System 55 (2015) 63–73

Contents lists available at ScienceDirect

System

journal homepage: www.elsevier.com/locate/system

# Input modality and working memory: Effects on second language text comprehension in a multimedia learning environment

Kadir Kozan [a, *], Gülcan Erçetin [a], Jennifer C. Richardson [b]

[a] Faculty of Education, Boğaziçi Üniversitesi, 34342 Bebek, Istanbul, Turkey
[b] College of Education, Purdue University, 100 N. Street, 47907 West Lafayette, IN, USA

ARTICLE INFO

*Article history:*
Received 21 January 2014
Received in revised form 23 August 2015
Accepted 4 September 2015
Available online 30 September 2015

*Keywords:*
L2 multimedia learning
Working memory
Visual-only presentation
Audiovisual presentation
Input modality

ABSTRACT

This study investigated the modality effect in relation to verbal working memory capacity and time of testing within a computerized second language multimedia learning environment. Twenty-nine advanced learners of English with Turkish as the first language were randomly assigned to audiovisual or visual-only presentations about an unfamiliar topic and completed immediate and delayed tests to assess retention and transfer of information performance. They also completed a reading span test as a measure of working memory capacity. Results revealed a significant combined effect of time, input modality and working memory capacity on participants' retention performance while the only significant effect observed on transfer performance was that of time. Analyses on the significant three-way interaction revealed that although input modality and working memory capacity play some role in retention performance, their effects emerged dependent upon one other and time of testing.



## 1. Introduction

A fundamental question regarding multimedia instruction is how to effectively present verbal and visual information to promote learning in multimedia environments. One answer to this question comes from the modality effect of cognitive load theory (e.g., Mousavi, Low, & Sweller, 1995; Sweller, 2005, 2010; Sweller, Ayres, & Kalyuga, 2011; Sweller, van Merriënboer, & Paas, 1998) or the modality principle of cognitive/generative theory of multimedia learning (e.g., Mayer, 2005, 2009; Mayer & Moreno, 2003). The modality effect or principle assumes that when written verbal information accompanies visual information (i.e., visual-only presentation), they, at least initially, compete for the same resources in the visual channel of working memory (WM), thereby possibly overloading it. Therefore, it would be better to present verbal information through the auditory channel in the presence of corresponding visual information (i.e., audiovisual presentation).

Most of the cognitive load theory (CLT) and cognitive theory of multimedia learning research on audiovisual presentation has used one's first language (L1) as the medium of instruction. From a second language (L2) learning perspective, the hypothesis of audiovisual presentation (e.g., a listening text with corresponding pictures) leading to enhanced performance compared to visual-only presentation (e.g., a reading text with corresponding pictures) is counterintuitive as L2 learners,

* Corresponding author. Tel.: +90 542 490 8131.
*E-mail addresses:* kkozan@purdue.edu, kadirkozan53@gmail.com (K. Kozan), gulcaner@boun.edu.tr (G. Erçetin), jennrich@purdue.edu (J.C. Richardson).

http://dx.doi.org/10.1016/j.system.2015.09.001
0346-251X/

especially those with lower levels of proficiency, typically consider listening to be a more difficult skill to master (Vandergrift, 2007). Previous L2 research generally compared comprehension performance and processes in reading-only versus listening-only conditions (e.g., Lund, 1991; Park, 2004) without considering the effects of visuals. These studies have shown that what readers and listeners recall from the text is qualitatively different in that readers recalled more information from text but they mainly recalled details while the listeners recalled main ideas (Lund, 1991). Moreover, listeners also performed better on global comprehension questions that required inferencing and synthesizing while readers performed better on local factual questions (Park, 2004).

To our knowledge, there has been little attempt to compare the effects of an entire computerized L2 reading text along with corresponding visuals (i.e., visual-only presentation) to its identical listening version along with corresponding visuals (i.e., audiovisual presentation). Thus, the purpose of the current study is to compare L2 reading and listening in the presence of corresponding visuals that would typically be the case in multimedia presentations. Specifically, this study investigated whether the modality effect holds for L2 text comprehension by comparing audiovisual and visual-only presentations. In addition, the role of WM is investigated in order to determine whether WM capacity moderates the effects of input modality as the modality effect assumes that the visual-only presentation would impose more WM load than the audiovisual presentation.

## 2. Background

### 2.1. Working memory

WM simultaneously stores and manipulates information for a short term prior to information becoming stored in long-term memory (Baddeley, 1992). WM consists of three domain-specific short-term storage subcomponents, namely the phonological loop, the visuospatial sketchpad, and the episodic buffer as well as a domain-general attention mechanism controlling these sub-systems called the central executive (e.g., Baddeley, 2000, 2003).

The phonological loop that stores speech-based verbal information plays a crucial role in language learning, particularly phonological processing, speech development and vocabulary learning while the visuospatial sketchpad that stores visual and spatial information is not essentially associated with language learning (Dehn, 2008). However, regarding the sketchpad, Baddeley (2003) suggests: “the system will be involved in everyday reading tasks, where it may be involved in maintaining a representation of the page and its layout that will remain stable and facilitate tasks such as moving the eyes accurately from the end of one line to the beginning of the next” (p. 200). Since learning with multimedia involves presentation of information both verbally and visually, the sketchpad also plays an important role in multimedia learning (Gyselinck, Jamet, & Dubois, 2008). The episodic buffer combines information from long-term memory (LTM) with the information from short-term storage systems. Finally, the central executive is a domain-general limited-capacity attention system that controls the short-term storage systems.

Research has shown that both the central executive and the phonological loop have central roles in L1 (e.g., Bowey, 2001; Daneman & Hannon, 2007; Engle, 2002) and L2 language processing (e.g., O'Brien, Segalowitz, Collentine, & Freed, 2006; Service, 1992). The role of WM in L1 (Daneman & Carpenter, 1983; Daneman & Merikle, 1996) and L2 reading (Harrington & Sawyer, 1992; Leeser, 2007; Walter, 2004) as well as L1 (Ivanova & Hallowell, 2014; McInnes, Humphries, Hogg-Johnson, & Tannock, 2003) and L2 listening (Kormos & Sáfár, 2008; Shanshan & Tongshun, 2007) has also been demonstrated.

### 2.2. Cognitive load theory

CLT assumes that learning occurs through a limited WM that temporarily manipulates and stores information as well as an unlimited LTM that functions as storage for what is learned (Sweller et al., 2011). Mayer (2009) stated: “The central work of multimedia learning takes place in working memory” (p. 62). Likewise, Sweller et al. (2011) argued that selecting, organizing and integrating visual and auditory information occurs in WM. As such, cognitive load (CL) refers to the load imposed on WM resources when a specific task is performed (Sweller et al., 1998).

CLT has identified three types of CL: intrinsic, germane, and extraneous (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Stemming from the inherent complexity of learning materials, intrinsic load is dependent on the number of interacting information elements a learning material contains (Sweller et al., 2011). An information element is one information unit that can be individually processed by WM (Leahy, Chandler, & Sweller, 2003). Intrinsic load also depends on prior knowledge or existing schemas (McCrudden, Schraw, Hartley, & Kiewra, 2004; Paas et al., 2003). CLT assumes that WM can become limited especially when dealing with novel information because it needs to be organized in LTM (Sweller, 2004). Thus, prior or familiar knowledge may facilitate learning by decreasing possible WM load since multiple information pieces organized as a schema in LTM are managed as a single unit by WM (Antonenko, Paas, Grabner, & van Gog, 2010; Kalyuga, Chandler, & Sweller, 1999; van Gog & Paas, 2008). In essence, familiar information imposes less intrinsic load than unfamiliar information (McCrudden et al., 2004).

Given that germane load emanates from the amount of WM resources used to deal with intrinsic load (Kalyuga, 2011; Sweller et al., 2011), it is reasonable to assume that as the level of intrinsic load increases due to a higher level of inner complexity and lower level of prior knowledge, learners need to invest more germane or effective load for learning to occur. Accordingly, intrinsic load relates to both learning materials and learning goals thereby covering germane load as the effective