Examining the Modality Effect in a Second Language Multimedia Learning Environment[[1]](#footnote-1)\*

Kadir Kozana, Gülcan Erçetina and Jennifer C. Richardsonb

aFaculty of Education, Boğaziçi Üniversitesi, 34342 Bebek, Istanbul, Turkey

bCollege of Education, Purdue University, 100 N. Street, 47907 West Lafayette, IN, USA

Author Note

Kadir Kozan is now at the College of Education, Department of Curriculum and Instruction, Purdue University, 100 N. Street, 47907 West Lafayette, IN, USA

Correspondence can be sent to Kadir Kozan, College of Education, Department of Curriculum and Instruction, Purdue University, 100 N. Street, 47907 West Lafayette, IN, USA. Tel.: +1 765 496 2978; fax: +1 765 494 5832; e-mail: kkozan@purdue.edu

E-mail addresses: kkozan@purdue.edu (K. Kozan), gulcaner@boun.edu.tr (G. Erçetin), jennrich@purdue.edu (J. Richardson).

**Abstract**

This study investigated the modality effect in relation to working memory capacity

within a computerized second language multimedia environment. Twenty-nine

advanced Nationality learners of English completed a reading span test and were

randomly assigned to read or listen to an English text about an unfamiliar topic.

Learners also completed immediate and delayed retention and transfer comprehension

tests. Results revealed inconsistent effects on retention and transfer performance:

There was a significant combined effect of time, modality of presentation and working memory capacity on participants' retention performance, with the effect of time having a significant effect on transfer performance. Analyses on the significant triple interaction revealed that although presentation mode and working memory capacity play some role in retention performance, their effects emerged dependent upon one other and time of testing.

*Keywords:* cognitive load; working memory; the modality effect

1. **Introduction**

Multimedia instruction means presenting information through words and pictures to enhance learning (Mayer, 2009; Mayer & Moreno, 2003). Words refer to written or spoken textual information accompanied by corresponding pictorial materials ranging from pictures to animations. A fundamental question is how to effectively present information or content in order to promote learning in computer-assisted learning environments. To this end, this study was framed by the cognitive/generative theory of multimedia learning (CTML: Mayer, 1997, 2001, 2005, 2009; Mayer & Moreno, 2003) and cognitive load theory (CLT: Chandler & Sweller, 1991; Sweller, 1988, 1994, 2005; Sweller, van Merriënboer, & Paas, 1998), focusing on the modality effect (Mousavi, Low, & Sweller, 1995; Sweller, Ayres, & Kalyuga, 2011) or the modality principle (Mayer, 2001, 2009).

**1.1. Cognitive/Generative Theory of Multimedia Learning**

CTML assumes that multimedia learning happens in limited visual and auditory channels of working memory (WM) through the active processes of selecting, organizing and integrating information. Mayer (2009) argued that “The central work of multimedia learning takes place in working memory” (p. 62). Likewise, Sweller et al. (2011) stated that selecting, organizing and integrating visual and auditory information occurs in WM. WM stores and manipulates information for the short-term prior to learned information becoming stored in long-term memory (Baddeley, 1992). Prior or familiar knowledge may facilitate learning by decreasing possible WM load since multiple information pieces organized as a schema in long-term memory are managed as a single unit by WM (Antonenko, Paas, Grabner, & van Gog, 2010; Kalyuga, Chandler, & Sweller, 1999; Sweller et al., 1998; Sweller et al., 2011; van Gog & Paas, 2008).

Furthermore, integrating Mayer`s (2001) CTML model and a second-language acquisition perspective, Plass and Jones (2005) created a second-language acquisition with multimedia model. Both models imply that learning can be facilitated by simultaneous presentation of auditory and visual information thereby decreasing WM load and leaving enough room for active processing. Plass and Jones (2005) suggested that cognitive load needs to be watched for while designing comprehensible input using multimedia environments in which the computer can function “as a partner in the learning process.” (p. 471). Therefore, although technology may not always lead to higher language learning outcomes (Salaberry, 1999), it is reasonable to assume that computer technology can support multimedia learning and cognitive processes associable with it thereby fostering learning. For instance, the computer can encourage audiovisual or multisensory information processing of learning materials through audiovisual presentation.

**1.2. Cognitive Load Theory**

Cognitive load (CL) is imposed on WM resources when a specific task is performed (Sweller et al., 1998). Three types of CL are covered by cognitive load theory (CLT): intrinsic, extraneous, and germane (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Previously, extraneous CL, imposed by poorly designed instructional materials, germane CL, the amount of cognitive capacity dedicated to learning, and intrinsic CL, created by element interactivity were thought to be additive (Paas et al., 2003; Sweller, 2005; Sweller et al., 1998). Recently, Sweller (2010) suggested that only intrinsic and extraneous loads are additive to explain empirically found changes in total CL. Similarly, Kalyuga (2011) asserted that intrinsic load relates to both learning materials and learning goals thereby covering germane load as the effective load leading to learning.

**1.2.1. Cognitive load theory and working memory.**

CLT assumes that learning occurs through a limited WM that performs learning tasks, and an unlimited long-term memory that functions as storage for what is learned. WM processes information through auditory and visual channels that are limited in terms of the number of input processed at a time (e.g., Brünken, Plass, & Leutner, 2004; Brünken, Steinbacher, Plass, & Leutner, 2002). Likewise, Carlson, Chandler, and Sweller (2003) suggested that if WM has to tackle multiple information elements simultaneously, learning is less likely to happen; if these elements are processed successively or serially, then, WM load is reduced.

Therefore, limited WM is important to both CLT and CTML. According to Tardieu and Gyselinck (2003), research on how WM relates to or moderates multimedia learning is necessary for validation of theoretical insights based on WM load (as cited in Schüler et al., 2011, p. 390). Consequently, Schüler et al. (2011) suggested that “it is worthwhile to investigate empirically how working memory is involved during multimedia learning.” (p. 392).

**1.2.2. Cognitive load theory and prior knowledge.**

CLT also assumes that WM can become limited especially when dealing with novel information because it needs to be organized in the long term memory (Sweller, 2004). According to Sweller (2003), as learners’ familiarity with the information increases, the manner in which they deal with information changes dramatically (as cited in Kalyuga & Sweller, 2004, p. 559). In other words, a specific format of instruction may promote learning for some learners while impeding for others depending on learner expertise (Yeung, Jin, & Sweller, 1997). Sweller (2005) and Sweller et al. (2011) also claimed that prior knowledge functions as the only central executive for further learning. In essence, familiar information imposes less intrinsic load than unfamiliar information (McCrudden, Schraw, Hartley, & Kiewra, 2004).

**1.3. The Modality Effect**

The modality effect encourages presenting instructional materials audio-visually. This way, both of the limited channels (auditory and visual) of WM are used. Mayer and Moreno (2003), and Mayer (2009) described this as off-loading the visual channel. Additionally, Tabbers (2002) claimed that superiority of audiovisual presentation over visual only presentation may stem from decreased visual search and more efficient use of time on task as the visual and auditory information can be processed simultaneously.

Research revealed that audiovisual presentation results in enhanced learning (e.g., Leahy, Chandler, & Sweller, 2003; Mayer & Moreno, 1998; Moreno & Mayer, 1999). However, several studies also found no modality effect in different contexts under certain conditions such as educational sciences (e.g., De Westelinck, Valcke, De Craene, & Kirschner, 2005) and reasoning (e.g., Goolkasian, 2000), and in some cases even showed the reverse modality effect (e.g., Tabbers, Martens, & van Merriënboer, 2004). Benefits associated with audiovisual presentation have been varied. For instance, it was found that audiovisual presentations result in less time invested in problem solving (Jeung, Chandler, & Sweller, 1997; Mousavi et al., 1995) and quicker reaction times on a visual secondary task (Brünken et al., 2002). It was also found that the audiovisual mode: (a) enhances knowledge acquisition (Brünken et al., 2004; Brünken & Leutner, 2001, 2002, as cited in Brünken et al., 2002); (b) promotes better performance on retention, transfer, and matching tests (Mayer & Moreno, 1998; Moreno & Mayer, 1999); and (c) leads to “lower reattempts at interactive exercises” (Kalyuga et al., 1999, p. 368). Research also revealed lower CL ratings by both experienced and inexperienced learners and better performance on a multiple-choice test for inexperienced learners (Kalyuga, Chandler, & Sweller, 2000). Finally, Tabbers (2002) revealed lower CL ratings in audiovisual instructional phase, and better retention and transfer scores in system-paced audiovisual presentations, which turned out to be either the reverse or equivalent case in learner-paced instruction.

Research further suggests that the modality effect happens when intrinsic CL is high (Sweller, 1994, 2010; Sweller & Chandler, 1994). For instance, learning materials that can be processed consecutively impose little intrinsic CL, therefore, decreasing extraneous CL may not be of great importance (Carlson et al., 2003). Likewise, it is important to reduce the total amount of CL when learners deal with learning materials having high element interactivity (Carlson et al., 2003; Paas, Renkl, & Sweller, 2003; Sweller et al., 1998). For instance, Leahy et al. (2003) contended that the superiority of audiovisual presentation over visual only presentation was greater under high element interactivity. Similarly, Tindall-Ford, Chandler, and Sweller (1997) showed that the modality effect appears through instruction carrying high intrinsic load only. Ginns’s (2005) meta-analysis also showed that learning materials having low element interactivity are associated with smaller average modality effect sizes than the ones with high element interactivity.

Moreover, despite modality studies favoring system-pacing compared to learner pacing (e.g., Ginns, 2005; Tabbers, 2002), the reverse modality effect (e.g., Tabbers et al., 2004), and no modality effect under self-pacing (e.g., Wouters, Paas, & van Merriënboer, 2009), some other studies under certain conditions either replicated the modality effect in self-paced learning environments (e.g., Harskamp, Mayer, & Suhre, 2007) or found almost no difference between system-pacing and self-pacing as to learning outcomes (e.g., Schmidt-Weigand, Kohnert, & Glowalla, 2010). While accepting that the modality effect may disappear under self-paced instruction, Sweller et al. (2011) asserted that the reverse modality effect can emanate from “lengthy, complex, auditory” information (p. 137). Varying the length of verbal information presented, Leahy and Sweller (2011) reached the reverse modality effect with longer textual information while the modality effect recovered when information was presented in shorter segments.

Consequently, most of the available studies are first language studies where the instruction was presented in a first language (mostly English). They used immediate tests only as the measure of comprehension or understanding (e.g., Mayer & Moreno, 1998; Moreno & Mayer, 1999) with a few exceptions (Segers, Verhoeven, & Hulstijn-Hendrikse, 2008; Witteman & Segers, 2010). Additionally, most of them were conducted in controlled laboratory settings where participants were engaged in a one-time brief learning experience (Moreno, 2006; Tabbers, 2002); instructions employed were either system-paced or paper-based and were limited to the total duration of the narration (Tabbers, 2002); and included learning materials from exact sciences (Tabbers, 2002). Therefore, it is reasonable to claim that the results obtained so far are limited to short-term learning outcomes and to L1 learning contexts to a great extent.

* + 1. The Present Study

This study investigated the modality effect in relation to second/foreign language (L2) reading and listening using a text since the application of the modality effect in L2 learning is mostly based on presentation of multimedia glosses and annotations to facilitate reading comprehension or vocabulary learning (e.g., Ariew & Erçetin, 2004; Chun & Plass, 1996; Şakar & Erçetin, 2005). Moreover, the present study incorporated a computerized L2 reading span task in order to see how WM contributes to multimedia learning in an L2 context. Finally, except for a few studies (e.g., Segers et al., 2008; Witteman & Segers, 2010) less is known concerning the long term effects of audiovisual presentation. As a result, the delayed effects were also examined in this study. Specifically, what are the immediate and delayed effects of presentation mode and WM capacity on retention and transfer performance in L2 reading versus listening?

1. **The Present Study**

There were three hypotheses examined. Firstly, time was hypothesized to have a deteriorating effect on both retention and transfer of information based on previous research (e.g., Banikowski & Mehring, 1999; Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008) (Hypothesis 1). Secondly, possible performance differences between audiovisual and visual only presentation modes predicted by the modality effect proposed by CLT (e.g., Sweller, 1988, 1994; Sweller et al., 1998) and CTML (Mayer, 1997, 2001, 2009) were not expected to be statistically significant over time due to the mediating effect of self-paced instruction (Tabbers, 2002; Tabbers et al., 2004; Ginns, 2005; Sweller et al., 2011) (Hypothesis 2).

Finally, as for Hypothesis 3, participants with high WM capacity were hypothesized to perform better on comprehension tests over time compared to those with low WM capacity based on research demonstrating the beneficial effects of WM capacity on language comprehension and other higher-level cognitive skills (Daneman & Carpenter, 1980; Gathercole & Baddeley, 1993; Hambrick & Engle, 2002; Hambrick & Oswald, 2005; Harrington & Sawyer, 1992; Juffs & Harrington, 2011; Miyake & Friedman, 1998; Walter, 2004) as well as the relationship between multimedia learning and WM (e.g., Austin, 2009; Brunye, Taylor, Rapp, & Spiro, 2006; Gyselinck, Cornoldi, Dubois, De Beni, & Ehrlich, 2002; Gyselinck, Jamet, & Dubois, 2008; Schüler, Scheiter, & van Genuchten, 2011).

**2.1. Method**

**2.1.1. Participants.**

29 university students (27 females and 2 males; *M* = 20.31 years, SD= 1) participated in the experiment. The participants were given partial course points for their contribution. They were enrolled in the Department of Foreign Language Education at a Turkish university where English is the medium of education. The participants were considered to be advanced English learners as they had all passed the English proficiency test of the university with a minimum score of C, which is considered to be equivalent to 7 on the IELTS or 550 on the TOEFL.

**2.1.2. Materials.**

The treatment used was an expository text in English on tornado formation. The text with corresponding pictures was an authentic text taken from the British Broadcasting Corporation (BBC) “science and natural disasters” webpage. It consisted of 254 words and 26 affirmative sentences and was delivered through a multimedia presentation in the form of a webpage. With 7 interacting elements, the text carried high intrinsic CL on the basis of Ginns (2005), who suggests that 6 to 8 interacting elements lead to high load while 1 to 2 lead to a low load. The number of interacting elements (7) refers to the number of the tornado formation steps some of which were stated in more than one sentence, and was determined by following Sweller and Chandler (1994). Using the same content, two experimental conditions were created: visual only and audiovisual presentations.

***2.1.2.1. Visual only and audiovisual presentations.***

The *visual only* presentation included an onscreen text with corresponding static pictures. The text was bold and represented in 12 font size using Times New Roman.The text appeared sentence by sentence on the screen along with the corresponding pictures. The size of each picture and that of the space between each sentence and its picture were the same. The pictures were placed right above the corresponding sentences and did not include any textual information or cues. This spatial separation of textual information and their corresponding pictures violates the spatial contiguity principle to some extent in that textual information and corresponding pictorial information were not placed next to each other. The reason was to establish “informational equivalence” (Brunye et al., 2006, p. 937) by exposing all the participants to as similar visual information as possible.

For the *audiovisual* presentation, pictures were presented on a webpage along with audio narration that was embedded. The duration of the narration was one minute and fifty-five seconds. An American native speaker of English was hired to do the recording at a normal pace (127 words per minute).

Both presentations were learner-controlled and included the same content, and the participants were asked to read or listen to the treatment text with an intention of learning. They were also informed that they would get comprehension tests after reading or listening to the passage. Moreover, general supportive information was provided at the beginning of the presentations so that the participants could construct a schema as suggested by van Merriënboer, Kirschner, and Kester (2003). Specifically, there was a pre-reading section on interesting information and myths about tornadoes. In addition, the participants watched a one-minute and twenty-two second-long video of a real tornado.

**2.1.3. Procedure.**

Instruments were piloted with 22 advanced Turkish learners of English before actual data collection to determine the comprehensibility of the text and comprehension tests in terms of vocabulary, accent, speed, volume, wording, difficulty of the concepts or ideas. Consequently, the wording of a question was changed only and there was no floor effect.

The data were collected in three sessions over three weeks. The participants first took the computerized reading span test (RST) in a computer lab. In the next session, one week later, they completed the participant profile and then took the prior knowledge test. Next, participants read or listened to the text, answered the comprehension tests, and completed Paas’s (1992) CL scale respectively. Follow-up tests were administered two weeks later.

**2.1.4. Measures.**

***2.1.4.1. Working memory test.***

A computerized English RST functioned as the verbal WM test. The test consisted of seventy affirmative sentences that were shown on the computer screen, one at a time, for seven seconds each. The participants were asked to (a) read each sentence and decide whether it was grammatically correct, and (b) try to remember the last word of each sentence. The sentences were divided into sets having different sizes ranging from two to five sentences. After viewing all the sentences in a particular set, the participants were prompted to recall and write the sentence-final words in a box shown on the screen in the order they remembered. Total numbers of the correct grammaticality judgments and the last word recollections were transferred into standardized values (*z* scores). A composite score consisting of the mean of the *“z”* scores was calculated as the WM index.

Then, high-WM and low-WM groups were formed based on a median split procedure. To determine whether this refers to true WM capacity differences, the mean *z-*scores were analyzed through a one-way analysis of variance (ANOVA) with WM grouping as the independent variable. There was a significant composite *z-*score difference between the groups (*F* (1, 27) = 64.40, *p* < .001, partial *η2*= .705).

***2.1.4.2. Prior knowledge test.***

The first part of the test asked the participants to indicate whether they had experienced a real tornado and to write down everything they knew about tornadoes. None of them had experienced a real tornado before the experiment, and answers to the second question were scored based on the 7 steps of tornado formation given in the treatment text (7 points). The second part included four open-ended conceptual questions. Each acceptable answer was given 1 point. Additionally, participants were not given any feedback during or after the prior knowledge test.

Using Spearman’s rho, an inter-rater reliability of .613 was found. The scores by the first rater had a range of 4 with maximum and minimum scores of 0 and 4 respectively (*M* = .72, SD = 1.16) while the scores by the second rater had a range of 3 with maximum and minimum scores of 0 and 3 respectively (*M* = .62, SD = .89). Moreover, 62 % (18) of the participants received 0, and only 3.4 % (1) received 4. Therefore, it is reasonable to claim that participants had “very low” or “no” prior knowledge before the treatment.

***2.1.4.3. Comprehension tests.***

The content validity of the comprehension tests was endorsed by an academician from the university with a specialization in L2 testing. The expert was asked to revise the test items in terms of content, wording, and design. The tests were administered immediately after and two weeks following the treatment. Immediate and delayed comprehension tests were scored by one of the researchers and an independent rater. Disagreements were resolved either by consensus or taking the average of the marks given. Additionally, there was a high positive correlation between the immediate and delayed retention test scores (*r =* .559, *p <* .05) and between the immediate and delayed transfer test scores (*r =* .555, *p <* .01).

*2.1.4.3.1. Retention test.*

The retention test required reproduction or recognition of the presented information. First, participants were asked to write down how a tornado forms; the number of the key steps provided in the answer was counted (1 point for each step; 7 points possible). The second question required matching of sample damages with their corresponding damage scales (1 point for each correct match; 6 points possible). The third and fourth questions were multiple-choice items (with four choices) each of which was worth 1 point if answered correctly. Retention performance was expressed as a raw score out of a total of 15. Using Spearman’s rank-order correlation, the inter-rater reliability was found to be .934 for the immediate retention test and .803 for the delayed retention test.

*2.1.4.3.2. Transfer test.*

The first question was a conceptual question about the causes of a tornado that was worth up to 3 points. The second question was a troubleshooting question that presented participants with a possible unexpected scenario and asked why that would be the case. The answer consisted of two alternatives given one point in the case of each correct answer (2 points). The third question was a redesign question that required participants to find a solution to a specific problem. There were two complementary steps to answer this question (2 points). The fourth question was a prediction question asking for a solution to a process of tornado formation that was not explicitly stated in the treatment passage; with four possible predictions (4 points). The fifth question was a conceptual question asking for the factors impacting the duration and strength of a tornado; with two possible answers (2 points). The last question, a conceptual question, asked participants to list the environmental clues to watch for that may result in a tornado; with seven clues (7 points). Hence, the transfer scores were calculated out of 20. Using Spearman’s rank order correlation, the correlation between the two raters’ scores was .845 for the immediate transfer test and .912 for the delayed transfer test.

***2.1.4.4. Paas’s (1992) subjective cognitive load scale.***

Paas’s (1992) scale has 9 points arranged from *very, very low mental effort* (1) to *very, very high mental effort* (9). Paas, van Merriënboer, and Adam(1994) showed the reliability and sensitivity of the scale, and that one-dimensional scales have sensitivity with respect to CL differences and such scales are valid, reliable, and unintrusive.

Paas`s (1992) scale was applied immediately after the comprehension tests. Overall CL ratings of the participants varied from 1 to 8 (*M* = 5, SD = 2). In order to rule out the possibility that performance differences may have resulted from the differences between the amount of mental effort spent by the participants, a two-way ANOVA with WM capacity (high, low) and presentation mode (audiovisual, visual only) on mental effort ratings was conducted. Due to violation of the normality and homogeneity of variance assumptions, “reflect and logarithm” (Pallant, 2001, p. 79) transformation was conducted. The results showed a non-significant interaction between WM and presentation mode (*F* (1, 25) = .184, *p* >.05). Nor were there significant main effects for presentation mode (*F* (1, 25) = 2.333, *p* >.05) or WM capacity (*F* (1, 25) = .115, *p* > .05).

1. **Results**

There were four experimental groups in the present study: high-WM participants in the visual only group (*n* = 6), low-WM participants in the audiovisual group (*n* = 6), high-WM participants in the audiovisual group (*n* = 9) and low-WM participants in the visual only group (*n* = 8). Participants were randomly assigned to the treatment groups.

Accordingly, two mixed-design two way ANOVAs with WM (high, low) and presentation mode (audiovisual, visual only) as the between-participants factors, and time (immediate, delayed) as the within-participants factor were conducted on immediate and delayed retention and transfer scores. The assumptions of the ANOVA (normality, homogeneity, independence of observations, absence of outliers) were checked (p`s >.05) as well as Box’s M statistics (*p’*s >.05).

**3.1. Retention performance**

Retention scores on immediate and delayed tests are displayed in Table 1. The scores on the delayed test were lower for all groups.

**Table 1** Descriptive Statistics for Retention Tests

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Groups | | Retention Tests | | | |
| Immediate | | Delayed | |
| Presentation mode | WM | *M* | *SD* | *M* | *SD* |
| Audiovisual | Low | 7.17 | 2.32 | 4.67 | 1.51 |
| High | 6.44 | 1.67 | 6.22 | 2.22 |
| Visual only | Low | 7.25 | 2.87 | 5.25 | 1.91 |
| High | 7.83 | 2.14 | 4.67 | 1.75 |

The distribution of mean immediate and delayed retention scores for experimental groups can also be seen in Figure 1:

**Fig. 1** Distribution of scores on retention tests

In order to determine whether the mean differences were statistically significant, the immediate and delayed retention test scores were submitted to a 2 (WM: high, low) X 2 (Presentation: Audiovisual, visual only) X 2 (Time: immediate, delayed) mixed ANOVA. Results indicated a significant combined effect of time, presentation mode, and WM on retention comprehension over time, *F* (1, 25)= 5.66, *p =* .025, partial *η2* = .18 and a significant main effect of time, *F* (1, 25)= 29.73, *p <* .001, partial *η2* = .54. On the other hand, the following effects were non-significant: (a) the interaction between time and WM, *F* (1, 25) =  .59, *p =* .450; (b) the interaction between time and presentation mode, *F* (1, 25) = 285, *p =* .104; (c) the interaction between presentation mode and WM, *F* (1, 25) = .090, *p =* .770; (d) the main effect for presentation mode, *F* (1, 25) = .031, *p =* .861; and (e) the main effect for WM, *F* (1, 25) = .090, *p =* .770. The significant effect of time suggested that the participants achieved significantly higher scores on the immediate test (*M* = 7.10, SD = 2.20) compared to the delayed test (*M* = 5.31, *SD =* 2). The significant triple interaction points to retention differences over time across the experimental groups.

In order to find the source of the significant triple interaction, orthogonal contrasts were conducted on retention difference scores since all groups got lower scores on the delayed retention test. Table 2 presents these contrasts:

**Table 2** Orthogonal Contrasts between the Experimental Groups

|  |  |  |  |
| --- | --- | --- | --- |
| Contrast | *T* | *Df* | *SE* |
| 1. Audiovisual presentation: High- WM vs. Low-WM | 2.25a | 25 | 1.01 |
| 2. Visual only presentation: High-WM vs. Low-WM | 1.13 | 25 | 1.03 |
| 3. Low-WM participants: Audiovisual vs. Visual only | .483 | 25 | 1.03 |
| 4. High-WM participants: Audiovisual vs. Visual only | 2.914b | 25 | 1.01 |

*ap =* .033, b*p* = .007

The significant first contrast and the non-significant second contrast suggest that WM facilitated retention of information under the audiovisual presentation only. In addition, the non-significant third contrast and the significant fourth contrast indicate that the audiovisual presentation enhanced retention performance of participants with high WM capacity only. These imply that benefits of the modality effect and high WM capacity may be additive in L2 multimedia comprehension when there is low or no prior knowledge.

These findings confirm our first hypothesis about the deteriorating effects of time on comprehension in terms of retention performance, indicating that the participants showed significantly lower performance on the delayed retention test. Additionally, our second hypothesis, which predicted that the participants in the audiovisual presentation would not achieve significantly better scores compared to those in the visual only presentation due to the self-pacing of the presentation, is also partially confirmed. The results point to a modality effect for the high-WM group in the audiovisual condition in that the time of testing had a significantly less deteriorating effect on retention performance of this group compared to the high-WM group in the visual only presentation. However, the low-WM participants in the audiovisual condition lost almost the same amount of retention scores over time as their counterparts in the visual only condition. Finally, our third hypothesis, predicting a significant effect of WM capacity regardless of the presentation mode was also partially confirmed since high WM capacity promoted less forgetting only in audiovisual condition suggesting that WM capacity and the modality effect can function as catalysts for each other in the retention of information over time.

**3.2. Transfer performance**

Table 3 provides transfer scores on immediate and delayed administrations. All groups obtained lower delayed test scores:

**Table 3** Descriptive Statistics for Transfer Tests

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Groups | | Transfer Tests | | | |
| Immediate | | Delayed | |
| Presentation mode | WM | *M* | *SD* | *M* | *SD* |
| Audiovisual | Low | 7.00 | .90 | 6.33 | .73 |
| High | 7.56 | .74 | 7.22 | .22 |
| Visual only | Low | 8.38 | .33 | 6.75 | .98 |
| High | 7.67 | .60 | 6.83 | .13 |

The distribution of mean immediate and delayed transfer scores for each experimental group can be seen in Figure 2:

**Fig. 2** Distribution of scores on transfer tests

A 2 (WM: high, low) X 2 (Presentation: audiovisual, visual only) X 2 (Time: immediate, delayed) mixed ANOVA was conducted to determine whether the differences were significant. The results revealed a significant deteriorating effect of time on transfer performance, *F* (1, 25)= 4.60, *p <* .05, partial *η2* = .15, suggesting that the participants achieved significantly higher scores on the immediate test (*M* = 7.70, SD = 2.25) than the delayed test (*M* = 6.83, SD *=* 2.16). However, the following effects were non-significant: (a) the interaction between time and presentation mode, *F* (1, 25)= .819, *p =* .374; (b) the interaction between time and WM, *F* (1, 25)= .487, *p =*.492; (c) the combined effect of time, presentation mode and WM, *F* (1, 25)= .081, *p =*.780; (d) the main effect for presentation mode, *F* (1, 25)= .245, *p =*.625; (e) the main effect for WM, *F* (1, 25)= .072, *p =*.80; and (f) the interaction between presentation mode and WM, *F* (1, 25)= .457, *p =*.505.

Results regarding transfer performance support our first hypothesis assuming a significant decrease in transfer scores over time. Furthermore, our second hypothesis predicting a non-significant transfer performance difference between the presentation modes is also confirmed as the main effect of presentation mode was non-significant. In contrast, the third hypothesis about the possible beneficial effects of WM capacity is not confirmed since no significant effects of WM were observed.

1. **Discussion**

This study investigated immediate and delayed effects of presentation mode (audiovisual, visual only) and WM capacity (high, low) on retention and transfer of information from an expository L2 text. The purpose was to determine to what extent previous modality effect findings could be extended to a self-paced L2 multimedia learning context under the condition of low prior knowledge. Another goal was to determine the effects of WM capacity in relation to the modality effect over time in a computer-assisted L2 multimedia learning environment.

Findings revealed a significant deteriorating effect of time on both retention and transfer performance. Participants lost scores on both delayed retention and transfer tests. However, the complementary facilitative effects of audiovisual presentation and WM capacity were found to be on retention performance only; high WM capacity may help learners retain more information under audiovisual presentation and audiovisual presentation may encourage retention of more information under high WM capacity over time. This interaction seems to occur especially in the absence of prior knowledge. This may suggest that low-WM participants spend so much of their WM capacity on trying to select and hold information that the remaining cognitive capacity may not be sufficient to organize and integrate verbal and pictorial models (i.e., Mayer`s [2009] essential and generative processing). However, high-WM participants could more easily select and hold presented information in WM, leaving more cognitive resources to integrate verbal and visual representations.

These results are predictable given that participants, especially those exposed to visual only presentation, might have spent more learning time (Harskamp, et al., 2007) due to slower pacing and possible unfamiliarity with some vocabulary items that were diagnosed as limiting conditions for the modality effect by Mayer (2009). Consequently, learners in the visual only group might have paid relatively more attention to the content compensating for the modality effect (Wouters et al., 2009).

These results contradict previous research most of which had system-paced instructions where time-on-task depended on the total time of the narration (e.g., Kalyuga et al., 1999; Mayer & Moreno, 1998; Moreno & Mayer, 1999). The present results partially support self-pacing as a boundary condition for the modality effect (Mayer, 2009; Tabbers, 2002; Tabbers et al., 2004). Therefore, it can be concluded that even though the modality effect may disappear in learner-paced instructions in the short-term, there may be a recovery of the effect in the long term in terms of less forgetting of retained information.

The present results also partially correspond with previous research showing facilitative WM effects on L2 reading (e.g., Alptekin & Erçetin, 2010; Harrington & Sawyer, 1992; Walter, 2004). For instance, Leeser (2007) posited that the effect of WM highly depends on prior knowledge in that high-WM capacity learners who also have prior knowledge can outperform learners with low WM capacity. Yet, Alptekin and Erçetin (2011) found independent and additive effects of WM capacity and prior knowledge on inferential comprehension in L2 reading. Current findings suggest that, under audiovisual presentation, WM may be facilitative even when prior knowledge is low. It should also be acknowledged here that it is the computer technology that made it possible to realize audiovisual presentation in the present study thereby helping WM as well. Moreover, Plass and Jones (2005) suggested investigating the effectiveness of authentic materials that are not originally designed for language learning or teaching. The computer can let learners access such materials, and help to design learning materials that align with multimedia learning principles.

These results should be read cautiously for several reasons: One argument could be that observed performance differences might have stemmed from possible mental effort differences. If this had been the case, we would expect mental effort rating differences among the experimental groups. In contrast, neither the main effect of presentation mode and that of WM capacity nor the combined effect on mental effort turned out to be significant. However, mental effort rating was employed after completing the comprehension tests. Such measures addressing the learning phase and testing phase yield different insights (van Gog & Paas, 2008). Additionally, Sweller et al. (2011) referred to the importance of determining efficiency during both learning and testing phases. Therefore, future research may need to employ these measures separately for the learning phase and testing phase.

Second, it is also possible that audiovisual participants spent more time-on-task. In this regard, a two way ANOVA with presentation mode (audiovisual, visual only) and WM (high, low) was conducted on immediate retention scores. Results revealed no significant effects, all *p’*s >.05. An analogous ANOVA conducted on the immediate transfer test results produced no significant effects either, all *p’*s >.05. Assuming that time-on-task was spent on learning, these findings would not have been reached if the amount of time-on-task had differed significantly across the experimental groups. Finally, results should be approached with caution given the low power indexes (< .80) for most of the effects obtained. However, Leahy et al. (2003) stated that “Obtaining significant effects using small sample sizes is difficult and only possible with very large effects.” (p. 414). As a result, the significant effects of the present study may refer to large effects as indicated by eta squared statistics as well. Still though, results warrant further research with larger sample sizes.

Third, the reading span test used in the present study is a complex WM test and it is assumed to tap the central executive (CE) component of WM as well as the phonological loop (Schüler et al., 2011). So, this test does not tap visuospatial sketchpad even though the learning material included static pictures as well. Schüler et al. (2011) warned that while addressing CE, a minimum of two complex span tests (one tapping CE and phonological loop, and the second one tapping CE and visuospatial sketchpad) need to be employed in order to interpret data more unambiguously. This point clearly warrants further research in order to get more precise insights into WM involvement in audiovisual information processing in L2 multimedia comprehension. Interestingly though, the pictorial material used in the present study mostly consisted of nameable elements such as the sun, clouds and rain. Baddeley (1992) stated that subvocal rehearsal employed by the phonological loop can also process “words and nameable pictures” (p. 558) and put them in the phonological store. As a result, most of the pictorial information used in the present study might have been processed mainly by the phonological loop. Therefore, future research should consider characteristics of the pictorial information while interpreting data concerning the involvement of WM components in multimedia learning.

**4.1. Conclusions**

The present results show that individual differences such as WM capacity may have effects on L2 multimedia comprehension in interaction with presentation mode, time of testing, and type of comprehension. In other words, the modality effect may hold true for high-WM participants in self-paced L2 multimedia learning environments regarding retention of information over time. Thus, this study shows that individual WM differences may mediate the effects of multimedia design principles under certain conditions, suggesting that these principles may not work effectively for all types of learners. These also support Plass and Jones` (2005) claim that some multimedia learning principles may not apply to L2 learning exactly, and that we need research examining applicability of such principles to L2 learning. Additionally, current results imply that even though computer technology can strongly support multimedia learning and cognitive processes accompanying it, effectiveness of computer technology on language learning might be moderated by the design of learning materials and individual learner differences. All these warrant further computer-assisted language learning studies focusing on the role of computers, instructional design, and learner characteristics in learning.

References

*Alptekin, C. and Erçetin, G. (2010) The role of L1 and L2 working memory in literal*

*and inferential comprehension in L2 reading. Journal of Research in Reading, 33,*

*(2): 206-219.*

*Alptekin, C. and Erçetin, G. (2011) Effects of working memory capacity and content*

*familiarity on literal and inferential comprehension in L2 reading. TESOL Quarterly,*

*45 (2): 235-266.*

*Antonenko, P., Paas, F., Grabner, R. and van Gog, T. (2010) Using electroencephalography to measure cognitive load. Educational Psychology Review, 22 (4): 425-438.*

*Ariew, R. and Erçetin, G. (2004) Exploring the potential of hypermedia annotations for second language reading. Computer Assisted Language Learning Journal, 17 (2): 237–259.*

*Austin, K. A. (2009) Multimedia learning: Cognitive individual differences and display design techniques predict transfer learning with multimedia learning modules. Computers & Education, 53: 1339-1354.*

*Baddeley, A. (1992) Working memory. Science, 255: 556-559.*

*Banikowski, A. K. and Mehring, T. A. (1999) Strategies to enhance memory based on brain-research. Focus on Exceptional Children, 32 (2): 1-16.*

*British Broadcasting Corporation. (n.d.) How a tornado is formed. Retrieved from*

*http://downloads.bbc.co.uk/news/nol/shared/spl/hi/sci-nat/06/tornado/pdf/how\_ tornado\_is\_formed.pdf*

*Brunye, T. T., Taylor, H. A., Rapp, D. N. and Spiro, A. B. (2006) Learning procedures: The role of working memory in multimedia learning experiences. Applied Cognitive Psychology, 20: 917-940.*

*Brünken, R., Plass, J. L. and Leutner, D. (2004) Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. Instructional Science, 32: 115-132.*

*Brünken, R., Steinbacher, S., Plass, J. L. and Leutner, D. (2002) Assessment of cognitive load in multimedia learning using dual-task methodology. Experimental Psychology, 49 (2): 109-119.*

*Carlson, R., Chandler, P. and Sweller, J. (2003) Learning and understanding science instructional material. Journal of Educational Psychology, 95 (3): 629-640.*

*Cepeda, N. J., Vul, E, Rohrer, D., Wixted, J. T. and Pashler, H. (2008) A temporal ridgeline of optimal retention. Psychological Science, 19 (11): 1095-1102.*

*Chandler, P. and Sweller, J. (1991) Cognitive load theory and the format of instruction. Cognition and Instruction, 8 (4): 293-332.*

*Chun, D. M. and Plass, J. L. (1996) Effects of multimedia annotations on vocabulary acquisition. The Modern Language Journal, 80: 183-198.*

*Daneman, M. and Carpenter, P. A. (1980) Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19: 450-466.*

*De Westelinck, K., Valcke, M., De Craene, B. and Kirschner, P. (2005) Multimedia learning in social sciences: limitations of external graphical representations. Computers in Human Behavior, 21: 555-573.*

*Gathercole, S. E. and Baddeley, A. D. (1993) Working memory and language. Hillsdale, NJ: Lawrence Erlbaum Associates.*

*Ginns, P. (2005) Meta-analysis of the modality effect. Learning and Instruction, 15: 313-331.*

*Goolkasian, P. (2000) Pictures, words, and sounds: From which format are we best able to reason? The Journal of General Psychology, 127 (4): 439-459.*

*Gyselinck, V., Cornoldi, C., Dubois, V., De Beni, R. and Ehrlich, M. F. (2002) Visuospatial memory and phonological loop in learning from multimedia. Applied Cognitive Psychology, 16: 665-685.*

*Gyselinck, V., Jamet, E. and Dubois, V. (2008) The role of working memory components in multimedia comprehension. Applied Cognitive Psychology, 22: 353-374.*

*Hambrick, D. Z. and Engle, R. W. (2002) Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. Cognitive Psychology, 44: 339-387.*

*Hambrick, D. Z. and Oswald, F. L. (2005) Does domain knowledge moderate involvement of working memory capacity in higher-level cognition? A test of three models. Journal of Memory and Language, 52: 377-397.*

*Harrington, M. and Sawyer, M. (1992) L2 working memory capacity and L2 reading skill. Studies in Second Language Acquisition, 14: 25-38.*

*Harskamp, E. G., Mayer, R. E. and Suhre, C. (2007) Does the modality principle for multimedia learning apply to science classrooms? Learning and Instruction, 17 (5): 465-477.*

*Jeung, H., Chandler, P. and Sweller, J. (1997) The role of visual indicators in dual sensory mode instruction. Educational Psychology, 17: 329-343.*

*Juffs, A. and Harrington, M. (2011) Aspects of working memory in L2 learning. Language Teaching, 44 (2): 137-166.*

*Kalyuga, S. (2011) Cognitive load theory: How many types of load does it really need? Educational Psychology Review, 23 (1): 1-19.*

*Kalyuga, S. and Sweller, J. (2004) Measuring knowledge to optimize cognitive load factors during instruction. Journal of Educational Psychology, 96 (3): 558-568.*

*Kalyuga, S., Chandler, P. and Sweller, J. (1999) Managing split-attention and redundancy in multimedia instruction. Applied Cognitive Psychology, 13: 351–371.*

*Kalyuga, S., Chandler, P. and Sweller, J. (2000) Incorporating learner experience into the design of multimedia instruction. Journal of Educational Psychology, 92: 126-136.*

*Leahy, W., Chandler, P. and Sweller, J. (2003) When auditory presentations should and should not be a component of multimedia instruction. Applied Cognitive Psychology, 17: 401-418.*

*Leahy, W. and Sweller, J. (2011) Cognitive load theory, modality of presentation, and the transient information effect. Applied Cognitive Psychology, 25: 943-951.*

*Leeser, M. J. (2007) Learner-based factors in L2 reading comprehension and processing grammatical form: topic familiarity and working memory. Language Learning, 52 (2): 229-270.*

*Mayer, R. E. (1997) Multimedia learning: Are we asking the right questions? Educational Psychologist, 32 (1): 1-19.*

*Mayer, R. E. (2001) Multimedia learning. New York: Cambridge University Press.*

*Mayer, R. E. (2005) Cognitive theory of multimedia learning. In Mayer, R. E. (Ed.), The Cambridge handbook of multimedia learning, 31-48. New York: Cambridge University Press.*

*Mayer, R. E. (2009) Multimedia learning (2nd ed.). New York: Cambridge University Press.*

*Mayer, R. E. and Moreno, R. (1998) A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. Journal of Educational Psychology, 90: 312-320.*

*Mayer, R. E. and Moreno, R. (2003) Nine ways to reduce cognitive load in multimedia learning. Educational Psychologist, 38 (1): 43-52.*

*McCrudden M., Schraw, G., Hartley, K. and Kiewra, K. A. (2004) The influence of presentation, organization, and example context on text learning. The Journal of Experimental Education, 72 (4): 289-306.*

*Miyake, A. and Friedman, N. P. (1998) Individual differences in second language proficiency: Working memory as language aptitude. In Healy, A. F. and Bourne, L. E. (eds.), Foreign language learning: Psycholinguistic studies on training and retention, 339-364. Mahwah, NJ: Lawrence Erlbaum.*

*Moreno, R. (2006) Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. Journal of Computer Assisted Learning, 22: 149-158.*

*Moreno, R. and Mayer, R. E. (1999) Cognitive principles of multimedia learning: The role of modality and contiguity. Journal of Educational Psychology, 91: 358-368.*

*Mousavi, S. Y., Low, R. and Sweller, J. (1995) Reducing cognitive load by mixing auditory and visual presentation modes. Journal of Educational Psychology, 87: 319-334.*

*NationalGeographic. (2007, May 29) Tornado destruction [Video file]. Retrieved from*

*http://www.youtube.com/watch?v=43VoMesUd2Q*

*Paas, F. (1992) Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. Journal of Educational Psychology, 84: 429-434.*

*Paas, F., Renkl, A. and Sweller, J. (2003) Cognitive load theory and instructional design: recent developments. Educational Psychologist, 38 (1): 1-4.*

*Paas, F., van Merriënboer, J. J. and Adam, J. (1994) Measurement of cognitive load in instructional research. Perceptual and Motor Skills, 79: 419-430.*

*Paas, F., Tuovinen, J. E., Tabbers, H. and Van Gerven, P. W. M. (2003) Cognitive load measurement as a means to advance cognitive load theory. Educational Psychologist, 38 (1): 63-71.*

*Pallant, J. (2001) SPSS Survival manual. UK: Open University Press.*

*Plass, J. L., and Jones, L. C. (2005). Multimedia learning in second language acquisition. In R.E. Mayer (Ed.), The Cambridge handbook of multimedia learning, 467-488. New York: Cambridge University Press.*

*Salaberry, M. R. (1996) The theoretical foundation for the development of pedagogic tasks in computer mediated communication. CALICO Journal, 14(1): 5-34.*

*Schmidt-Weigand, F., Kohnert, A. and Glowalla, U. (2010) A closer look at split visual attention in system- and self-paced instruction in multimedia learning. Learning and Instruction, 20 (2): 100-110.*

*Schüler, A., Scheiter, K. and van Genuchten, E. (2011) The role of working memory in multimedia instruction: Is working memory working during learning from text and pictures? Educational Psychology Review, 23 (3): 389-411.*

*Segers, E., Verhoeven, L. and Hulstijn-Hendrikse, N. (2008) Cognitive processes in*

*in children`s multimedia text learning. Applied Cognitive Psychology, 22: 375-387.*

*Sweller, J. (1988) Cognitive load during problem solving: Effects on learning. Cognitive Science, 12: 257-285.*

*Sweller, J. (1994) Cognitive load theory, learning difficulty and instructional design. Learning and Instruction, 4: 295-312.*

*Sweller, J. (2004) Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. Instructional Science, 32: 9-31.*

*Sweller, J. (2005) Implications of cognitive load theory for multimedia learning. In R. E. Mayer, R. E. (Ed.), The Cambridge handbook of multimedia learning, 19-30. New York: Cambridge University Press.*

*Sweller, J. (2010) Element interactivity and intrinsic, extraneous and germane cognitive*

*load. Educational Psychology Review, 22: 123-138.*

*Sweller, J., Ayres, P. and Kalyuga, S. (2011) Cognitive load theory. New York: Springer.*

*Sweller, J. and Chandler, P. (1994) Why some material is difficult to learn? Cognition and Instruction, 12 (3): 185-233.*

*Sweller, J., van Merriënboer, J. J. G. and Paas, F. (1998) Cognitive architecture and instructional design. Educational Psychology Review, 10 (3): 251-296.*

*Şakar, A. and Erçetin, G. (2005) Effectiveness of hypermedia annotations for foreign language reading. Journal of Computer Assisted Learning, 21: 28–38.*

*Tabbers, H. K. (2002) The modality of texts in multimedia instructions: refining the design guidelines. Unpublished doctoral dissertation, Open University of the Netherlands, Heerlen. Retrieved from http://hdl.handle.net/1820/1025*

*Tabbers, H. K., Martens, R. L. and van Merriënboer J. J. G. (2004) Multimedia instructions*

*and cognitive load theory: Effects of modality and cueing. British Journal of*

*Educational Psychology, 74: 71-81.*

*Tindall-Ford, S., Chandler, P. and Sweller, J. (1997) When two sensory modes are better than one. Journal of Experimental Psychology: Applied, 3: 257-287.*

*van Merriënboer, J. J. G., Kirschner, P. A. and Kester, L. (2003) Taking the load off a learner’s mind: Instructional design for complex learning. Educational Psychologist, 38 (1): 5-13.*

*van Gog, T. and Paas, F. (2008) Instructional efficiency: Revisiting the original construct in educational research. Educational Psychologist, 43 (1): 16-26.*

*Walter, C. (2004) Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. Applied Linguistics, 25 (3): 315-339.*

*Witteman, M. J. and Segers, E. (2010) The modality effect tested in children in a user-paced*

*multimedia environment. Journal of Computer Assisted Learning, 26: 132-142.*

*Wouters, P., Paas, F. and van Merriënboer, J. J. G. (2009) Observational learning from animated models: Effects of modality and reflection on transfer. Contemporary Educational Psychology, 34: 1-8.*

*Yeung, A. S., Jin, P. and Sweller, J. (1997) Cognitive load and learner expertise: Split-attention and redundancy effects in reading with explanatory notes. Contemporary Educational Psychology, 23: 1-21.*

1. \* The data were collected at the computer laboratory of the Department of Foreign Language Education, Bogazici University, Turkey. We are thankful to them. [↑](#footnote-ref-1)