2019) states that language learning is different from domain learning as language is still considered the main mode of communication, adding that multimodality is about recognising modes in connection with language and is not necessarily concerned with replacing it. In the light of the aforementioned, more research is required to enhance our understanding of multimedia principles and their applicability in L2 learning contexts.

1.2 Multimodality and second language learning

Due to the advent of technology, the current teaching and learning environment has become increasingly multimodal and teaching materials, such as textbooks, employ multiple modes of presentation including images and texts which will hopefully support reading comprehension and facilitate learning (Pellicer-Sánchez 2022). This gradual change in knowledge representation, from language being the dominant representational mode to a greater use of various kinds of images (Bezemer & Kress 2008), necessitates research on various resources readers use and on how they move between visual and verbal information to read multimodal texts. Furthermore, studies must be conducted on the main challenges in using and combining different meaning making resources.

Although most of Mayer’s work has focused on L1 and domain learning, due to the absence of a more relevant theoretical account, his principles of multimedia learning have been applied in the interpretation of the findings from L2 multimodal studies (Pellicer-Sánchez 2022). For instance, while redundant information in multimodal materials is believed to create cognitive overload and inhibit learning by such concurrent presentations (Mayer 2009), Lee and Mayer (2018) observe a reversed redundancy effect in L2 context. For instance, according to their findings, learners who are exposed to the narrated video and onscreen text learn better than students who only watch the narrated video due to the fact that the printed text provides a more long-lasting display which contributes to their full processing of each word. Hence, it is essential to develop efficient models for L2 multimodal and multimedia language learning because its primary goal is to master a new language and not to learn about a specific domain (Schnotz & Baadte 2008).

Recent research on multimodality in L2 appears to discuss different aspects of the teaching and learning process. So far studies in multimodality and L2 learning have primarily focused

on multimodal discourse analysis of language textbooks (Chen & Qin 2007; Salbego et al. 2015; Tan 2012). The findings suggest that vigorous layout and image ease literacy learning, encourage students to participate in learning practice, support students' understanding of the activity and facilitate learning. On the other hand, a few studies investigate how teachers teach multimodal texts and relate visual forms of knowledge to learning. Ajayi (2012), for instance, states that teachers integrate visual images to contextualize instruction, however, they also have different attitudes towards multimodal texts depending on whether the images enhance second language acquisition or distract students from understanding the linguistic message. The effect on students' language acquisition is also in the focus of research. Choi and Yi (2016) mention that multimodal approach develops students' English skills and help them apply their knowledge to real life situations (Choi & Yi 2016). Hernandez (2004) examines the effect of captioned text and video on listening comprehension and vocabulary acquisition. The results suggest that the application of AVT (audio, video and text) aids vocabulary acquisition and listening comprehension most. Students also attend to the modes (video or text) which support comprehension and facilitate learning. Better performance on short-term comprehension tasks and in both short- and long-term memory vocabulary tasks result from the semantic match between audio-video-text resources (Bianchi & Ciabattini 2008). Textual enhancement in captions is also a useful pedagogical tool to develop L2 grammatical knowledge (Lee & Révész 2018). It can be concluded that empirical research into the effect of static images on reading and listening comprehension reveals that static pictures in multimodal texts improve reading comprehension and facilitate decoding words and learning vocabulary (Bisson et al., 2015; Elley & Mangubhai 1983). However, compared to the images, both young and adult L2 learners rely on written texts more frequently (e.g., Serrano & Pellicer-Sánchez 2022; Tragant & Pellicer-Sánchez 2019).

To sum up, we can claim that most L2 multimodal studies concentrate on multimodal discourse analysis of language textbooks and the effect of captioned videos on listening comprehension, vocabulary and grammar learning. Thus, more research is required to explore the effect of age and intersemiotic relations on L2 multimodal reading processes.

2 Main objectives of the research

The present piece of research is an attempt, on the one hand, to investigate whether ageing affects reading and recalling multimodal texts and, on the other hand, to shed some light on what type of image-text relation facilitates reading and maximizes comprehension. In our work we seek responses to the following questions:

1. How does the generality of a particular mode (e.g. visual or verbal) affect reading multimodal texts?
2. What mode (visual or verbal) do second language learners rely on to answer comprehension questions?
3. How does ageing affect multimodal reading and information retrieval?

3 Research design and procedure

In this section, the development of the research instrument along with the theoretical frameworks are explained. Second, the two multimodal texts considered for the present research are presented. Third, the demographic background and learning context of the participants are discussed along with the sampling procedure. Finally, the research procedure is provided in details.

3.1 Research instrument

3.1.1 Online multimodal reading test

The primary multimodal reading test instrument was developed following Martinec and Salway's (2005) and Engebretsen's (2012) frameworks. Martinec and Salway (2005) believe in a mutual relationship between visual and verbal modes. Accordingly, they do not discriminate between modes or categorise their relational types based on the type of the mode, but rather focus on interactions between visual and verbal modes together in a multimodal relation. Their model of "a generalised system of image-text relations" can be used to describe relations in which the image is subordinate to the text or the text is subordinate to the image or both modes are equally dependent or independent of each other.

According to their framework, in an unequal relationship one mode is not only subordinate but serves the other. However, in an equal relationship the modes are either independent of or complementary to each other. In a complementary relationship both modes equally depend on or modify each other. And when neither mode depends on each other, and they can exist parallel yet separately, they have an independent relationship. When one mode is subordinate, it relies on and is only related to a part of the other mode. The enhancement type heightens the other mode by providing circumstantial information of a place, time, purpose or reason. Tension between image and text also happens when the visual information is different from the verbal information (Engebretsen 2012).

The primary multimodal reading test includes twenty multimodal texts with ten different image-text relations. Two texts are designed for each type of image-text relation.

- Image-text-independent
- Image-text-complementary
- Image-text-tension
- Text-more-general-than-image
- Image-more-general-than-text
- Text-subordinate-to-image
- Image-subordinate-to-text
- Image-enhances-text-place
- Image-enhances-text-time
- Image-enhances-text-reason

In order to develop the multimodal reading test, the researcher browsed available ESL textbooks widely used in language classrooms, such as English File (3rd Edition) and Headway, and analyzed the visual and verbal relations in multimodal texts applying the frameworks described

above. The multimodal texts, which carry one of the ten image-text relations, were selected to create the test.

After each multimodal text one question with two answers was developed (see Figures 1 and 2). Among the answers the first and the second answers include verbal and visual representation of the multimodal text (visual and verbal responses), respectively. Flexiquiz website, an online quiz and assessment maker, was used to present the test, where the multimodal text was shown on the first and the question on the second slide. The backward button was also disabled, so the participants could not go back to the text when answering the question. This method was applied to enable us to investigate what mode the participants relied on more and also to identify what information they could recall more quickly in order to answer the questions.

3.1.2 Multimodal texts with IGT and TGI relations

In the present research, two types of the image-text relations of the framework are included from the study, namely image-more-general-than-text (IGT) and text-more-general-than-image (TGI) relations. According to Martinec and Salway (2005: 352), “when the level of the generality of the image and the text is different, either the image or the text can be more general”. In figure 1, which demonstrates an IGT relation, there is a general picture of a lake, while the text makes the topic more specific (Loch Ness).

The George Inn is a friendly hotel.
It has a fantastic view of Loch Ness.



What can you see from the window of the hotel?

- a. Loch Ness**
- b. The lake**

Source: retrieved from <http://whapub.com/pubs/craigdarroch-inn-am-fuaran-bar-foyers> on August 20, 2021

Figure 1. Image-more-general-than-text relation

In contrast, in Figure 2, a representation of TGI relation, the general term ‘sport’ is used in the text while the image specifies the sport through the concept of ‘volleyball’.

They like to do sports.
 They do sports in the
 park.



How do they usually spend their free time?

- a. They usually play sports.**
- b. They usually play volleyball.**

Source: retrieved from <https://www.meetup.com/denvervolleyball-103/> on August 20, 2021

Figure 2. Text-more-general-than-image relation

Originally, two multimodal texts were designed for each type of visual-verbal relation (See section 3.1.1 for more information.). However, based on the results from the online multimodal reading test, two texts were removed from the pilot questionnaire. Therefore, the results of two multimodal texts, one each type (Figure 1 and 2) are discussed in this paper.

3.2 Sampling procedure

3.2.1 English language tests and participant selection

The final sample of participants in our research was established through the application of two types of proficiency tests administered in two distinct groups of language learners. The main aim of the testing procedure was to identify the participants of the research with a proficiency level of B1.

English level B1 is the third level of English in the Common European Framework of Reference (CEFR), a definition of different language levels written by the Council of Europe. In everyday speech, this level would be called “intermediate” and is the official level descriptor in the CEFR. At this level, students are beyond the basics but are still not able to work or study exclusively in English. They have the ability to express themselves in a limited way in familiar situations and to deal in a general way with non-routine information.

On the one hand, some students from a secondary school in Veszprém, Hungary completed a set of test tasks from B1 Euroexam practice test book, which was selected and presented on the online platform Flexiquiz. The test included 3 reading tasks and 20 questions in total. The

tasks are explained in details in Table 1. No images were included in the texts. The students had 60 minutes to complete the test.

	Type of task	Number of questions
Task 1	Match the headings with the paragraphs.	8
Task 2	Read the texts and decide if the information is in text A, B, C or D.	7
Task 3	Read the text and answer the multiple-choice questions about it.	5

Table 1. Type of task and number of questions in Euroexam test

According to Euroexam website (euroexam.com) candidates need to score 60% of the total marks with a minimum of 40% in each skill (reading, writing and listening, and speaking skills) to reach B1 level. Because only the reading section of the Euroexam test was used to evaluate the participants' English language proficiency, 60% of the total marks was considered as the pass mark in our research.

Initially, 20 students in a classroom at a secondary school completed the B1 Euroexam test. The teacher had already informed the researcher that the students demonstrated around B1 English language proficiency. Based on their results, 9 students who answered 60% (or above) of the questions correctly were selected for the research. Their mean result was 82.11 with a standard deviation of 12.03, which shows that their exam scores were within the minimum of 71 and maximum of 100 (Table 2).

	N	Minimum	Maximum	Mean	Std. Deviation
Euroexam test results	9	71.00	100.00	82.11	12.03

Table 2. Descriptive statistics of Euroexam test results

On the other hand, 61 B1 second language learners from a language school in Veszprém, Hungary attended the research. The language school students had been required to complete a placement test run by the language school before enrolling for their courses. The written placement test included 60 grammar, vocabulary and language function questions in total from A1 to C1 level and the students were ranked based on the number of the questions they could answer. The language learners were later interviewed by the head teacher to evaluate their speaking skill.

3.2.2 Participants

70 Hungarian L2 learners of English attended the research. The selected sample represents two age groups of 13–30 (young adult and adult) (Group A) and 31–46 (middle-aged) (Group B) participants. There were 30 and 40 participants in Group A and B, respectively (Table 3). All the participants demonstrated the same level of proficiency (B1) at the time of the research. None of the participants studied English as their major of study.

Sheida Marzban & Gyöngyi Fábán:
Second language learners' reading strategies.
The case of intersemiotic relations
Argumentum 18 (2022), 392–409
Debreceni Egyetemi Kiadó
 DOI: 10.34103/ARGUMENTUM/2022/22

Age group	Number of participants
Group A (13–30)	30 (42.9%)
Group B (31–46)	40 (57.1%)
Total	70 (100%)

Table 3. Number and percentage of participants in Group A and Group B

Considering the gender of the participants, there were more male participants in Group A (63.3%) than group B (47.5%). In group A, male participants dominated while in Group B males and females almost shared the same percentage (Table 4).

Age group	Male	Female
Group A (13–30)	19 (63.3%)	11 (36.7%)
Group B (31–46)	19 (47.5%)	21 (52.5%)

Table 4. Number and percentage of male and female participants

3.3 Procedure of data collection and analysis

As the research was conducted during the COVID-19 lockdown and school closures period, in-person meeting with the students was not possible, so data collection was achieved completely online. To collect data from the secondary school students, first the researcher acquired the permission from the school principal to run the research. The researcher was not allowed to contact the students so the course teacher was informed about the aim and procedure of the research via email. The teacher was asked to explain the aims and procedures of the research to the students during her online lessons with the students. She also emphasized that participation in the research was completely voluntarily. The link to the Euroexam test was sent to the teacher, which she forwarded to 20 students. After finalizing the Euroexam test results, the teacher was informed about the students' achievements. 9 students who passed the exam at B1 level later received the second link to the online multimodal reading test sent by the teacher.

In order to collect data from the language school, the researcher discussed the details of her research, its aim and procedures with her B1 level students during online lessons and asked students to contribute to her research by filling in the online multimodal reading test. Later, the researcher sent the link of the online test to 72 students, which was completed by 61 persons.

By clicking on the link, all the participants were directed to Felixquiz website where they could first read the consent form consisting of a description of the study, the research process and the statement of confidentiality of the data collected during the study. The participants were required to tick the consent box and click the 'submit button'. Next, they were asked to complete the online multimodal reading test including 40 slides (20 multimodal reading texts and 20 response texts). The participants were required to register before the test entering their email address, gender, and age. After the registration, they were directed to the instruction page which asked the participants to read the images and texts thoroughly and answer the relevant questions. It was made clear that there were no right or wrong answers to the questions. The registration process and the instruction for the test was written in Hungarian in order to avoid

any misunderstandings of the language. The data were later exported to Microsoft Excel and IBM SPSS Statistics version 25 for further analysis.

4 Results

In this section, the general results from the online multimodal reading test instrument are discussed and later, the findings per age group are presented.

4.1 Multimodal reading

In this section, the preliminary results from the online multimodal reading test are set forth. Comparing the responses in TGI and IGT texts in Table 5, the results reveal that the participants select verbal responses more readily to answer the questions in both types of multimodal texts, however, the percentage of verbal responses is much higher in the TGI (88.57%) text than the IGT text (57.14%).

Image-text relation	Verbal response	Visual response
TGI	62 (88.57%)	8 (11.43%)
IGT	40 (57.14%)	30 (42.84%)

Table 5. Number and percentage of visual and verbal responses

Figure 3 shows the average time (mean duration) the participants spend on reading and responding to the texts. The results indicate that the participants spend less on average reading a TGI text (10s) compared to an IGT text (13s). The mean duration of response also indicates that the longer time they spend on reading texts the shorter the information retrieval process is (13s to 10s for the IGT text) and vice versa (10s to 14s for the TGI text).

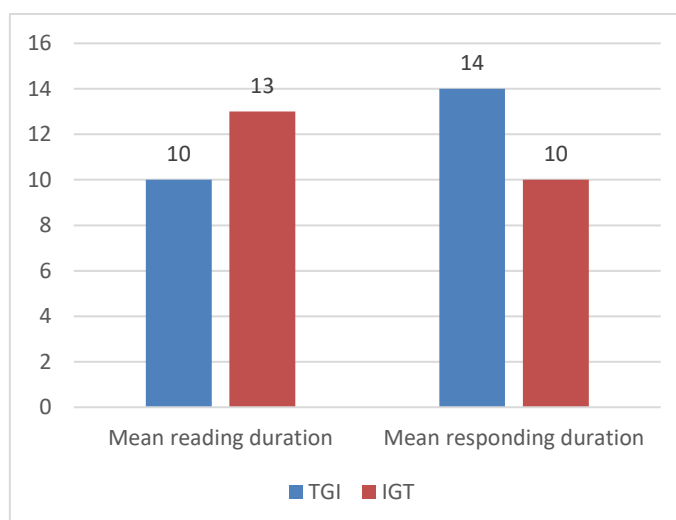


Figure 3. Mean reading and response duration (in seconds)

It can be inferred from the results that in general participants trust more in the written text than the image. However, the effect of the generality of the mode on the reading and information recall process cannot be ignored. The preliminary findings reveal that the general mode provides a ground for the participant's knowledge and comprehension, thereby speeding up and facilitating the reading process. Consequently, participants spend shorter time on reading TGI as their reliable source of information is also more general than the image which accelerates the reading task and discourages the participants from searching the image. However, the specific information in the verbal information of IGT text slows down the comprehension and increases the reading time and, in consequence, might direct the attention to the image in search of some scaffolding for their comprehension. The findings also confirm that the longer time participants spend reading the text, the shorter will the response time be; conversely, the long information recall duration is preceded by a short scrutiny of the multimodal text. The results may allow for the conclusion that longer reading duration indicates higher involvement with the multimodal text and easier subsequent information recall. It can also be inferred that the presence of a general image in IGT provides some background information which increases the participants' interaction and involvement with the image, and thereupon contributes to a stronger mental representation of the multimodal text, faster response time and a higher percentage of visual responses. However, the general verbal information in TGI does not encourage the readers to interact with the image, which results in a looser mental representation of the multimodal text, longer recall information duration and a higher percentage of verbal responses.

4.2 Age and multimodal reading

In this section, the effect of age on multimodal reading is outlined. The results show that Group A prefer verbal responses more frequently than Group B in both texts, however, the difference is more considerable in the IGT text. Group A also spend shorter time on average reading and responding to the texts than Group B. The findings also reveal that the reading time affects the response duration and that recalling verbal information takes a shorter time than that of the visual information.

According to Table 6, in total, Group A (83.33%) choose verbal responses in multimodal texts more frequently than Group B (65%). Considering the type of the image-text relation, the difference is even more striking in the IGT text where 70% of the participants in Group A rely on verbal information, which is much higher than in group B (47.50%). Regarding the TGI text, 96.66% of participants in Group A select the verbal response while the percentage for Group B is 82.5%.

Image-text relation	13–30 (group A)		31–46 (Group B)	
	Visual	verbal	Visual	verbal
TGI	3.33	96.66	17.5	82.5
IGT	30	70	52.50	47.50
Total	16.66	83.33	35	65

Table 6. Age groups and percentage of visual and verbal responses

An independent samples t-test was also run to compare the responses for Group A and Group B. The results found a significant difference ($t(68)=2.246$, $p=.028$) between the mean scores of responses in Group A ($M=1.66, SD=.54667$), which chose verbal responses more often than Group B ($M=1.30, SD=.75786$).

Considering the average reading and response duration (Mean of means), Group A, on average, spend less time on reading and responding to texts (10s) than Group B (13.25s) (Table 7). Running the independent samples t-test also confirms that this difference is significant ($t(68)=-2.213$, $p=.030$) between Group A ($M=0.1014$, $SD=.04516$) and Group B ($M=.1309$, $SD=.06146$).

The results also corroborate that the reading time affects the response duration. The longer the reading time, the shorter and easier the information recall is, and vice versa, especially for Group B who spent 11s and 16s to read and respond to the TGI text. Conversely, it took them 15s and 11s to read and respond to the IGT, respectively.

Image-text relation	13–30 (Group A)		31–46 (Group B)	
	Mean reading duration	Mean response duration	Mean reading duration	Mean response duration
TGI	10	10	11	16
IGT	11	9	15	11
Mean	10.5	9.5	13	13.5
Mean of means	10		13.25	

Table 7. Age group and mean reading and response time duration (in seconds)

Moreover, there is a strong negative correlation between the type of responses (visual and verbal) and the response duration ($r=-.318$, $n=70$, $p=.007$) which shows that recalling verbal takes a shorter time than recalling visual information.

In sum, Group A spends less time reading and responding to multimodal texts and they also highly trust verbal information to answer the questions. However, Group B has a longer duration of reading and response time and rely on images more frequently to respond to the texts. The reading time also affects the duration of response time which means the longer the reading takes, the shorter the information retrieval process will be. Moreover, the findings explain that the time period for recalling verbal information is shorter than recalling visual information.

5 Conclusion

The present research attempts to investigate how L2 learners of different age range read and process information in multimodal texts with different intersemiotic relations. The findings indicate that participants consider language as the main source of information which they trust more often. This could result from the fact that English language proficiency for L2 learners has traditionally meant and encompassed mastering of linguistic knowledge and skills (e.g., vocabulary, grammar and reading) for comprehension and communication (Hung, Chiu & Yeh 2013: 400). This conventional notion considers literacy only as a set of decoding and encoding skills which L2 learners need to master. Jakobsen (2019) believes that language learning is different from domain learning as language is still considered to act as the focal mode.

Moreover, in educational settings students are usually assessed based on the end product rather than the process of learning. Prevailing monomodal methods of assessment and the lack of multimodal assessment framework (Unsworth 2008) hinder students from developing critical multimodal reading ability and perceptions. The presence of the questions in the multimodal reading test might create a washback effect (Alderson & Wall 1993) by reminding participants of school and university exams and, as a result, prevent them from tapping into higher-order reading skills. This can be confirmed by the fact that some participants had expressed their concerns to the researcher about their possible grades before the test.

The results also reveal that visual and verbal information is processed in different cognitive subsystems (Paivio 1986). The participants move between the visual and verbal information to evaluate their modal loads, identify the relation between the modes, build visual and verbal mental representations of the text and later map them onto each other to construct meaningful reading paths. Paivio (1979) believes the information which is memorised through words (left hemisphere of the brain) and a picture (in the right hemisphere of the brain) sets up a powerful combination in the memory. In this research, the generality of the mode and the degree of the participant's visual involvement with the text affect the reading and information recall processing and speed. The participants spend shorter time on reading TGI where the general written text does not invite the reader to search the detailed image for extra information. Proposing an example of sport (volleyball), the component image narrows down and does not expand the concept of sport which increases the uncertainty condition for the readers. The distracting details in the image confuse the readers (Pan & Pan 2009) and partly result in a looser mental representation of the multimodal text and also in longer duration of response. Regarding the IGT text, detailed information (e.g., Loch Ness) in the verbal discourse slows down the comprehension and increases the reading time, which results in learners' searching for the visual support. The general image builds a familiar ground (a lake) and acts as a schema for the readers to build their knowledge on and develop their comprehension, which consequently contributes to a stronger mental representation of the multimodal text and faster response time. The picture promotes reading comprehension and has a more facilitative effect by providing information that is difficult for the readers to comprehend through the text (Gyselinck & Tardieu 1999; Pan & Pan 2009). Another reason could be that in IGT text the image repeats and expands the information in the verbal information, and consequently aids the task of recall and recognition, called redundancy principle by Mayer (2001).

Pictures can improve reading comprehension and recall when they do not include too much information from the content of the reading text, represent the language complexity of the text, and primarily, supplement the representation of the information which the text itself also invites the readers to process (Levin 1983; Omaggio 1979). To conclude, although pictures play a positive role in supporting the readers' comprehension as additional sources of information, not all intersemiotic relationships lead to a stronger mental model of multimodal texts. Pictures primarily supplement the information where the text itself invites the readers to read and process. Nevertheless, images can regulate their readers' mnemonic effectiveness.

Considering the age of the participants, the findings reveal that it takes longer for Group B (31–46) to read and respond to texts due to their age (İyigün, Bekircan, Maviş 2018) and in addition, they rely on images more frequently to respond to the texts. According to Chen et al (2019), reading speed increases with age but deteriorates when adults reach 40 and above, which could be due to the decline across different domains of cognitive function (Cullum et al. 2000), memory decline (Grady & Craik 2000), poorer working memory (Park et al. 2002) and

reduced information processing speed (Eckert et al. 2010). Thus, they more frequently look for visual cues which reinforce information provided in the text and help improve comprehension.

Another reason could be that Group A (13–30) are digital natives (Marsh 2005), who are surrounded by multimodal information, videogames and social media and interact with and process the knowledge faster than their prior generation (Prensky 2001). Using the internet also involves reading a lot of creative and non-linear multimodal (e.g., hypertext and hyperlinks) texts with eye-catching visual effects in wide varieties of formats, lengths and topics, which constantly compete for the attention of the reader (Afflerbach & Cho 2010). Therefore, students might power down and lose interest when they read static linear multimodal texts which are different from the digital multimodal world and may end up skipping several words. This demonstrates shorter reading duration and lower visual engagement.

Hardiess and Weissert (2021: 2) also mention that “people have a habit of reading text because our prior experiences have taught us that text often conveys important information and incorporate text as high-level (feature) object”. In addition to language schools, what other educational settings (e.g. schools, colleges and universities) offer students and prepare them for in domain learning differs from the visual and pictorial world outside school (Kress 2003) and students are mostly habituated into learning the language content. This habit formation during education – starting from early childhood and continuing in adulthood at universities – might gradually fade away as students graduate. This might explain why Group B (31–46) trust visual information more frequently to answer the question as they are no longer in educational settings where language and its content are assessed and overvalued.

6 Implications of the study and future research

The present research can help textbook designers and material and app developers to analyse available multimodal materials in educational settings to see if they are pacing up with the needs of students in different age groups. While young adults may prefer more engaging and attractive multimodal texts, middle-aged adults would prefer simple images which develop their reading comprehension. Teachers and curriculum designers' knowledge and awareness of multimodal reading contributes to students' engagement, interpretation and attention through an effective combination of semiotic modes of teaching materials. Teachers are also encouraged to develop meaningful multimodal texts and create attractive and useful interactive resources which maximize learning potentials and exercise students' metacognitive and multimodal literacy skills.

In the present article, only the results from the online multimodal reading test instrument of the pilot study are presented. Eye tracking measurements can also offer valuable empirical insights on users' interaction with multimodal texts and deliver rich and precise data on perceptual and cognitive processing during multimodal reading. Future research can include a larger sample of texts with different degree of visual and verbal generality and a larger group of participants to refine and generalise our findings related to the multimodal reading processes of L2 learners. Also, more research is required to use the eye-tracking tool and to examine the effect of other visual-verbal relations (e.g., image-text tension, image-text enhancement) on reading processes. Moreover, interviews and questionnaires can be adopted to investigate the young and middle-aged adult L2 learners' attitudes to multimodal texts and the role of visual information in reading comprehension. Collecting further empirical evidence will surely

provide a more comprehensive and efficient model of L2 multimodal and multimedia language learning environment.

References

- Abraham, P. & Fariás, M. (2017): Reading with eyes wide open: Reflections on the impact of multimodal texts on second language reading. *Íkala: Revista de Lenguajey Cultura* 22. 1, 57–70. <https://doi.org/10.17533/udea.ikala.v22n01a04>
- Afflerbach, P. A. & Cho, B. Y. (2010): Determining and describing reading strategies: Internet and traditional forms of reading. In: Waters, H. S. & Schneider, W. (eds.): *Metacognition, Strategy Use, and Instruction*. New York: Guilford Press, 201–255.
- Ajayi, L. (2012): How teachers deploy multimodal textbooks to influence English language learners. *TESOL Journal* 6, 16–35.
- Alderson, J. & Wall, D. (1993): Does washback exist? *Applied Linguistics* 14, 115–129. <https://doi.org/10.1093/applin/14.2.115>
- Bezemer, J. & Kress, G. (2008): Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication* 25. 2, 166–195. <https://doi.org/10.1177/0741088307313177>
- Bianchi, F. & Ciabattoni, T. (2008): Captions and subtitles in EFL learning: An investigative study in a comprehensive computer environment. In: Baldry, A., Pavesi, M. & Taylor Torsello, C. (eds.): *From Didactas to Ecolingua*. Trieste: Edizioni Università di Trieste, 69–90.
- Bisson, M. J., van Heuven, W. J., Conklin, K. & Tunney, R. J. (2015): The role of verbal and pictorial information in multimodal incidental acquisition of foreign language vocabulary. *Quarterly Journal of Experimental Psychology* 68. 7, 1306–1326. <https://doi.org/10.1080/17470218.2014.979211>
- Bucher, H. J. (2011): Multimodalität – eine Universalie des Medienwandels: Problemstellungen und Theorien der Multimodalitätsforschung. In: Bucher, H.-J. et al. (eds): *Neue Medien – Neue Formate. Ausdifferenzierung und Konvergenz in der Medienkommunikation*. Frankfurt am Main: Campus, 41–79.
- Chandler, P. & Sweller, J. (1991): Cognitive load theory and the format of instruction. *Cognition and Instruction* 8. 4, 293–332. https://doi.org/10.1207/s1532690xci0804_2
- Chen, A. H., Khalid, N. M. & Buari, N. H. (2019): Age factor affects reading acuity and reading speed in attaining text information. *International Journal of Ophthalmology* 12. 7, 1170–1176. <https://doi.org/10.18240/ijo.2019.07.19>
- Chen, Y. M. & Qin, X. Y. (2007): Multimodal Engagement Resources and Voice Interaction in Textbook Discourse, *Foreign Languages and Their Teaching* 12, 15–18.
- Choi, J. & Yi, Y. (2016): Teachers' integration of multimodality into classroom practices for English language learners. *TESOL Journal* 7. 2, 304–327. <https://doi.org/10.1002/tesj.204>
- Clark, J. M. & Paivio, A. (1991): Dual coding theory and education. *Educational Psychology Review* 3, 149–210. <https://doi.org/10.1007/BF01320076>
- Cullum, S., Huppert, F. A., McGee, M., Denning, T., Ahmed, A., Paykel, E. S., & Brayne, C. (2000): Decline across different domains of cognitive function in normal ageing: Results of a longitudinal population-based study using CAMCOG. *International Journal of Geriatric Psychiatry* 15, 853–862. [https://doi.org/10.1002/1099-1166\(200009\)15:9<853::AID-GPS211>3.0.CO;2-T](https://doi.org/10.1002/1099-1166(200009)15:9<853::AID-GPS211>3.0.CO;2-T)

- Eckert, M. A., Keren, N. I., Roberts, D. R., Calhoun, V. D. & Harris, K. C. (2010): Age-related changes in processing speed: Unique contributions of cerebellar and prefrontal cortex. *Frontiers in Human Neuroscience* 4, 10. <https://doi.org/10.3389/neuro.09.010.2010>
- Elley, W. B. & Mangubhai, F. (1983): The impact of reading on second language learning. *Reading Research Quarterly* 19. 1, 53–67. <https://doi.org/10.2307/747337>
- Engelbrechtsen, A. (2012): Balancing cohesion and tension in multimodal rhetoric. An interdisciplinary approach to the study of semiotic complexity. *Learning, Media and Technology*, 37. 2, 145–162. <https://doi.org/10.1080/17439884.2012.655745>
- Grady, C. L. & Craik, F. I. (2000): Changes in memory processing with age. *Current Opinion in Neurobiology* 10. 2, 224–231. [https://doi.org/10.1016/S0959-4388\(00\)00073-8](https://doi.org/10.1016/S0959-4388(00)00073-8)
- Gyselinck, V. & Tardieu, H. (1999): The role of illustrations in text comprehension: What, when for whom, and why? In: Oostendorp, H. Van. & Goldman, S. R. (eds.): *The Construction of Mental Representations During Reading*. Mahwah, NJ: Lawrence Erlbaum, 195–218.
- Hardiess, G. & Weissert, C. (2021): Interaction between image and text during the process of biblical art reception. *Journal of Eye Movement Research* 13. 2. <https://doi.org/10.16910/jemr.13.2.14>
- Hernandez, S. S. (2004): *The effects of video and captioned text and the influence of verbal and spatial abilities on second language listening comprehension in a multimedia learning environment*. Unpublished PhD dissertation. New York University, New York, United States.
- Hung, H., Chiu, Y. J. & Yeh, H. (2013): Multimodal assessment of and for learning: A theory-driven design rubric. *British Journal of Education Technology* 44. 3, 400–409. <https://doi.org/10.1111/j.1467-8535.2012.01337.x>
- İyigün, E. , Bekircan, M. & Maviş, İ. (2018): Yetişkin Bireylerde Yaş, Eğitim ve Cinsiyet Değişkenlerinin Okuma Hızına Etkisi. *Dil Konuşma ve Yutma Araştırmaları Dergisi* 1. 2, 162–178.
- Jakobsen, I. K. (2019): Inspired by image: A multimodal analysis of 10th grade English school-leaving written examinations set in Norway (2014–2018). *Acta Didactica Norge* 13. 1, 1–27. <https://doi.org/10.5617/adno.6248>
- Jewitt, C. (2008): Multimodality and literacy in school classrooms. *Review of Research in Education* 32. 1, 241–267. <https://doi.org/10.3102/0091732X07310586>
- Johnson, C. I. & Mayer, R. E. (2012): An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied* 18. 2, 178–191. <https://doi.org/10.1037/a0026923>
- Kalyuga, S., Chandler, P. & Sweller, J. (1999): Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology* 13. 4, 351–371. [https://doi.org/10.1002/\(SICI\)1099-0720\(199908\)13:4<351::AID-ACP589>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6)
- Kress, G. (2003): *Literacy in the New Media Age*. London: Routledge. <https://doi.org/10.4324/9780203299234>
- Lee, H. & Mayer, R. (2018): Fostering learning from instructional video in a second language. *Applied Cognitive Psychology* 32, 648–654. <https://doi.org/10.1002/acp.3436>
- Lee, M. & Révész, A. (2018): Promoting grammatical development through textually enhanced captions: An eye-tracking study. *The Modern Language Journal* 102, 557–577. <https://doi.org/10.1111/modl.12503>
- Levin, R. (1983). Pictorial strategies for school learning: practical illustrations. In: Pressley, M. & Levin, R. (eds.): *Cognitive strategy research: Educational applications*. New York, NY: Springer-Verlag, 213–237. https://doi.org/10.1007/978-1-4612-5519-2_8

- Marsh, J. (2005): *Popular Culture, New Media and Digital Literacy in Early Childhood*. London: Routledge/Falmer. <https://doi.org/10.4324/9780203420324>
- Martinec, R. & Salway, A. (2005): A system for image-text relations in new (and old) media. *Visual Communication* 4. 3, 337–371. <https://doi.org/10.1177/1470357205055928>
- Mayer, R. E. (1997): Multimedia learning: Are we asking the right questions? *Educational Psychologist* 32, 1–19. https://doi.org/10.1207/s15326985ep3201_1
- Mayer, R. E. (2001): *Multimedia Learning*. New York, NY: Cambridge University Press.
- Mayer, R. E. (2009): *Multimedia Learning*. 2nd ed. Cambridge: Cambridge University Press.
- Mayer, R. E. & Moreno, R. (2003): Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist* 38, 43–52. https://doi.org/10.1207/S15326985EP3801_6
- Omaggio, C. (1979): Pictures and second language comprehension: Do they help? *Foreign Language Annals* 12, 107–116. <https://doi.org/10.1111/j.1944-9720.1979.tb00153.x>
- Paivio, A. (1979): *Imagery and Verbal Processes*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Paivio, A. (1986): *Mental Representations: A Dual Coding Approach*. Oxford: Oxford University Press.
- Pan, Yi-Chun, & Pan, Yi-Ching (2009): The effects of pictures on the reading comprehension of low-proficiency Taiwanese English foreign language college students: An action research study. *VNU Journal of Science, Foreign Languages* 25, 186–198.
- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D. & Smith, P. K. (2002): Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging* 17. 2, 299–320. <https://doi.org/10.1037/0882-7974.17.2.299>
- Pellicer-Sánchez, A. (2022): Multimodal reading and second language learning. *International Journal of Applied Linguistics* 173. 1, 2–17. <https://doi.org/10.1075/itl.21039.pe1>
- Prensky, M. (2001): Digital natives, digital immigrants, part 1. *On the Horizon* 9. 5, 1–6. <https://doi.org/10.1108/10748120110424816>
- Salbego, N., Heberle, V. M., & Balen, M. G. S. da S. (2015): A visual analysis of English textbooks: Multimodal scaffolded learning. *Calidoscopio* 13. 1, 5–12. <https://doi.org/10.4013/cld.2015.131.01>
- Schnotz, W. (2002): Towards an integrated view of learning from text and visual displays. *Educational Psychology Review* 14, 101–120. <https://doi.org/10.1023/A:1013136727916>
- Schnotz, W. & Baadte, C. (2008): Domain learning versus language learning with multimedia. In: Farías, M. & Obilinovic, K. (eds.): *Aprendizaje Multimodal/Multimodal Learning*. Santiago de Chile: Publifahu USACH, 21–49.
- Serrano, R. & Pellicer-Sánchez, A. (2022): Young L2 learners' online processing of information in a graded reader during reading-only and reading-while-listening conditions: A study of eye-movements. *Applied Linguistics Review* 13. 1, 49–70. <https://doi.org/10.1515/applirev-2018-0102>
- Tan, Y. (2012): *A multimodal discourse analysis of the senior high school English textbooks Advance with English* (Unpublished MA Thesis). Hunan University of Science and Technology, China.
- Tragant, E. & Pellicer-Sánchez, A. (2019): Young learners' engagement with multimodal exposure: An eye-tracking study. *System* 80, 212–223. <https://doi.org/10.1016/j.system.2018.12.002>

Sheida Marzban & Gyöngyi Fábán:
Second language learners' reading strategies.
The case of intersemiotic relations
Argumentum 18 (2022), 392–409
Debreceni Egyetemi Kiadó
DOI: 10.34103/ARGUMENTUM/2022/22

- Unsworth, L. (2007): Image/text relations and intersemiosis: Towards multimodal text description for multiliteracies education. In: Barbara, I. & Sardinha, T. B. (eds.): *Proceedings of the 33rd International Systemic Functional Congress (33rd ISFC)*. Pontificia Universidade Católica De Sao Paulo (PUCSP), Sao Paulo, Brazil, 1165–1205.
- Unsworth, L. (2008): Multiliteracies and metalanguage: Describing image/text relations as a resource for negotiating multimodal texts. In: Coiro, J. et al. (eds.): *Handbook of Research on New Literacies*. New York: Lawrence Earlbaum, 377–405.

Sheida Marzban
University of Pannonia
Doctoral School of Multilingualism
Egyetem u. 10
H-8200 Veszprém
sheida.marzban@phd.uni-pannon.hu

Dr. habil Gyöngyi Fábán
University of Pannonia
Egyetem u. 10
H-8200 Veszprém
fabian.gyongyi@mftk.uni-pannon.hu