EXPLORING THE LEARNING POTENTIAL OF MULTIMODAL INPUT-BASED TASKS: THE EFFECTS OF CAPTIONING, TEXTUAL ENHANCEMENT AND WORKING MEMORY ON GRAMMATICAL DEVELOPMENT

Minjin Lee

UCL Institute of Education, University College London

Thesis submitted for the degree of Doctor of Philosophy

2018

‘I, Minjin Lee, confirms that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.’

Word count (exclusive of front matter, reference list and appendices):

71695

Signature________________________________

ABSTRACT

In the field of second language acquisition (SLA), there is a general consensus that attention is necessary for second language (L2) development. Therefore, SLA researchers have shown considerable interest in exploring ways to draw learners' attention to L2 constructions. Of various attention-getting techniques, textual enhancement has attracted much interest from researchers, given its presumed capacity to direct attention to the target linguistic constructions implicitly during meaning-based written comprehension activities. So far, however, few studies have examined the pedagogical potential of textual enhancement when it is included in captions, that is, in the context of multimodal activities combining aural, textual and visual input. This study aims to fill this gap by assessing the potential of typographically enhanced captions to draw learners' attention to L2 constructions and assist in L2 grammatical development. The study also explored whether these relationships were influenced by individual differences in the phonological short-term memory, visuospatial short-term memory and executive control functions of working memory.

The present thesis reports on two empirical studies. Study 1 examined the extent to which increased salience of target linguistic constructions achieved through textual enhancement affected learners' allocation of attention and development in the use of L2 grammatical knowledge. Study 2 additionally investigated whether individual differences in working memory had mediated the effects of textual enhancement in captions on the allocation of attentional resources and development in L2 grammatical knowledge. In both studies, the participants were Korean learners of English. Attention allocation was measured by eye-tracking methodology and multiple measures were employed to assess L2 development and the functions of working memory.

Overall, the results indicated that textual enhancement succeeded in directing learners' attention to target linguistic constructions and promoting learning gains. However, only marginal effects were observed for working memory in the allocation of attentional resources and developing L2 grammatical knowledge.

DISSEMINATION

Peer-reviewed Journal

1. Data from Study 1 have been accepted for publication as a journal article:
   Lee, M. & Révész, A. (in press). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *The Modern Language Journal*.

2. A manuscript including data from Study 2 is currently in preparation.

Conference Presentations

1. Data from Study 1 were presented at the London Second Language Acquisition Research Forum (L-SLARF) 2015.
   Title: The effects of textual enhancement in input-based tasks on L2 development: An eye-tracking study. November, 2015, London, UK.

2. Data from Study 1 were presented at the Annual Meeting American Association for Applied Linguistics (AAAL) 2016.
   Title: Input enhancement in multimodal input-based tasks: An eye-tracking study. April 2016, Orlando, USA.

3. Data from Study 2 were presented at the UCL-Birkbeck SLA doctoral seminar 2017.
   Title: Effects of captions and textual enhancement on L2 grammatical development. June 2017, London, UK.

4. Data from Study 2 were presented at the Annual Meeting American Association for Applied Linguistics (AAAL) 2017.
   Title: Exploring the pedagogical potential of multimodal input-based tasks: A study of captioning, textual enhancement, and working memory using eye-tracking. April 2017, Portland, USA

ACKNOWLEDGEMENTS

My deepest gratitude goes to my supervisor and my mentor, Dr. Andrea Révész. She has been a real support throughout the course of my studies. With her profound insights and guidance during every step of the process, I was able to complete this thesis. I have the greatest respect for her expertise, commitment to research in the field of second language acquisition and devotion to her students. I have been genuinely blessed and honoured to have her as my supervisor and to be a perfect role model for all aspects of my life. No words can express how truly I am grateful for her support.

I must sincerely thank the Viva examiners, Dr. Pellicer- Sánchez and Dr. Carmen Muñoz for their constructive feedback and insights that were crucial to enriching my thesis. Their comments and suggestions helped me to refine the arguments and to engage in an extended discussion. I am genuinely grateful for their encouragement and generosity in sharing their knowledge and experience.

I would also like to express my appreciation to Dr. Rogers Gillabert and Dr. Themelis Karaminis for reviewing my upgrade document and giving me insightful feedback. Their thoughtful comments and keen perception were invaluable in improving the quality of my thesis. Furthermore, my sincere appreciation goes to Dr. Ronald P. Leow. I sincerely appreciated his interest in my study at conferences. His incisive comments and insightful feedback helped me to sharpen my thinking and arguments.

Furthermore, my sincere appreciation goes to the professors at Ewha Womans University, Prof. Yeon Hee Choi, Prof. Jong Im Han, Prof. Sang-Keun Shin, Prof. Eun-Joo Lee, Prof. See Young Park, and Prof. Josephine Lee, for their continuous support and valuable advice.

I would also like to express my special thanks to Dr. John Rogers, Dr. Jookyoung Jung, and Dr. Nektaria Kourtali for their thoughtful comments and invaluable assistance with different parts of the study. My special thanks also go to all my friends and colleagues. I am especially indebted to Dr. Laura Park, Dr. Sunghee Park, Dr. Hyo Jin

Cho, and Dr. Minsun Sung. Their continuous support and thoughtful consideration have been a great encouragement to me. I also want to thank YunKyeong Cho, Shin-Hye Lee, Sukyoung An and Suhae An for their warm encouragement and friendship.

I would like to thank Language Learning and The International Research Foundation for English Language Education (TIRF) for supporting the research projects reported in this thesis with a dissertation grant.

My most heartfelt gratitude goes to my beloved family for their unconditional support. I would like to express my appreciation to Jaeuk Jung, Yunjin Lee, Joong Hee Lee, Uijin Lee, Hyeseung Jung, and Jiyun Lee. I was able to complete my thesis with their prayers, love and faith in me. Finally, I want to give my wholehearted thanks to my parents, Yuksoo Lee and Ilsun Yang. With their endless support and love, I was able to pursue my study. This thesis is dedicated to my father and mother. Last but most importantly, I thank the Lord who walks beside me every step of the way. He has granted me strength and grace throughout all the years of my study.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

In the area of second language acquisition (SLA), a general consensus has been reached on the importance of input (Ellis, 1994; Gass, 1997; Krashen, 1982, 1985; Long, 1983, 1985, 1996). Exposure to input, defined as "potentially processible language data which are made available by chance or by design to the language learner" (Sharwood Smith, 1993, p. 167), is regarded as a necessary condition for second language (L2) development to occur (Gass, 1997). However, it has been argued that not all the input that is available to learners is processed by them; only part of the input to which learners have access is processed and subsequently learned. From this perspective, a distinction is made between input as target linguistic resources that learners are exposed to and intake as what is registered in the learner's mind (Corder, 1967). Thus, a theoretical question of interest arises as to how input can be converted into intake, which is seen as a prerequisite for subsequent processing to take place. In this regard, the role of attention has been of great interest to many SLA researchers, in that attention is suggested as a principal cognitive mechanism mediating the process of selecting input for further processing (Robinson, 2003; Schmidt, 2001).

The notion that input needs to be attended to and processed by L2 learners for subsequent second language development is theoretically supported by Schmidt's (1990, 1993b, 2001) noticing hypothesis, Tomlin and Villa's (1994) model of attention, Robinson's (1995, 2003) model of attention and memory and Leow's (2015) model of the L2 learning process in instructed SLA. Although there exist some discrepancies regarding how various frameworks view the relationship between attention, awareness and learning, the important role of attention in promoting second language acquisition is generally accepted by researchers (Doughty, 2001; Robinson, 1995, 2003; Schmidt, 1990, 1993b, 2001; Tomlin & Villa, 1994).

Given the centrality of attention in processing L2 input, the potential of instructional interventions to draw learners' attention to target L2 constructions has been the subject of much research. In an attempt to increase the chances of L2 learners attending to a target linguistic construction, various attention-getting techniques have been proposed to make input more salient to language learners (e.g., Schmidt, 1990, 1993b, 2001; Sharwood Smith, 1993). Among several methods, input enhancement (also known as textual enhancement), a type of implicit "focus on form" instruction, has been suggested as having the capacity to direct learners' attention to input (Doughty, 2003). The fundamental assumption underlying textual enhancement is that visually salient target linguistic constructions will attract learners' attention and, further, increased attention will result in promoting development in the use of L2 knowledge (Sharwood Smith, 1991, 1993).

The extent to which textual enhancement can attract learners' attention and promote learning has been examined by many researchers (e.g., Alanen, 1995; Bowles, 2003; Izumi, 2002; Jourdenais, Ota, Stauffer, Boyson, & Doughty, 1995; Lee, 2007; Leow, 1997, 2001; Leow, Egi, Nuevo, & Tsai, 2003; Overstreet, 1998; Park, 2004; Shook, 1994; White, 1998; Wong, 2003). The majority of these textual enhancement studies appear to have adopted a reading-based approach. That is, the effectiveness of textual enhancement, included in a reading text, on directing learners' attention to target L2 constructions has been empirically attested when learners receive input through a single modality, reading. However, relatively little attention has been devoted to examining whether and how perceptually salient input influences learners' allocation of attentional resources when they are exposed to input through multiple modes, such as aural, visual and/or textual modes. For second language learners, one way to provide linguistic input simultaneously through different modes is by means of multimodal materials, such as captioned video.

The effectiveness of captions in promoting L2 developmental processes is well documented, with extensive research on whether the provision of captions facilitates or hinders L2 listening comprehension and L2 vocabulary knowledge development. Overall, positive effects of captions have been evidenced for both L2 listening comprehension (e.g., Chai & Erlam, 2008; Danan 2004; Garza, 1991; Huang & Eskey, 2000; Rodgers & Webb, 2017; Winke, Gass, & Sydorenko, 2010) and vocabulary learning (e.g., Bird & Williams, 2002; Chai & Erlam, 2008; Danan, 1992; Markham, 1999; Markham, Peter, & McCarthy, 2001; Sydorenko, 2010; Winke, Gass, & Sydorenko, 2010), and positive results have also been confirmed in a meta-analysis addressing the effects of captioning (Montero Perez, Van den Noortgate, & Desmet, 2013). Captioning is presumed to help learners better engage in form-meaning mapping, which is a necessary process for second language acquisition (Doughty, 2004). Thus, in this line of inquiry, further investigation is needed to assess whether learners can also benefit from captions in developing their L2 grammatical knowledge.

Additional research is also warranted on how task-based approaches might profit from the potential acquisitional benefits offered by exposure to captions in consideration of increasing the prominence of task-based language teaching (TBLT) within the fields of instructed second language acquisition and L2 pedagogy (e.g., Bygate, Skehan, & Swain, 2001; Ellis, 2003; Samuda & Bygate, 2008). Interest in tasks has been motivated by the fact that carrying out communicative tasks prepares learners for real-life activities and engages psycholinguistic processes that are thought to be beneficial for L2 learning (Long, 2000). Captioned video-based activities can take the form of a task and thus be incorporated into task-based teaching, but so far little attention has been paid to the utility of captions in the context of TBLT.

Tasks using captioned videos can be characterized as input-based in the context of TBLT. Input-based tasks differ from output-based tasks; the main difference between these two task types lies in whether the requirement of producing output is imposed on language learners or not while performing a task (Ellis, 2013; Shintani, 2012). Output-based tasks require language learners to engage in production, either speaking or writing; whereas input-based tasks do not require learners to produce output (Ellis, 2013; Shintani, 2012). However, it is important to note that generating output is not proscribed in input-based tasks; learners can elect to engage in language production (Ellis, 2013). Furthermore, given that captioned video provides input to learners through multiple modes (aural, textual and/or visual modes), tasks with captioned videos can be described as multimodal input-based tasks, diverging from the traditional view of input-based tasks involving either listening or reading (Ellis & Shintani, 2014). Although the use of both output-based tasks and input-based tasks is advocated (Ellis, 2009, 2013), comparatively more attention has been paid to the role of output-based tasks in second language learning based on the false assumption that successful completion of a task entails production (Shintani, 2012). The construct of input-based tasks, on the other hand, seems to be under-researched. Thus, further exploration of the role of input-based tasks, particularly multimodal input-based tasks in SLA research and language pedagogy, is warranted.

In examining the extent to which a combination of multimodal input-based tasks using captions and textual enhancement can increase the likelihood that learners will attend to and learn L2 grammatical constructions, one individual difference that should be taken into consideration is learners' working memory capacity. Working memory is generally defined as a cognitive process that is responsible for temporarily storing and manipulating the information needed to carry out complex tasks (Baddeley, 1992). According to Baddeley's (2000) influential model, working memory comprises four different components: phonological loop, visuospatial sketchpad, central executive and episodic buffer. Two components that have been of great interest to SLA researchers are the phonological loop and the central executive. The phonological loop is responsible for storing and manipulating verbal information, whereas the central executive controls and regulates complex cognitive processes, such as allocating attentional resources, switching between tasks, inhibiting processing routines, updating information or regulating subsidiary memory systems (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). Considering the limited capacity of working memory and its function to regulate attentional resources while performing complex tasks, it seems reasonable to assume that individual differences in working memory will affect the extent to which learners pay attention to and, further, process input, textually non-enhanced or enhanced, included in multimodal input-based tasks. These relationships, however, have been the object of little empirical research to date.

1.1. Aims of the Thesis

To fill the research gaps outlined in the previous section, two studies were conducted as part of this thesis. The overriding goal of the two studies was to examine the effectiveness and pedagogical value of textual enhancement included in multimodal input-based tasks, by means of captions, in drawing learners' attention to target linguistic constructions and, further, promoting development in the use of L2 grammatical knowledge. Study 1 aimed to investigate the potential of textually enhanced captions in comparison to non-enhanced captions to promote attention to and learning of target linguistic constructions with the following research questions:

1. To what extent do non-enhanced captions versus textually enhanced captions draw learners' attention to target linguistic constructions?
2. To what extent do non-enhanced captions versus textually enhanced captions affect L2 development in the knowledge of target linguistic constructions, as measured by a written and an oral grammaticality judgement test (GJT)?
3. To what extent is attention to target linguistic constructions in captions related to L2 development? Is this relationship influenced by textual enhancement?

In other words, the first research question addressed the extent to which the textual enhancement included in captions led language learners to attend to target linguistic constructions. The second research question asked whether textually enhanced captions promoted the use of L2 grammatical knowledge. The relationship between the attention allocated to target linguistic constructions and learning gains was examined with research question three. As a follow-up, Study 2 was conducted, which was guided by the following research questions:

1. To what extent do multimodal input-based tasks without captions versus those with captions affect development in L2 grammatical knowledge?

2. To what extent do textually non-enhanced captions versus enhanced captions in multimodal input-based tasks affect development in L2 grammatical knowledge?
3. To what extent do textually non-enhanced captions versus enhanced captions in multimodal input-based tasks draw learners' attention to the target linguistic construction?
4. To what extent does the learner attention allocated to the target linguistic construction relate to development in L2 grammatical knowledge? Is this relationship influenced by whether learners are exposed to textually non-enhanced or enhanced captions?
5. To what extent do individual differences in working memory capacity moderate the effects of captions, textually non-enhanced or enhanced, in multimodal input-based tasks on L2 development?
6. To what extent does learners' attention allocated to the target linguistic construction correlate with their working memory capacity? Is this relationship influenced by whether learners are exposed to textually non-enhanced or enhanced captions?

In Study 2, along with the three questions discussed in Study 1, additional research questions were addressed to explore (a) the mediating role of working memory in the relationship between textual enhancement and the development of L2 grammatical knowledge and (b) the relationship between learners' working memory and their allocation of attentional resources to the target linguistic constructions.

To address these questions, a number of methodological advances were made. First, an eye-tracking methodology was used to measure the amount of attention allocated to a target linguistic construction. Learners' eye-movement data were expected to provide not only information about learners' online cognitive processes but also quantifiable indices that could be used to examine the relationship between attentional processing and learning gains. In addition, considering the theoretical assumption that the central executive functions in working memory might be separable (Miyake et al., 2000), an attempt was made to examine the contribution of each function of the central executive (shifting, inhibition and updating) to second language acquisition in Study 2. Thus, separate measures attempting to assess distinctive functions of the central executive were used: a colour-shape task for shifting (Miyake, Emerson, Padila, & Ahn, 2004), a stop signal task for inhibition (Logan, 1994; Verbruggen, Logan, & Stevens, 2008) and an automated operation span task for updating (Turner & Engle, 1989).

1.2. Structure of the Thesis

This thesis comprises five chapters. In Chapter 1, the rationale and aims of the thesis and operationalisation of the key terms are explained. Chapter 2 describes the theoretical framework for this study and gives an overview of previous relevant research. First, theoretical accounts that address the importance of input are introduced, followed by a discussion of different models that explain the role of attention and awareness in input processing. In the next section, a brief introduction of task-based language teaching and a definition of tasks as well as input-based tasks are provided, along with a review of previous empirical studies on input-based tasks in L2 learning. As an input-based task is operationalized as a multimodal-input based task using captions, definitions of captions and previous empirical studies which have examined the effects of captions on L2 learning are also reviewed.

The following section presents the notion of textual enhancement as an implicit focus on form intervention to draw learners' attention to target linguistic constructions, and it discusses the results of related empirical studies. In addition, different points of view on measures of attention are addressed, which entails discussion of the use of eye-movement data as an alternative measure of 'noticing' as attention. Regarding this, a review of studies that have used eye-tracking to gauge learners' attention paid to target linguistic constructions is conducted. The last section of this chapter is reserved for an explanation of the construct of working memory, Baddeley's working memory model, and measures of working memory. A theoretical account of the unitary nature and diversity of central executive functions is also discussed. Furthermore, a review of previous studies investigating the potential effects of individual difference in working memory on second language acquisition is included.

Chapter 3 reports on Study 1, which examined the effectiveness of textual enhancement in drawing learners' attention to target linguistic constructions and promoting development in the use of L2 grammatical knowledge. In this chapter, the methodology, including the participants, target linguistic constructions, materials, data collection procedures and statistical analyses, is explained, followed by the results and a discussion of Study 1. This chapter ends with a discussion of the perceived limitations of the study and an explanation of the modifications made for Study 2. Then, Chapter 4 reports on Study 2, beginning with the methodology along with a brief explanation of the linear mixed-effects model, which is the main statistical analysis tool used in Study 2. This chapter also presents the results and discusses the findings of the study. Chapter 5 synthesizes the overall findings from Study 1 and Study 2, followed by a discussion of the theoretical, methodological and pedagogical implications. This chapter ends with the limitations of the studies and suggestions for future studies.

1.3. Definitions and Operationalisation of Key Terms

1.3.1. Input

Input is linguistic data that are potentially available to learners (Corder, 1967). It is the language that learners read and/or hear, which may be subject to further processing for meaning.

1.3.2. Attention

Attention is "the process that encodes language input, keeps it active in working and short-term memory, and retrieves it from long-term memory" (Robinson, 2003, p. 631). Robinson (1995) explains the concept of attention with reference to the use of attentional resources: (a) learners' ability to select information for processing, (b) learners' limited capacity of attention and (c) learners' mental effort involved in information processing. In this thesis, therefore, attention is operationalized as a cognitive mechanism that learners use to select information, further process the information. Furthermore, on the basis of an "eye-mind link" assumption (Reichle, Pollatsek, & Rayner, 2006, p. 4), which claims a close association between "overt attention" (as manifested by the exact eye location) and "covert attention" (mental attention), a quantitative measure of the attention, that is the time spent on target linguistic constructions, was used with the help of eye-tracking methodology.

1.3.3. Noticing

According to Schmidt's (1990, 1993b, 2001) noticing hypothesis, noticing is defined as attention plus awareness. That is, learners first need to consciously 'notice' the input, via focal attention and awareness, for subsequent processing of it. Schmidt (1995) further explains that there are two levels of awareness: awareness at the level of noticing and awareness at the level of understanding. Awareness at the level of noticing includes focal attention, and noticing occurs with conscious awareness, which allows subsequent learning. Awareness at the level of understanding "implies recognition of a general principle, rule or pattern" (Schmidt, 1995, p. 29), in that learners are conscious and engaged in deeper processes. Schmidt argues that only what has been noticed in the input can become intake.

In Robinson's model of attention and awareness, built on Schmidt's noticing hypothesis, noticing is defined as "detection plus rehearsal in short-term memory, prior to encoding in long-term memory" (Robinson, 1995, p. 296); that is, detection first assists input to be encoded in short-term memory, which is further processed by means of rehearsal before it is transferred to long-term memory. Robinson hypothesizes that detection occurs before rehearsal, which is noticing with awareness in the acquisition process. The present study, on the basis of Robinson's model, measured attention, which is part of the construct of noticing, using eye-tracking indices that quantify the amount of attention paid to target linguistic constructions (Godfroid, Boers, & Housen, 2013).

1.3.4. Task

Within the framework of Task-based Language Teaching (TBLT), several different definitions of task have been proposed by researchers (e.g., Ellis, 2003;

Long, 1985; Nunan, 2004; Prabhu, 1987; Skehan, 1998; Willis, 1996). The common characteristics that encompass the definitions of task include a focus on meaning, language use and achievement of outcome. The present task follows the general criteria features defining a task summarized by Ellis (2003): (a) a primary focus on meaning, (b) the presence of some kind of 'gap', (c) use of learners' own linguistic and non-linguistic resources and (d) achievement of an outcome.

1.3.5. Input-based Task

An input-based task refers to a task that does not require production, such as speaking or writing, while performing it (Ellis, 2013). However, learners are not prohibited from producing output through speaking or writing.

1.3.6. Multimodal Input-based Task

A multimodal input-based task is a task presenting learners with audio, visual and/or textual input simultaneously. This multimodal input-based task, conforming to the basic definition of an input-based task, does not require learners to produce output through speaking or writing.

1.3.7. Captions

In previous studies, the terms "captions" and "subtitles" have often been used interchangeably to refer to on-screen text in either the L1 or L2. In the current study, captions are operationalised as on-screen text in the target language (L2) (Markham & Peter, 2002-2003; Vandergrift, 2007).

1.3.8. Textual Enhancement

Textual enhancement, as an implicit type of focus on form, is defined as a pedagogical intervention to make specific linguistic features more perceptually salient to learners in order to attract their attention (Sharwood Smith, 1990, 1993). To promote the salience of particular linguistic features, different kinds of typographical modifications can be employed, such as <u>underlining</u>, **boldfacing**, *italicization*, CAPITALIZATION or other techniques, such as colouring or using different font types.

1.3.9. Eye-tracking Indices

Among the various eye-tracking indices proposed in previous literature (Roberts & Siyanova-Chanturia, 2013), four indices are used: first pass reading, second pass reading, total fixation duration and number of visits. First pass reading time is the sum of all fixation durations during a first visit to the area of interest. This measure is regarded as an index of initial processing. Second pass reading time is defined as the sum of fixation durations when the eyes return to an area of interest after a first visit. In other words, second pass reading time captures rereading in the area of interest, which is associated with re-analysis of the input. Total fixation duration is the sum of all fixation durations, i.e. the sum of all fixation durations made within an area of interest. A visit includes all the fixations made within an area of interest (AOI) from the time a participant's eyes first enter the AOI and until they leave.

1.3.10. Working Memory Capacity

Following Baddeley's definition, working memory is defined as a limited cognitive resource that stores and manipulates information for a short period of time.

According to the most widely discussed model, namely the multi-component working memory model (Baddeley & Hitch, 1974), working memory comprises four main components: a phonological loop, a visuospatial sketchpad, a central executive and an episodic buffer.

1.3.11. Phonological Loop / Phonological Short-term Memory

The phonological loop is responsible for the temporary storage and manipulation of verbal and acoustic information. Two subcomponents are a phonological store and an articulatory rehearsal process. The phonological store holds verbal information for short periods of time, while the articulatory rehearsal process is responsible for translating nonauditory information into phonological form and refreshing the verbal input through retrieving and rearticulating (Gathercole & Baddeley, 1993). In the present study, phonological short-term memory was measured using a nonword span task (Jung, 2017).

1.3.12. Visuospatial Sketchpad / Visuospatial Short-term Memory

Another slave system of the central executive is the visuospatial sketchpad. This system is known to temporarily store and manipulate visual and spatial information. As a measure of visuospatial short-term memory, a forward corsi block task (Corsi, 1972; Kessels, Van Zandvoort, Postma, Kappelle, & De Haan, 2000) was used in this study.

1.3.13. Central Executive or Complex Working Memory Capacity

The central executive is the main component of working memory, it controls and regulates complex cognitive operations, such as allocating attentional resources, switching between tasks, inhibiting processing routines, regulating subsidiary memory systems and retrieving information from long-term memory. Among these processes, three frequently referenced executive functions are switching (or shifting), inhibition and updating (e.g., Baddeley, 1996; Logan, 1985; Miyake et al., 2000; Smith & Jonides, 1999). Switching or shifting refers to the ability to switch or shift flexibly between tasks. Inhibition concerns the ability to deliberately inhibit responses when required, whereas updating is the ability to monitor, revise and update incoming information constantly (Miyake et al., 2000). In the present thesis, particularly in Study 2, each distinctive function of the central executive was measured using: a colour-shape task for shifting (Miyake, Emerson, Padila, & Ahn, 2004), a stop signal task for inhibition (Logan, 1994; Verbruggen, Logan, & Stevens, 2008) and an automated operation span task for updating (Turner & Engle, 1989).

1.3.14. Episodic Buffer

The fourth component of working memory is called the episodic buffer, which is suggested to integrate verbal and visual information from subsidiary systems – the phonological loop and the visuospatial sketchpad – and from long-term memory into unitary episodic representations.

CHAPTER 2

LITERATURE REVIEW

This chapter provides the theoretical framework and reviews previous relevant research. First, the importance of input in second language acquisition is addressed with an explanation of Krashen's (1985) comprehensible input hypothesis, followed by a discussion of the role of attention and awareness in input processing. Different theoretical perspectives on attention and level of awareness in relation to the construct of 'noticing' are discussed, including Schmidt's (1990) noticing hypothesis, Tomlin and Villa's (1994) model of attention in SLA, Robinson's (1995) model of attention and memory and Leow's (2015) model of the L2 learning process in instructed SLA. Building on the notion that attention is an important aspect of second language acquisition (Doughty, 2001; Robinson, Mackey, Gass, & Schmidt, 2012; Schmidt, 2001), the effectiveness of input-based tasks in terms of creating an input-rich environment and drawing learners' attention to input within the framework of Task-based Language Teaching (TBLT) is addressed; previous empirical studies that have investigated the relative effectiveness of input-based tasks on second language acquisition are also reviewed.

Next, an input-based task is operationalized as a multimodal-input based task by means of captions in the present study; thus, the definition of captions and the theoretical rationale for using captions for instructional purposes are presented, along with a review of a number of empirical studies examining the effectiveness of captioning on different aspects of L2 learning. Premised on that learners may not attend to all available input, textual enhancement, as an implicit focus on form intervention to draw language learners' attention to target linguistic constructions, is introduced in the following section. More specifically, the definition, underlying assumptions and different types of textual enhancement are explained. In the next section, related empirical studies, which have investigated the effects of textual enhancement on drawing learners' attention to target constructions and promoting L2 development, are reviewed. In addition, different approaches to the measurement of attention are addressed, entailing a discussion of the use of eye-movement data as an alternative measure of 'noticing' as attention. Then, a review of studies that have used eye-tracking to gauge learners' attention paid to target linguistic constructions is conducted. The last section of this chapter is dedicated to a discussion of the construct of working memory, Baddeley's working memory model, and measures of working memory. A theoretical framework of the unitary nature and diversity of central executive functions is also provided, followed by a review of previous studies investigating the potential effects of individual difference in working memory on second language acquisition.

2.1. Role of Input in Second Language Acquisition

The significant role of input in second language acquisition has been addressed by many researchers (e.g., Ellis, 1994; Gass, 1997; Krashen, 1982, 1985; Long, 1983, 1985, 1996), highlighting the input as "the most important concept of second language acquisition" (Gass, 1997, p. 1). In the mid-1980s, Krashen (1985) introduced his comprehensible input hypothesis as part of the Monitor model, with particular emphasis on the role of input in second language acquisition. According to this hypothesis, comprehensible input – defined as input containing language forms slightly beyond the language learner's current level of interlanguage proficiency (i+1) – is integral to second language acquisition. Krashen (1985) claims that the provision of comprehensible input is a sufficient condition for L2 acquisition to occur. The comprehensible input hypothesis has been influential in the field of SLA, as it was the first to underscore the significance of comprehensible input in second language acquisition.

However, this hypothesis has been criticised for its vagueness in defining "comprehensible input" and its exclusive focus on the role of comprehensible input in SLA (e.g., Robinson, 1995, 2003; Schmidt, 1990, 1993b, 2001; Swain, 1985; VanPatten, 1996). Researchers have argued that the mere provision of input does not guarantee successful L2 learning (e.g., Schmidt, 1990, 1993b, 2001). In fact, there is a general consensus that not everything present in the input is processed by the learner (Corder, 1967), and attention is a principal cognitive mechanism mediating the process of selecting input for further processing (Robinson, 2003; Schmidt, 2001). Hence, drawing learners' attention to input has been emphasized for input to become intake for subsequent processing (e.g., Schmidt, 1990, 1993b, 2001).

2.2. Attention and Awareness in Second Language Acquisition

The notion that attention is necessary for learning to take place is generally accepted in the field of instructed SLA and cognitive psychology (e.g., Robinson, 1995, 2003; Schmidt, 1990, 1993b, 2001). Schmidt (2001) argues that "there is no doubt that attended learning is far superior and, for all practical purposes, attention is necessary for all aspects of learning" (p. 3). Learners' attention, therefore, should be directed to the input if it is to be processed and, further, become intake (Leow, 1999; Robinson, 1995, 2003; Schmidt, 1990, 2001; Sharwood Smith, 1991; Tomlin & Villa, 1994). However, while many SLA researchers agree on the importance of attention in language learning, differences have been perceived as to the type and level of attention and/or awareness involved in ‘noticing’ (e.g., Robinson, 1995, 2003; Schmidt, 1990, 1993b, 2001; Tomlin & Villa, 1994).

2.2.1. Schmidt’s Noticing Hypothesis

Schmidt (1990, 1993b, 2001) emphasises the importance of consciousness in language learning by proposing the noticing hypothesis, which claims that “noticing is the necessary and sufficient condition for the conversion of input to intake for learning” (Schmidt, 1994, p. 17). According to this hypothesis, learners first need to consciously ‘notice’ the input, with focal attention and awareness, for subsequent processing of the input. Thus, Schmidt defines noticing as attention entailing a certain level of awareness. Regarding awareness, Schmidt (1995) explains that there are two levels of awareness: awareness at the level of noticing and awareness at the level of understanding. More specifically, awareness at the level of noticing includes focal attention, and noticing occurs with conscious awareness, which allows subsequent learning. Noticing is characterized as awareness at a low level of abstraction; only surface features, not underlying grammatical rules and patterns, of the input are noticed by learners (Schmidt, 2001). On the other hand, awareness at the level of understanding “implies recognition of a general principle, rule or pattern” (Schmidt, 1995, p. 29), in that learners are conscious and engaged in deeper processes. Schmidt notes that both problem-solving and all forms of metacognition are indicative of awareness at this level (Schmidt, 1990). Furthermore, if noticing is related to the transfer of information to long-term memory and item learning, then awareness at the level of understanding is connected to “the organization of material in long term memory, to restructuring, and to system learning” (Schmidt, 1993b, p. 213). A clear distinction between the two levels of awareness (i.e., awareness at the level of noticing and awareness at the level of understanding) is explained as follows:

> I use noticing to mean registering the simple occurrence of some event, whereas understanding implies recognition of a general principle, rule, or pattern. For example, a second language learner might simply notice that a native speaker used a particular form of address on a particular occasion, or at a deeper level the learner might understand the significance of such a form. (Schmidt, 1993a, p. 26)

Given the importance of noticing and attention for language acquisition, Schmidt originally introduced a strong version of the noticing hypothesis, arguing that “noticing is the necessary and sufficient condition for the conversion of input to intake” (Schmidt, 1990, p.129). This version, however, has been modified into a weaker version, suggesting a facilitative rather than a necessary role for noticing in L2 development.

2.2.2. Tomlin and Villa’s Model of Attention

Tomlin and Villa (1994) present a somewhat different view of the role of awareness, arguing for a dissociation between awareness and learning. According to Tomlin and Villa’s fine-grained analysis, attention has three components: (a) alertness, (b) orientation and (c) detection. The first component, alertness, is “an overall, general readiness to deal with incoming stimuli or data” (Tomlin & Villa, 1994, p. 190) and the level of alertness can affect learners’ processing or performance of tasks (Tomlin & Villa, 1994). This level of attention is important for second language acquisition, as learners are more likely to be prepared to deal with the input for further processing. At the second level, orientation, attentional resources are directed to a certain type of stimuli, e.g., sensory information. The underlying concept of orientation is that “the specific aligning of attention ("orienting") on a stimulus has facilitative or inhibitory consequences for further processing depending on whether information occurs as expected or not as expected" (Tomlin & Villa, 1994, p. 190). It is suggested that these two components, alertness and orientation, increase the likelihood that detection will occur, which is the last component of Tomlin and Villa's model. Detection refers to "cognitive registration of the stimuli" (Tomlin & Villa, 1994, p.192), which enables learners to select and process particular information. That is, according to Tomlin and Villa, detection is a necessary condition for further processing of the input and language acquisition must take place at this level. The main argument made by Tomlin and Villa underlines that none of the three components – alertness, orientation and detection – require awareness, which is viewed as "the subjective experience of any cognitive content or external stimulus" (p. 194). In other words, Tomlin and Villa argue for a dissociation of attention and awareness and, moreover, awareness is not necessary for input to be available for language development.

2.2.3. Robinson's Model of Attention and Memory

To further determine the roles of attention and awareness in language learning, Robinson (1995, 2003) introduced a model of attention and memory by reconciling two different positions addressed by Schmidt's noticing hypothesis and Tomlin and Villa's model of attention. In Robinson's model, adapted from Cowan's (1993) model, noticing is defined as "detection plus rehearsal in short-term memory, prior to encoding in long-term memory" (Robinson, 1995, p. 296), as illustrated in Figure 1.

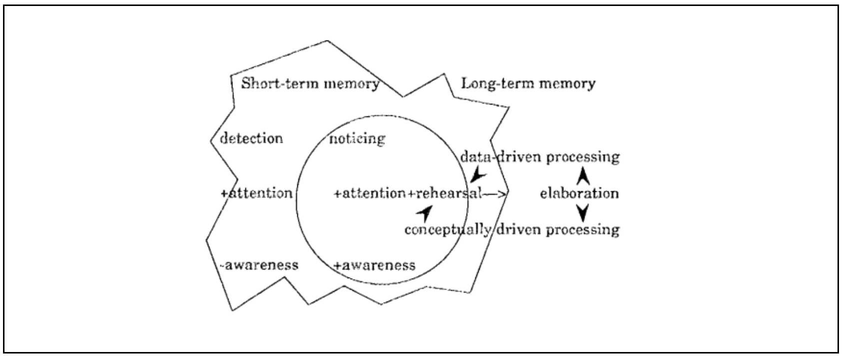


Figure 1. Noticing as Detection with Awareness in Short-term Memory (Robinson, 1995, p. 297)

That is, detection first assists input to be encoded in short-term memory, which is further processed by means of rehearsal before it is transferred to long-term memory. According to Robinson (1995),

> . . . activation in short-term memory must exceed a certain threshold before it becomes part of awareness (Cowan, 1988, p. 165; Shiffrin, 1993, p. 195). Further, short-term memory is the subset of long-term memory in a currently active state. Thus, noticing can be identified with what is both detected and then further activated following the allocation of attentional resources from a central executive. Rehearsal following detection would be a consequence of the allocation of resources to fulfil task demands…. (p. 297)

Robinson explains that detection occurs before rehearsal, which can be of two kinds: maintenance rehearsal or elaborated rehearsal. Maintenance rehearsal is data-driven processing, involving “rehearsal and maintenance in memory of isolated instances” (Leow, 2015, p. 81). Elaborative rehearsal, on the other hand, is conceptually-driven processing, in that it is “a more elaborated form or rehearsal that distributes instances into abstract configurations” (Leow, 2015, p. 81). Hence, rehearsing detected information entails awareness and occurs as learners allocate attention resources to complete tasks. In sum, Robinson’s theoretical position, following Schmidt’s noticing hypothesis, is that detection is important in input processing; however, it is at the level of noticing (i.e., attention with a low level of awareness) that linguistic data are available for further processing.

2.2.4. Leow’s Model of the L2 Learning Process in Instructed SLA

Leow (2015) introduced a new model that depicts how L2 linguistic information is processed. In this model, three processing stages are included: input processing stage, intake processing stage and knowledge processing stage. The first stage, namely the input processing stage, is “between the input and the intake of specific linguistic information and what is taken in (intake) is initially stored in working memory” (pp. 241–242). At this stage, the peripheral attention, selective attention or focal attention paid by learners predicates three sub-phases, attended intake, detected intake and noticed intake, respectively. More specifically, attended intake, comparable to the initial stages of perception of input suggested by Chaudron (1985), is peripherally attended input, which is not for further processing or storing in working memory. The next phase, detected intake, results from learners’ selective attention accompanied by a very low level of processing paid to input. The theoretical assumption underlying this is that detected intake occurs without learners’ awareness, i.e. learners may detect target linguistic forms in the input without being aware of them (Leow, 2015). Thus, this detected intake is similar to ‘detection’ as explained by Tomlin and Villa (1994). Noticed intake, comparable to Schimdt’s (1990) ‘noticing’, entails focal attention, a low level of awareness and a low level of processing. Leow explains that noticed intake, which requires focal attention along with a certain level of awareness, is more likely to be stored and processed in working memory and, further, attributes to development of the L2 grammar system.

In the second stage, the intake processing stage, attended intake, detected intake and noticed intake are further processed with differences in the depth of processing. For instance, data-driven processing is considered a lower depth of processing, in which the preliminary intake is processed with a low level of cognitive effort. At a greater depth of processing, which is referred to as conceptually-driven processing, learning occurs with higher levels of awareness, including consciously encoding and decoding linguistic information. The final stage, the knowledge processing stage, takes place between the learner's L2 developing system and their production of output. According to Leow, this stage is part of the learning process in which learners can monitor and modify their output based on the feedback they may receive. At this stage, a greater depth of processing and higher level of awareness learners may also be needed for learners to develop their current L2 grammatical system.

2.2.5. Summary

Researchers, Schmidt, Tomlin and Villa, Robinson and Leow, have proposed different theoretical perspectives regarding the type of attention and level of awareness involved in 'noticing' input. The main difference discussed by these researchers is related to whether the learner's awareness is required or not in 'noticing' the input (Hama & Leow, 2010; Leow, 2001). Although there is disagreement on the role of awareness, a general consensus has emerged that attention is necessary, that is, it is a precondition for language learning to occur (Leow, 2001). Against this theoretical account, an overarching question that arises is how language learners' attention can be drawn to input linguistic forms with minimal interruption to the learning process. Among the various techniques discussed, the use of input-based tasks, as a type of task used in Task-based Language Teaching (TBLT), has been proposed to effectively direct learners' attention to the input (Ellis, 2003).

2.3. Task-based Language Teaching

Task-based Language Teaching (henceforth TBLT) has attracted much attention in the second language acquisition field (e.g., Bygate, Skehan, & Swain, 2001; Ellis, 2003; Long, 1985; Skehan, 1998). TBLT was developed as an alternative to traditional form-focused instruction to provide learners with opportunities to actively engage in meaning-focused communication. According to Ellis (2013), TBLT is "an approach to teaching a second/ foreign language that seeks to facilitate language learning by engaging learners in the interactionally authentic language use that results from performing a series of tasks" (p. 1).

2.3.1. Definition of Task

In TBLT, a task is determined to be a central unit of curricular and syllabus planning and language learners acquire the target language by performing a set of tasks (Robinson, 2011; Samuda & Bygate, 2008). With the role of tasks gaining in prominence, various definitions have been proposed by many researchers that help to conceptualize the construct of task (Table 1). On the basis of the diverse definitions proposed, Ellis (2009) summarizes the general criteria features defining a task as follows:

1. The primary focus should be on 'meaning' (by which is meant that learners should be mainly concerned with processing the semantic and pragmatic meaning of utterances).
2. There should be some kind of 'gap' (i.e. a need to convey information, to express an opinion or to infer meaning).
3. Learners should largely have to rely on their own resources (linguistic and non-linguistic) in order to complete the activity.
4. There is a clearly defined outcome other than the use of language (i.e. the language serves as the means for achieving the outcome, not as an end in its own right). (Ellis, 2009, p. 223)

Table 1. Definitions of Task

Long (1985, p. 89)	A piece of work undertaken for oneself or for others, freely or for some reward. Thus, examples of tasks are painting a fence, dressing a child. 'Tasks' are the things people will tell you they do if you ask them and they are not applied linguists.
Candlin (1987, p. 10)	One of a set of differentiated, sequenceable, problem-posing activities involving learners and teachers in some joint selection from a range of varied cognitive and communicative procedures applied to existing and new knowledge in the collective exploration and pursuance of foreseen or emergent goals within a social milieu.
Prabhu (1987, p. 24)	An activity which required learners to arrive at an outcome from given information through some process of thought and which allows teachers to control and regulate that process.
Bachman & Palmer (1996, p. 44)	An activity that involves individuals in using language for the purpose of achieving a particular goal or objective in a particular situation.
Willis (1996, p. 23)	Activities where the target language is used by the learner for a communicative purpose (goal) in order to achieve an outcome.
Skehan (1998, p. 95)	An activity in which: meaning is primary; there is some communication problem to solve; there is some sort of relationship to comparable real-world activities; task completion has some priority; the assessment of the task is in terms of outcome.
Ellis (2003, p. 16)	A work plan that requires learners to process language pragmatically in order to achieve an outcome that can be evaluated in terms of whether the correct or appropriate prepositional content has been conveyed. To this end, it requires them to give primary attention to meaning and to make use of their own linguistic resources, although the design of the task may predispose them to choose particular forms. A task is intended to result in language use that bears a resemblance, direct or indirect, to the way language is used in the real world. Like other language activity, a task can engage productive or receptive, and oral or written skills, and also various cognitive processes.
Nunan (2004, p. 4)	A piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is primarily focused on mobilizing their grammatical knowledge in order to express meaning, and in which the intention is to convey meaning rather than to manipulate form.

While TBLT has emerged as a promising language teaching approach, the general principles of TBLT have been challenged by some researchers (Widdowson, 2003). One of the criticisms concerns the definition of 'task'; it has been claimed that there seems to be an unclear distinction between tasks and traditional activities

(Littlewood, 2007; Widdowson, 2003). To counter these arguments, Ellis (2009, 2013) explains that tasks are different from traditional activities in the sense that learners are required to produce their own message using their own linguistic resources rather than simply manipulating language. Another distinctive feature of tasks is that tasks require communicative outcomes, including both linguistic and pragmatic outcomes, to be achieved.

In addition, TBLT has been misconstrued as emphasizing the output or production of language learners so that, as a result, insufficient linguistic input is available for second language acquisition to occur. Swan (2005), for instance, argues that "TBLT provides learners with substantially less new language than 'traditional' approaches" (p. 392). TBLT is thus presumed to provide only limited input to learners, and with a particular emphasis on learners' engagement in production, such as speaking or writing. Such a claim is predicated on the false assumption that tasks used in TBLT must entail a certain type of learners' interaction and production. However, tasks do not necessarily have to be output-based; rather, input-based tasks, involving listening or reading, are often encouraged to be used in TBLT (Ellis, 2009).

2.3.2. Input-based Tasks

To address the need to create an input-rich environment for L2 learners, some TBLT researchers have recommended that more extensive use be made of input-based tasks in task-based contexts (Ellis, 2009, 2013; Shintani, 2012). Unlike output-based tasks, which require learners to produce output, input-based tasks "promote interlanguage development by directing learners' attention to second language (L2) input through listening or reading without requiring them to produce the L2"

(Shintani, 2012, p. 254). Thus, input-based tasks offer contexts for natural processing of the target language, without the added pressure to produce output, which is particularly advantageous in the initial stages of learning a language (Ellis, 2013). The difference between input-based tasks and output-based tasks is summarized as follows:

> The difference lies in whether production (speaking or writing) is or is not required on the part of learners. Input-based tasks can be performed by learners listening or reading the information provided by the task. Speaking or writing is not required but it is also not prohibited. Learners can elect to produce if they choose to do so. Output-based tasks require production. Learners have to speak or write to achieve the task outcome. (Ellis, 2013, p. 5)

The basic underlying assumption is that input-based tasks draw learners' attention to input and, further, input attended to can become intake.

Motivated by these theoretical frameworks, several studies have been undertaken to compare the relative effectiveness of input-based and output-based instruction in developing L2 knowledge (e.g., Erlam, 2003; Izumi, 2002; Nagata, 1998a, 1998b; Shintani & Ellis, 2010; Takimoto, 2007). However, on account of the common misperception that output and production skills are underscored in TBLT, comparatively more attention has been paid to the role of output-based tasks in second language learning (Shintani, 2012). Consequently, to date, input-based tasks have received relatively little attention (Shintani, 2012); and therefore, the construct of input-based tasks seems to be under-researched. An input-based task, however, has substantiated its position as an effective language teaching practice in TBLT, with its potential to provide learners with extensive input and expose them to authentic language use (Ellis, 2013). Thus, more empirical evidence seems to be needed to better understand the application of input-based tasks in SLA research and language pedagogy, given that input-based tasks serve as an important source of rich and comprehensible input, which is essential to the success of second language learning (Shintani, 2012).

2.3.3. Empirical Studies on Input-based Tasks

The main areas that previous studies have been examined include: (a) the differential effects of input-based tasks and output-based tasks on L2 learning and (b) the effects of different conditions of input-based task on L2 learning. A majority of studies have been conducted to compare the relative effectiveness of input-based and output-based tasks on different aspects of L2 learning. Nagata (1998a, 1998b), for instance, examined how computer-assisted input-based and output-based practice affected the acquisition of a target grammatical structure in second language. In this study, input-based practice was provided as the format of a comprehension exercise which asked students to select the correct answer from the choices given. Output-based practice required the students to produce sentences according to the questions presented. The results confirmed an advantageous role for output-based practice for L2 learners to comprehend and produce the target structure.

Erlam (2003) also compared the effects of structured-input instruction and output-based instruction on the acquisition of a target grammatical structure in L2. Erlam explained that structured-input instruction was a type of explicit input-based instruction, which was different from implicit input-based instruction (e.g., enriched-input instructional technique and enhanced-input instructional technique). In this study, structured-input instruction included explicit instruction on target grammatical items and input-based activities, such as choosing a statement corresponding to a given picture and identifying errors. Output-based instruction also provided explicit information regarding target constructions; however, it was followed by production-based activities, including rewriting sentences, gap-fill activities and oral practice. The results favoured output-based instruction over input-based instruction.

A comparative study conducted by Shintani (2011) investigated the extent to which input-based and production-based instruction influenced the L2 vocabulary acquisition of young EFL learners. The participants, 36 Japanese children learning English as a foreign language, were assigned to three groups: an input-based group, a production-based group and a control group. Input-based instruction was designed as a 'listen-and-do' task, whereas a production-based task included certain types of oral production (e.g., repeating, naming an item etc.). The results indicated that both input-based and production-based tasks contributed to gains in vocabulary knowledge of young L2 learners. An additional important finding that Shintani reported was that the input-based task elicited more negotiation and interaction as students performed the task. In a subsequent study, Shintani (2012a) investigated whether beginner learners of English could benefit from input-based tasks in learning L2 vocabulary and grammar. In this study, input-based tasks were again operationalized as listen-and-do tasks, which required the participants to listen to the teacher's commands and respond to them. The results demonstrated that input-based tasks were beneficial for young L2 learners in terms of acquiring both receptive and productive knowledge of vocabulary. As for grammar, input-based tasks fostered the acquisition of receptive grammatical knowledge.

A study on the effects of input-based and output-based practice on the development of pragmatic knowledge was conducted by Li and Taguchi (2014). They examined whether different practice modalities (input- or output-based tasks) influenced the development of accuracy and speed in recognizing and producing request-making forms in L2 Chinese. In this study, the participants were randomly assigned to an input-based practice group, an output-based practice group and a control group. Input-based practice was designed to require making judgements on grammaticality, on the level of imposition of a target request and on the level of appropriateness of a given utterance. Output-based practice, on the other hand, required production, including sentence translation and dialogue completion. The study reported mixed results; input-based tasks were found to be more effective in promoting recognition accuracy and speed, whereas output-based tasks boosted production speed.

In a study by Takimoto (2007), the effects of different types of input-based tasks on L2 learning were compared. More specifically, Takimoto examined whether the development of language learners' pragmatic proficiency was affected by three different types of input-based tasks: (1) structured input tasks with explicit information, (2) problem-solving tasks and (3) structured input tasks without explicit information. A teacher-fronted explanation of the target feature was only included in 'structured input tasks with explicit information'. The results indicated that all input-based tasks, regardless of instructional condition, were equally effective in developing L2 pragmatic proficiency. The extent to which different amounts of input-based practice influenced the development of accurate and speedy recognition and the production of request-making forms in L2 Chinese was examined by Li (2012). In this study, computerized structured input-based activities were provided to both an intensive training group and a regular training group, differing only in the amount of training time. That is, the intensive training group received twice as much structured input practice on the target linguistic construction (request-making forms) as the regular training group. The results showed that input-based practice was effective in promoting accuracy in an oral discourse completion task and in enhancing speed in a pragmatic listening judgement task. Regarding the differential amount of practice, a positive association between the amount of practice and improvement in pragmatic performance accuracy for production tasks was found.

One noticeable feature of the aforementioned studies is that various terms are used, such as ‘input-based practice’, ‘input-based activities’, ‘input-based instruction’ and ‘input-based tasks’. There seems to be a general consensus reached in conceptualizing ‘input-based’ or ‘comprehension-based’, indicating that target linguistic constructions are provided in the input without a requirement for production, i.e. speaking or writing. However, some caution needs to be exercised when making comparisons across studies and interpreting findings, because there seem to be some differences in defining and operationalizing ‘practice’, ‘instruction’, ‘activity’ and ‘task’. More importantly, although the term ‘task’ is used in some studies, it appears that the way a ‘task’ is operationalized is different from the definition of a ‘task’ within the framework of TBLT, which must entail a primary focus on meaning, the presence of a gap, the use of non-linguistic resources and the existence of non-linguistic outcomes (Ellis, 2009). In this regard, to date, only a few studies have specifically examined the effects of ‘input-based tasks’ on L2 learning.

In one of the few extant studies, Shintani (2012b) investigated whether beginner learners of English could benefit from input-based tasks in learning L2 vocabulary and grammar. According to Shintani (2012b), input-based tasks can be classified into ‘enriched input tasks’ and ‘comprehension-based tasks’. With ‘enriched input tasks’, learners are provided with extensive input on target features, such as lexical or grammatical items, without any requirement to demonstrate that they have processed the input. ‘Comprehension-based input tasks’, on the other hand, expose learners to input and require them to display their understanding by responding to a certain type of task (e.g., listen-and-do task). In Shintani's study, an 'input-based task', as a type of comprehension-based input task, was operationalized as a 'listen-and-do task', which required the participants to listen and respond to the teacher's commands without the obligation to produce output. An investigation was carried out to determine whether the input-based task affected the acquisition of vocabulary items including 24 nouns, 12 adjectives and plural *–s* as grammatical knowledge. The results demonstrated that input-based tasks were beneficial for young L2 learners acquiring both receptive and productive knowledge of vocabulary. As for grammar learning, input-based tasks assisted in the acquisition of receptive knowledge of plural *-s*.

While previous studies have yielded mixed findings, mostly due to some discrepancies in their design, treatment conditions and operationalization, Shintani, Li and Ellis (2013) conducted a meta-analysis to provide a synthetic review of previous studies examining the relative effectiveness of comprehension-based instruction (CBI) and production-based instruction (PBI). In this meta-analysis study, however, it is important to note that "CBI" and "PBI" were used as umbrella terms encompassing a range of operationalisations of input-based and output-based instruction proposed by researchers (e.g., Erlarm, 2006; Nagata, 1998a, 1998b; Shintani, 2011, 2012a, 2012b; Tanaka, 2001). In the review, it was found that both comprehension-based and production-based instruction were effective in the development of L2 grammatical knowledge, particularly both receptive and productive knowledge.

2.3.4. Summary

Taken together, the advantages of input-based tasks are well documented in the literature. However, the fact that some discrepancies exist in operationalizing terms, as aforementioned, seems to contribute to mixed results; and consequently, the empirical findings to date have been inconclusive. A brief review of empirical previous studies indicates that very little research has operationalized ‘input-based tasks’ that meet the criteria for defining a ‘task’ within the TBLT framework (Ellis, 2009) and investigated their effects on second language acquisition; therefore, the construct and potential of ‘input-based tasks’ as a pedagogical intervention have yet to be empirically ascertained. In addition, input-based tasks are traditionally defined as involving either listening or reading (Ellis & Shintani, 2014). Input-based tasks can, however, also be conceptualised as multimodal tasks entailing various modes, such as audio, written and visual input modes. Hence, in the context of TBLT, one way to operationalise multimodal input-based tasks is by means of captioning.

2.4. Captions and Language Learning

Captioning, as one way to enhance the input provided to language learners, has received increasing attention in the area of instructed second language acquisition (e.g., Garza, 1991; Huang & Eskey, 2000; Markham, Peter & McCarthy, 2001; Markham & Peter, 2002–2003; Taylor, 2005; Winke, Gass, & Sydorenko, 2010). Captions are generally defined as “redundant text that matches spoken audio signals and appears in the same language as the target audio” (Vandergrift, 2007, p. 79). Although captions are different from subtitles, in that they refer to text provided on the screen in the viewer’s native language (Markham & Peter, 2002–2003), a distinction between captions and subtitles is not clearly made in some previous studies.[1] Captioning has been identified as a potentially effective tool for promoting language learning (Borras & Lafayette, 1994; Danan, 2004; Garza, 1991; Markham & Peter, 2002–2003). In particular, captioning assists learners to visualize word boundaries and parse structural patterns by connecting auditory and visual input, which may help learners remember and develop L2 knowledge (Bird & Williams, 2002; Garza 1991; Winke, Gass, & Sydorenko, 2013).

2.4.1. Theoretical Framework for the Use of Captions for L2 Teaching and Learning

The theoretical foundation underlying the research discussing the use of captioned video for pedagogical purposes is generally based on two theories, namely, Dual coding Theory (Paivio, 1986, 2007) and Cognitive Theory of Multimedia Learning (Mayer, 2001, 2005).

2.4.1.1. Dual coding Theory

Dual coding Theory (Paivio, 1986, 2007) provides a theoretical background for the use of captions for second language learning. According to Dual coding Theory, human information-processing is composed of two separate systems: a verbal system (written, auditory and articulatory verbal codes) and a non-verbal system (images for environmental sounds, activities and nonlinguistic features), as shown in Figure 2.

The two systems, verbal and nonverbal, are interconnected to each other through referential connections. Thus, representations in one system can activate those in the

[1] In previous studies, the terms “captions” and “subtitles” have often been used interchangeably to refer to on-screen text in either the L1 or the L2. In the current study, captions are operationalised as on-screen text in the target language (the L2). However, in this thesis, the original terms used in previous studies are maintained when reviewing them.

other one; that is a particular word in the verbal system can evoke related images and feelings in the nonverbal system.

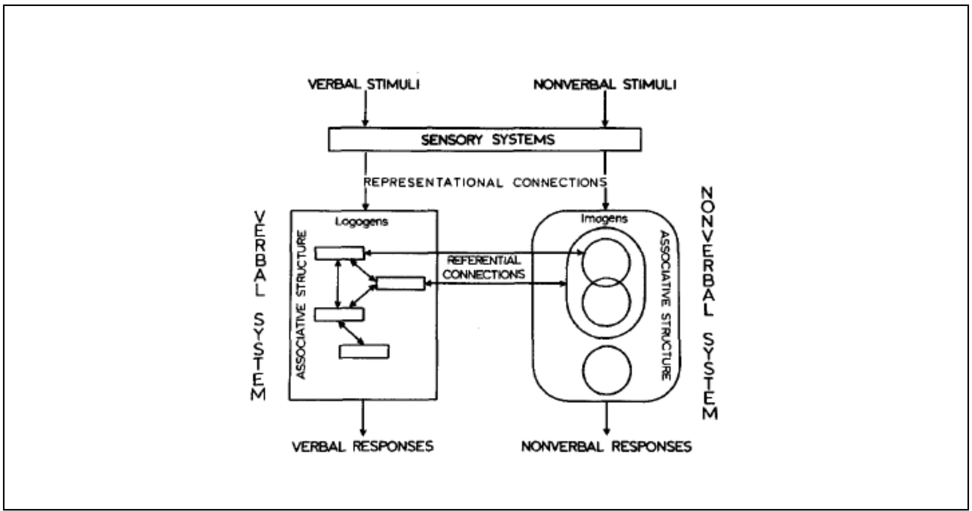


Figure 2. Dual coding Theory (Paivio, 1986, p. 67)

Another link, associated connections, connects representations within each system, verbal and nonverbal. For instance, a particular word used in the verbal system may elicit other related words; and likewise, a particular image or feeling may elicit related images and feelings in the nonverbal system. Based on this assumption, Paivio (2007) hypothesizes that activation of both the verbal and nonverbal systems can improve information processing, which may lead to better recall. Thus, providing input through multimodalities, including aural, visual and/or textual modes, may stimulate both the verbal and imagery systems, which is expected to facilitate information processing, and hence result in better learning (Paivio, 1986).

2.4.1.2. Cognitive Theory of Multimedia Learning

The potential for captioned video as an instructional intervention to facilitate language learning can also find theoretical support in educational theory, namely, the Cognitive Theory of Multimedia Learning (Mayer, 2001, 2005). Influenced by

Baddeley's (1986, 1992) Model of Working Memory and Paivio's (1986) Dual coding Theory, Mayer (2005) proposed the Cognitive Theory of Multimedia Learning to explain how learners process information in a multimedia environment. Three fundamental assumptions that underlie this theory are: (a) a dual-channel assumption, (b) a limited capacity assumption and (c) an active-processing assumption. The dual-channel assumption presupposes that there are separate information-processing channels for auditory/ verbal information and visual information. The assumption that the amount of information that can be processed in each channel at once is limited is called the limited capacity assumption; humans are believed to have limited processing capacity. The active-processing assumption refers to the idea that learners are engaged in active processing to construct a coherent mental representation, such as selecting relevant information, organizing information into coherent verbal and visual representations, and integrating information with prior knowledge.

As depicted in Figure 3, this theory presumes that learners first select relevant pictorial and/or linguistic from multimedia input; then, this is organized into coherent visual and verbal mental representations. The mental representations from two channels – visual and auditory– are integrated along with prior knowledge in working memory. The cognitive theory of multimedia learning further predicates that the learner must engage in five cognitive processes for meaningful learning to occur in a multimedia environment: (a) selecting relevant words for processing in verbal working memory, (2) selecting relevant images for processing visual working memory, (3) organizing selected words into a verbal model, (4) organizing selected images into a pictorial model and (5) integrating verbal and pictorial representations with each other and also with prior knowledge (Mayer, 2005).

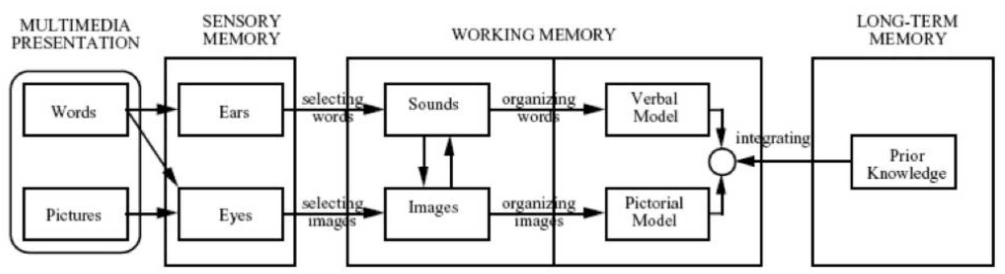


Figure 3. Cognitive Theory of Multimedia Learning (Mayer, 2005, p. 37)

Motivated by the theoretical assumptions discussed above, a number of previous studies have attempted to empirically test the effectiveness of captions for enhancing L2 learning (e.g., Başaran, & Köse, 2013; Bird & Williams, 2002; Chai & Erlam, 2008; Chang, 2009; Danan, 2004; Garza, 1991; Hayati & Mohmedi, 2011; Huang & Eskey, 2000; Markham, 1989, 1999; Markham, Peter, & McCarthy, 2001; Sydorenko, 2010; Taylor, 2005; Winke, Gass, & Sydorenko, 2010; Yang & Chang, 2014).

2.4.2. Captions as Pedagogical Intervention in L2 Teaching

The effects of captions on language learning have been researched extensively, most studies focusing on listening comprehension and vocabulary learning (e.g., Başaran & Köse, 2013; Bird & Williams, 2002; Chai & Erlam, 2008; Chang, 2009; Danan, 2004; Garza, 1991; Hayati & Mohmedi, 2011; Huang & Eskey, 2000; Markham, 1999; Markham, Peter, & McCarthy, 2001; Sydorenko, 2010; Taylor, 2005; Winke, Gass, & Sydorenko, 2010; Yang & Chang, 2014). The provision of captions has been evidenced to promote L2 learners' listening comprehension in a number of studies (e.g., Chai & Erlam, 2008; Danan 2004; Garza, 1991; Huang & Eskey, 2000; Rodgers & Webb, 2017; Winke et al., 2010). Garza (1991) evaluated the effects of captioning on the comprehension of content of video materials. The participants were advanced learners of two different target languages, namely, Russian and English. In each target language group, half of the participants were asked to watch video segments with subtitles, whereas the other half viewed video segments without subtitles. Content-based multiple-choice question tests were administered to assess how well the participants comprehended the content of video materials. The results indicated that the use of captions helped the learners comprehend the content of video materials. Garza states that "captioning may help teachers and students of a foreign language bridge the often sizable gap between the development of skills in reading comprehension and listening comprehension, the latter usually lagging significantly behind the former" (p. 246).

Huang and Eskey (2000) reported similar results, suggesting that closed-captioned television positively affected the listening comprehension of learners studying English as a second language at intermediate level. In this study, two treatment groups were included: one group watching traditional TV without captions and the other group watching TV with captions. The same video episode was viewed twice by both groups. A listening test followed the format of the listening comprehension subtest of the TOEFL, i.e. consisting of short spoken conversations. The results of the study suggested positive effects for closed-captioned TV (CCTV) on the listening comprehension of ESL students at intermediate level.

A recent study conducted by Rodgers and Webb (2017) attempted to use authentic video materials to examine the extent to which the presence of captions facilitated or hindered learners' comprehension of video episodes. Japanese university students (n = 372) learning English as a foreign language were divided into two groups, a captions group and a no captions group. After watching ten episodes of an American television programme (the American drama "*Chuck*"), each having an average length of 42 minutes, the participants' comprehension of the ten episodes was tested. Rodgers and Webb reported that participants who watched the video with captions achieved slightly higher scores than their counterparts across all ten episodes, with significantly higher scores only on three of them. They concluded that captions were more beneficial for learners when the content of episodes was more difficult.

While it is generally acknowledged that captioning improves L2 listening comprehension (e.g. Baltova, 1999; Huang & Eskey, 2000), some researchers have paid attention to whether captioning assists learners across proficiency levels to better comprehend listening materials. The potential relationship between the effects of using captions on L2 listening comprehension and language learners' proficiency levels was examined by Markham (1989). Markham explored the effects of subtitled TV on the listening comprehension of beginner, intermediate and advanced learners of English. Each group watched two different videos, one with subtitles and the other one without subtitles, which were followed by multiple-comprehension tests. It was found that all three groups performed better on a comprehension test when provided with subtitles. However, somewhat different results were presented in Taylor's (2005) study. He examined whether beginning learners of Spanish benefited from the use of captions. A total of 85 university students, beginning-level learners of Spanish, participated in this study. They were randomly assigned into two different groups, one group watched video with captions, the other group viewed video without captions. When comparing the scores on a free recall task and a multiple-choice comprehension test, no significant differences were found between the captioning group and the no captioning group. Taylor concluded that the use of caption might not be effective in enhancing the comprehension of beginning language learners. In addition, from reports collected from the learners, 35 per cent of the first-year students reported that the captions were distracting and confusing. The findings yielded from Taylor's study might indicate that the use of captions could have differential effects according to language learners' proficiency levels.

Building on the premise that presence of captions may promote learners' comprehension, some researchers have attempted to examine whether captions offered in learners' native language or target language have differential effects on L2 listening comprehension. (e.g., Başaran & Köse, 2013; Guichon & McLornan, 2008; Hayati & Mohmedi, 2011; Markham et al., 2001). Markham et al. (2001) explored the effects of Spanish captions, English captions and no captions on Spanish learners' comprehension of video materials. A total of 168 university students learning Spanish as a foreign language were assigned to one of three groups: a Spanish captions group, an English captions group and a no captions group. They were required to view Spanish DVD materials and complete a written summary as well as a multiple-choice test. The results demonstrated positive effects of using captions, in that both the Spanish captions group and the English captions group outperformed the no captions group. Furthermore, as for using either L1 or L2 captions, the English captions group performed better than the Spanish captions group on both measures; using L1 captions was found to be advantageous in improving language learners' comprehension. Based on their findings, Markham et al. suggested that language learners might benefit from multilingual captions, varying from L1 to L2.

Markham et al.'s (2001) study was replicated by Markham and Peter (2002–2003). Similar to Markham et al. (2001), participants viewed a video in one of three treatment conditions: English captions, Spanish captions or no captions. The results indicated that students who watched video in their L1 (English captions) performed significantly better on a multiple-choice test than the Spanish captions (L2) group and the no captions group. However, different results emerged in Guichon and McLornan's(2008) study. The primary aim of their study was to investigate the effects of multimodality on second language comprehension. A total of 40 French undergraduate students were asked to watch BBC news. They were assigned to: (a) a sound only group, (b) an image and sound group, (c) an image, sound and L1 subtitles group or (d) an image, sound and L2 subtitles group. As a measure of comprehension, the students were asked to produce a written summary. The results revealed that students' comprehension was better when they were exposed to multimodal input; in addition, L2 subtitling had a greater effect than L1 subtitling on learners' comprehension.

The superiority of L2 captions over L1 captions on listening comprehension was further ascertained in the study by Hayati and Mohmedi (2011). They compared the effects of using no captions, L1 captions (Persian) and L2 captions (English) with 90 university students who were intermediate learners of English. A documentary film in English was segmented into six different parts; the participants viewed one segment per week. Immediately after watching each part of the film, a multiple-choice comprehension test was administered. Similar to the results from Guichon and McLornan's (2008) study, the total average scores of the comprehension tests of the English captions group were higher than those of both the Persian captions group and the no captions group; in other words, the participants were found to benefit most from L2 (English) captions.

A similar study was conducted by Lwo and Lin (2012), focusing on children. Thirty-two fourteen-year-old Taiwanese students were allocated into four different groups: a no captions group, an L1 captions (Chinese) group, an L2 captions (English) group and an L1 plus L2 captions (Chinese plus English) group. Of the eight students in each group, half of them were more proficient and the other half were less proficient in English. From the results of a comprehension test administered after watching the videos, Lwo and Lin concluded that the effects of different modes of captions varied depending on students' proficiency levels. That is, L2 (English) captions and L1 (C hinese) plus L2 (English) captions were found to be helpful only for less proficient students to recall sentences. A further comparative study on the effects of L1 and L2 captions was carried out on younger children by Başaran and Köse (2013). A total of 30 Turkish primary school students learning English as a foreign language participated in this study. The participants were asked to watch a video without captions, with Turkish captions or with English captions. They were also asked to complete a multiple-choice test, which was included to assess their listening comprehension. The results of the study demonstrated that there were no significant differences in the listening comprehension test scores of all three groups. There seem to be some inconsistencies in previous studies on the effects of L1 or L2 captions on the comprehension of language learners. Hence, further research is necessary, which includes language learners who have different proficiency levels and come from diverse L1 backgrounds, as well as various types of tests to assess the comprehension of language learners.

However, the empirical studies reviewed above appear to adopt a ‘listening for comprehension’ approach (Richards, 2005), which refers to when the main objective of listening is to comprehend the meaning by adapting different processes and appropriate listening strategies according to the purpose of listening. Taking this perspective, these studies have attested to whether listening through captioned video assisted learners’ overall comprehension of listening materials using a certain type of comprehension test (e.g., a multiple-choice test or a written summary). The other approach, ‘listening for acquisition’ (Richards, 2005), on the other hand, pertains to listening instruction focusing on developing L2 proficiency (Richards, 2005; Rost, 2002). Within the framework of ‘listening for acquisition’, linguistic resources are provided, through listening, for L2 learners to learn target constructions. In this respect, L2 learners are asked to attend to particular linguistic features, such as grammatical, lexical or pragmatic features, while engaging in listening. Compared to the extensive research conducted within the view of ‘listening for comprehension’, relatively insufficient attention has been paid to how listening through captioned video can serve as a means to facilitate second language learning. There exist studies taking a ‘listening for acquisition’ approach, mainly focusing on investigating the extent to which listening using captioned video might assist L2 vocabulary acquisition (e.g., Bird & Williams, 2002; Chai & Erlam, 2008; Danan, 1992; Markham, 1999; Markham et al., 2001; Sydorenko, 2010; Winke, Gass, & Sydorenko, 2010). Early studies attempted to examine the effectives of captioned video on word recognition. Neuman and Koskinen (1992), for instance, investigated the effects of captioned television on incidental vocabulary learning from context. The participants’ vocabulary knowledge was identified by using a word recognition test. Of four different treatment conditions – captioned television, traditional television viewing without captions, reading along and listening to text, textbook only – captioned television was found to be most effective to enhance language learners' ability to recognize words.

The relationship between the use of captioned videotapes and improvement in aural word recognition of ESL students was also investigated by Markham (1999). A total of 118 advanced, university-level ESL students were asked to view excerpts from two different episodes of educational television programmes, either with or without captions. To identify the students' ability to recognize words, a listening multiple-choice test was administered for each episode. The results of the study indicated that the advanced ESL learners benefited from captions to recognize words. Similarly, Bird and Williams (2002) investigated whether single modal presentation (sound or text) or bimodal presentation (sound and text) facilitated learning of both familiar and unfamiliar words, as measured by improvements in auditory word recognition and word retention. Based on two consecutive experiments, Bird and Williams concluded that bimodal presentation, sound and text, positively affected both auditory word recognition and word retention.

In this line of inquiry, Chai and Erlam (2008) examined whether captioned video impacted receptive vocabulary knowledge development. The main purpose of this study was to investigate the effects of captioned video on learning L2 phrases and words. Twenty native speakers of Chinese ESL learners were divided into two groups: a no captions group and a captions group. At the outset of the study, a pretest was administered to the participants in order to examine their prior knowledge of targeted words and phrases. In addition, prior to watching two segments of video, the participants were provided with an opportunity to comprehend the content by viewing the video in their L1 (Chinese). Chai and Erlam explained that since their study focused mainly on examining the effects of captions on a micro-level of learning, such as phrases and words, achieving a macro-level understanding of the content was a necessary step. After ensuring that the participants comprehended the content of the video, they watched two segments of video, either without or with captions in L2 (English). An immediate posttest was administered for each segment of video and a delayed post-test was conducted after two weeks. In the tests, which aimed to measure receptive knowledge, 45 items were classified into three categories: 15 items for testing words, 19 items for testing phrasal verbs and 11 items for testing phrases. The participants were asked to match words with correct definitions, to complete sentences by choosing the correct word, to choose the appropriate preposition to be used for a given phrasal verb and to give definitions. Chai and Erlam reported that the participants benefited from captions to learn phrases in English (e.g., *suit yourself, for once, you are a natural, don't give me that*); however, the captions had negligible effects on the acquisition of L2 words or phrasal verbs.

Sydorenko (2010), however, presented somewhat different results. This study examined the effects of different input modalities (video, audio and captions) on developing L2 vocabulary knowledge. After viewing videos in one of three conditions, i.e., (a) a video with audio and captions, (b) a video with audio and (c) a video with captions, L2 learners' vocabulary knowledge was assessed using four different measures: (a) a comprehension test, (b) a written and aural recognition test, (c) a written and aural translation test and (d) a word knowledge test. In this study, differential effects of modality of input were observed depending on types of word recognition tests, either written or aural recognition tests. The group with audio input performed better on an aural than a written recognition test, whereas the group with captions scored higher on a written than an aural recognition test. In addition, Sydorenko observed that videos with both audio and captions were more beneficial for L2 learners to learn word meanings than videos with either audio or captions.

Winke et al. (2010) also investigated L2 learners' use of captions while watching videos in a foreign language. With native speakers of English who were second- and fourth-year learners of Arabic, Chinese, Spanish and Russian, an investigation was undertaken to examine whether captioning facilitated comprehension and L2 vocabulary learning. Furthermore, the researchers also looked into the effect of the order of caption presentation. The participants were required to watch three short videos twice, once with and once without captioning; in each language, one group of participants was presented with captions during a first viewing of the video and the other group was offered with captions during a second viewing. After watching a series of videos, the participants were asked to take vocabulary and listening comprehension tests based on the videos. Winke et al. revealed that students who had watched videos with captioning performed significantly better on written vocabulary tests, aural vocabulary tests and listening comprehension tests. The ordering effect of captioning was found to be language-specific. More specifically, learners of Spanish and Russian generally benefited more from captioning when it was presented in the first video than in the second one. However, captioning offered in the second video was more effective for Arabic and Chinese. In addition, whether the ordering effect of captioning was mediated by learners' proficiency levels was also explored. However, learners' proficiency levels did not have any influence on the effects of captioning; in other words, captioned video was beneficial for L2 learners, regardless of their proficiency levels.

Overall, on the basis of accumulated evidence, the positive effects of captioning on L2 listening comprehension and vocabulary learning were confirmed in a quantitative meta-analysis (Montero Perez, van den Noortgate, & Desmet, 2013b). In a comprehensive review of 18 empirical studies on the effectiveness of captioning, separate analyses were performed to estimate the magnitude of effect sizes of listening comprehension (including data from 15 studies) and vocabulary learning (including data from 10 studies). The results indicated a large effect size for both L2 listening comprehension (g = .99) and L2 vocabulary learning (g = .87), demonstrating positive effects for captioning on L2 learning. Nonetheless, no studies have explored whether learners might benefit from listening with captioned video in developing L2 grammatical knowledge from the perspective of 'listening for acquisition'. Furthermore, the observed benefits of supplying captions have been explained by the fact that captions can assist learners in segmenting speech into words (Bird & Williams, 2002; Vanderplank, 1988). Access to segmented speech, in turn, is likely to facilitate word recognition (Bird & Williams, 2002; Markham, 1999), which is a key determinant of successful L2 listening (Rost, 2011) and reading comprehension (Grabe, 2012). More successful word recognition also enables learners to identify new lexical items in the input with greater ease, promoting attention to and learning of new vocabulary (Winke, Gass, & Sydorenko, 2010). Hence, it seems reasonable to assume that the availability of captions also helps in drawing learners' attention to grammatical constructions. Having less advanced processing skills, L2 listeners (Rost, 2011) and L2 readers (Grabe, 2012) often need to draw on controlled, conscious rather than automatic processing when decoding aural and written input (Segalowitz, 2003). With the availability of captions, the demands on word recognition processes will decrease, which is likely to allow learners to pay more attention to morphosyntactic features and engage in more in-depth processing of grammar. Thus, an exploration of how and whether different types of captioning influence the acquisition of a grammatical feature is expected to extend this line of research.

In another strand of research, attempts have been made to compare the effects of diverse types of captions, such as no captions, keyword captions or full captions, on L2 vocabulary learning and L2 listening comprehension (e.g., Guillory, 1998; Montero Perez, Peters, Clarebout, & Desmet, 2014; Yang & Chang, 2014). The few studies that have looked into the effects of different types of captions on listening comprehension and vocabulary acquisition yielded differential gains depending on captioning conditions. Guillory (1998), for instance, investigated how no captions, keyword captions and full captions included in an educational video affected learners' overall comprehension. The participants were randomly assigned to three different groups: a no captions group, a keyword captions group and a full captions group. Guillory found that learners benefited most from full captions provided in the video. However, somewhat different results were reported by Yang and Chang (2014), who compared the influence of different modes of captions – full captions, keyword only captions and annotated keyword captions – on the listening comprehension of EFL learners. In this study, 44 EFL university students viewed 65-minute-long video clips and completed a listening comprehension test, which consisted of four different types of questions: dictation cloze, short dialogue comprehension questions, reduced-form recognition questions and reduced form marking. Yang and Chang found that the annotated keyword captions group performed better than the other two groups. A recent study conducted by Bensalem (2016) also investigated the effect of two types of captioning – full and keyword captioning – on listening comprehension with 36 university level EFL students at beginning level. The students, who were divided into a full captions group (*n* = 13), a keyword captions group (*n* = 13) and a control group (*n* = 13), were asked to watch three video clips (i.e., documentaries from an educational programme) according to the treatment conditions to which they were assigned. The full captioning group was found to outperform both the keyword captioning and the no captioning group on listening comprehension tests.

Montero Perez et al. (2014) also explored the effectiveness of different types of captioning in improving L2 learners' incidental vocabulary learning and comprehension of video material. The participants, 133 undergraduate students, were assigned to four groups: (a) a control group watching video clips without captioning, (b) a second group watching fully captioned clips, (c) a third group watching keyword captioned clips and (d) a fourth group watching fully captioned clips with highlighted keywords. The researchers used four tests to measure lexical gains: form recognition, meaning recognition, meaning recall and clip association (testing whether participants could associate words with corresponding video clips). In addition, three tests were used to measure comprehension (i.e., short open-ended questions, true and false questions, and a combination of the previous two tasks). The results revealed that the provision of captions, regardless of caption type, had a positive effect on the learners' performance on form recognition and clip association tests. The type of captioning, however, had a significant impact on the meaning recognition test: the groups with keyword captioning and full captioning with highlighted keywords achieved greater gains than the control group. This study by Montero Perez and colleagues suggests that increasing the visual salience of target areas in captions has the capacity to capture learners' attention to target lexis and generate vocabulary gains.

In a follow-up study, Montero Perez, Peters and Desmet (2015) further investigated the usefulness of different types of captions in promoting vocabulary knowledge, this time not only examining the impact of captioning on vocabulary learning but also on the allocation of attention to target lexis. The type of captioning was operationalised as access to either keyword captions or full captions. Another independent variable in the study was the presence versus absence of test announcement, resulting in an incidental (no announcement) and an intentional (announcement) condition. The participants were assigned to four treatment groups: full captioned video and incidental, full captioned video and intentional, keyword captioned video and incidental, and keyword captioned video and intentional groups. Development in vocabulary knowledge was gauged with the same type of assessments as those used in Montero Perez et al. (2014): form recognition, meaning recognition, meaning recall and clip association. Attention allocation to target words was measured with three eye-tracking indices: gaze duration (i.e., sum of fixations before leaving the target word area), which captured initial processing (Rayner, 1998); second pass reading time (i.e., time spent rereading the target word area), reflecting reanalysis of information; and total fixation duration. The results revealed an advantage for keyword captioning in terms of gaze duration and performance on a form recognition test. The keyword captions group also showed greater second pass reading times and total fixation duration under the intentional condition. However, significant links between attentional allocation and learning gains were only attested for the full-captions groups on the form recognition test: longer total fixation time and second pass reading time led to better scores when learners were made aware of a forthcoming test, whereas longer initial gaze duration was associated with greater gains when a vocabulary test was not announced. Interestingly, in the absence of test announcements, those who had higher second pass reading times displayed lower form recognition scores in the full captions group.

A review of previous studies indicates that mixed findings have emerged; and thus, the answer to the question as to which type(s) of captions would be of benefit to language learners remains uncertain. More importantly, the basic assumption underlying these studies seems to be that learners do pay attention to input provided through multimodalities, such as aural, visual and textual modes. However, as discussed above, not all available input is likely to be processed by learners (e.g., Robinson, 1995; Schmidt, 1990, 1993b, 2001; Sharwood Smith, 1991, 1993; Tomlin & Villa, 1994); in particular, when learners receive input simultaneously in multimodal input-based tasks, by means of captions, they may not attend to all the input provided. Given the importance of the attention directed to L2 input, it would be interesting to examine a type of captioning that is specially developed to draw learners' attention to linguistic features. Among various types of attention-getting techniques, input enhancement, as a relatively implicit and proactive type of focus of form intervention, can possibly be incorporated into captions to make particular linguistic features more salient to language learners so as to draw their attention to target linguistic constructions with the attention remaining on understanding the meaning of input (Sharwood Smith, 1991, 1993).

In fact, in Montero Perez et al.'s (2014) study, which investigated whether types of captions had differential effects on vocabulary learning and comprehension of video material, full captions with highlighted keywords were included as one type of captioning. However, the focus of this study remains on examining the effectiveness of perceptually salient target linguistic constructions in captions for facilitating incidental vocabulary acquisition. Considering that textually enhanced input elicits vocabulary gains, a further investigation on whether similar results would be yielded for L2 grammatical knowledge development is warranted.

2.4.3. Summary

There exist two main strands of research focusing on the extent to which captions promote either L2 listening comprehension or L2 vocabulary acquisition. Accumulated empirical evidence suggests the potential of captions as a pedagogical intervention (Montero Perez et al., 2013b). Hence, further investigation of whether captions would have the same beneficial effects on L2 grammatical knowledge development is expected to extend this line of research.

Table 2. Summary of Research on Captions

Study	Purpose of the Study	Participants	Input	What they were asked to do	Findings
Garza (1991)	To evaluate the effects of captions on learners using captioned video materials	• Students of Russian as a foreign language (N =40) • ESL students (N =70)	5 English and 5 Russian video segments on different genres	Watch a video	Positive effect on global comprehension of captioned video segments.
Danan (1992)	To examine how subtitled video programmes could enhance foreign language learning	[Study 1] 30 college students learning French • French audio only (N =11) • French audio with English subtitles (N =10) • English audio with French subtitles (N =9) [Study 2] 57 first-year French students • French audio with French titles (N =21) • French audio (N =13) • English audio with French titles (N =23)	Video excerpt combining narration and dialogue Video excerpt	Watch a video excerpt Watch a video excerpt	Positive effects of reversed subtitling (English dialogue with French titles) on vocabulary learning Positive effects of reversed subtitling (English dialogue with French titles) on vocabulary learning

Neuman & Koskinen (1992)	To examine whether comprehensible input in the form of captioned television might affect bilingual students' acquisition of vocabulary and conceptual knowledge	129 bilingual 7–8 graders (middle-school students) • Captioned TV (N =32) • Traditional TV without captions (N =37) • Reading along while listening to text (N =32) • Textbook only (N =28)	9 television segments	Watch a TV segment	Positive effects of captioned television on word recognition and recall information (remembered more science information than others.)
Guillory (1998)	To investigate the optimum amount of a second language text (French) in captions by comparing full-text captions, keyword captions and no captions.	202 university students learning French • No captions (N =70) • Full text captions (N =68) • Keyword captions (N =64)	2 video clips	Watch video clips	Positive effect of both keyword and full text captions on comprehension
Chung (1999)	To compare listening comprehension rates for video texts using a variety of techniques: advanced organizers; captions; a combination of both;	170 university students • Advanced organizer group • Captions group • Combination group • Control group	2 episodes from a television series created to promote English learning	Watch video segments	Combination group outperformed other groups on a comprehension test.

	none of the foregoing.				
Markham (1999)	To examine the effects of captioned videotapes on ESL students' listening word recognition [word recognition]	118 university level ESL students • Civil rights with captions (N = 26) • Civil rights without captions (N = 33) • Whales (N =35) • Whales (N = 24)	2 educational television programmes concerning whales and the civil rights movement	Watch clips with or without captions	The availability of captions significantly improved ESL students' listening ability to recognize words on the video tapes
Huang & Eskey (2000)	To investigate the effects of closed-captioned TV (CCTV) on listening comprehension	30 ESL students with intermediate-level proficiency • Traditional TV without captions (N=15) • CCTV (N=15)	Television series (26 episodes) designed for ELS classroom teaching	Watch a television series	CCTV helped ESL students' general comprehension, vocabulary acquisition and listening comprehension.
Markham (2001)	To examine the effects of caption availability on advanced university level ESL students' comprehension of contrasting, religion-specific video material	79 advanced ESL students • Religion-neutral (N =44) • Muslims (N =16) • Buddhists (N =9)	Videotaped episodes consisted of two excerpts from a series of public television broadcasts concerning world religions 1 – Muslim 2 – Buddhism	Watch videos	Positive effect of captions on comprehension for religion-neutral students
Markham, Peter, &	To examine the effects of using Spanish captions,	169 intermediate university-level Spanish as foreign language students	DVD episode presenting information	Watch a DVD episode	English captions group performed and at a substantially higher level than Spanish captions group.

McCarthy (2001)	English captions or no captions with a Spanish-language soundtrack on the comprehension of DVD passage material	• No captions (N =68) • Spanish captions (N =43) • English captions (N =58)	concerning preparation for the Apollo 13 NASA space-exploration mission		
Markham & Peter (2003)	To investigate the effects of using Spanish captions, English captions or no captions with a Spanish language soundtrack on listening/ reading comprehension	213 intermediate university-level Spanish as foreign language students • No captions (N =63) • Spanish captions (N =85) • English captions (N =65)	DVD episode presenting information concerning preparation for the Apollo 13 NASA space-exploration mission	Watch a DVD episode	English captions group performed at a considerably higher level than Spanish captions group.
Taylor (2005)	To investigate the effectiveness of captioned video on beginning Spanish students' comprehension	71 beginning second-semester students of Spanish • Captioned video (N =35) • No captioned video (N=36)	A video clip	Watch a video clip	No difference between captioning and no-captioning groups on either free recall or a multiple-choice test.
Chai & Erlam (2008)	To investigate how video plus captions impacted the	20 Chinese learners of English • No captions group	Movie excerpt	Watch a movie excerpt (L1) → Look at a word	No statistically significant effect was found between the two treatment groups.

	learning of second language words and phrases	(N =10) • Captions group (N =10)		list → Watch segment one in English (L2) with or without captions according to a group → posttest (same procedure for the segment 2)	Captions group achieved statistically higher scores on phrases than did no captions group.
Guichon & McLornan (2008)	To examine the impact of different types of input on L2 learners' comprehension of spoken English	40 French-speaking students learning English • Audio alone (N =10) • Video with audio (N =10) • Video with audio and L2 subtitles (N =10) • Video with audio and L1 subtitles (N =10)	A news clip	Watch a news clip (note-taking was allowed in L1 or L2)	Groups with subtitles (L1 and L2) outperformed groups without subtitles.
Sydorenko (2010)	To examine the effect of input modality (video, audio and captions) on (a) learning of written and aural word forms, (b) overall vocabulary gains, (c) attention to	26 learners of Russian • Video with audio with captions (N=8) • Video with audio (N=9) • Video with captions (N=9)	3 video clips	Watch video clips	Group with captions (video+audio+captions / video+captions) scored higher on written than on aural recognition of word forms, while reverse applied to video with audio group.

	input and (d) vocabulary learning strategies of beginning L2 learners.				Video+audio+captions group learned more word meanings than video+audio group.
Winke, Gass, & Sydorenko (2010)	To investigate the effects of captioning during video-based listening activities	150 foreign language learners • Spanish (2nd: N=47/ 4th: N=20) • Russian (2nd: N=24 / 4th: N=17) • Arabic (2nd: N=29) • Chinese (2nd: N=13)	Three short English-language documentaries about three animals: salmon, bears and dolphins	Watch three videos twice (once with captions / once without captions)	Positive effect of captioned videos on overall comprehension. Order of viewing had an effect on subsequent recognition of vocabulary presented in aural mode.
Hayati & Mohmedi (2011)	To examine the efficacy of subtitled movies on listening comprehension	90 EFL students • English subtitles group (N=30) • Persian subtitles group (N=30) • Without subtitles group (N=30)	6 episodes of a DVD	Watch DVD movie episodes	English subtitles group outperformed Persian subtitles group, which in turn performed at a substantially higher level than no subtitle groups on a listening test.
Lwo & Lin (2012)	To explore the impact of different captions on L2 learning in a computer-assisted multimedia context	32 eighth-graders No captions (N=8) Chinese captions (N=8) English captions (N=8) Chinese and English captions (N=8)	Animation with English narration	Watch animation with English narration in one of the conditions	Effects of different modes of captions varied depending on students' proficiency levels. L2 (English) captions and L1 (Chinese) plus L2 (English) captions were found to be helpful only for less proficient students to recall sentences.

Montero Perez, Peters & Desmet (2013b)	To investigate (a) the effects of two types of captioned video on listening comprehension and (b) L2 learners' perceptions of the usefulness of captions while watching L2 video	226 university level students • Control group (N=70) • Fully captioned clips group (N=81) • Keyword captioned clips group (N=75)	3 video clips in French	Watch videos in French	Full captioning group outperformed both no captioning and keyword captioning groups on global comprehension questions. No differences found between keyword captioning and no captioning groups.
Montero Perez, Peters, Clarebout, & Desmet (2014)	To examine how three captioning types can assist L2 learners in the incidental acquisition of target vocabulary words and in the comprehension of L2 video	133 undergraduate students • Control group (N=32) • Full captioning group (N=30) • Keyword captioning (N=34) • Full captioning with high-lighted keywords (N=37)	3 short clips	Watch clips	All captioning groups, regardless of type, outperformed control group on form recognition and clip association. Keyword captioning and full captioning with highlighted keywords groups outperformed control group on meaning recognition. Captioning did not affect comprehension or meaning recall.
Yang (2014)	To explore how subtitles and advance organizers affect EFL learners' listening	71 college students • Subtitles group (N=36) • No subtitles group (N=35)	CNN news clips	Watch news clips	Positive effect of subtitles on EFL learners' listening comprehension

	comprehension of authentic videos				
Frumuselu, De Maeyer & Donche (2015)	To explore informal and conversational speech (slang, phrasal verbs and colloquial expressions) through the use of a subtitled TV series (interlingual and intralingual)	40 undergraduate students studying for a BA in English • Interlingual mode (English sound + Spanish subtitles) (N=18) • Intralingual mode (English sound + English subtitles) (N=22)	13 subtitled episodes (American series ‘Friends’)	Watch episodes.	Learners performed better under EE (intralingual) than ES (interlingual) mode.
Bensalem (2016)	To investigate the effect of two types of captioning (full and keyword captioning) on listening comprehension	36 university-level EFL students (beginning level) Full captions (N=13) Keyword captions (N=12) Control group (N=12)	3 video clips (documentaries from an educational programme)	Watch video clips	Full captioning group significantly outperformed both keyword captioning and no captioning groups on listening comprehension tests.
Rodgers & Webb (2017)	To investigate the effects of captions on EFL learners’ comprehension of English language television programmes	372 university students learning English • Captions group (N=51) • No captions group (N=321)	Ten television episodes	Watch episodes	Positive effect of captions on comprehension of episodes, especially when episodes more difficult.

2.5. Textual Enhancement

The term “input enhancement” was first proposed by Sharwood Smith (1991, 1993). It refers to a pedagogical intervention that makes specific linguistic features in the input more salient so as to draw learners' attention to them. Given the centrality of attention to the process of acquisition (Leow, 1997, 1999, 2001; Robinson, 1995; Schmidt, 1990, 1993b, 1994, 1995; Sharwood Smith, 1991, 1993; Tomlin & Villa, 1994), a basic assumption underlying input enhancement is that textually enhanced input, as a form of implicit instruction, makes input more perceptible and thereby increases the chance that visually prominent input will be noticed by learners, which may then result in progress in learning (Doughty & Williams, 1998). Input enhancement is considered an external attention-drawing technique (Izumi, 2002), as the input is manipulated using external means. In oral mode, stress and intonation can be used to enhance the input (Cho & Reinders, 2013; Gascoigne 2006). In written mode, generally, visual or textual enhancement is utilised to promote the salience of particular linguistic features. These typically involve some kind of typographical modification, such as underlining, **boldfacing**, *italicization*, CAPITALIZATION, colouring or using different font types (Sharwood Smith, 1991, 1993).

Learners are presumed to be less likely to pay attention to morphosyntactic features in the input, especially if these are non-salient and communicatively redundant (Long & Robinson, 1998), while vocabulary has generally been shown to be more susceptible to noticing (e.g., Mackey, Gass, & McDonough, 2000). Thus, a number of studies have examined the effectiveness of textual enhancement in drawing learners’ attention to grammatical features in the field of instructed SLA (Lee & Huang, 2008). Textual enhancement – a type of implicit and unobtrusive “focus-on-form” instruction – has been proposed as an effective means of making linguistic features salient to language learners while retaining an overall focus on attention to the meaning of the input provided (Sharwood Smith, 1991, 1993).

2.5.1. Focus on Form

Focus on Form, as a pedagogical intervention used within the framework of TBLT, is defined thus:

> Focus on form . . . overtly draws students’ attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication. (Long, 1991, pp. 45–46).

Given the assumption that a focus on form can draw learners’ attention to target linguistic constructions without diverting their focal attention to meaning, considerable interest has been placed on determining the effectiveness of focus on form techniques in promoting second language learning. One issue that has intrigued researchers is the degree to which focus on form interventions should be explicit to capture learners’ attention (Doughty & Williams, 1998). The aim of an explicit focus on form “is to direct learner attention and to exploit pedagogical grammar” (Doughty & Williams, 1998, p. 232), whereas an implicit focus on form seeks to “attract learner attention to avoid metalinguistic discussion, always minimizing any interruption to the communication of meaning” (Doughty & Williams, 1998, p. 232). In other words, according to Doughty and Williams, an explicit focus on form provides explicit learning conditions to learners through, for example, dictogloss, conscious-raising tasks and input processing. An implicit focus on form, on the other hand, aims to draw learners’ attention to rules and forms without overt explanation; examples of an implicit focus on form include input flood, input enhancement and recasts. Extensive research has been conducted to identify the effectiveness of different types of focus on form as pedagogical interventions to trigger learners' attention and, among the various techniques, input enhancement, an implicit and unobtrusive type of focus on form, has proved to be of particular interest to many researchers in the field of instructed second language acquisition.

2.5.2. Empirical Studies on Textual Enhancement

Given the important role of attention in second language acquisition, there has been a plethora of empirical studies exploring the effect of textual enhancement (TE) on subsequent comprehension and L2 development (e.g., Alanen, 1995; Bowles, 2003; Izumi, 2002; Jourdenais, Ota, Stauffer, Boyson, & Doughty, 1995; Lee, 2007; Leow, 1997, 2001; Leow, Egi, Nuevo, & Tsai, 2003; Overstreet, 1998; Park, 2004; Shook, 1994; White, 1998; Wong, 2003). Overall, findings in this strand of research lack congruence, with some studies reporting positive effects, some negative effects and others no effects of textual enhancement. For instance, the main purpose of Doughty's (1991) study was to investigate the effectiveness of second language instruction on the acquisition of interlanguage (IL) grammar. In this study, Doughty compared three different types of instruction: (a) meaning-oriented instructional treatment (b) rule-oriented instructional treatment and (c) a control treatment. In meaning-oriented instructional treatment, a highlighted target language structure was included to attract learners' attention. The findings of the study revealed that increasing the perceptual saliency of target language structures that was intended to capture learners' attention resulted in developing learners' grammatical knowledge. Similar results were obtained in a series of experiments conducted by Williams (1999). In one of the studies by Williams, the input enhancement was included as one variable to examine whether the increase in the saliency of the grammatical morphemes would result in learning gains of the target forms. Overall, the effectiveness of visual input enhancement in improving the learners' knowledge of the target linguistic constructions was evidenced.

However, the fact that different target linguistic constructions are included in the textual enhancement studies should be taken into consideration. In fact, the perceived effectiveness of instructional intervention may be susceptible to the constructive nature of target constructions (e.g., Goldschneider & Dekeyser, 2001; Spada & Tomita, 2010). Thus, researchers have carried out investigations on whether textual enhancement promotes the learning of target constructions to a different extent depending on the nature of the construction (e.g., Alanen, 1995; Park & Nassif, 2014; Shook, 1994). In a study by Shook (1994), for instance, the effects of input enhancement on the development of L2 grammatical knowledge were examined with target linguistic forms, i.e., Spanish present perfect and relative pronouns. Shook explained that use of the present perfect was a more meaningful, aspectual decision by speakers compared to relative pronouns which were less meaningful and governed by syntactic choice. The participants (*n* = 125) were divided into three groups: (a) an unenhanced text group, (b) an enhanced text group with no specific instructions to pay attention to target forms and (c) an enhanced text group with explicit instruction focusing on target forms. A multiple-choice recognition task and a fill-in-the-blank task were presented to the participants. The results showed that the two groups receiving enhanced input outperformed the group exposed to unenhanced input, with significant gains for the present perfect, but not for relative pronouns. Shook interpreted the results as meaning that the distinctive features of the two target forms, in terms of their degree of meaningfulness, could be attributed to differential gains in the two forms.

A similar study was conducted using semi-artificial language as target forms by Alanen (1995). The extent to which textual enhancement had differential effects on two target linguistic forms, one being "clearly definable semantic content" (p. 269) and the other being "semantic empty" (p. 269), was examined. In this study, 36 English speaking learners of Finnish were randomly assigned to four groups: (a) an input enhancement group, (b) an explicit rule instruction group, (c) an input enhancement and explicit rule instruction group and (d) a control group. The target linguistic constructions were locative suffixes (i.e., clearly definable semantic content) and consonant alternation (i.e. semantically empty) in semi-artificial Finnish. The results of this study showed that the explicit instruction group exhibited significantly higher gains than other groups. Input enhancement was found to have some facilitating effect on learners' recall and use of locative suffixes, but not on consonant alternation.

More recently, Park and Nassif (2014) also addressed a possible relationship between the effectiveness of textual enhancement and the nature of target forms. The main purpose of this study was to examine the impact of textual enhancement on the comprehension of text and on the production of target linguistic forms varying in their degree of communicative value: Arabic comparative forms (meaning-bearing structure / high communicative value) and Arabic dual pronouns (grammatical structure / low communicative value). English-speaking students learning Arabic as a foreign language were assigned to either an enhanced group (n = 7) or an unenhanced group (n = 9). Two measures were used to assess the learners' comprehension of a reading passage, namely, a free recall task and a comprehension test. Learners' immediate production of target forms was assessed with a fill-in-the-blanks task and a free production task. The results revealed differential effects of textual enhancement according to the nature of the target forms; enhancing a linguistic form with high communicative value facilitated the learners' local comprehension. Enhancing a form with low communicative value, on the other hand, hindered learners' local as well as global comprehension. A conclusion was drawn that textual enhancement had a negative effect on learners' comprehension and a negligible effect on the immediate production of target forms.

Conceivably, the constructive nature of target linguistic constructions is an important variable that needs to be accounted for when interpreting the results of these studies; however, another question that needs to be addressed concerns variations in the typographical techniques that were used to manipulate the input. More specifically, in the font to enhance target linguistic constructions, Shook (1994) used upper-case letters and bold, whereas Alanen (1995) used italics. In Park and Nassif's (2014) study, a combination of three types, namely an enlarged font, bolding and underlining, was employed. Along with the different nature of target constructions, the fact that differences exist in the types of typographical modifications across studies appears to demand further discussion.

In other studies, the effect of textual enhancement was also examined in relation to other variables, such as the length of the text, topic familiarity, output or simplified input (e.g., Lee, 2007; Leow, 1997; Overstreet, 1998; Wong, 2003). Leow (1997), for instance, investigated the effects of written input enhancement and text length on second language learners' comprehension and acquisition of target forms. College students learning Spanish were assigned to one of four conditions: an enhanced short text, an enhanced long text, an unenhanced short text and an unenhanced long text. The results showed that reading short authentic materials had a positive effect on comprehension but not on the acquisition of forms. However, input enhancement was found to have no effect on comprehension or acquisition of the target structure.

The issues of textual enhancement and topic familiarity were addressed in the studies by Overstreet (1998) and Lee (2007). Both studies explored the relationship between topic familiarity and textual enhancement in terms of promoting learners' knowledge and use of target linguistic forms. In each study, four groups, with a combination of [+/-] enhancement and [+/-] topic familiarity, were included. In Overstreet's study, a larger font size and underlining were used to enhance target structures, which were the preterit and imperfect verbs; however, preterit verbs were also shadowed while imperfect verbs were bolded. The assessment comprised circle-the-verb pretest and posttest tasks, a written narration task and a comprehension quiz. Lee conducted a study with 259 Korean high-school students, focusing on learning the passive form in English. The enhanced target form was presented in bold. After reading a text according to their treatment condition, the participants were asked to complete a form-correction task and a free-recall task. Regarding the effect of textual enhancement on developing L2 grammatical knowledge, opposite results emerged; a positive effect of textual enhancement was observed in Lee's study, whereas no effect was reported in Overstreet's study. However, both studies reported a negative effect of textual enhancement on comprehension. The results could thus be interpreted as suggesting that learners may not be able to focus their attentional resources on both content and form at the same time, and thus textual enhancement directed their attention away from comprehension, often referred to as a 'trade-off effect' (Winke, 2013).

Wong (2003) investigated the effects of textual enhancement (TE) and simplified input (SI) on the comprehension and acquisition of a target linguistic form, i.e. French past participle agreement in relative clauses. Adult English speakers learning French (*n* = 81) participated in this study and were asked to read a text according to treatment conditions they were assigned to: +TE +SI, +TE -SI, -TE +SI and -TE -SI. An error correction task was used to measure acquisition of the target form and a free recall task was administered to assess comprehension. It was revealed that neither TE nor SI had an effect on acquisition of the target form.

As mentioned previously, an important issue that needs to be addressed in interpreting the results of previous studies concerns the types of textual enhancement used in each study. The variations in the typographical cues used to enhance target forms (see Table 3), sometimes combining two or more types, presumably contributed to the mixed results. To explore this issue further, two studies attempted to examine the effects of different types of textual enhancement on L2 grammatical knowledge development. In Simard (2009), the effects of eight different types of textual enhancement on English plural markers were examined: (a) italics only, (b) underlining only, (c) bold only, (d) highlighted only, (e) capitalized only, (f) five cues, including italics, underlining, bold, highlighting and capital letters, (g) three cues, including bold, capital and underlining and (h) unenhanced. Simard assessed the effects of different forms of textual enhancement on L2 intake using a multiple-choice recognition test. It was revealed that enhanced input using a combination of three cues and capital letters was found to produce the best results for L2 intake.

However, another type of textual enhancement was found to be more effective in a study conducted by LaBrozzi (2016). Seven different groups were included in the study, a control group with six experimental groups differing in the type of textual enhancement of the target structure, the Spanish present and preterit tense morphemes (i.e., underline, bold, italics, font size, capital letters and a different font). An L2 to L1 translation task was used to measure the participants' recognition of the L2 form. In addition, differing from Simard's study, a 40-item multiple choice test was also employed to assess the participants' reading comprehension. The results showed that changing the font size was more effective than other types of textual enhancement to promote L2 form recognition; however, no negative effects of textual enhancement on reading comprehension were reported, regardless of enhancement type. The results obtained from Simard's study and LaBrozzi's study indicate that the effectiveness of textual enhancement may vary depending on the types of typographical cues used to make target linguistic constructions visually salient to learners. Given the mixed findings that emerged from these previous studies, the types of typographical cues that are most effective to attract learners' attention are subject to debate.

Drawing on the mixed results reported in textual enhancement studies, Lee and Huang (2008) conducted a meta-analysis to address the relative effectiveness of textual enhancement as a pedagogical intervention for grammar instruction. A comprehensive review of 16 previous empirical studies that were designed to examine the effects of visual input enhancement (VIE) on L2 grammar learning through reading tasks was conducted. The analysis concluded that input enhancement had only a marginal impact on grammar learning with a small effect size (d = .22). However, some caution should be exercised before drawing any firm conclusions based on this result because of the divergence in methodologies used in previous studies (Han, Park, & Combs, 2008; Lee & Huang, 2008). Han et al. (2008)

summarized ten main differences across studies in a comprehensive review of this strand of research:

> (a) employing simple versus compound enhancement
> (b) employing isolated words versus sentences versus discourse as stimuli
> (c) enhancing a meaning-bearing versus a non-meaningful form
> (d) employing learners with or without prior knowledge of the target form
> (e) enhancing the target form many versus one or a few times
> (f) using a longer versus a shorter text
> (g) employing single versus multiple short sessions over an extended period of time
> (h) enhancing one form versus multiple forms
> (i) providing (or not) comprehension support prior to the treatment and
> (j) providing (or not) explicit instruction on what to focus on prior to the treatment (p. 600).

Besides studies investigating the effects of textual enhancement on reading texts, recently, Montero Perez et al. (2014) launched an examination into the effects of input enhancement on L2 vocabulary learning and L2 listening comprehension in relation to captioned video clips (for a detailed review, see captioning section in Montero Perez et al., 2014). The purpose of this study was to explore the effects of different types of captions, including no captions, full captions, keyword captions and full captions, with highlighted keywords on L2 learning. The results revealed that the enhanced condition, full captions with highlighted key words, yielded no significant advantage over the form recognition and clip association tests: all participants who watched video clips with captions, regardless of caption type, outperformed the control group. However, the group exposed to full captioning with highlighted keywords achieved greater gains than the control group on the meaning recognition test.

Although somewhat different findings were obtained in previous studies, which could be due to methodological discrepancies, the previous studies reviewed have provided considerable insights into the potential of textual enhancement as an instructional intervention for second language acquisition. However, one important issue that requires further discussion is an implicit assumption that underpinned these studies, which is that “enhancement would implicitly draw learners’ attention to these highlighted forms, which, in turn, should theoretically promote superior noticing and further processing of the attended forms when compared to unenhanced input” (Leow et al., 2003, p. 2). Without measuring whether learners actually pay attention to textually enhanced target linguistic constructions in the input, as assumed by proponents of the technique, the results of these studies do not yet necessarily provide support that increasing the perceptual salience of a target linguistic construction will attracts learners’ attention and, further, increased attention to the construction will lead to more learning (Han et al., 2008; Leow et al., 2003). Thus, an attempt has been made by some researchers in the field of instructed SLA to identify the extent to which enhancing the saliency of targeted constructions attracts learners’ attention to the construction and, further, its relationship to overall L2 development (e.g., Izumi, 2002; Izumi & Bigelow, 2000; Jourdenais, et al., 1995; Leow, 1997, 2001; Uggen, 2012).

2.5.3. Measurements of Attention

A number of empirical studies have been conducted to determine whether enhanced input assists language learners to notice certain linguistic forms and, further, facilitates L2 learning. In these studies, various types of techniques have been used to measure the effects of enhanced input on ‘noticing’ (e.g., Izumi, 2002; Izumi & Bigelow, 2000; Jourdenais, et al., 1995; Leow, 1997, 2001; Uggen, 2012). Broadly, there are offline and online measures of noticing (Robinson, 2003). Offline measures involve the collection of retrospective data using questionnaires (e.g., Robinson, 1997), diaries (Schmidt & Frota, 1986) or stimulated recall protocols (Mackey, 2006; Mackey, Philip, Egi, Fujii & Tatsumi, 2002), whereas online measures refer to concurrent methods, such as underlining (e.g., Izumi & Bigelow, 2000; Uggen, 2012), note-taking while reading (Izumi, 2002) or think-aloud verbal protocols (e.g., Alanen, 1995; Jourdenais, et al., 1995; Leow, 1997, 2000; Rosa & Leow, 2004). For instance, in Robinson's (1997) study, an offline measure was used to measure noticing. A primary goal of the study was to examine the extent to which ESL learners were able to acquire an L2 grammatical structure when they were trained under conditions with no focus on form (implicit and incidental conditions) and with a focus on form (enhanced and instructed conditions). In this study, a debriefing questionnaire was used after learners participated in training sessions in order to determine whether they noticed any rules during the sessions. In Mackey's (2006) study, although the focus of the study was not to examine the effectiveness of textual enhancement but to explore whether learners noticed interactional feedback (i.e., recasts) and whether it promoted L2 learning outcomes, a stimulated recall technique was included to obtain introspective data about learners' noticing of interactional feedback. These offline measures of noticing offer some advantages, in that they do not impose an additional task on learners; however, the main problems identified relate to the possibility of memory decay and the difficulty of verifying whether learners report what they were thinking at the time of recording or at the time of recall sessions (Gass & Mackey, 2016). Leow (2007) notes a further problem with introspective data in that they only allow a researcher to "make inferences as to whether learners either paid attention to or became aware of targeted forms or structures in the input" (p. 23).

Compared to offline measures, online measures are generally considered to be more valid as they produce concurrent data on noticing while learners are engaged with an L2 task in comparison to offline measures (Godfroid, Boers, & Housen, 2013; Winke, 2013). However, some researchers have argued that online measures should be used with some caution due to the reactivity problem, which refers to the possible influence of verbalization on cognitive processes (Godfroid, Housen, & Boers, 2010; Leow & Morgan-Short, 2004). In the study by Izumi (2002), notetaking was used as a measure of noticing. The main purpose of this study was to compare the effects of two different pedagogical approaches, namely, external (input enhancement) and internal (output) attention-drawing devices on the noticing and acquisition of a grammatical form by adult L2 learners. In this study, four different conditions, to which 61 adult ESL learners were randomly assigned, were included based on a combination of exposure to visual textual enhancement and/or the requirement of output (i.e., +O +TE, +O -TE, -O +TE, -O -TE). The participants were asked to read a given text and instructed to "take notes of any and every word" (Izumi, 2002, p. 552) in an attempt to identify whether learners noticed L2 target grammatical forms. The participants' knowledge of the target form (relative clauses) was measured with a sentence combination test, a picture-cued sentence test, an interpretation test and a grammaticality judgement test. The results derived from notetaking and performance on the tests showed that output production served as a priming device. In other words, output production was more effective than textual enhanced or control conditions. The data collected from notetaking revealed that textual enhancement was more effective at attracting learners' attention (i.e. noticing forms) than the control condition, but not in developing L2 grammatical knowledge. Izumi suggested that noticing might not be directly linked to learning.

Izumi and Bigelow (2000) used another type of online measure, underlining, as a measure of noticing. The main purpose of this study, in fact, was not to explore the effects of textual enhancement, but to examine to what extent output promoted noticing and second language acquisition. However, this study provides some insights, as underlining was used to measure the participants' noticing of the target form in written input passages; more specifically, the participants were instructed to "underline the word, words, or parts of words that you feel are particularly necessary for your subsequent production (or reconstruction)" (Izumi & Bigelow, 2000, p. 250). In a replication of Izumi and Bigelow (2002) conducted by Uggen (2012), underlining was also employed as a measure of noticing, in that all participants were asked to underline what they thought important for the following writing activity.

In other studies, noticing was gauged with verbal think-aloud protocols, which are online verbal reports of learners' thoughts while performing a task (e.g., Bowles, 2003; Hama & Leow, 2010; Jourdenais et al., 1995; Leow, 2000; Rosa & Leow, 2004). In an early study by Jourdenais et al. (1995), the effect of input enhancement on noticing target linguistic forms was examined using think-aloud protocols. In this study, 14 learners of Spanish were randomly assigned to two groups: (a) an enhanced group provided with a sample text with target forms (Spanish preterit and imperfect verbs) highlighted by means of underlining, bolding and changing the font and (b) a comparison group provided with the same text but without textual enhancement. Data were collected through think-aloud reports produced by the learners while they were engaged in a picture-based writing task. The results of the study showed that learners in the enhanced group provided significantly more instances of the target linguistic forms in their think-aloud reports, suggesting that input enhancement promoted learners to notice the L2 target form and had an effect on learners' subsequent output.

A think-aloud protocol, as a measure of noticing, was used in the study by Leow (2001), which examined whether enhanced input had an effect on noticing and the intake of target forms. A total of 38 college-level students learning Spanish were divided into two groups: an enhanced group (*n* = 21) and an unenhanced group (*n* = 17). The targeted linguistic form was the formal imperative/ command in Spanish. In this study, three different measures were used: (a) a multiple-choice recognition task for participants' intake of targeted forms, (b) a fill-in-the-blank task for written production and (c) a comprehension task for participants' comprehension of a reading text. To assess whether participants noticed the forms, they were asked to verbalize their thoughts throughout the entire experiment. On the basis of data gathered through think-aloud protocols and tasks, it was revealed that enhanced input did not lead learners to notice substantially more targeted forms and did not contribute to better comprehension.

In a subsequent study, Leow et al. (2003) used the same methodological approach to further examine the effects of textual enhancement on noticing and acquiring two different linguistic forms: a relatively more salient form (Spanish present perfect) and a relatively less salient form (Spanish present subjunctive). The participants were asked to think aloud while reading a text. An immediate recognition and a comprehension task were administered to assess the participants' intake of forms and overall comprehension of the text, respectively. The results showed no significant beneficial effects of enhanced input for noticing, acquisition of the form or comprehension. However, as for linguistic forms, more salient forms were more noticed than less salient forms by both the enhanced and unenhanced groups.

From a review of extant studies, one noticeable aspect that needs to be addressed is that 'noticing' seems to be used as an umbrella term to refer to attention or attention plus awareness (Godfroid et al., 2013). For instance, in Izumi's study (2002), 'noticing', which was measured by notetaking, appears to be defined as attentional processes on the basis of an interpretation of the finding made by Izumi that "if the minimum requirement of noticing, as defined by Schmidt (1990, 1995, 2001), is to pay attention to key grammatical elements in the input with greater than a threshold level of subjective awareness (i.e., reportable subsequent to the experience), then these data indicate that the basis requirement of noticing has been met, in Schmidt's sense of the term" (p. 568). In a follow-up study (Izumi & Bigelow, 2002), however, 'noticing', which was measured by underlining, appeared to be operationalized as attention plus awareness (Schmidt, 1990, 1993b, 2001) in consideration of Izumi and Bigelow's (2000) explanation, stating "[b]ecause underlining was assumed to involve at least a minimum level of awareness, we believe that it tapped noticing in Schmidt's (1994) sense and not detection as discussed by Tomlin and Villa (1994)" (p. 250).

Along with the discrepancies in theoretical perspectives viewing the construct of 'noticing' in relation to 'attention' and 'awareness' (Robinson, 1995; Schmidt, 1990, 1993b, 2001; Tomlin & Villa, 1994), additional complexities are associated with the diversity in the measures, such as underlining, notetaking or think-aloud protocols, that were intended to investigate the extent to which learners noticed target linguistic constructions. In fact, it has been argued that different types of noticing measures may tap into different cognitive processes. For instance, the production of think-aloud verbal reports, which was used as evidence of learners' noticing in previous studies, has been claimed to require a higher level of awareness (Godfroid et al., 2013) and represent a conscious process that learners are engaged in. In fact, some researchers (Leow, 1997, 2001; Rosa & Leow, 2004) who used verbal think-aloud protocols to measure noticing interpreted the results in relation to awareness (e.g., Godfroid et al., 2013; Hama & Leow, 2010; Leow, 1997, 2001; Rosa & Leow, 2004; Sachs & Polio, 2007, Winke, 2013).

Additional problem with using think-aloud protocols as a measure of noticing relates to the issue of reactivity as mentioned previously (Bowles & Leow, 2005). Given that participants are required to verbally report their thoughts while performing a main task, they may thus be engaged in dual-tasking. This can be an important issue in relation to the internal validity of a study as "participants' internal processes may differ from what they would have been had they not performed the verbalization" (Leow & Morgan-Short, 2004, p. 38). Among the various types of online measures of noticing used in the studies reviewed so far, a general consensus had been reached that only underlining offers a measure of noticing as attention (Godfroid et al., 2013). However, although underlining captures noticing as attention, one perceived shortcoming that has been noted concerns the difficulty of quantifying the attention allocated to target linguistic constructions, which may be important for further analyses to determine whether learners' attention directed to target linguistic constructions results in promoting L2 learning.

There is an ongoing discussion on the measurements to be used to gauge learners' noticing of target linguistic constructions, mainly prompted by different theoretical perspectives presented in defining the construct of 'noticing' with reference to attention and awareness. In addition, as mentioned previously, the empirical evidence accumulated from previous studies appears to be elusive regarding whether noticing was defined as attention, awareness or a combination of the two. However, given the general consensus over the critical role of attention in language learning (Schmidt, 2001) compared to controversy over awareness, measuring noticing as attention may be an initial step to understand the constructive nature of noticing.

An inquiry into whether learners pay attention to target linguistic constructions and the amount of attention allocated to it would allow further investigation of the relationship between learners' attention and second language acquisition. In this regard, the use of an alternative method, i.e. eye-tracking, has been introduced in the field of SLA as this can not only quantify attention but also resolve the reactivity issue (Godfroid et al., 2013). Considering the main advantage of eye-tracking methods, which can capture moment-to-moment changes in eye-gaze, eye-movement registration is expected to provide a more comprehensive picture of whether textual enhancement actually draws learners' attention to target linguistic constructions and results in L2 grammatical knowledge development. This is also expected to contribute to a better understanding of the construct of 'noticing' with reference to attention and awareness.

2.5.4. Summary

Textual enhancement, as a relatively implicit and proactive type of focus on form intervention, has been subject to extensive research in the field of SLA. However, somewhat mixed findings have emerged from previous studies, which could be explained by methodological variations across them (Han et al., 2008). Furthermore, different theoretical accounts' view of the construct of 'noticing' add some complications to operationalizing the measure of noticing, as attention, as awareness or a combination of the two, and, further, interpreting the results of textual enhancement studies. Given the need for learners' attention to be drawn to input for language acquisition to occur, it is thus recommended that further studies are conducted that employ a measure that can gauge the attention allocated to the perceptual salience of target linguistic constructions, and also examine whether increased attention contributes to learning gains.

Table 3. Summary of Research on Textual Enhancement

Study	Participants	Target Construction	Typographical cue	Measurement	Noticing measure	Results
Doughty (1991)	20 adult intermediate-level ESL learners with diverse L1 backgrounds	English relative clauses	Highlighted Capitalized	(a) Comprehension question (b) Free recall-task (c) Grammaticality judgement task (d) Sentence-combination task (e) Guided sentence-completion task (f) Oral task	None	Positive effect on acquisition
Shook (1994)	125 first and second-year college students of Spanish	Spanish Present Perfect / relative pronouns	Bold Uppercase letters	(a) Written-recognition tasks (b) Written production task (fill in the blanks)	None	Positive effects on recognition and production
Alanen (1995)	36 adult English speakers	Finnish locative suffixes / consonant alternation	Italics	(a) Sentence completion test (b) Comprehension test (c) Word translation test (d) Grammaticality judgment test (e) Rule statement	Think-aloud protocol	Some positive effects on recall and the use of target forms
Jourdenais et al. (1995)	10 second-semester college students of Spanish	Spanish preterit / imperfect	Bold Shadowed Underlined Enlarged letters Different font	(a) Picture-based writing task	Think-aloud protocol	Positive effect on noticing and the use of target linguistic features

Leow (1997)	84 second-semester college students of Spanish	Spanish impersonal and imperative forms	Bolded Underlined	(a) Multiple-choice form Recognition task (b) Short-answer comprehension task	None	No effect on acquisition
Overstreet (1998)	50 college-level students of beginning Spanish	Spanish preterit and imperfect	Shadowed Underlined Enlarged letters	(a) T/F comprehension test (b) Circle-the-verb task (c) Written narration task	None	No effect on intake but a negative effect on comprehension
White (1998)	86 sixth-grade ESL learners	English possessive determiners	Bold Underlined Italic Enlarged letters	(a) Passage-correction task (b) Multiple-choice test (c) Oral production description task	None	Partial effects on acquisition
Williams (1999)	16 learners of Italian	Italian possessive adjectives/ inflectional verb endings for subjects	Different color	(a) Verbatim memory task (b) Translation task	None	Positive effect on acquisition
Leow (2001)	38 first-semester college students of Spanish	Spanish formal imperative / command	Underlined bolding	(a) Multiple-choice recognition task (b) Fill-in-the-blank production task (c) Comprehension task (short-answer and multiple-choice tasks)	Think-aloud protocol	No effect on intake, noticing or comprehension
Izumi (2002)	61 adult ESL learners with diverse L1 backgrounds	English Relative clauses	Bolding Shadowing Different font	(a) Sentence-combination test Picture-cued sentence test (b) Interpretation test	Notetaking	Positive effect on noticing, no effect on learning

			Different font size	(c) Grammaticality judgement test		
Bowles (2003)	15 native English speaking college students	Spanish imperatives	Underlined (only morphemes bolded)	(a) Fill-in-the-blanks production task (b) Multiple-choice recognition task (c) Comprehension task	Think-aloud protocol	No effect on noticing or intake
Leow et al. (2003)	72 first-year college-level learners	Spanish present perfect and present subjunctive	Bold	(a) Multiple-choice recognition task (b) Multiple-choice comprehension task	Think-aloud protocol	No effect on intake or noticing
Wong (2003)	81 second-semester college students	French past-participle agreement in relative clauses	Increased font size Bolded Italicised Underlined	(a) Error identification and correction (b) Free-recall task	Debriefing questionnaire	No effect on acquisition
Lee (2007)	259 high-school students	Passive voice	Bold	(a) Form correction tasks (b) Free-recall task	None	Positive effect on acquisition but unfavourable effect on comprehension
Simard (2009)	188 French-speaking secondary one students enrolled in a regular English as a second language course	English plural markers	Italic Underline Bold Highlighted Capital letters 5 cues 3 cues	(a) Information-transfer task (b) Multiple-choice recognition test	None	Positive effect of 2 types of textual enhancement (3 cues and capital letters) on L2 intake

Park & Nassif (2014)	16 English-speaking graduate students learning Arabic as a foreign language	Arabic comparative and dual pronouns	Enlarged font Bold Underlined	(a) Recall task (b) Comprehension questions (c) Fill-in-the blank task (d) Free production task	None	No effect on noticing the target forms Negative effect on comprehension
Jahan & Kormos (2015)	97 first-year university students	Modal auxiliaries (will, be going to)	Bold	(a) Multiple-choice comprehension test (b) Fill-in-the-blanks task (c) Form recognition task (d) Metalinguistic task	Noticing questions	Positive effect on noticing No effect on acquisition or comprehension
LaBrozzi (2016)	125 third-semester college students learning Spanish as a second language	Spanish preterit tense of -er	Underline Italics Bold Capital Font size Font type	(a) Multiple-choice test (e) Translation task	None	Positive effect of one type of textual enhancement (font size) on form recognition No effect on comprehension

2.6. Eye-Tracking

Eye-tracking records the movement of participants' eye-gaze behaviour in real time while they are engaged in a certain visual task (Rayner, 1998, 2009). The assumption underlying eye-tracking is that the location, length and sequence of eye movements are a close reflection of attentional processes, and thereby where and when the eyes move can supply information about the nature, order and timing of cognitive operations while individuals interact with a visual stimulus (Just & Carpenter, 1976). This is premised on a principle called the "eye-mind link" (Reichle, Pollatsek, & Rayner, 2006, p. 4), which assumes that there is a close relationship between the eyes and the mind, in that overt attention is a reflection of covert attention (Reichle, Pollatsek, & Rayner, 2006). The use of eye-movement data is, thus, advocated by SLA researchers as they provide moment-to-moment information of cognitive processes (Rayner, 2009).

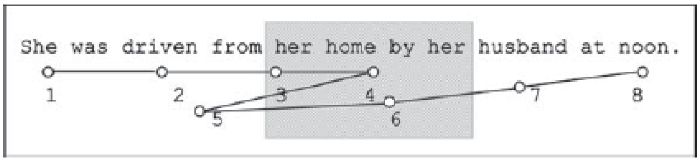
The components of eye movement data include *saccades* and *fixations*. *Saccades* refer to the movement of the eyes from one location to the next. Considering that the nature of reading is progressive, saccades are generally moving forward from left to right (e.g., English); however, in some cases, it could be from right to left depending on the language (e.g., Arabic). According to Rayner (2009), approximately 10–15 per cent of saccades show movements in the opposite direction, from right to left, which are referred to as *regressions. Fixations* occur when the eyes remain at a particular position to process a certain visual input (Rayner, 2009). In the literature on eye-tracking, fixation durations are often discussed in terms of early and late measures (e.g., Clifton, Staub, & Rayner, 2007; Rayner, 2009). Early measures (e.g., first pass reading) capture the initial processes that occur in the early stages of the comprehension of a text (Clifton et al., 2007;

Roberts & Siyanova-Chanturia, 2013), whereas late measures (e.g., second pass reading) reflect later processes in which learners engage, for example, in reanalysing or integrating information; therefore, the late measures may be indicative of problems encountered during initial processing.

Eye-tracking measures that are frequently used in L2 studies are summarized by Roberts and Siyanova-Chanturia (2013) (see Figure 4), and measures relevant to this line of research are described here. First fixation duration is an index of the duration of the first fixation within the area of interest, and first pass reading refers to the sum of all fixation durations during the first visit to the area of interest. Total reading time is the sum of all fixation durations. Regression path duration is "a measure of the time spent on the word itself and any prior parts of sentence before the reader moves past the critical word to the right" (Conklin, Pellicer-Sánchez, & Carrol, 2018, p. 66). Rereading is estimated by subtracting first pass reading time from regression path duration for the region. Second pass reading time is defined as the sum of fixation durations when the eyes return to an area of interest after the first visit. In other words, second pass reading time captures rereading in the area of interest, which is associated with re-analysis of the input. Fixation count is defined as the number of fixations a participant's gaze enters the area of interest (for a detailed explanation, see Roberts & Siyanova-Chanturia 2013). The number of visits includes all the fixations made within an area of interest from the time a participant's eyes first enter the arear of interest and until they leave.

A review of previous studies reveals that there seem to be some differences in types of eye-tracking measures that L2 researchers have employed to analyse eye-movement behaviours (e.g., Godforid et al, 2013; Godfroid & Uggen, 2013; Indrarathne & Kormos, 2017; Issa, Morgan-Short, Villeges, & Raney, 2015;

Pellicer-Sánchez, 2016; Smith, 2012); furthermore, there seem to be dissenting opinions on which measures best represent the cognitive processes involved (Rayner, 1998). However, the inclusion of several different measures is generally recommended to obtain a more valid picture of cognitive processes, particularly considering early and late measures may tap into different processes (Rayner, 1998; Roberts & Siyanova-Chanturia, 2013).



1. First Fixation Duration = 3[a]
2. First Pass Reading Time = 3 + 4[b]
3. Total Reading Time = 3 + 4 + 6
4. Regression Path Duration = 3 + 4 + 5 + 6
5. Rereading = 5 + 6
6. Second Pass Reading Time = 6
7. Fixation Count = 3 + 4 + 6 (3 fixations)

[a] The first fixation duration measure is used when the region of interest is a single word. This measure is included in Figure 1, where the area of interest is larger than the word, for illustrative purposes only.

[b] First pass reading time is known as gaze duration when the area of interest is a single word.

Figure 4. Eye-movement Measurements (Roberts & Siyanova-Chanturia, 2013, p. 218)

Although eye-tracking can provide no information about levels of awareness, it offers a number of advantages over more traditional techniques for capturing attentional processes as discussed previously. Therefore, a growing number of studies have begun to use eye-tracking to measure attention in the field of second language acquisition.

2.6.1. Empirical Studies Using Eye-Tracking in the Field of Second Language Acquisition

Eye-tracking has been used in a number of studies, mainly to explore (a) the cognitive validity of items included in language tests, (b) L2 learners' sentence processing and (c) the attention allocated to L2 constructions and its relationship to L2 development. For instance, in the field of L2 assessment research, learners' eye-movements have been analysed to examine the validity of test items (Bax & Weir, 2012) or to investigate the cognitive processes that test-takers employ while performing assessment tasks (Brunfaut & McCray, 2015). Some researchers have analysed test takers' eye-movements to understand their processing behaviour while performing a test. For instance, an investigation of the processing behaviours of test takers while engaged in the video-based academic listening test (VALT) was conducted by Suvorov (2014) using an eye-tracking technique. In a number of studies, eye-tracking technology has also been employed to examine L2 learners' sentence processing in the field of second language acquisition (e.g., Alptekin & Erçetin, 2015; Dussias, Valdés Kroff, Tamargo, & Gerfen, 2013; Jackson, Dussias, & Hristova, 2012; Kaushanskaya & Marian, 2007; Keating, 2009; Siyanova-Chanturia, Conklin, & Schmitt, 2011). Keating's eye-tracking experiment (2009) was conducted to determine whether L2 learners could acquire nativelike knowledge of gender agreement and whether L2 learners were sensitive, like native speakers, to violations of Spanish gender agreement during online sentence comprehension. In another study by Dussias et al. (2013), eye-tracking was used to examine whether grammatical gender facilitated noun recognition during L2 sentence processing.

The line of research that has used eye-tracking methodology to identify learners' caption-reading behaviour is more relevant to the present study. For instance, Winke, Gass, and Sydorenko (2013) examined caption-reading behaviour of L2 learners and whether the time spent on captions varied depending on the language (i.e., Arabic, Chinese, Russian and Spanish) being learned. The students' eye movements were tracked while watching videos, one with more familiar content and another with less familiar content. Two measures were used to analyse learners' eye-movements: (a) the sum of all fixation durations on captions divided by the total time the captions were shown on screen and (b) the average duration of fixations on captions. Caption-reading behaviour, assessed as the time spent on captions, was found to vary significantly by the target language learned. In a recent study, Muñoz (2017) also carried out an investigation of foreign language learners' behaviour when reading L1 (Spanish) and L2 (English) subtitles and the possible influence of age and proficiency on their reading behaviours using eye-tracking methodology. While watching two clips, one with L1 subtitles and the other with L2 subtitles, the participants' eye-movement data were collected and analysed using eight different measures: (a) percentage of Spanish subtitles skipped, (b) percentage of English subtitles skipped, (c) total fixation count on Spanish text, (d) total fixation count on English text, (e) average fixation duration on Spanish text, (f) average fixation duration on English text, (g) total fixation duration on Spanish text, and (h) total fixation duration on English text. The results indicated that two learner variables, age and proficiency, had some influence on the subtitle reading behaviour of language learners.

Another strand of research particularly relevant to the present research has used eye-tracking to investigate the extent to which learners pay attention to target linguistic constructions and, further, the extent to which the amount of attention allocated to constructions is related to L2 acquisition (e.g., Godforid et al, 2013;

Godfroid & Uggen, 2013; Indrarathne & Kormos, 2017; Issa, Morgan-Short, Villegas & Raney, 2015; Pellicer-Sánchez, 2016; Smith, 2012). However, the possibility of using different eye-movement measures, as discussed above, seems to add some complications to reviewing and interpreting the results elicited from previous studies. In addition, the operationalization of each of these eye-movement measures may differ by researcher. There also exist methodological variations across the studies in terms of the target linguistic constructions and measures used to assess learning gains. Thus, caution should be exercised in interpreting and drawing conclusions from the empirical studies reviewed below.

In one of the studies exploring the amount of attention allocated to target linguistic constructions and its relation to L2 learning, Godfroid and Uggen (2013) conducted an eye-movement study to examine whether beginning second language learners paid attention to irregular verb morphology. Learners' eye movements were collected while processing sentences in order to investigate whether learners paid attention to irregular verb features and whether the amount of attention paid to features had a relationship with acquisition. The learners were required to read 12 German sentence pairs containing stem-changing verbs and 12 German sentence pairs with regular verbs. Producing sentences using action verbs to describe given pictures was used as pre- and posttests to measure learners' development in using stem-changing verbs. Learners' eye movements were analysed using three fixation time measures, i.e., first fixation duration, gaze duration, and total time. The results indicated that there were significant differences in the total time spent on stem-changing verbs and on regular verbs; in other words, learners paid more attention to the irregular verb form (stem-changing verbs). In addition, an increased total time had a modest, favourable effect on the subsequent production of stem vowels.

Godfroid et al. (2013) used eye-tracking to examine the relationship between L2 learners' fixation duration on pseudowords in a text and subsequent vocabulary recognition. In the study, 28 EFL students' eye movements were recorded while reading 20 short paragraphs containing target words (i.e., pseudowords). Eye-movement data collected were analysed for first fixation duration, gaze duration, second pass time and total time. As posttreatment measures, a multiple-choice gap-filing exercise was administered. There was a significant relationship between total reading time and vocabulary recognition. The results thus indicated that increased total reading time was associated with better vocabulary recall, underscoring the positive effects of increased attention on L2 learning.

Another empirical study which employed eye-tracking to determine the relationship between incidental vocabulary acquisition and reading was conducted by Pellicer-Sánchez (2016). In her study, both offline measures (i.e., form recognition, meaning recall, and meaning recognition) and an online measure (i.e., eye-movement data) were employed. The participants' eye movements, which were recorded while reading a story including unknown items, were analysed using four different measures: first fixation duration, gaze duration, number of fixations, and total reading time. The results showed a significant relationship between total reading time and performance on recalling the meaning of words, with longer total reading times resulting in better performance on meaning recall.

Building on the premise that the amount of time spent on viewing linguistic input represents learners' cognitive engagement with the input and can serve as a measure of the amount of attention paid (Rayner, 2009), eye-tracking has also been used to identify the effectiveness of textual enhancement in drawing learners' attention. For instance, in a study carried out by Simard and Foucambert (2013), which examined the effectiveness of textual enhancement in noticing target language forms, eye-movements were used as an online measure of noticing, along with an offline measure by means of verbal reports. Simard and Foucambert recorded the participants' eye movements while they were reading a given text containing target language features (i.e., complex relative pronouns). Enhanced target forms were provided using a bolded font and underlining. The eye-movements were analysed using three different measures: a) total reading time, (b) first-fixation duration, and (c) regression duration. The results showed that textual enhancement successfully provoked more noticing, as indicated by all three eye-movement measurements; that is, participants were more likely to fixate longer on the enhanced target forms.

Winke (2013) conducted a study using an eye-tracking technique to investigate the role of textual enhancement in directing learners' attention to the English passive construction. Winke replicated Lee's (2007) study, which was designed to examine the effectiveness of input enhancement in drawing learners' attention to targeted linguistic forms and thus promoting L2 learning. However, in Lee's study, an eye-tracking methodology was not employed to measure the attention paid to targeted forms. In Winke's study, a total of 55 participants were asked to read two texts, one with enhanced (underlined and in red) passive forms and the other with normal passive forms. While the participants read the texts, data on their eye movements were collected using four measures: (a) total fixation time, (b) number of visits, (c) first pass reading time, and (d) rereading time. The participants' gains in understanding passive form constructions were measured using a pre-test and a post-test, and a free recall test was used to assess overall reading comprehension. The findings showed that textual enhancement for L2 reading successfully drew learners' attention to the target form as reflected in rereading time and total fixation duration, but failed to promote learning of the target form.

Similar results were obtained in subsequent studies by Issa, Morgan-Short, Villegas and Raney (2015) and Loewen and Inceoglu (2016). Issa et al. (2015) compared the effectiveness of external and internal manipulations of attentional conditions on directing learners' attention to the target form, Spanish direct-object pronouns. The participants were asked to complete the tasks according to the treatment conditions they were assigned to: a control condition, an external attentional condition (textual input enhancement), and an internal attentional condition (structured input practice). In this study, a different colour (red) was used to enhance the target forms. Learners' attention paid to the forms was examined using eye-tracking and learners' development in L2 grammatical knowledge was assessed with a sentence interpretation task. Two different eye-tracking measures were employed, fixation duration and skipping rate for pronouns. Participants who were exposed to textually enhanced input (external attentional condition) exhibited a decreased skipping rate for target forms, indicating positive effects of increasing the perceptual salience of target forms by drawing learners' attention. No significant relationship was observed between the measure of attention and learning gains. The results yielded from this study can be interpreted in line with the findings of Winke's study, since both studies showed textual enhancement increased the amount of attention paid to target forms. However, the eye-tracking measures that were used to represent the amount attention paid to target linguistic constructions were different; thus, the findings need to be interpreted with caution.

Another study conducted by Loewen and Inceoglu (2016) aimed to investigate the effectiveness of visual input enhancement in facilitating noticing and second language development using eye-tracking. A total of 30 college-level students, assigned to an enhanced group and an unenhanced group, were asked to read a given text according to their treatment conditions. The targeted forms, Spanish preterit and imperfect verbs, were highlighted using different colours. The participants' eye movements were recorded and analysed in terms of (a) the number of fixations for each targeted item, (b) the amount of time spent on each targeted item, (c) the duration of first fixation, and (d) the average total time of each of the targeted verbs; however, the researchers reported only the results of the overall average total times. Pretests and posttests, including a cloze test and an oral production test, were administered to measure the participants' knowledge of the targeted forms. The results of the studies showed no significant differences between the enhanced group and the unenhanced group in terms of the total reading time spent on the target forms and the development in their knowledge of L2 forms.

Indrarathne and Kormos (2017) investigated the extent to which L2 learners paid attention to a target syntactic construction, English causative had, in written L2 input in four different conditions: input flood, input enhancement, instruction to pay attention to the target construction and explicit metalinguistic explanation. In addition, the changes in learners' knowledge of the targeted construction and the relationship between the change in knowledge and attentional processing were also examined. Three short stories were presented to the participants to read in conditions to which they were assigned. Eye-tracking was used to collect data on the attentional processing of 45 participants. To assess their development in L2 grammatical knowledge, a sentence construction task and a grammaticality judgement task were used. Two eye-tracking measurements were used to gauge the amount of attention paid to the target items; (a) total fixation duration and (b) the difference between observed total fixation duration and expected total fixation duration on the areas of interest. In this study, textually enhanced input (i.e., highlighted in bold in a reading text) alone was not effective in drawing participants' attention to the target syntactic construction and in improving their L2 grammatical knowledge; textual enhancement combined with more explicit focus on form interventions (i.e., an instruction to pay attention to the target construction and explicit metalinguistic explanation) successfully increased the amount of attention paid to the target structure and improved related knowledge.

In the eye-tracking studies reviewed so far, the effectiveness of textual enhancement was investigated for drawing learners' attention to grammatical structures. However, there is one study, by Alsadoon and Heift (2015), which included English vowels as the target form. The purpose of the study was to examine whether textual enhancement assisted Arabic L2 learners of English to notice, decode and encode English vowels. The eye-gazes of 30 beginning English L2 learners were recorded during the treatment phase, which asked them to complete a reading task either with or without textual input enhancement, depending on the group they were assigned to. In the reading task, a sentence included target words with vowels presented on a screen; as for textual enhancement, target words were underlined and vowels were presented in bold and red. Learners' eye-movement data were collected using four measurements, including first fixation duration, refixation duration, rereading duration, and total duration. As a measure of participants' intake of the target forms, a multiple-choice recognition task was administered. Textual enhancement resulted in longer durations, as determined by all four measurements. In addition, a strong correlation between refixation duration and a reduction in error rates in word forms was observed for learners who were exposed to textual enhancement.

Choi (2017) conducted an experimental study to explore the effects of textual enhancement on allocating attention to and on learning of English collocations as well as on recalling of unenhanced textual information. Korean college students were recruited and asked to read a text according the treatment conditions they were assigned to, a text either with enhanced or unenhanced target collocations. Learning gains were measured with a collocation test, and a recall cloze test was used to gauge learners' recall of unenhanced textual information. The amount of attention paid to the target collocations was measured with eye-tracking methodology and analysed using two indices, total reading time and fixation count. The results revealed that visually enhanced linguistic constructions attracted more learner attention and led to increased collocational knowledge. However, textual enhancement impaired learners’ recall of unenhanced textual information. Choi interpreted this finding as suggesting that there were trade-off effects, that is, the extra cognitive resources allocated to the processing of visually enhanced target collocations might have drawn learners' attention away from the unenhanced textual information.

Closely related to the present research, besides the eye-tracking studies conducted in a reading context, there exists a research strand on captioned video using eye-tracking to measure learners’ attentional resources devoted to lexical items (see Montero Perez et al., 2015, Captioning section, for a detailed review). The primary purpose of the study was to investigate whether the type of captioning (full and keyword captioning) and text announcements (with or without announcements of an upcoming test) might facilitate vocabulary learning. In addition, a possible relationship between the amount of attention allocated to target constructions and vocabulary learning was identified. In this study, three eye-tracking indices to indicate the amount of attention directed to target words were used, including gaze duration (i.e., sum of fixations before leaving the target word area), which captured initial processing (Rayner, 1998); second pass reading time (i.e., time spent rereading the target word area), reflecting re-analysis of information; and total fixation duration. The analysis of eye-movement data revealed that the keyword captions group exhibited significantly longer gaze durations than the full captions group. The keyword-captions group also showed greater second pass reading times and total fixation duration under the intentional condition. However, interestingly, significant links between attentional allocation and learning gains were only observed for the full captions groups on the form recognition test. More specifically, for the full captions group, who were informed about the forthcoming vocabulary test, the total fixation time and second pass reading time led to better scores on the form recognition test. However, for the full captions group, being unaware of the test, longer gaze durations were associated with greater gains. In addition, in the absence of a test announcement, those who had longer second pass reading times displayed lower form recognition scores in the full captions group.

Overall, Montero Perez et al.'s (2015) study may suggest that increasing the visual salience of target areas in captions has the capacity to capture learners' attention and generate vocabulary gains. Notably, however, only when captions were not visually enhanced did the researchers found significant links between attention and L2 development in vocabulary knowledge. More specifically, in the results of General Estimating Equation (GEE) analysis for form recognition and total fixation duration, a positive b-value was observed for the group assigned to full captions plus intention condition. That is, the results indicated that the longer the learners fixated on target words, the higher scores they achieved on a form recognition test, particularly when learners were exposed to the full captions and made aware of a forthcoming test. Visual salience, operationalized as keyword captions, increased the amount of attention allocated to target words, but it did not lead to better form recognition. The results were interpreted as meaning that the amount of attention allocated might not necessarily result in elaborated processing. For the students in the keyword captions group, target words that were presented in isolation on the screen might have led them to fixate for longer on target words. However, the students might not have engaged in processing them in depth during fixation. Thus, an increased amount of attention directed to target words had no significant effect on vocabulary learning. The results of Montero Perez et al.'s (2015) study provide insights into the role of attention in L2 acquisition, yet scant research exists to draw a firm conclusion.

2.6.2. Summary

The studies by Winke (2013), Alsadoon and Heift (2015), Issa et al. (2015), Indrarathne and Kormos (2017) and Loewen and Inceoglu (2016) reviewed here used eye-tracking to explore the relationship between textual enhancement included in the input and learners' attention drawn to target forms. As discussed above, the somewhat contradictory findings from these studies can be explained by methodological variations across studies, including a variety of eye-movement measures, the types of assessment used to measure learning gains and the nature of target linguistic constructions. In addition, these studies were mainly reading-based research; that is, the focus of the studies remained on investigating the effects of textual enhancement on learners' attention to and learning of target L2 constructions while the learners were engaged in L2 reading. Although Winke, Gass, and Sydorenko (2013) and Montero Perez et al (2015) used eye-tracking to measure learners' attention directed to captions while learners were viewing captioned videos, the focus of their research was to explore either caption-reading behaviour or the effects of different types of captions on L2 vocabulary knowledge development, without including input enhancement as an independent variable. Furthermore, to the best of my knowledge, no research exists that used eye-movement data to investigate the effects of enhanced input in captions on L2 grammar learning. Considering the proposal that input enhancement in multimedia listening environments can be beneficial to achieving better comprehension and possibly promoting L2 learning (Chapelle, 2003), examining whether a combination of captioning and textual enhancement might influence development in the use of L2 grammatical knowledge seems to be a worthwhile research endeavour.

Table 4. Summary of Eye-tracking Studies

Study	Purpose of the Study	Participants	Outcome Measures	Eye-movement Measures	Results
Godfroid and Uggen (2013)	To examine to what extent beginning second language learners pay attention to irregular verb morphology	40 beginning learners of German	(a) Picture-cued production test	(a) first fixation duration, (b) gaze duration and (c) total time	Increased total time had a modest, favourable effect on the subsequent production of stem vowels.
Godfroid et al. (2013)	To examine whether more attention allocated to novel language elements led to more learning	28 EFL students	(a) Gap filling exercise	(a) first fixation duration, (b) gaze duration, (c) second pass time and (d) total time	A significant relationship between total time and vocabulary recognition.
Simard & Foucambert (2013)	To examine the effect of textual enhancement on noticing, taking into account learners' attentional capacity, reading skills and individual sensitive to TE	20 university students	none	(a) total reading time (b) first-fixation duration and (c) regression duration	Positive effect of textual enhancement on inducing noticing as indicated by all three measures.
Winke (2013)	To investigate the role of textual enhancement in directing learners' attention to target forms (passive constructions)	55 university students learning English as a second language	(a) A form-correction task (b) A free recall for overall reading comprehension	(a) total fixation time, (b) number of visits, (c) first pass reading time and (c) rereading time.	Positive effect of textual enhancement on drawing learners' attention to the target form, as reflected in rereading time and total fixation duration, but no effect on learning of the target form.
Winke, Gass, & Sydorenko (2013)	To investigate caption-reading behaviour in a foreign language	33 second-year, college-level, foreign language learners • Arabic (N=7)	(a) A multiple-choice comprehension test	(a) The sum of all fixation durations on captions divided by the total time the captions were shown on the screen	Foreign language learners fixated on the captions area 68% of the time that the captions were on screen.

		• Chinese (N=7) • Russian (N=8) • Spanish (N=10)		(b) The average duration of fixations on captions	Arabic spent significantly higher percentage of time reading captions than did learners of Spanish and Russian. Chinese language learners' use of captions depended on their familiarity with the content of the video, while for others it did not.
Alsadoon & Heift (2015)	To explore the impact of textual input enhancement on the noticing and intake of English vowels	30 ELS learners	(a) A multiple-choice recognition task	(a) First fixation duration (b) Refixation duration (c) Rereading duration and (d) Total duration	A significant relationship between (a) textual input enhancement and longer eye fixations on the target words and their vowel(s) as indicated by all four measures
Issa et al. (2015)	To examine the effectiveness of external and internal manipulations of attention on directing learners' attention to a target form in the input	43 novice L2 learners of Spanish	(a) A sentence interpretation task	(a) Total time and (b) Skipping rate on pronouns	Positive effect of textual enhancement on directing attention to the target forms (a decrease in skipping rate on target forms). No relation between measures of attention and learning gains.
Montero-Perez et al. (2015)	To investigate the effects of type of captions and test announcements on	51 university students learning French	(a) form recognition (b) meaning	(a) gaze duration, (b) second pass reading time and (c) total fixation duration.	Keyword captions group exhibited significantly longer

	promoting vocabulary learning and allocation of attention		recognition (c) meaning recall (d) clip association		gaze durations than the full captions group. The keyword-captions group also showed greater second pass reading times and total fixation duration under the intentional condition. Significant links between attentional allocation and learning gains on the form recognition test for full captions group under the intentional condition.
Indrarathne and Kormos (2017)	To investigate how language learners' attention processing of a target syntactic construction in written L2 input in four different conditions, the change in learners' knowledge of the targeted construction and the relationship between a change in knowledge and attentional processing	100 English learners	(a) sentence-construction task (b) grammaticality judgement task	(a) total fixation duration and (b) the difference between observed total fixation duration and expected total fixation duration on the areas of interest.	No significant effect textual enhancement on drawing participants' attention to the target syntactic construction and improving L2 grammatical knowledge.
Loewen and Inceoglu (2016)	To examine the effectiveness of visual input enhancement on facilitating noticing forms and second	30 college-level students enrolled in second semes-ter Spanish	(a) cloze test (b) oral production test	(a) the number of fixations for each targeted item, (b) the amount of time spent on each targeted item, (c) the	No significant difference between the enhanced group and the unenhanced group in terms of both the amount of

	language development	courses		duration of the first fixation and (d) average total time	attention paid to the target form (i.e., average total time)and development in their knowledge of L2 forms
Pellicer-Sánchez (2016)	To examine incidental acquisition from reading and online reading of unknown lexical items	37 L2 speakers of English	(a) offline measures (i.e. form recognition, meaning recall and meaning recognition) (b) online measure (i.e., eye-movement data	(a) first fixation duration, (b) gaze duration, (c) number of fixations and (d) total reading time.	A positive relationship between new vocabulary learning outcomes and online reading with longer reading times (i.e., total reading time) associated with higher vocabulary recall test score.
Muñoz (2017)	To investigate foreign language learners' behaviour in reading L1 and L2 subtitles and possible influence of age and proficiency on their reading behaviour	40 Spanish-Catalan learners of English	(a) comprehension questions about the content of a clip	(a) percentage of Spanish subtitles skipped, (b) percentage of English subtitles skipped, (c) total fixation count on Spanish text, (d) total fixation count on English text, (e) average fixation duration on Spanish text, (f) average fixation duration on English text, (g) total fixation duration on Spanish text and (h) total fixation duration on English text.	Effect of age and proficiency, on the subtitle reading behaviour of language learners.

2.7. Working Memory

Working memory generally refers to an individual's cognitive ability to temporarily store and manipulate information to carry out a wide range of tasks (Baddeley, 1992; Juffs & Harrington, 2011). The role of working memory has received considerable attention from researchers in the field of SLA, and individual differences in working memory capacity have been demonstrated to account for variable outcomes in developing various types of L2 knowledge and skills.

2.7.1. Baddeley's Working Memory Model

Of the different working memory models suggested by researchers, Baddeley and Hitch's (1974) seminal work on working memory has been widely referred to in second language acquisition research. Working memory, conceptualized as "an integrated system for temporarily storing and manipulating information" (Baddeley, 2003, p. 837), is differentiated from what is traditionally called short-term memory, which is mainly responsible for passive storage of information for a short period (Baddeley, 2012; Harrington, 1992). Baddeley (1986, 2003) explains working memory as a limited-capacity storage and manipulating information system – usually auditory, visual or spatial – that is necessary for carrying out a range of tasks. The multicomponent model of working memory was proposed to explain the constructs of working memory (Baddeley & Hitch, 1974). In the original model, three different components were included: a central executive, a phonological loop and a visuo-spatial sketchpad (see Figure 5).

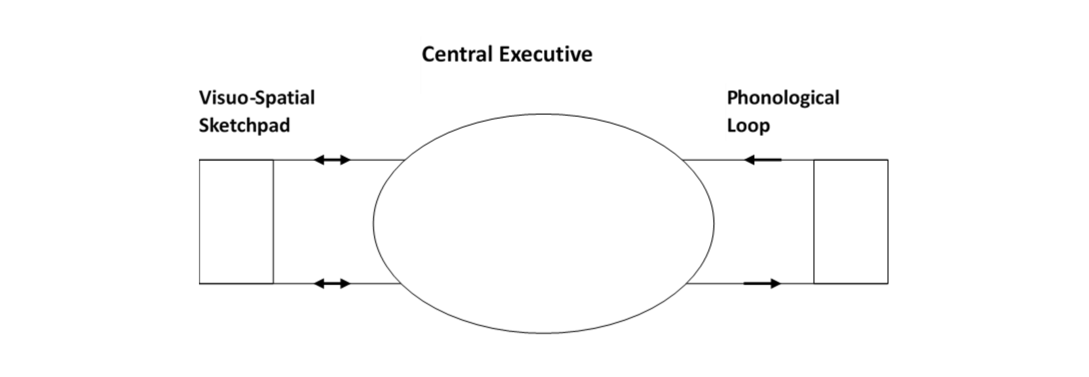


Figure 5. Original Baddeley and Hitch’s (1974) Working Memory Model (Baddeley, 2012, p. 6)

The central executive is the main task control centre responsible for directing attentional processes and allocating cognitive resources. Baddeley (1986, 1996) specifies the four major component functions of the central executive: (1) capacity to operate and coordinate two different tasks; (2) capacity to switch and allocate attention to two tasks; (3) capacity to focus on one task; and (4) capacity to store and process information in long-term memory. Along with the central executive, there are two slave systems included in working memory: the phonological loop and the visuo-spatial sketchpad.

The phonological loop, comparable to verbal short-term memory, stores and rehearses auditory information temporarily. The phonological loop is further divided into two sub-processes: auditory information storage and articulatory rehearsal process. In general, the auditory information store holds verbal information for approximately one to two seconds in the phonological loop, whereas the articulatory rehearsal process is related to the function of refreshing verbal input on command (Gathercole & Baddeley, 1993). The other slave system is called the visuo-spatial sketchpad; it retains and processes visual images and spatial relations. In addition to the three main components, a new sub-component, called the episodic buffer, was added to a revised model of working memory proposed by Baddeley (2000), as shown in Figure 6.

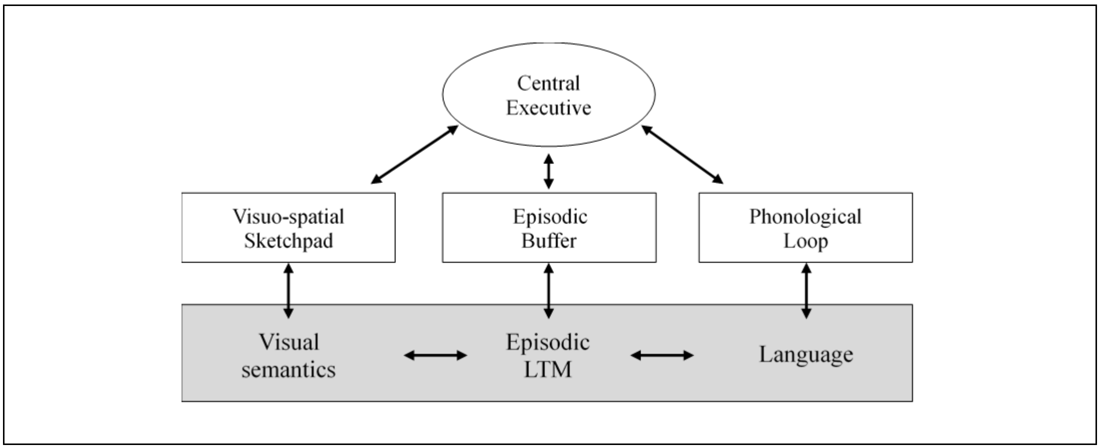


Figure 6. Current Multicomponent Model of Working Memory (Baddeley, 2003, p. 203)

The role of the episodic buffer is related to long-term memory, with the main function being to combine auditory and visual codes and integrate information from a variety of systems, as well as from long-term memory (Baddeley, 2003).

2.7.2. Unity and Diversity Framework of the Central Executive

The central executive system is the main component of working memory, which is responsible for controlling and regulating cognitive processes (Miyake et al., 2000), such as allocating attentional resources, switching between tasks, inhibiting processing routines or regulating subsidiary memory systems. Among these processes, three frequently referenced executive functions are switching (or shifting), inhibition and updating (e.g., Baddeley, 1996; Logan, 1985; Miyake et al., 2000; Smith & Jonides, 1999). Switching or shifting refers to the ability to switch or shift flexibly between tasks. Inhibition concerns the ability to deliberately inhibit responses when it is required, whereas updating is the ability to monitor, revise, and update incoming information constantly (Miyake et al., 2000). An integrated framework for working memory in SLA (Wen, 2012), which attempts to conceptualize the construct of working memory in SLA research, describes these subprocesses of the central executive (see Figure 7).

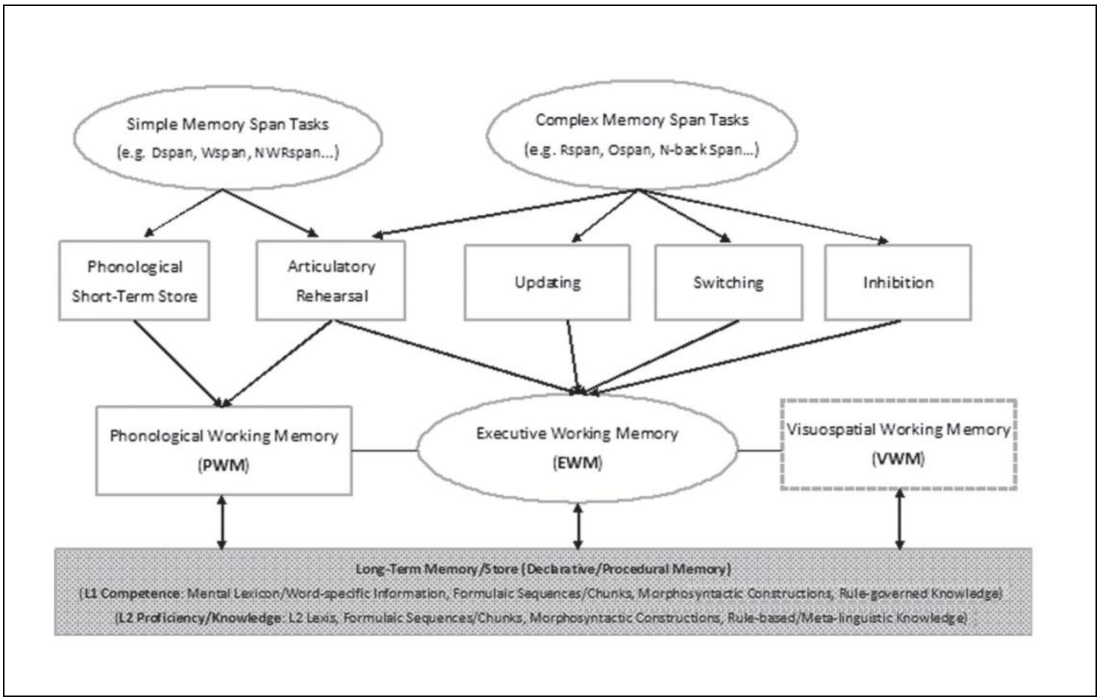


Figure 7. An Integrated Framework of Working Memory in SLA (Wen, 2015, p. 51)

The central executive's functions, however, seem traditionally to be conceptualized as a single cognitive process (e.g., Duncan, Emslie, Williams, Johnson, & Freer, 1996; Engle, Kane, & Tuholski, 1999). Against this view, an argument has been made over whether the functions of the central executive are unitary or separable (Miyake et al., 2000). That is, a question is raised as to whether three subprocesses share the same construct, which can be measured using a single working memory measure, or if the three subprocesses are separable, in that a particular measure should be used for each function of the central executive. A seminal study conducted by Miyake et al. (2000) explored the extent to which the three executive functions are unitary or separable using confirmatory factor analysis (CFA). Miyake et al. (2000) used nine tasks that are presumed to tap into one target executive function: (a) plus-minus task (Jersild, 1927), number-letter task (Rogers & Monsell, 1995) and local-global task for shifting function; (b) keep track (Yntema, 1963), letter-memory task (Morris & Jones, 1990) and tone-monitoring task for updating function; and (c) stroop task (Stroop, 1935), anti-saccade task (Hallett, 1978) and stop-signal task (Logan, 1994) for the inhibition function. These individual tasks are specified as manifest variables for each latent variable, switching, updating and inhibition, in the models of CFA seeking to determine the separability of the three postulated functions. The results of CFA indicated both unity and diversity of executive functions. That is, it is claimed that the three functions, shifting, inhibition and updating, are separable; however, they also share some common underlying construct on account of moderate correlations among the functions.

To date, a number of studies, which have attempted to examine the extent to which individual differences in working memory impact second language learning, operationalized the central executive function as a single process, premised on the unity of executive functions. Accordingly, these studies have generally used one type of task, such as a reading span task (Daneman & Carpenter, 1980), an operation span task (Turner & Engle, 1989) or a listening span task (Martin & Ellis, 2012), to measure an individual's complex working memory, and these have been presumed to provide information about the overall functions of the central executive. Research, adopting the notion that executive functions are separable, on the potential effects of each function (i.e., switching, updating and inhibition) is not yet a fully developed area of inquiry. In the field of instructed SLA, therefore, insufficient attention has been paid to the possibility of the contribution of each function of the central executive to second language acquisition. Considering the important role of the central executive in second language learning, particularly supported by Baddeley (2012), who states that "[o]ne component of WM that one might expect to influence second language learning is the central executive" (p. 25), it appears worthwhile to identify the extent which each distinctive function of the central executive affects the second language learning using suggested measures that attempt to tap into each function.

2.7.3. Working Memory Measures

To measure the working memory capacity of individuals, a wide variety of working memory span measures have been developed (e.g., Daneman & Carpenter, 1980; Turner & Engle, 1989). Based on the scope of these measures, working memory span measures can be classified into two broad categories, simple span tasks and complex span tasks, as shown in Figure 8 (Linck, Osthus, & Koeth, 2014).

	Verbal	Nonverbal
Simple	Word span	Digit span
	Nonword span	Counting span
	Letter span	Backward digit span
		Letter rotation
		Size judgment
Complex	Reading span	Operation span
	Listening span	Math span
	Speaking span	*N*-back
	English opposites span	AMIPB

Figure 8. Classification System for Coding Working Memory Span Tasks (Linck et al., 2014, p. 866)

Simple span tasks are likely to tap into simple storage functions; they measure the capacity of the phonological loop or the visuo-sketchpad. That is, simple span tasks are designed to assess how much information can be stored for a short period of time. For instance, a nonword repetition test is a type of simple span task that is supposed to measure phonological loop capacity. According to Gathercole and Baddeley (1993), the nonword repetition test requires participants to recall nonsense words presented to them. Nonwords ‘coined’ for the task are recommended to adhere to the phonotactic rules of the language (Gathercole, 2006). Digit span is a nonverbal type of simple span task. The digit span task, for example, developed by Yuill, Oakhill and Parkin (1989), does not involve any linguistic processing ability; lists of digits that incorporate groups of three digits (e.g., 835–402) are presented to participants. They are asked to read the groups of digits aloud and remember the final digit of each group (e.g., 5 and 2). Examples of tasks designed to measure the capacity of the visuo-sketchpad include the Corsi Block Task and the Visual Pattern Test (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999). The Corsi Block Task is intended to assess the spatial component of the visuo-spatial sketchpad by requiring participants to remember spatial sequences. In the Visual Pattern Test, visual components are measured through recalling visual patterns presented to participants.

Complex span tasks, on the other hand, are designed to measure the central executive’s capacity. In these tasks, participants are required to store information while performing some kind of cognitive activity. A reading span task (Daneman & Carpenter, 1980) and an operation span task (Turner & Engle, 1989) are the most commonly used complex span tasks to measure the capacity of the central executive. A sentence-based reading span task, developed by Daneman and Carpenter (1980), generally consists of unrelated simple sentences in the active voice, each sentence ending with a different noun. The number of sentences in the sets gradually increases and this is used as an individual’s working memory span. As participants view the sentences, they are asked to read the sentences presented aloud and recall the final word of each sentence. The reading span task is widely considered to be a well-established and influential task for assessing working memory capacity (Wen, 2012). However, some researchers have expressed doubts about using reading span task in consideration of the fact that it demands a certain degree of linguistic competence on the part of participants. The procedure of the reading span task, incorporating reading unrelated simple sentences and recalling final words, has inspired researchers to validate the use of the reading span task as a measure of working memory capacity. As an alternative, non-linguistic working memory span measures, such as an operation span task, have been proposed (Turner & Engle, 1989). In the operation span task, generally, a simple arithmetic equation (e.g., 3 + 4 = 7 or 2 + 6 = 1) is presented, followed by a letter (e.g., F, T, K, L, Q). In the task, participants are asked to make a judgement as to whether a given answer to the arithmetic equation is correct or incorrect and to remember the letter. At recall, participants need to choose the letters in the order they appeared. The number of correctly recalled letters is defined as the set size for each participant. While performing the task, participants are instructed to make judgements as accurately and quickly as possible for the maths operations, while remembering the letters to be recalled.

Given the theoretical assumption of the separability of central executive functions, i.e., switching, inhibition and updating, diverse measures have been suggested to attempt to tap into each function of the central executive. The tasks suggested to measure the updating function include the number-letter task (Rogers & Monsell, 1995), the plus-minus task (Jersild, 1927) or the colour-shape task (Miyake, Emerson, Padilla & Ahn, 2004). In general, these types of tasks require the participant to shift between two tasks according to the instructions given. For instance, the colour-shape task asks participants to identify either the colour (e.g., red or green) or shape (e.g., circle or triangle) of a stimulus, following an instruction given. Regarding the inhibition function, the stop signal task (Logan, 1994; Verbruggen, Logan, & Stevens, 2008), the antisaccade task (Hallett, 1978) and the stroop task (Stroop, 1935) can be used. To tap into the inhibition ability, participants are required to deliberately withhold a response while performing a task. The updating function can be measured with the letter memory task (Morris & Jones, 1990) or the automated operation span task (Turner & Engle, 1989). In these tasks, participants need to constantly monitor and update the information in working memory.

2.7.4. Working Memory and Second Language Acquisition

Motivated by the premise that working memory is an important cognitive resource accounting for individual differences in language learning, working memory has been a subject of a number of studies focusing on: reading (e.g., Friedman & Miyake, 2004; Harrington & Sawyer, 1992; Leeser, 2007; Linderholm & van den Broek, 2002; Walter, 2004; Waters & Caplan, 1996), sentence processing (e.g., Felser & Roberts, 2007; Juffs, 2004), writing (e.g., Adams & Guillot, 2008), speaking (e.g., Fortkamp, 1998; O'Brien, Segalowitz, Feed, & Collentine, 2007), vocabulary development (e.g., Cheung, 1996) and grammatical knowledge (e.g., Martin & Ellis, 2012; Williams & Lovatt, 2003). In these empirical studies, among the four components of working memory, the functions of the phonological loop and the central executive are most frequently investigated. The phonological loop is generally operationalized as a simple span task, such as a nonword span task and a digit span task, whereas the central executive is examined via a complex span task including a reading span task and an operation span task. Considering the notion that each component has distinctive functions, it seems important to acknowledge how working memory is defined in each study for the interpretation of results.

In the field of SLA, extensive research has focused on the relationship between individual difference in working memory and L2 reading comprehension (e.g., Alptekin & Erçetin, 2009; Friedman & Miyake, 2004; Harrington & Sawyer, 1992; Leeser, 2007; Linderholm & van den Broek, 2002; Walter, 2004; Waters & Caplan, 1996). In an early study, for instance, Harrington and Sawyer (1992) explored the possible role played by working memory capacity in the L2 reading comprehension of Japanese college students who were advanced learners of English. The participants' working memory span was measured using a reading span task in both L1 and L2 for complex working memory, along with a digit span task and a word span task for phonological short-term memory. As a measure of L2 reading comprehension ability, TOEFL grammar and reading sections and a cloze passage were used. In a correlation analysis, none of the measures of phonological short-term memory had a significant correlation with the reading comprehension test. However, it was discovered that learners who had higher scores on the L2 reading span task performed better on the TOEFL grammar and reading sections. Complex working memory, although different types of measures might have been used, was found to have a strong relationship with L2 reading comprehension ability in follow-up studies. Leeser (2007), for instance, used a computerized version of Waters and Caplan's (1996) reading span task to investigate the role played by working memory capacity in L2 reading comprehension. With a total of 94 adult learners of Spanish, the effects of topic familiarity and working memory on L2 reading comprehension, and their processing of future tense morphology were examined. The participants' working memory span was measured using their L1, and their reading comprehension ability was assessed through a recall protocol task after reading the passages provided. The findings from the study suggested that participants with higher working memory capacities generally performed better in comprehending the text, though this result was only applicable when participants were familiar with the passage topic.

Similar results were obtained in Alptekin and Erçetin's (2009) study, which reported a significant relationship between complex working memory and L2 reading comprehension. The researchers investigated the role of working memory span in reading comprehension, specifically literal and inferential comprehension. The participants' working memory capacity was assessed in their L2 using a modified version of Daneman and Carpenter's (1980) reading span task. In the task, participants were asked to read sentences aloud and make grammaticality judgements as a processing task. In addition, in order to measure storage capacity, two different tasks were employed, namely a recall task and a recognition task. The results of the study revealed a significant role played by working memory in inferential comprehension, on condition that storage was measured through a recall-based procedure.

Besides studies examining the relationship between working memory and L2 reading comprehension, there exist studies which empirically have attested to the role of working memory in different areas of L2 learning. Kormos and Sáfár (2008) investigated whether phonological short-term memory and complex working memory influenced L2 listening ability. In this study, a non-word span test and a backward digit span test were used to measure individuals' working memory. In line with the findings of Harrington and Sawyer, phonological short-term memory did not significantly correlate to L2 listening ability; however, complex working memory, when it was measured with a backward digit span test, had a strong correlation with L2 listening ability.

In a study by Martin and Ellis (2012), the extent to which phonological short-term memory and complex working memory were related to vocabulary and grammar learning was investigated. In this study, phonological short-term memory was measured using nonword repetition and nonword recognition, and complex working memory was identified with a listening span task. Fifty monolingual university students learned vocabulary and grammatical aspects of an artificial language, consisting of nouns, verbs, adjectives and prepositions. Participants' comprehension and production knowledge were tested. The results showed that phonological short-term memory had a strong correlation with both vocabulary comprehension and production; however, complex working memory only correlated with vocabulary production. Both phonological short-term memory and complex working memory were found to correlate positively with L2 grammar learning.

The possible effects of working memory on L2 oral fluency were also examined by Mota (2003) and Gilabert and Muñoz (2010). In the study by Mota (2003), a speaking span task was used to measure participants' complex working memory. The speaking span task was similar to a reading span task, differing in the modality. That is, the participants were asked to listen to a sentence, make a judgement on its grammaticality and remember the last word for recall. Speech elicited from a picture description task and a narrative task was analysed in terms of fluency, accuracy and structural complexity. Working memory was found to have significant but moderate correlations with fluency and complexity; in addition, working memory had a significant, but negative, correlation with accuracy. Gilabert and Muñoz (2010) examined the influence of working memory capacity on the overall proficiency and L2 oral fluency, structural and lexical complexity and accuracy, of 59 high intermediate/ advanced learners of English. Participants' working memory capacity was measured using an L1 reading span task and their overall proficiency was determined with an Oxford Placement Test (OPT). The participants' fluency, structural and lexical complexity and accuracy were assessed through a complex narrative task, namely, a film retelling task. It was found that working memory did not have a significant correlation with overall L2 proficiency, but it did have a strong relationship with fluency and complexity.

Another line of research has focused on investigating the relationship between learners' working memory capacity and their L2 sentence processing ability (e.g., Ardila, 2003; Juffs, 2004, 2005; Williams, 2011); however, these morphosyntactic studies have reported somewhat mixed results. In studies conducted by Juffs (2004, 2005), for example, learners' sentence processing ability was determined by asking them to read sentences presented to them and make judgements as to whether the sentences were correct or not. A reading span task was used as a measure of working memory (Daneman & Carpenter, 1980). The results revealed that there was no interactional effect between the sentence processing ability of Chinese, Japanese and Spanish L2 learners of English and working memory capacity. Thus, Juffs concluded that working memory capacity is not a determining factor in L2 learners' ability to process complex syntactic structures. Foote (2011) reported similar results, in that no relationship was found between advanced Spanish L2 learners' sensitivity to gender and the number of agreement violations as a function of the length of agreement dependency and L2 learners' working memory capacity, as measured by a reading span task, designed by Waters and Caplan (1996), in Spanish. In contrast, studies on the processing of agreement morphology suggest that working memory capacity can be good predictor of L2 learners' ability to process agreement dependencies. Sagarra (2007), for instance, examined the relationship between low-proficiency Spanish L2 learners' working memory capacity and their sensitivity to gender agreement violations. It was found that L2 learners with higher working memory scores processed sentences with gender agreement violations more accurately than L2 learners with lower working memory scores. Havik et al. (2009) also reported working memory effects on processing subject-object relative clause ambiguities by German L2 learners of English in their study. The results of the study suggested that working memory capacity influenced the use of agreement morphology for the processing of syntactic structure.

In assessing L2 learners' working memory in order to examine its role in second language learning, one important methodological issue addressed by some researchers concerns the validity of using a reading span task (or listening span task). The reading span task incorporates a series of sentences which learners are required to read. Thus, some researchers have raised a question as to whether working memory is language dependent or language independent; that is, a claim has made regarding the appropriateness of measuring working memory capacity of participants using their L2 on account of the possibility that participants' language proficiency can affect the outcome of working memory capacity measures (Rai, Loschky, Harris, Peck, & Cook, 2010; Van den Noort, Bosch, & Hugdahl, 2006). Whether empirical findings on working memory should be measured in participants' L1 or L2 is still open to debate (e.g., Alptekin & Erçetin, 2010; Osaka & Osaka, 1992; Osaka, Osaka, & Groner, 1993; Van den Noort et al., 2006). In an early study conducted by Osaka and Osaka (1992), remarkably high correlations were observed between scores on L1

(Japanese) and L2 (English) reading span tasks. They interpreted the outcomes as suggesting that working memory resources are, in large part, shared across languages, and working memory capacity, therefore, is independent of linguistic knowledge. A different point of view is presented by the findings of Van den Noort et al.'s (2006) study. In this study, the working memory capacity of participants who were multilinguals was measured in their L1 (Dutch), L2 (German) and L3 (Norwegian). Interestingly, the results of the study showed that working memory span interacted with the participants' language proficiency, in that significant differences were found in working memory span when measured in the participants' L1, L2, and L3. The participants had significantly higher scores on a working memory span task when measured in their L1 compared to in their L2 and L3.

More recently, in Alptekin and Erçetin's (2010) study, a positive relationship was evidenced between scores on L1 and L2 processing tasks as well as those on L1 and L2 storage tasks. The results were interpreted as meaning that the "cognitive resources underlying working memory capacity in the L1 are analogous to those in the L2" (p. 213). However, only L2 working memory span was found to have a significant relationship with L2 reading comprehension. Consequently, measuring working memory span in the participants' L2 has been recommended if the purpose of a study is to investigate the potential role of working memory span in L2 reading comprehension (Alptekin & Erçetin, 2010; Wen, 2012).

To resolve this methodological issue, an operation span task that does not incorporate a reading related task was suggested as an alternative measure to a reading span task. For instance, Rai et al. (2010) investigated the effects of working memory capacity on foreign language learners' inferential processing ability during reading comprehension. A computerized version of the operation span task was used to assess the working memory capacity of learners of Spanish. A reading comprehension test composed of short stories in L2 was accompanied by three different types of questions: non-inference, bridging inference and pragmatic inference questions. Rai et al. reported that the performance of participants with higher working memory capacity was better on all three levels of inferential complexity.

A study conducted by Yi and Luo (2013) focused on examining how working memory and L2 lexical knowledge affected L2 learners' argumentative writing. To measure 31 university students' working memory capacity, an operation span task was employed (Unsworth, Heitz, Schrock, & Engle, 2005). Their writing ability was assessed using two argumentative writing tasks. In addition, the participants' L2 lexical knowledge was determined by three different tests, i.e., a productive vocabulary knowledge test, a vocabulary size test, and a word association task. The results revealed that working memory capacity had a moderate correlation with syntactic complexity and fluency, while it had no effect on the accuracy and lexical complexity of written language production.

To examine the relationship between complex working memory and L2 learners' oral fluency, Fortkamp (1998) included an operation-word span test (Cantor & Engle, 1993; Engle, Cantor, & Carullo, 1992) along with a speaking span test (Daneman, 1991; Daneman & Green, 1986) to measure the participants' working memory capacity. The participants' oral fluency was assessed through a picture description test and a narrative task. The results showed that there was no statistically significant correlation between the participants' working memory capacity and L2 oral fluency.

The empirical studies reviewed above seem better able to support the theoretical assumption that the central executive functions in working memory are unitary. That is, in these studies, learners' working memory was determined by a type of complex working memory measures via, for example, a reading span task, a listening span task or an operation span task. Given the possibility that the functions of the central executive may be separable (Miyake et al., 2000), further investigation of whether each function may play a different role in second language acquisition seems warranted. This issue was addressed in a recent study conducted by Révész, Marije and Lee (2017). The main purpose of the study was to determine L2 learners' cognitive processes and writing behaviours while engaged in writing tasks. The researchers further explored whether learners' writing behaviours and text quality were influenced by individual differences in phonological short-term memory and executive control functions. In this study, five different working memory measures were used, which were suggested to tap into phonological short-term memory, visuospatial short-term memory and each function of the central executive, respectively. The results demonstrated that phonological short-term memory, visual-spatial short-term memory and the switching function of the central executive were related to some of the measures of text quality or writing behaviours.

However, to date, scant L2 research has empirically attested to the diversity of central executive functions (i.e., switching, inhibition and updating) and, further, their distinctive role in second language acquisition. Thus, further working memory studies based on the theoretical assumption that the central executive functions may be separable are expected to shed more light on the construct and role of working memory in second language acquisition.

2.7.5. Working Memory and Attention

Working memory is considered an important individual difference factor that is closely related to learners' control of attentional resources (Robinson, Mackey, Gass, & Schmidt, 2012). Given that one particular function of the central executive in working memory is to control the allocation of attentional resources (Baddeley, 1986, 1996), a link between attention and working memory is well established (Robinson, Mackey, Gass, & Schmidt, 2012). Such a relationship between attention and working memory, which is also of importance to conceptualizing 'noticing' in relation to the involvement of awareness, can be well explained with reference to Robinson's (2003) model of attention and memory. According to Robinson (2003), noticing "involves that subset of detected information that receives focal attention, enters short-term working memory, and is rehearsed" (p.655). That is, as depicted in Figure 9, detected information first enters short-term memory; and further, this detected information requires focal attention and needs to be rehearsed in working memory for it to be encoded in long-term memory (Robinson, 2003). Rehearsal processes, either maintenance rehearsal or elaborate rehearsal, provoke noticing with higher levels of awareness; as a consequence, information that is rehearsed is transferred to long-term memory.

Following this model, which conceives noticing as detection plus rehearsal in working memory, individual difference in working memory capacity is presumably an important determinant of how language learners control their attentional resources in processing the input. In fact, Wen and Skehan (2011) state that "a larger working memory will make noticing more likely to occur, which will greatly facilitate L2 learners' attention to focus on form (FonF) in the dominantly meaning-focused classrooms" (Wen & Skehan, 2011, p. 25).

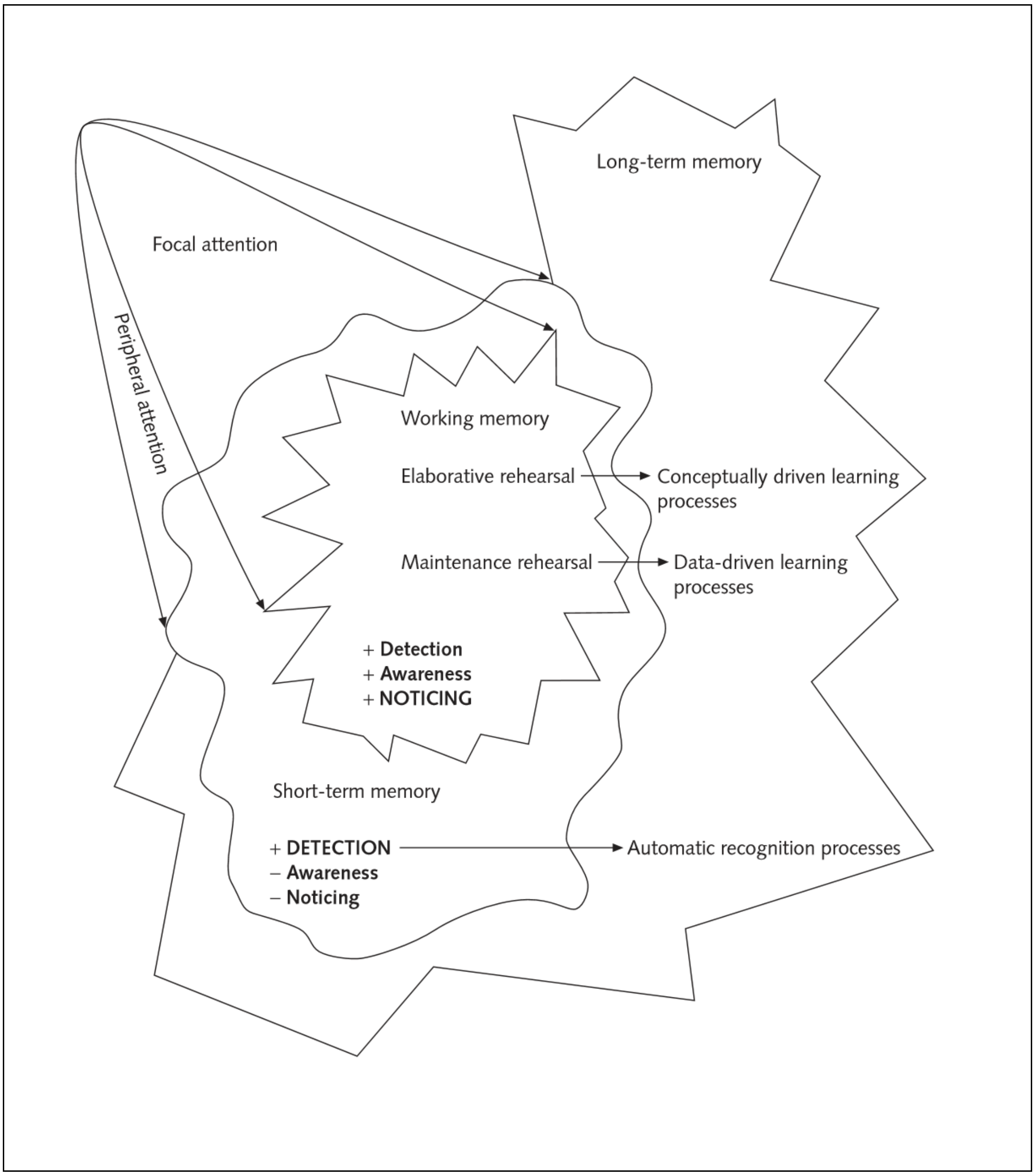


Figure 9. "Noticing" as Selective Focal Attention and Rehearsal in Working Memory: "Detection" as Recognition outside of Awareness in Passive Short-term Memory (Robinson, 2003, p. 655)

2.7.6. Working Memory and Focus on Form Pedagogical Interventions

The extent to which working memory mediates the effects of focus on form techniques on second language learning has been researched in SLA studies (e.g., Mackey et al., 2002; Révész, 2012). As mentioned earlier, the fundamental principle of focus on form is to draw learners' attention to linguistic elements while they are engaged in meaning or communicative focused tasks. Considering that the main function of the central executive is focusing, dividing and switching attentional resources (Baddeley, 1986, 1996), it seems reasonable to predict an important role played by working memory in individual learners' responses to implicit focus on form techniques. An investigation into how learners' working memory influenced the noticing of recasts and, further, the use of a target feature, *wh*-question forms, was conducted by Mackey et al. (2002). Second language development was observed using the developmental stages of English question formation suggested by Pienemann and Johnston (1987). Learners' noticing was determined using verbal responses to input. In this study, following Robinson's model, noticing was defined as "detection with awareness and rehearsal in short-term memory" (Robinson, 1995, p. 318). To assess learners' working memory, a listening span task and a non-word recall test were used as a measure of verbal working memory and phonological short-term memory, respectively. The researchers found that learners' phonological working memory capacity was positively related to their ability to notice recasts and, further, to development in the use of target forms.

Révész (2012) also examined the extent to which working memory mediated the effects of recasts, an implicit type of focus on form, determined by different types of outcome measures. In this study, 90 learners of English, who were randomly assigned to a recast, a nonrecast and a control group, were asked to complete contextualized treatment tasks. Three different outcome measures, including a grammaticality judgement test, a written description task and an oral description task, were administered as a pretest, an immediate posttest and a delayed test to determine development in the use of a target construction, the *past progressive* in English. Three different working memory measures were used: a digit span and a nonword span task for phonological short-term memory and a reading span task for complex working memory capacity. The results indicated that the working memory capacity of the participants who received recasts had significant correlation with developmental measures. Participants with complex verbal working memory were likely to show substantial improvement on a written test, whereas participants who had high scores on a phonological short-term memory test tended to exhibit greater growth on an oral test.

Individual difference in working memory was also found to be strongly related to the attention learners paid to a textually enhanced target grammatical construction in Indrarathne and Kormos's (2018) study. In this study, 80 Sri Lankan learners of English were exposed to the target construction in explicit (i.e., enhanced + instructions or enhanced + instructions + explanation) and implicit learning (i.e., unenhanced (input flood) or enhanced only) conditions. As for a test to determine grammaticality knowledge of the target construction, *causative had*, a receptive test (a grammaticality judgement test) and a productive test (sentence reconstruction) were administered. The learners' eye movements were recorded as they read the input text and their working memory was measured using four different working memory tests: a digit span task, a keep track task, a stroop task and a plus minus task. Indrarathne and Kormos calculated composite scores of working memory tests. It was found that there was a strong relationship between learners' working memory capacity and gains in knowledge of the target construction. In terms of receptive knowledge, working memory had a close association with gains in all conditions; however, they had an insignificant relationship with improvement in productive knowledge in the implicit learning conditions. Indrarathne and Kormos hypothesized that individual learners' working memory storage and processing abilities might not have played important roles in the productive use of grammatical knowledge that was acquired in implicit input conditions. In addition, individual differences in working memory had a strong relationship with the amount of attention paid to textual enhancement, indicating learners with "high WM storage and efficient attention regulating abilities engage in more attentional processing of the input if their awareness of the existence of a target syntactic construction in the input is experimentally manipulated" (p. 16).

2.7.7. Summary

In the field of SLA, considerable attention has been paid to the construct of working memory and, further, to the relationship between individual differences in working memory, attention allocation and L2 learning. Previous empirical studies have provided significant insights into the extent to which individual differences in working memory capacity predict abilities in various aspects of L2, such as reading comprehension (e.g., Harrington & Sawyer, 1992; Leeser, 2007; Walter, 2004), oral production (e.g., Gilabert & Muñoz, 2010; Mota, 2003), the acquisition of grammar and morphosyntactic processing (e.g., Juffs, 2004). More importantly for the present study, working memory has also been proposed to influence learners in responding to implicit focus on form techniques, such as recasts (e.g., Mackey et al., 2002; Révész, 2012).

Generally, however, working memory has been operationalized as either phonological short-term memory or complex working memory, measured by simple span tasks (e.g., nonword span task, digit span task) or complex span tasks (e.g., reading span task, operation span task), respectively. In particular, with respect to complex working memory, the fundamental underlying assumption is the unity of cognitive functions of the central executive. Considering the theoretical account that argues for the diversity of central executive functions, one important issue that should be of interest to further studies concerns whether the functions of the central executive are separable and whether each function influences L2 acquisition to a different extent.

CHAPTER 3

STUDY 1

This chapter reports on Study 1, which compared the capacity of two types of captions, non-enhanced and textually enhanced, to foster attention to targeted language features and second language development. Against the theoretical background and previous empirical work, the novelty of the study lay in the fact that I focused on multi- rather than unimodal input-based tasks using captioned videos, and I examined the effects of different types of captioning on attention to and acquisition of grammatical rather than lexical features. With the help of an eye-tracking methodology, a further attempt was made to expand on existing research by assessing the extent to which attention paid to target grammatical constructions is related to L2 learning (e.g., Godfroid, Boers, & Housen, 2013) and to begin to explore whether this link may be influenced by the type of captioning. Importantly, unlike the majority of previous research on captioning which used video clips with dynamic images, the multimodal input-based tasks in this study included static visual input (i.e., pictures), along with corresponding aural input (i.e., audio recording) and textual input (i.e., captions). Study 1 was guided by the following research questions:

1. To what extent do non-enhanced captions versus textually enhanced captions draw learners' attention to target linguistic constructions?
2. To what extent do non-enhanced captions versus textually enhanced captions affect L2 development in the knowledge of target linguistic constructions, as measured by a written and an oral grammaticality judgement test (GJT)?
3. To what extent is attention to target linguistic constructions in captions related to L2 development? Is this relationship influenced by textual enhancement?

This chapter is composed of three main sections, methodology, results and discussion. In the methodology section, an explanation of the overall design, participants, target linguistic constructions, materials, data collection procedure and statistical analyses is given. The next section presents the results derived from data analyses for each research question. In the discussion section, the findings are discussed and interpreted in relation to previous literature and empirical studies.

3.1. Methodology

3.1.1. Overall Design

The study employed a pretest-posttest experimental design (see Figure 13). The participants were 48 English as a Foreign Language (EFL) students, who were randomly assigned to two comparison groups: a captions group (*n* = 24) and an enhanced captions group (*n* = 24). First, all participants completed a background questionnaire, the Oxford Placement Test (OPT), and a pretest. Then, the participants engaged in three treatment sessions, each involving the completion of nine multimodal input-based tasks. The format of the multimodal input-based tasks was the same for both groups. The groups, however, differed as to whether they completed tasks with regular captions or captions with textually enhanced input. While the participants worked on the treatment tasks, their eye movements were recorded using a Tobii X2-30 mobile eye-tracker. Finally, a posttest was administered. The pretest and the posttest each included a written and an oral grammaticality judgement task (GJT).

3.1.2. Research Ethics

Research ethics approval for Study 1 was received from the Research Ethics Committee at the UCL Institute of Education, University College London. At the beginning of the study, the participants were given a general introduction and provided with an information sheet explaining (a) the main purpose of the study, (b) overall procedures and length of the study, (c) the participants' right to withdraw from the study at any time, and (d) the measures taken to maintain privacy and confidentiality. Each participant's agreement to participate in the study was obtained by them signing a consent form (see Appendix A).

3.1.3. Participants

All 48 participants were Korean undergraduate students learning English as a foreign language in Seoul, South Korea. They were recruited from three different universities located in Seoul, South Korea. There were 31 female and 17 male students, with an age range of 19 to 25 years (M = 22.53, SD = 1.89). The participants' proficiency levels fell into the B1–B2 bands according to the Common European Framework for Reference, as determined by an Oxford Placement Test. The captions and enhanced captions groups achieved comparable scores on both the listening (captions group: M = 73.88, SD = 4.49; enhanced captions group: M = 74.21, SD = 5.07) and grammar (captions group: M = 64.08, SD = 6.91; enhanced captions group: M = 63.38, SD = 7.44) sections of the OPT.

3.1.4. Target Linguistic Constructions

3.1.4.1. Anaphoric References in English

Anaphoric references were chosen as the target linguistic construction for the present study. The term anaphoric reference, or anaphora, describes the "relation between two linguistic elements, in which the interpretation of one (called an anaphor) is in some way determined by the interpretation of the other (called an antecedent)" (Huang, 2005, p. 231). In English, there are six types of anaphora: (a) pronominal anaphora, (b) lexical noun phrase anaphora, (c) noun anaphora, (d) verb anaphora, (e) adverb anaphora, and (f) zero anaphora. Of the different types of anaphora, pronominal anaphora is known to be frequently used in English (Mitkov, 2002). Examples of pronominal anaphora include personal pronouns (*he, she*), possessive pronouns (*his, her*), reflexive pronouns (*himself, herself*), demonstrative pronouns (*this, that*) and relative pronouns (*who*).

a. John had to go to a meeting so he decided to have a shave. (personal pronoun)
b. John grabbed his old razor. (possessive pronoun)
c. John cut himself while shaving. (reflexive pronoun)
d. Dali, however, used photographic precision to transcribe the images of his dreams. This would become one of the constraints of his work. (demonstrative pronoun)
e. Dali, a Catalan who was addicted to fame and gold, painted a lot and talked a lot. (relative pronoun)

Among these categories of anaphoric references, the use of third-person pronouns (he, she and they) was selected as the target linguistic construction for the current research. The processing of pronominal anaphoric references requires integrating information from two sources, a referential antecedent (e.g., Mark) and a pronoun (e.g., he). The integration of information online from multiple sources has been shown to cause difficulty for many L2 users (e.g., Roberts, Gullberg, & Indefrey, 2008), thus the processing of pronominal anaphoric references is also likely to pose a challenge.

3.1.4.2. Anaphoric References in Korean

In Korean, there are different ways of expressing anaphora, including deletion (also referred to as zero anaphora), personal pronouns, repetition of the antecedent nominal, use of definite noun phrases, reflexive pronouns and demonstrative plus noun anaphora. In comparison with English, anaphora may also be signalled by personal pronouns in some cases in Korean.

a. *Minhaka* *on ta*. *Cey-ka* *wul-ko iss-ta*!
Minha come. She is crying.

"Minha is coming … she is crying!" (personal pronoun)

However, the use of personal pronouns is not as extensive as it is in English. Third-person pronouns are used less frequently in Korean than in English (An, 2008; Choi, 2007; Im, 1993; Kim, 1997). In many cases, deletion is considered to be the most frequent anaphora to indicate the antecedent.

b. *Minha* *haiking* *ka-ss-ta*. *Keki-se* *[ ]* *seonsangnimul* *manna ss ta*.
Minha hiking went There [she] teacher met

'Minha went hiking. There, [she] met his teacher." (deletion)

Along with deletion, the repetition of noun phrases or the use of demonstrative plus noun anaphora (e.g., *Ku yeoja-tul* (those women) *i-haksaeng* (this student)) are also commonly used in Korean. In addition, the use of third-person pronouns in Korean is a very recent phenomenon; third-person pronouns are used less frequently in Korean than in English (An, 2008). Given that the two languages have different anaphoric systems, processing pronominal anaphoric references is expected to be especially demanding for Korean learners of English.

3.1.5. Materials

3.1.5.1. Background Questionnaire

A background questionnaire was used to obtain basic information about the participants, such as their ages, majors, experience of living in English speaking countries and TOEIC scores (see Appendix F).

3.1.5.2. General English Proficiency Test

As previously mentioned, the OPT was used to assess the participants' general English proficiency. The OPT comprises a listening section and a grammar section, each containing 100 questions. During the test, the participants were asked to choose the correct answers from the choices given.

3.1.5.3. Treatment Task

The multimodal input-based tasks that participants completed during the three treatment sessions were developed using items included in Listening Part 1 of a series of practice Cambridge Preliminary English Tests (PET). The original PET, which is available in both paper- and computer-based formats, has four parts: part 1: multiple choice; part 2: multiple choice; part 3: a gap-filling exercise; and part 4: a true or false exercise. In part 1, particularly, the items provide test-takers with three pictures (A, B and C), a question and an audio-recording; and the test-takers' task is to listen to the recording and answer the question by choosing the correct picture based on the information provided in the recording. Multimodal input-based tasks were adapted from these items. For the purposes of the study, items containing third-person pronominal anaphora references were selected. Then, software Camtasia 8.0 was used to add captions to the audio-recordings of the PET original items, with the captions appearing below the pictures on a computer screen. The captions were synchronised to the audio recordings. It was ensured that the resulting multimodal input-based tasks met the following criteria for a ‘task’:

> 1. The primary focus should be on ‘meaning’ (by which is meant that learners should be mainly concerned with processing the semantic and pragmatic meaning of utterances).
> 2. There should be some kind of ‘gap’ (i.e. a need to convey information, to express an opinion or to infer meaning).
> 3. Learners should largely have to rely on their own resources (linguistic and non-linguistic) in order to complete the activity.
> 4. There is a clearly defined outcome other than the use of language (i.e. the language serves as the means for achieving the outcome, not as an end in its own right). (Ellis, 2009, p. 223)

The most familiar type of input-based task that meets the above criteria is a ‘listen-and-do task’, defined as a “one-way information gap task that requires learners to listen to commands or descriptions and then perform actions (e.g., a physical action or pointing to a picture)” (Shintani, 2014, p. 282). Following this, in this study, multimodal input-based tasks were operationalised as tasks providing input in multiple modes: static visual stimuli (pictures), audio recordings and captions. The participants were presented with three pictures (A, B and C) on a computer screen and asked to choose the correct one based on information provided in the recording and/or the captions. The primary focus of the task was on meaning, in that the participants were encouraged to infer meaning from what they heard and/or read. It required the participants to use their own non-linguistic resources with a defined outcome so that the learners could choose the correct pictures based on their understanding of the audio, textual and/or visual stimuli.

For the enhanced captions group, the target linguistic constructions (antecedent and pronoun) were additionally enhanced and presented in boldface (see Figure 10). A total of 27 multimodal input-based tasks were developed for three treatment sessions, with nine tasks in each set (see Appendix C for an example set with captions). The duration of the tasks was approximately 30 seconds to 1 minute. Cronbach's α, calculated based on the response options, was found to be acceptable (α = .68).

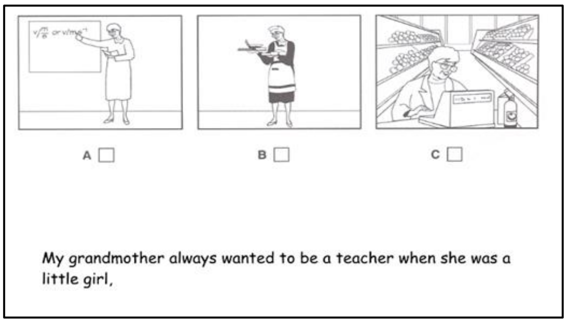
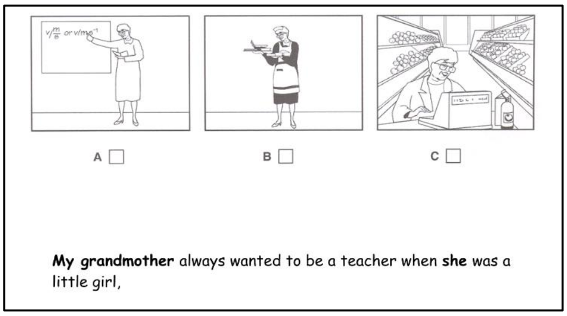
Non-enhanced Caption	
Enhanced Caption	

Figure 10. Textually Non-enhanced Caption versus Enhanced Caption for Study 1

3.1.5.4. Grammaticality Judgement Tests

A written GJT and an oral GJT were used to assess the participants' pretest-posttest gains in the use of third-person pronominal anaphoric references. The inclusion of both written and oral GJTs was expected to determine whether the modality of the outcome measures mediated the relationship between textual input enhancement and development in the use of the target construction. The written GJT and the oral GJT had the same format; they only differed in modality. Before making a grammaticality judgement, the written GJT asked participants to read a given item on a computer screen, whereas the oral GJT involved participants listening to the item. The GJTs were developed following guidelines offered by Keating and Jegerski (2015).

Thirty-two target items, including an antecedent and a third-person pronoun anaphora, were constructed. The items were identical in terms of sentence structure, numbers of syllables of words in the same position and numbers of syllables in the items. Care was taken to include only high-frequency words by selecting words from the New General Service List (Brezina & Gablasova, 2015). Each item was formulated in four versions. Two versions were grammatical, including (a) a singular antecedent with a singular pronoun or (b) a plural antecedent followed by a plural pronoun. The other two versions were ungrammatical, one with (c) a singular antecedent and a plural pronoun and another with (d) a plural antecedent and a singular pronoun (see Appendix D). In addition, the target items allowed for testing learners' sensitivity to gender agreement between the pronouns and preceding antecedents. Given that participants' L1, Korean, uses different pronominal anaphoric resolutions in terms of number and gender agreement, processing the target sentences was expected to be challenging for the participants. It is worth noting, however, that the items did not require participants to resolve ambiguity in anaphora resolution. An example of each version is given below:

a. My sister saw the market on the street. She bought some cookies. (grammatical, singular antecedent – singular pronoun)
b. My sisters saw the market on the street. They bought some cookies. (grammatical, plural antecedent – plural pronoun)
c. *My sister saw the market on the street. They bought some cookies. (ungrammatical, singular antecedent – plural pronoun)

d. *My sisters saw the market on the street. She bought some cookies.

(ungrammatical, plural antecedent – singular pronoun)

In addition to the target items, 48 distractors were developed using three constructions: (a) the passive (16 items), (b) verb + gerund or verb + to infinitive (16 items) and (c) participial adjectives ending in either -ed or -ing (16 items) (see Appendix E). The purpose of including these constructions, which do not have direct equivalents in Korean, was to distract learners' attention from the target construction and thereby prevent them from easily identifying the focus of the assessment (Keating & Jegerski, 2015). Similar to the target items, four versions were developed for each distractor. All distractors had a similar syntactic structure and were designed to be of equal length. They were also designed to be of similar length to the target items. A total of 128 target sentences (32 items, 4 versions each) and 192 distractors (48 items, 4 versions each) were created, which were then distributed into four sets of 80 items, counterbalanced across the four sets as follows:

Set A: 1a, 2b, 3c, 4d etc.

Set B: 1b, 2c, 3d, 4a etc.

Set C: 1c, 2d, 3a, 4b etc.

Set D: 1d, 2a, 3d, 4c etc.

Thus, each set contained 32 target sentences and 48 distractors. All the sentences used in the GJTs were reviewed by a native American speaker.

Both the oral GJT and the written GJT were administered using Eprime 2.0, which allowed recording reaction times (RTs). For the oral GJT, a sentence was delivered aurally, which were recorded by a native American female speaker. After listening to each sentence, the participants were required to make a judgement (Figure 11). The written GJT was untimed, in that a sentence was presented on a computer screen for the participants to read and make a judgement without time pressure (Figure 12). For both GJTs, participants were first required to press “z” or “m” to indicate whether they judged an item to be grammatical or ungrammatical, respectively. In each test version, all items (target and distractors) were randomly presented and fixation crosses were used to indicate transitions between items (see Figure 11 and Figure 12).

The four sets of GJTs (A, B, C and D) were counterbalanced across modality and testing sessions among participants using a Latin Square design. The instructions were provided in Korean, the participants’ native language. Before the GJTs were administered, participants completed a set of practice items in order to familiarise themselves with the format of the written and oral GJTs.

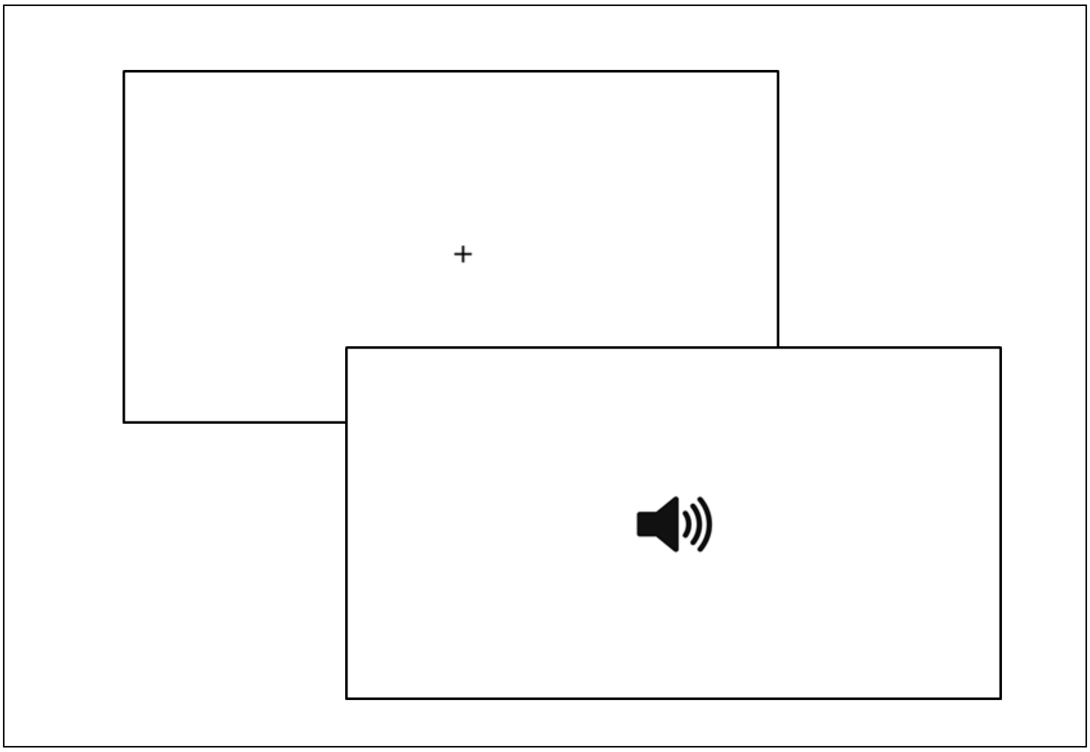


Figure 11. Oral Grammaticality Judgement Test

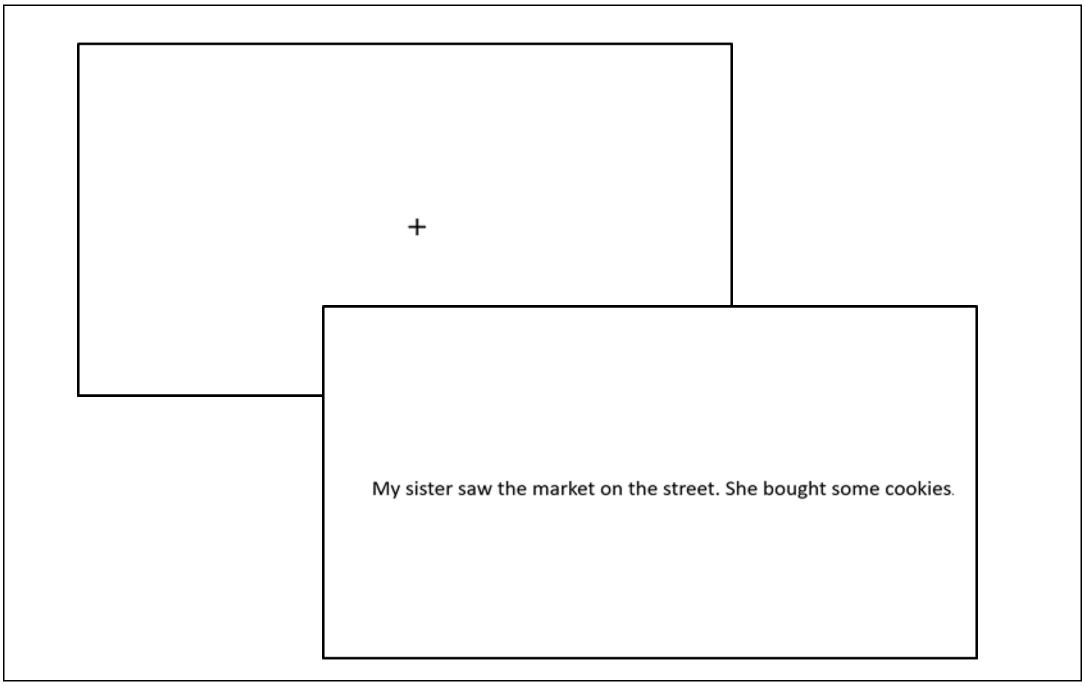


Figure 12. Written Grammaticality Judgement Test

3.1.5.5. Eye-movement Data

A Tobii X2-30 mobile eye tracker with a temporal resolution of 30 Hz was used to collect data on the participants' eye movements while performing the tasks. This remote eye-tracking system was mounted onto a laptop computer with a 15-inch screen. The participants were seated approximately 60 cm from the computer. Before performing each set of treatment tasks, the eye-tracker was individually calibrated to each participant to ensure optimal results. Using 9-point calibration, each participant was asked to follow a red circle on the computer screen with his/her eyes. After the calibration process, the participants were asked to complete the three sets of treatment tasks. The experiment was presented using Tobii Studio 3.3.0 software (Tobii Technology, 2015). Each answer was spoken aloud to enable the participants to remain still while completing the tasks. The participants' eye-movement data were collected during three sets of nine treatment tasks, giving a total of 27 tasks distributed across three treatment sessions.

3.1.6. Data Collection Procedure

For each participant, the experiment took place on three days over one week. On day one, a background questionnaire, an OPT and a pretest were administered. On day two, the participants completed the first and second treatment sessions. On day three, they took part in the third treatment session, followed by a posttest (see Figure 13). The OPT lasted for approximately 40 minutes: 10 minutes was allocated to the listening section and 30 minutes to the grammar section. Each GJT took approximately 20–25 minutes. The duration of each set of nine treatment tasks was between 9 and 10 minutes. All three sessions were individual and administered by the researcher.

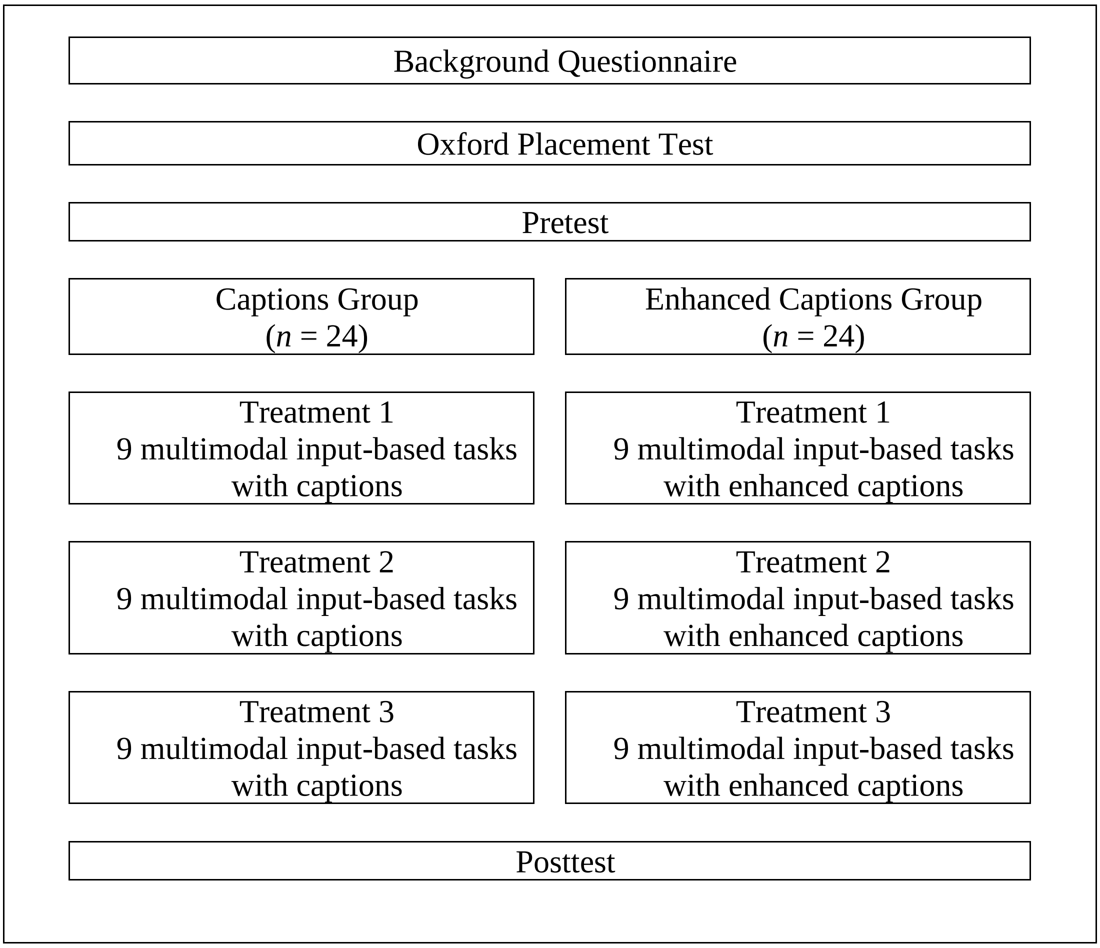


Figure 13. Data Collection Procedure for Study 1

3.1.7. Scoring and Data Analysis

3.1.7.1. Oxford Placement Test

The OPT comprised 200 questions: 100 questions in the listening section and 100 questions in the grammar section. Following the OPT scoring guidelines, one point was given for each correct answer, resulting in a maximum score of 100 for each section.

3.1.7.2. Grammaticality Judgement Test

To score the written and oral GJTs, each correct response was awarded one point, thus the total score for each GJT (excluding distractors) was 32 points. Participants' reaction times (RT: the time between the appearance of a sentence on the computer screen and the participant's response) were measured in milliseconds (Jiang, 2012). For each participant, mean RTs and SDs were calculated for correctly judged sentences only, and potential outliers were identified using the resulting means and SDs (Jiang, 2012). RTs that differed from a participant's mean by more than two standard deviations were considered as outliers (Jiang, 2012), and these were trimmed to two standard deviations above or below the mean.

3.1.7.3. Treatment Tasks

In the treatment tasks, which asked students to choose correct pictures based on their comprehension, participants were awarded one point for every correct picture choice, resulting in a maximum score of 27 (9 points per set).

3.1.7.4. Eye-movement Data

When analysing the eye-movement data, first the quality of the recordings was inspected using the gaze samples measure provided by Tobii Studio 3.3.0. This index is expressed as a percentage, and is calculated "by dividing the number of eye tracking samples that were correctly identified by the number of attempts" (Tobii Studio 3.3.0. User Manual, 2014, p. 39). A value of 100% would mean that the movement of both eyes was found during the full recording (a highly unlikely outcome as participants blink, etc.). For this study, 63.7% was set as a cut-off point for acceptable level of recording quality. This value was calculated by subtracting one standard deviation (18.2%) from the mean percentage (81.9%) of time participants spent viewing the screen. Next, two areas of interest, one for antecedents and the other for pronouns, were selected as shown in Figure 14.

Figure 14. Areas of Interest for Study 1

Drawing on previous eye-tracking research (e.g., Conklin & Pellicer–Sánchez, 2016; Pellicer-Sánchez, 2016; Roberts & Siyanova-Chanturia, 2013; Wink, 2013), four measurements – first pass reading time, second pass reading duration, total fixation duration and number of visits – were used to examine the amount of attention paid to the two areas of interest: antecedents and associated personal pronouns. As described in Chapter 2, first pass reading time is the sum of all fixation durations during a first visit to the area of interest. This measure is regarded as an index of initial processing. Second pass reading time is defined as the sum of fixation durations when the eyes return to an area of interest after the first visit. In other words, second pass reading time captures rereading in the area of interest, which is associated with re-analysis of the input. Total fixation duration is the sum of all fixation durations made within an area of interest. Finally, a visit includes all the fixations made within an area of interest from the time a participant's eyes first enter that area of interest and until they leave.

As recommended by Conklin and Pellicer-Sánchez (2016), the data generated were cleaned before calculating reading pass time and fixation duration. Following previous literature, two different procedures were used. First, fixation durations of individual target linguistic constructions that were shorter than 80ms were removed. However, skipped areas that were recorded as 0ms were included in fixation duration analyses, as it was an important indicator of the amount of attention learners allocated to the target linguistic constructions. Next, mean fixation durations and SDs were calculated for each participant. Fixation durations that differed from a participant's mean by more than three standard deviations were considered as outliers; three standard deviations from a participant's mean was used as an upper limit to trim fixation durations that were exceptionally long.

3.1.8. Statistical Analyses

SPSS 24.0 (Statistical Package for the Social Sciences) for Windows was the main statistical analysis software used to calculate both descriptive and inferential statistics in Study 1.

3.1.8.1. Descriptive Statistics

Descriptive statistics were computed for the participants' performance on the OPT, the written GJT, the oral GJT and eye-movement data.

3.1.8.2. Preliminary Analyses

Preliminary analyses were undertaken to ensure the reliability of the tests and the validity of the results. The internal consistency reliability of the OPT and the GJTs was examined using Cronbach's alpha. A Kolmogorov-Smirnov test was carried out to examine whether the data were normally distributed. The test confirmed that, for both groups, the distributions for each measure were not significantly different from normal.

3.1.8.3. Main Statistical Analyses

A series of independent samples *t*-tests were run to compare the eye-gaze behaviours of the captions and enhanced captions groups during the treatment sessions (RQ1). Next, mixed-model analysis of variance (ANOVA) tests were carried out to examine the effects of textual input enhancement on pretest-posttest GJT gains (RQ2). The relationships between eye-tracking measures and GJT gain scores were established with Pearson correlational analyses (RQ3). Considering the relatively small sample size included in this study, an alpha level of $p < .05$ was set for all tests to avoid Type II error. Effect size estimates were obtained by calculating Cohen's *d* for the independent samples *t*-tests, and eta-squared (η^2) and partial eta-squared (ηp^2) values for the mixed-model ANOVAs (Norouzian & Plonsky, 2017). The eta-squared values (η^2) were computed using the sums of squares for within- and between-subject variables combined. Following Plonsky and Oswald (2014), *d* values of .40, .70, and 1.00, η^2 values of .06, .16, and .36, and *r* values of .25, .40, and .60 were considered as small, medium, and large, respectively.

3.2. Analyses and Results

3.2.1. Preliminary Analyses

3.2.1.1. Test of Normality

The results of the Kolmogorov-Smirnov test are displayed in Table 5. All data were found to be normally distributed.

Table 5. Test of Normality

	Captions Group			Enhanced Captions Group		
	Statistic	*df*	*Sig.*	*Statistic*	*df*	*Sig.*
Oxford Placement Test						
Listening	.109	24	.200	.127	24	.200
Grammar	.086	24	.200	.097	24	.200
Eye-movement data						
First pass reading - ANT	.157	24	.132	.127	24	.200
Second pass reading - ANT	.114	24	.200	.114	24	.200
Total fixation duration - ANT	.106	24	.200	.164	24	.095
Number of visits - ANT	.140	24	.200	.173	24	.061
First pass reading - PRO	.169	24	.073	.128	24	.200
Second pass reading - PRO	.109	24	.200	.125	24	.200
Total fixation duration - PRO	.144	24	.200	.121	24	.200
Number of visits - PRO	.132	24	.200	.158	24	.126
PRE GJT						
Written – Score	.131	24	.200	.128	24	.200
Written – Reaction Time	.164	24	.096	.085	24	.200
Oral – Score	.147	24	.196	.155	24	.140
Oral – Reaction Time	.168	24	.079	.132	24	.200
POST GJT						
Written – Score	.148	24	.189	.174	24	.058
Written – Reaction Time	.151	24	.169	.109	24	.200
Oral –Score	.147	24	.194	.174	24	.059
Oral – Reaction Time	.136	24	.200	.109	24	.200

3.2.1.2. Internal Consistency Reliability

Table 6 summarizes the Cronbach's alpha values computed to identify the reliability of the OPT and each version of the GJT. Each section of the OPT and all four versions of the GJTs were found to be reliable.

Table 6. Internal Consistency Reliability

	N	*M*	*SD*	*α*
Oxford Placement Test – Listening	48	74.04	4.73	.70
Oxford Placement Test – Grammar	48	63.73	7.11	.70
GJT – Set A	48	22.17	5.70	.82
GJT – Set B	48	21.73	5.07	.77
GJT – Set C	48	22.19	5.77	.84
GJT – Set D	48	22.18	5.10	.77

3.2.1.3. Comparability of the Groups

An independent samples *t*-tests confirmed that there were no significant differences between the two groups on either the listening or grammar components of the OPT, as shown in Table 7.

Table 7. Results of Independent samples *t*-tests for Oxford Placement Test

	t	*df*	*p*	*d*
Listening	−.241	46	.810	.07
Grammar	−.241	46	.734	.10

3.2.2. Effects of Textual Enhancement on Treatment Task Performance

The descriptive statistics for the task completion scores achieved by the participants on the 27 multimodal treatment tasks are presented in Table 8. An independent samples *t*-test revealed no significant differences between the captions and enhanced captions groups on the participants' treatment task performance, $t(46) = -.58$, $p = .56$, $d = .17$.

Table 8. Descriptive Statistics for Treatment Task Performance by Group

	Captions Group (*n*=24)				Enhanced Captions Group (*n*=24)			
			95% CI				95% CI	
	M	*SD*	Lower	Upper	*M*	*SD*	Lower	Upper
Task completion score	22.54	2.75	21.38	23.70	23.08	3.62	21.55	24.61

The maximum score was 27 points.

3.2.3. Effects of Textual Enhancement on Learners' Attention to Target Linguistic Construction (RQ1)

To assess the effectiveness of textual enhancement in drawing learners' attention to the target construction, the participants' fixation durations and total number of visits to both areas of interest – antecedents and respective personal pronouns – were compared.

3.2.3.1 Fixation Duration

Table 9 provides descriptive statistics by group for the three indices of fixation duration, that is, first pass reading duration, second pass reading duration and total fixation duration. The means indicate the sum of fixation durations for the 27 treatment tasks combined.

Table 9. Descriptive Statistics for Fixation Durations on Areas of Interest by Group

	Captions Group (n = 24)				Enhanced Captions Group (n = 24)			
			95% CI				95% CI	
	M	*SD*	Lower	Upper	*M*	*SD*	Lower	Upper
Antecedent								
First pass reading(ms)	175	74	144	207	223	110	177	270
Second pass reading (ms)	191	98	150	233	270	134	214	327
Total fixation (ms)	367	138	309	426	494	217	402	586
Pronoun								
First pass reading (ms)	65	37	49	81	87	48	67	107
Second pass reading (ms)	26	18	18	33	27	20	18	35
Total fixation (ms)	91	45	71	110	114	65	87	142

For the antecedents, a series of independent samples *t*-tests found no significant differences in first pass reading duration between the captions and enhanced captions groups, but did reveal that the two groups differed significantly in terms of second pass reading duration and total fixation duration. The effect size for second pass reading time and total fixation duration was in the small range (Table 10). This means that the enhanced captions group spent significantly longer time rereading and reading antecedents overall than the captions group.

Table 10. Results of Independent samples *t*-tests for Fixation Duration – Antecedent

	t	*df*	*p*	*d*
First pass reading	−1.75	46	.09	.51
Second pass reading	−2.33	46	.02	.67
Total fixation	−2.41	46	.02	.70

Another series of independent samples *t*-tests revealed that, for pronouns, none of the fixation duration measures differed significantly across the two groups as presented in Table 11.

Table 11. Result of Independent samples *t*-tests for Fixation Duration – Pronoun

	t	*df*	*p*	*d*
First pass reading	−1.80	46	.08	.51
Second pass reading	−.20	46	.84	.05
Total Fixation	−.14	46	.15	.41

3.2.3.2 Visit Counts

To assess the effectiveness of enhanced input in drawing learners' attention to target constructions, the participants' total numbers of visits to both areas of interest – antecedents and pronouns – were also calculated. Table 12 provides descriptive statistics for visit counts by group. The means capture the total number of visits to the antecedents and pronouns for the 27 treatment tasks combined.

The independent samples *t*-test carried out to compare the total numbers of visits to the antecedents revealed a significant, medium-size difference between the captions and enhanced captions groups, indicating that participants in the enhanced captions groups visited antecedents more frequently. In contrast, the independent samples *t*-test conducted to compare the numbers of visits to pronouns did not yield a significant difference between the two groups (Table 13).

Table 12. Descriptive Statistics for Total Number of Visits to Areas of Interest by Group

	Captions Group (n = 24)				Enhanced Captions Group (n = 24)			
			95% CI				95% CI	
	M	*SD*	Lower	Upper	*M*	*SD*	Lower	Upper
Antecedent number of visits	26.83	9.44	22,85	30.82	36.99	14.85	30.61	43.14
Pronoun number of visits	12.17	6.05	9.61	14.72	14.58	7.54	11.40	17.77

Table 13. Results of Independent samples *t*-tests for Total Number of Visits

	t	*df*	*p*	*d*
Antecedents	−2.80	46	.01	.81
Pronouns	−1.22	46	.23	.35

3.2.4. Effects of Textual Enhancement on L2 Development (RQ2)

To examine the extent to which textual enhancement facilitated development in receptive knowledge of pronominal anaphoric references, the captions and enhanced caption groups' pretest and posttest performances were compared for written and oral GJTs.

3.2.4.1. Written GJT Results

Descriptive statistics for participants' performance on the written GJT are presented in Table 14. To test whether there were any differences between the captions and enhanced captions groups at the time of the pretest, independent samples *t*-tests were conducted using the written GJT pretest accuracy scores and RTs. The results indicated that the two groups did not differ significantly in terms of either total accuracy scores $t(46) = -.06$, $p = .95$, $d = .02$ or RTs: $t(46) = .34$, $p = .73$, $d = .10$ at the outset of the study.

Table 14. Descriptive Statistics for Written GJT Accuracy Scores and RTs by Group

	Captions Group (n = 24)				Enhanced Captions Group (n = 24)			
			95% CI				95% CI	
	M	*SD*	Lower	Upper	*M*	*SD*	Lower	Upper
Pretest								
Accuracy	23.92	4.44	22.04	25.79	24.00	4.60	22.06	25.94
RT	7563.85	2466.61	6522.29	8605.41	7361.19	1522.64	6718.23	8004.14
Posttest								
Accuracy	26.38	3.73	24.80	27.95	28.96	2.26	28.01	29.91
RT	5382.17	1450.47	4769.69	5994.65	5331.31	1389.13	4744.73	5917.88

Maximum score was 32 points.

Next, to examine the effects of textual enhancement on participants' gains on the written GJT, separate mixed-model ANOVAs were conducted for the written GJT accuracy scores and RTs. The within-subjects variable in the analyses was time (pretest versus posttest), and the between-subjects factor was group (captions versus enhanced captions groups). As shown in Table 15, a significant interaction effect emerged between time and group for the accuracy scores, with the interaction accounting for 2 per cent of the variation in the overall model, including within and between-subject variables. That is, the enhanced captions group achieved slightly greater gains in accuracy than the captions group on the written GJT. However, no significant interaction effect was found for reaction times, that is, the two groups did not show significantly different pretest-posttest decreases in the speed with which they responded to the written GJT items.

Table 15. Mixed-Model ANOVA for Written GJT

		F	p	η^2	η_p^2
Time * Group	Written GJT Score	9.454	.004	.021	.170
	Written GJT RT	.089	.767	<.001	.002

3.2.4.2 Oral GJT Results

Table 16 provides descriptive statistics for the oral GJT scores and reaction times. The independent samples *t*-tests which were carried out to examine whether there were initial differences between the captions and enhanced captions groups yielded no significant differences for either pretest oral GJT accuracy scores, $t(46) = .46$, $p = .65$, $d = .13$, or RTs, $t(46) = 1.24$, $p = .22$, $d = .36$.

Table 16. Descriptive Statistics for Oral GJT Accuracy Scores and RTs by Group

	Captions Group (*n* = 24)				Enhanced Captions Group (*n* = 24)			
			95% CI				95% CI	
	M	*SD*	Lower	Upper	*M*	*SD*	Lower	Upper
Pretest								
Accuracy	18.00	3.75	16.42	19.58	17.50	3.74	15.92	19.08
RT	7911.82	1004.29	7487.74	8335.86	7551.51	1005.54	7126.91	7976.12
Posttest								
Accuracy	18.92	3.91	17.27	20.57	25.21	3.15	23.88	26.54
RT	7264.47	1008.27	6838.71	7690.22	6617.73	654.46	6341.38	6894.09

The maximum score was 32 points.

A separate mixed-model ANOVA was also performed for the participants' oral GJT accuracy scores and RTs in order to see whether textual enhancement had an influence on students' oral GJT gains (see Table 17). Similar to the findings obtained for the written GJT, a significant time-by-group interaction emerged for accuracy scores, with the interaction explaining 13 per cent of the variation in the overall model, including within and between-subject variables. On the other hand, no interaction effect was found for reaction times (see Table 17). This means that the participants exposed to enhanced captions achieved significantly greater gains in accuracy on the oral GJT than participants who viewed captions without textual enhancement. The effect size for this difference was in the small range. Textual enhancement, however, did not influence the extent to which participants demonstrated a decrease in the time taken to make grammaticality judgements.

Table 17. Mixed-Model ANOVA for Oral GJT

		F	p	η^2	η_p^2
Time * Group	Oral GJT Score	47.083	.000	.128	.506
	Oral GJT RT	.644	.426	.004	.014

3.2.5. Relationship between Attention and L2 Development (RQ3)

To address the relationship between attention and L2 development, a series of Pearson correlations were computed between the eye-tracking measures and gain scores on the written and oral GJT tests for the antecedents and pronouns combined and separately. As shown in Table 18, for the antecedents and pronouns combined, a significant correlation was only identified between the GJT written scores and total reading time. The direction of this correlation was negative, and its size was in the medium range.

Table 18. Pearson Correlations between Gain Scores, Total Fixation Duration and Number of Visits for Antecedents and Pronouns Combined

	Captions		Enhanced Captions	
	Total Time	Visit Count	Total Time	Visit Count
Written GJT Gain	$-.46^{*}$	−.35	−.11	.05
	(.02)	(.09)	(.60)	(.82)
Oral GJT Gain	.13	.32	−.06	−.09
	(.53)	(.13)	(.77)	(.69)

* $p < .05$

Table 19 displays the results for the antecedent and pronoun separately. For the captions group, medium to large negative correlations were identified between the written GJT gain scores and all the eye-tracking indices calculated for the antecedents (second pass reading, total fixation duration, visit counts), except for first pass reading time. The oral GJT scores of the captions group, however, were found to have medium-size positive correlations with the total duration and total counts of eye-fixations on personal pronouns. Overall, these results mean that, in the captions group, participants achieved lower gains on the written GJT when they reread the anaphora antecedents more often and for longer periods and spent more time gazing at the antecedent and anaphora combined. On the other hand, participants who visited the pronouns more frequently and spent more time reading them displayed greater development on the oral GJT test. No significant relationships were found between the eye-tracking measures and gain scores of the enhanced captions group.

Table 19. Pearson Correlations between Gain Scores and Eye-tracking Measures

	Written GJT Gain *r* (*p*)	Oral GJT Gain *r* (*p*)	First Pass Ant *r* (*p*)	Second Pass Ant *r* (*p*)	Total Fixation Ant *r* (*p*)	Visit Count Ant *r* (*p*)	First Pass Pro *r* (*p*)	Second Pass Pro *r* (*p*)	Total Fixation Pro *r* (*p*)
Captions Group									
First pass	−.26	.24							
Antecedent	(.22)	(.24)							
Second pass	−.51*	−.18	.28						
Antecedent	(.01)	(.41)	(.19)						
Total fixation	−.50*	.00	.73**	.86**					
Antecedent	(.01)	(.99)	(.00)	(.00)					
Visit count	−.54*	.15	.82**	.48*	.78**				
Antecedent	(.01)	(.48)	(.00)	(.02)	(.00)				
First pass	−.17	.38	.54*	.11	.37	.60**			
Pronoun	(.42)	(.07)	(.01)	(.59)	(.07)	(.00)			
Second pass	.14	.38	.29	.01	.17	.04	.26		
Pronoun	(.52)	(.06)	(.16)	(.95)	(.43)	(.86)	(.22)		
Total time	−.09	.46*	.56**	.10	.37	.50*	.92**	.61**	
Pronoun	(.69)	(.02)	(.00)	(.64)	(.07)	(.01)	(.00)	(.00)	
Visit count	.07	.46*	.48*	−.25	.08	.45*	.68**	.18	.63**
Pronoun	(.73)	(.02)	(.02)	(.23)	(.71)	(.03)	(.00)	(.39)	(.00)
Enhanced Captions Group									
First pass	−.15	−.02							
Antecedent	(.49)	(.92)							
Second pass	−.09	−.04	.56**						
Antecedent	(.67)	(.84)	(.00)						
Total fixation	−.13	−.04	.86**	.91**					
Antecedent	(.54)	(.86)	(.00)	(.00)					
Visit count	−.04	−.11	.69**	.60**	.72**				
Antecedent	(.85)	(.62)	(.00)	(.00)	(.00)				
First pass	.06	−.10	.58**	.54*	.63**	.78**			
Pronoun	(.78)	(.64)	(.00)	(.01)	(.00)	(.00)			
Second pass	−.24	−.17	.53**	.71**	.71**	.66**	.76**		
Pronoun	(.26)	(.42)	(.00)	(.00)	(.00)	(.00)	(.00)		
Total fixation	−.03	−.13	.60**	.63*	.69**	.79**	.98**	.88**	
Pronoun	(.89)	(.55)	(.00)	(.01)	(.00)	(.00)	(.00)	(.00)	
Visit count	.21	−.03	.41*	.32	.41*	.75**	.81**	.56**	.77**
Pronoun	(.31)	(.88)	(.04)	(.12)	(.04)	(.00)	(.00)	(.00)	(.00)

** $p < .01$, *$p < .05$

Table 19 also shows the correlations between the various eye-tracking indices. For the captions group, fewer correlations were observed, the analyses yielding no significant links between the second pass and total reading durations calculated for antecedents and those for pronouns. For the enhanced captions group, on the other hand, a large majority of the eye-tracking indices computed for antecedents correlated to a medium or large degree with those computed for pronouns, including second pass and total reading times. That is, participants in the enhanced captions group who reread and fixated longer on anaphora antecedents also devoted more time gazing at associated personal pronouns. In the captions group, however, those who fixated longer on antecedents/ pronouns did not pay more attention to related pronouns/ antecedents

3.2.6. Summary of Results

RQ1 To what extent do non-enhanced captions versus textually enhanced captions draw learners' attention to target linguistic constructions?

For antecedents, a series of independent samples *t*-tests showed that there was a statistically significant difference between the textually non-enhanced captions group and the enhanced captions group in terms of second pass reading duration and total fixation duration with small effect sizes, .67 and .70 for each. In addition, the two groups also exhibited medium size differences for numbers of visits to the antecedents. However, no significant difference was observed between the two groups for all measures in relation to the other target linguistic construction, pronouns.

RQ2 To what extent do non-enhanced captions versus textually enhanced captions affect L2 development in the knowledge of target linguistic constructions, as measured by a written and an oral grammaticality judgement test (GJT)?

To answer research question 2, mixed-model ANOVAs, including within-subjects variables (pretest and posstest) and between-subjects variables (non-enhanced captions group versus enhanced captions group), were performed. The results indicated a significant time-by-group interaction for the accuracy scores of both written and oral grammaticality judgement tests; the enhanced captions group achieved greater gains in accuracy than the captions group on both written and oral grammatical judgement tests. Textual enhancement, however, did not have a significant influence on the time taken to make grammaticality judgements, whether measured by a written or oral test.

RQ3 To what extent is attention to target linguistic constructions in captions related to L2 development? Is this relationship influenced by textual enhancement?

For the non-enhanced captions group, somewhat different findings were reported for the written and the oral grammaticality judgement tests (GJT). That is, three eye-tracking indices (i.e., second pass reading, total fixation duration and total number of visits) were found to have medium to large negative correlations with written GJT scores, whereas two of the eye-tracking indices (i.e., total fixation duration and total number of visits) had positive relationships with oral GJT scores. For the enhanced group, on the other hand, no significant relationships were found between the eye-tracking indices and gain scores on both the written and the oral GJTs.

3.3. Interim Discussion

3.3.1. Effects of Textual Enhancement on Attentional Allocation

The first research question asked the extent to which textual enhancement in captions can draw learners' attention to pronominal anaphoric references, a grammatical construction. The eye-tracking data which were utilised to assess attentional processing revealed that textual enhancement was successful in directing learners' attention to the referential antecedents highlighted in the captions. When the antecedents were presented in boldface, the participants devoted more time to rereading forms, and fixated longer on them overall. Compared to the unenhanced group, learners exposed to enhanced input also visited antecedents more frequently. The effect sizes for these differences were in the small (second pass reading time, total fixation) and medium range (visit counts). Textual enhancement, however, did not lead to longer first pass reading times. Contrary to the results for the antecedents, the enhanced visual salience of personal pronouns did not generate more attention; none of the eye-tracking indices yielded significant differences between the enhanced and unenhanced groups for this target area. Overall, these results suggest that textual enhancement in captions was able to trigger more reflection and re-analysis of target antecedents, but it did not lead to increased initial processing of them. Nor did textual enhancement yield increased attention to associated personal pronouns.

The results for referential antecedents indicate that raising the visual salience of target features in captions cannot just facilitate attention to lexis, as observed in Montero Perez et al. (2015), but also promote attention to grammar. It is important to note, however, that Montero Perez et al. found an advantage for enhanced captions (key words versus full captioning) only when a forthcoming test was announced.

Although no test announcement was made in the present study, it is likely that the participants anticipated forthcoming posttests, given that they had taken a battery of pretests prior to the treatment. This, in turn, might have made them more attentive to textually enhanced constructions, as in Montero Perez et al. (2015).

It is also worth comparing the results to textual enhancement studies utilising unimodal input. As mentioned earlier, this line of research has so far yielded mixed findings, with some studies observing a benefit for attentional allocation under an enhanced condition (Simard & Foucambert, 2013; Winke, 2013), others generating null effects (Indrarathne & Kormos, 2017; Issa et al., 2015; Loewen & Inceoglu, 2016). Indrarathne and Kormos explain these mixed patterns by the differential visual salience created by the various textual enhancement techniques across studies, suggesting that underlining (Simard & Foucambert, 2013; Winke, 2013) might be more effective in creating an isolation effect than other forms of textual enhancement, such as boldfacing (Indarathne & Kormos, 2016) or the use of different coloured fonts (Issa et al., 2015; Loewen & Inceoglu, 2016). Although this study employed boldfacing, as did Indrarathne and Kormos, this technique might have been more successful in generating an isolation effect here. The enhanced constructions appeared in sentences, which probably made the highlighted input more salient, as compared to when target features are boldfaced in larger text.

The fact that the input was presented bimodally could have further promoted the salience of the textually enhanced features in the captions. It is possible that the captioned texts, at least for some of the participants, were delivered faster than their normal reading speed. Thus, due to a lack of time to read the captions in full, the learners might have devoted increased attention to highlighted words, assuming that they contained key information. A fast presentation speed could also explain why there was no difference in the amount of attention allocated to enhanced referential pronouns between the two groups. Even though the pronouns were made more salient to the enhanced group, participants probably had less time to revisit them after rereading the antecedents, due to the short-lived nature of the captions.

3.3.2. Effects of Textually Enhanced Captions on L2 Development

The second research question was concerned with the extent to which textually enhanced captions facilitated development in the knowledge of pronominal anaphoric references, as measured by a written and an oral grammaticality judgement test. On both GJTs, participants exposed to enhanced captions demonstrated greater gains in accuracy than participants who viewed captions without enhanced input. Notably, the advantage for the enhanced condition was more pronounced in the oral than written GJT scores, with the effect size values falling into and below the small range for the oral and written GJTs, respectively. Textual enhancement, however, did not affect the degree of decrease in reaction times from pretest to posttest. These findings suggest that the increased attention that participants paid to pronominal anaphoric references under the enhanced condition led to further processing of the target construction, resulting in greater longer-term gains in accuracy.

The results of the present study align well with the findings of Montero Perez et al. (2015) for lexis, indicating that enhanced visual salience in captions cannot just lead to better recognition of lexical forms but also improve receptive knowledge of grammar. The results, however, differ from the conclusion of Lee and Huang's (2008) meta-analytic review that textual input enhancement has only a marginal impact on grammar learning. The somewhat larger effect sizes found in the present study could be attributed to the fact that participants had some prior knowledge of the target construction, as reflected in the considerably higher than chance accuracy scores on the written GJT pretest. Several researchers (e.g., Han, Park, & Combs, 2008; Lee & Huang, 2008; Park, 2004; Winke, 2013) have noted that prior knowledge is likely to be a key determinant of whether textual enhancement succeeds in promoting grammatical knowledge. An additional, or alternative, explanation for the positive findings obtained here could be that textual enhancement was used in combination with captioning. As pointed out by Leow and Martin (2017), the joint use of textual enhancement with other attention-getting tools (e.g., captioning) is likely to trigger greater depth of processing of enhanced input, and thereby result in more developmental benefits than using textual enhancement alone.

3.3.3. Relationship between Attention and L2 Development

The third research question addressed the relationship between attention, operationalised in terms of eye-tracking measures, and development in the knowledge of pronominal anaphoric references. While no significant links were observed between the eye-movement indices and GJT gain scores for the enhanced captions group, the correlational analyses yielded a number of significant medium to large relationships between the gain scores of the captions group and eye-movement measures. In the captions group, participants who visited anaphora antecedents more often and reread and fixated longer on them exhibited less development on the written GJT, whereas participants who looked at the pronouns more frequently and devoted more time to reading them overall achieved greater gains on the oral GJT.

In light of previous research (e.g., Godfroid & Uggen, 2013; Indrarathne & Kormos, 2017), it was expected that a positive relationship would emerge between oral GJT gains and the overall length and number of eye fixations on target pronouns. This result, however, was anticipated for both groups, not just for the unenhanced captions condition. It is more puzzling that, in the unenhanced group, negative links were observed between written GJT gain scores, second pass reading times and visit counts for antecedents. Possibly, those participants who reread antecedents (but not pronouns) more frequently failed to pay attention to anaphora construction as a whole and, as a result, displayed less improvement on the written GJT. For these participants, the lack of increased attention to pronouns might have resulted from the fixed rate of input delivery. Due to time constraints, participants might only have had time to revisit antecedents but not personal pronouns in the captions. This account is supported by the fact that, for the unenhanced captions group, no significant correlations emerged between how much time participants spent rereading antecedents and associated pronouns.

The question also arises as to why there was no relationship between eye-tracking indices and their GJT gain scores for the enhanced group. Probably, participants in this group naturally orientated their attention to both components of the highlighted anaphora constructions, as indicated by the strong correlations between the amount of attention participants paid to antecedents and pronouns. Thus, despite the set speed of delivery, they might have had enough time to select both components of anaphora construction for further processing. For this group, the degree of gains in knowledge of pronominal anaphoric references might have been more related to the extent to which participants engaged in a higher level of processing after the initial selection of information, rather than differences among participants in the amount of lower-level processing they performed, which was captured in the eye-tracking measures (Leow, Grey, Marijuan, & Moorman, 2014).

Montero Perez et. al., when explaining a similar lack of a significant link between attention and learning in their enhanced condition, also speculated that longer fixations might not necessarily be a reflection of more elaborate processing. Future studies using verbal protocols such as the stimulated recall procedure could help shed light on the correctness of this interpretation.

Another possible explanation for this finding is that learners who were exposed to unenhanced target linguistic constructions had processing difficulty and thus, fixated longer at the target constructions. That is, increased attention allocated to the target linguistic constructions might not have reflected learners' attempt to learn and integrate information but might have been a reflection of the level of difficulty they encountered when processing the input (Montero-Perez et al., 2015). In fact, in this study, both positive and negative relationships emerged between the allocation of attentional resrouces and learning gains, which could possibly be interpreted as suggesting that longer fixations may reflect both increased cognitive effort to learn or increased effort to resolve a processing problem.

3.4. Limitations of Study 1 and Modifications in Study 2

When interpreting the findings of this research, it is also important to take into account the limitations of the study. Study 1 included participants who were considered to be at B1–B2 proficiency level according to the Common European Framework of Reference (CEFR). It would be interesting to examine the potential of an increase in visual salience of target linguistic constructions to facilitate learning across proficiency levels. In Study 2, therefore, more proficient students were recruited in an attempt to determine whether non-enhanced or textually enhanced captions had the same or differential effects on development of L2 grammatical knowledge in comparison with the positive effects evidenced for less proficient users of English. Consequently, different target linguistic constructions (i.e., present perfect and past simple) were selected in consideration of the proficiency level of the participants.

There also exist some methodological limitations identified in Study 1. First, a control group, which was provided with only aural and visual input without captions, was not included as a baseline for comparison. The inclusion of a control group would allow examination of whether the provision of captions only assisted in promoting use of the target linguistic constructions. This limitation was acknowledged in Study 2 by having three groups, a no captions group, a textually non-enhanced captions group, and a textually enhanced group. Another limitation that needs to be addressed concerns an experiment treatment task used in Study 1. The task was developed using a static image and captions. However, considering that many language learners watch news, movies and/or dramas to learn a second and foreign language, tasks incorporating video clips were considered to have more face validity for students. Thus, real news clips – without captions, with non-enhanced captions or with textually enhanced captions – were adapted for Study 2. The absence of a delayed posttest in Study 1 was an additional weakness. Administering a delayed posttest would allow determining whether increasing the visual salience of target linguistic constructions has a long-term effect on L2 grammatical knowledge development. Thus, this limitation was addressed in Study 2 by incorporating a pretest / immediate posttest / delayed posttest experimental design. The results derived from Study 2 were expected to identify not only the effects of textual enhancement on immediate gains in knowledge of L2 target linguistic constructions but also its impact on learners’ longer-term retention of their learning gains.

In relation to the treatment task, special attention should be drawn to the typographical modification used to enhance the target linguistic constructions in Study 1. The target linguistic constructions were presented in a bolded font as in Simard and Foucambert (2013) and Indrarathne and Kormos (2017). Bolding has been frequently used to increase the perceptual salience of target constructions in previous studies, in which textual enhancement was included in a reading context to draw learners' attention and facilitate L2 learning (e.g., Izumi, 2003; Jourdenais et al., 1995; Lee, 2007; Leow, 1997; Leow et al., 2003; Shook, 1994). However, in studies using eye-movement data as evidence of learners' attention directed to the input, a question has been raised regarding the use of bolding to manipulate the salience of the input, since there is a possibility of a slight change in the size of the font (Winke, 2013), which can be problematic for selecting the area of interest. In eye-tracking studies, areas of interest (target linguistic constructions) need to be selected first so that researchers can collect and analyse eye-movement data using various measurements (e.g., first pass reading, second pass reading, number of visits). In this study, great caution was exercised to ensure that identical size rectangles were used to select areas of visual interest for each target linguistic construction, regardless of bolding, across groups. However, recognizing this issue, a different colour (i.e., yellow) was used to manipulate the salience of the target linguistic constructions in Study 2.

Furthermore, in Study 1, as a pretest and posttest to evaluate gains in L2 grammatical knowledge, written and oral grammaticality judgement tests (GJTs) were employed, which are claimed to be a form-oriented post-instructional measure (Han et al., 2008). Such form-oriented tests (e.g., grammaticality judgement, sentence completion, sentence combination, multiple choice and fill in the blanks) are presumed to show developmental change in L2 grammatical knowledge. However, they may fall short in terms of representing learners' overall competence in using the target linguistic constructions for more meaningful purposes. In this sense, further studies could expand this line of research by including meaning-oriented measurements (e.g., Bowles, 2003; Doughty, 1991; Izumi, 2002; Leow, 2001) to assess L2 learners' grammatical knowledge. Considering that only form-oriented measurements were used in Study 1, both form-oriented and meaning-oriented tests were included in Study 2. As a type of form-oriented measurement, the format of filling in gaps was used to measure receptive knowledge of target linguistic constructions. Meaning-oriented measurement, which required the learners to use the target linguistic constructions in context based on their understanding, were created as a form of both oral and written productive test.

The statistical analysis used in Study 1 is also a subject for discussion. More specifically, in Study 1, parametric statistics, such as independent samples *t*-tests and mixed-model analysis of variance (ANOVA), were used to analyse the data. However, some researchers have pointed out some shortcomings of parametric statistics, including their inability to generalize the results beyond the participants and items included in a study and the requirement of meeting certain assumptions. Thus, in Study 2, more advanced statistical analysis was employed, which is a series of linear mixed-effects models using the statistical package R (R development core team, 2011). Mixed-effects models take random variance of participants and items into consideration. Thus, the results were expected to be more valid and generalizable beyond the participants and items included in the current study. A detailed explanation of the notions and procedure of the mixed-effects models is provided in the following chapter.

Another shortcoming of this research concerns the lack of control for individual differences among participants. Working memory, in particular, is likely to moderate the impact of textually enhanced captions on attentional allocation, given the multiple sources of input to which participants were exposed and the relatively limited L2 knowledge and processing skills they possessed. Working memory has also been demonstrated to be significantly linked to gains in receptive knowledge in other studies investigating the effects of textual enhancement on attentional allocation and development (e.g., Indrarathne & Kormos, 2018). Against this background, an attempt was made to investigate how individual differences in working memory mediated the effects of captions, non-enhanced captions and textually enhanced captions on promoting L2 grammatical knowledge, as well as on drawing learners' attention to the target linguistic constructions in Study 2. Given the argument stating that there is a close link between an individual's working memory and attention (Robinson et al., 2012), an empirical examination of the relationship between individual differences in working memory and learners' allocation of attentional resources in responding to perceptually salient input seems important from both theoretical and pedagogical perspectives.

CHAPTER 4

STUDY 2

This chapter reports on Study 2, which further investigated the extent to which textually enhanced captions in multimodal input-based tasks can draw learners' attention to and promote development in the use of L2 grammatical knowledge. Considering the limitations of Study 1, several modifications were made. One of the major modifications involved recruiting more proficient participants to test whether the effects observed for textually enhanced captions in Study 1 would transfer to different proficiency levels. The research design was improved by including a control group (i.e., a no captions group), administering a delayed posttest and employing both form-oriented and meaning-oriented tests as assessment tasks. In this study, a more advanced statistical analytical procedure, namely, a series of linear mixed-effects models, was used to analyse the data. Individual differences among the participants were also taken into consideration by investigating the extent to which working memory moderated the impact of captions, textually non-enhanced or enhanced, on attentional allocation and the development of L2 grammatical knowledge. A further methodological contribution of Study 2 was that six different working memory measures were included to assess different constructs of working memory capacity: (a) nonword span task, (b) forward corsi block task (c) backward corsi block task, (d) stop signal task, (e) colour shape task, and (f) automated operation span task.

This chapter begins by presenting the research questions addressed in this study. The methodology used to conduct the study is explained in the following section, including a description of the participants, research ethics, target linguistic constructions, experimental treatment tasks, materials, data collection procedure and statistical analyses. Next, the results of the study are provided for each research question, followed by a discussion of the findings.

4.1. Research Questions

The following research questions were addressed in Study 2:

1. To what extent do multimodal input-based tasks without captions versus those with captions affect development in L2 grammatical knowledge?
2. To what extent do textually enhanced versus non-enhanced captions in multi-modal input-based tasks affect development in L2 grammatical knowledge?
3. To what extent do textually non-enhanced versus enhanced captions in multi-modal input-based tasks draw learners' attention to the target linguistic construction?
4. To what extent does learner attention allocated to the target linguistic construction relate to development in L2 grammatical knowledge? Is this relationship influenced by whether learners are exposed to non-enhanced or enhanced captions?
5. To what extent do individual differences in working memory capacity moderate the effects of captions, textually non-enhanced or enhanced, in multimodal input-based tasks on L2 development?
6. To what extent does learners' attention allocated to the target linguistic construction relate to their working memory capacity? Is this relationship influenced by whether learners are exposed to textually non-enhanced or enhanced captions?

4.2. Methodology

4.2.1. Overall Design

This study employed a pretest-immediate posttest-delayed posttest experimental design. Seventy-two participants, learning English as a foreign language, were recruited from three different universities located in Seoul, South Korea. They were randomly assigned into three groups: a no captions group (n = 24), a captions group (n = 24) and an enhanced captions group (n = 24). All three groups were administered a general proficiency test, a pretest, a series of treatment tasks, an immediate posttest, working memory measures, a delayed posttest and an exit questionnaire. The groups only differed in what type of input they received, videos without captions, with textually non-enhanced captions, or with enhanced captions.

4.2.2. Participants

Of the total number of 72 undergraduate students, 45 were female and 27 were male. They were all native speakers of Korean learning English as a foreign language. Their ages ranged between 20 and 25 years (M = 21.86, SD = 1.42). To identify the general English proficiency level of the participants, an Oxford Placement Test (OPT) was administered at the outset of the study. On the basis of total scores on the OPT, between 166 and 187, the English proficiency level of the participants was determined to be highly proficient, comparable to C1 band and above according to the Common European Framework for Reference.

Table 20 provides descriptive statistics for the participants' performance on the OPT. Using the scores of the OPT, the comparability of the three groups was examined.

Table 20. Descriptive Statistics for Participants' Performance on the Oxford Placement Test

	Listening Section			Grammar Section		
	M	*SD*	95% CI	*M*	*SD*	95% CI
No Captions	89.04	4.72	[87.05, 91.04]	87.08	4.68	[85.11, 89.06]
Non-enhanced Captions	89.38	6.14	[86.78, 91.97]	89.00	4.75	[86.99, 91.01]
Enhanced Captions	91.17	4.06	[89.45, 92.88]	88.63	4.68	[87.13, 89.34]

One-way ANOVA was run on both the listening section and grammar section scores. There were no significant differences among the three groups: thus, the no captions group, the non-enhanced captions group and the enhanced captions group were considered comparable in terms of proficiency (listening section: F (2, 69) = 1.229, p = .229, η^2 = .034; grammar section: F (2, 69) = 1.119, p = .333, η^2 = .031).

4.2.3. Research Ethics

Research ethics approval for Study 2 was received from Research Ethics Committee at the UCL Institute of Education, University College London. Prior to the beginning of the experiment, each participant was asked to read an information sheet which explained the purpose and procedures of the study, the duration of the study and the measures taken to ensure their privacy and confidentiality. In addition, it was emphasized that participation was voluntary and withdrawal from the study was possible at any stage. For each participant, his or her agreement to participate in the study was obtained by signing a consent form (see Appendix B for the information sheet and consent form).

4.2.4. Target Linguistic Constructions

The target linguistic construction in the present study was the use of the present perfect versus the past simple in English to report news.

4.2.4.1. Present Perfect in English

In English, a distinction is made between tense and aspect: tense "relates the time of the situation referred to some other time, usually to the moment of speaking" (Comrie, 1976, p. 2), whereas aspect refers to "different ways of viewing the internal temporal constituency of a situation" (Comrie, 1976, p. 3).There are two verb tenses (i.e., present and past) and future time reference that are used to express the time of action in relation to the moment of speaking. Two aspects of construction, which express how the speaker views the action of the verb, include perfective and progressive (Cowan, 2008). Among various perfective aspects, the present perfect construction has the form *have/has + past participle,* and has been commonly interpreted as expressing "the continuing relevance of a past situation" (Comrie, 1976, p. 52). Leech (1971) makes a distinction between the past simple and the present perfect by stating that the present perfect refers to "past with present relevance" or "past involving the present" (p. 36). McCawley (1971) defines the present perfect as "an interval stretching from the past into the present" (p. 105), with an introduction to universal and existential terminology to explain uses of the present perfect in English:

a. to indicate that a state of affairs prevailed throughout some interval stretching from the past into the present (Universal)
   I've known Max since 1960.
b. to indicate the existence of past event (Existential)
   I have read Principia Mathematica five times.
c. to indicate that the direct effect of a past event still continues (Stative)
   I can't come to your part tonight – I've caught the flu.
d. to report hot news (Hot news)
   Malcom X has just been assassinated. (McCawley, 1971, p.104)

Of the various uses of the present perfect in context, the present study particularly focused on the use of the present perfect to report news.

e. The US space shuttle Atlantis has returned safely to earth. It landed in Florida this morning.

In news reports, the present perfect is often used first to mention a topic, with more details provided with the past simple in a following section (Eastwood, 1994). The rationale for including use of the present perfect versus the past simple to report news in English as the target linguistic construction is that English tense and aspect properties are difficult features for learners of English to master if morphosemantic discrepancies exist between L1 and L2 (e.g., Bardovi-Harlig, 2001; Gabriele, 2009; Montrul & Slabakova, 2002). Given that Korean has a different present perfect construction in comparison to English, use of the English present perfect in context is expected to be especially demanding for Korean learners of English.

4.2.4.2. Present Perfect in Korean

Along with tenses including the present tense, past tense and future tense, Korean has two aspects: a progressive aspect that denotes ongoing action and a perfective aspect that expresses resulting states (Chang, 1996). The perfective aspect, taking the form *V + e + issta*, indicates the continuation of a state resulting from the completion of an action. However, unlike English, Korean has relatively limited use of the present perfect in the form *V + e + issta*; that is, it is used with only a few stative and intransitive verbs, such as *anta* (to sit), *seta* (to stand), *nupta* (to lie down), *alieojita* (to be known), *jueojita* (to be given), *palpyotoita* (to be announced). In other cases, verbs with a past simple suffix, i.e. -*(e/a)ss,* are more frequently used and these substitute for verbs constructed with a present perfect suffix, i.e. -*e issta,* (Chang, 1996; Han & Hong, 2015; Sohn, 1995). In Korean, therefore, verbs with the past suffix -*(e/a)ss* can indicate both the past simple and the present perfect.

(1) State the past
“*Na-nun ku ttay yel-sal-i-ess-eyo.”* (I was 10 years old then.)

(2) Event in the past
“Ecey na-nun yenghwa-kwukeyng-ul ka-ss-eyo.” (I went to the cinema

yesterday.)

(3) Habitual action in the past
"Mayil, ku salam-ul manna-le ka-ss-supnita." (I went to see him every day.)

(4) State/event leading to the present
"Yong-I bangkum wass-eyo." (Yong has just come/arrived.)

In Korean, rather than using a distinct suffix to denote the perfective aspect, the difference between the past and the present perfect is often indicated with an appropriate discourse context, speech situation, time adverbial or other time-indicating word. For instance, when reporting news, co-occurring time adverbials, such as *bangkum* (just now) and *barojeone* (right before), are used along with the past suffix -*(e/a)ss* to indicate recent news. Probably because there is no perfect aspect suffix distinct from the past tense suffix in Korean, Korean learners of English find it difficult to acquire this construction. Han and Hong (2015) found that Korean students use the past simple when the present perfect is expected in English, even though they have adequate knowledge of the rules of the English present perfect (Han & Hong, 2015).

4.2.5. Experimental Treatment Task

A total of 24 multimodal input-based tasks were developed using news clips on diverse topics (Table 21). The news clips were carefully selected from various sources, including the BBC, CNN and ODN, and modified in terms of length to ensure that each news clip lasted between 20 and 50 seconds. They were selected in consideration of two criteria: appropriateness of the content and use of the target constructions (present perfect and past simple). The topics of the news clips fell into one of four categories: (a) natural disaster, (b) crime, (c) politics, and (d) accident and safety.

Table 21. Titles of News Clips for Treatment Tasks

Category	Title
Natural Disaster	3-year-old and elderly woman buried in China earthquake rubble rescued (task 1) Super typhoon Dujuan hits Taiwan & China (task 8) Peru's Ubinas volcano erupts again (task 14) Super typhoon Haiyan hits Philippines (task 16) Giant sinkhole opens up on Australia beach (task 18)
Crime	Body of US woman found in a suitcase in Bali (task 2) Starving baby found in car boot in France (task 19) Virginia journalists killed during live TV interview (task 23)
Politics	Donald Trump declares his love and respect for women (task 3) Chinese President Xi Jinpping meets Queen on UK visit (task 12) Police arrest man following grisly shooting spree in Texas (task 15) Tony Blair sorry for Iraq War 'mistakes' (task 17) Thailand prime minister Yingluck Shinawatra ordered to step down after 'abusing power' (task 22)
Accident and Safety	Eleven people, including eight school children, killed in China bus crash (task 4) Truck crashes into Arizona apartment (task 5) Escaped tiger kills man in Georgia (task 6) Fifth large object from AirAsia QZ8501 found (task 7) Firefighters rescue toddler left dangling from a third floor window in China (task 9) Florida man rescued after spending hours trapped in deep underground hole (task 10) Bullfight gone wrong – Ten people injured in Peru (task 11) Man killed in shark attack in Australia (task 13) Six dead and four injured in Kentucky car crash (task 20) Taiwan gas explosion kills 24, injures 271 (task 21) Escalator swallows man's foot in China (task 24)

The present perfect was used to introduce the topic of each news item and then the past simple tense was used to give details in the rest of the clips. The video clips were selected in such a way that active and passive uses of the present perfect tense were equal in numbers of occurrences; therefore, the participants were presented with 12 cases of each voice while engaging in the treatment tasks.

The multimodal input-based task, as defined in Study 1, was operationalized as a form of captioned video which incorporates different input (audio, textual and/or visual input). The task also met the main criteria for defining a task (Ellis, 2009): a primary focus on meaning, the existence of a gap, the use of learners' own linguistic and non-linguistic resources, and the generation of an outcome. It was contextualized in the following scenario: the participant played the role of an editor in a newsroom whose job was to review the content, title and graphic elements and the categorization of a news item. After viewing each news clip, the title and category of the news clip were presented to the participants on a computer screen which asked them to make a judgement on their appropriateness. If both were appropriate, the participants were instructed to press ‘z’ on the keyboard; if one of them – either the title or the categorization – was inappropriate, they had to press ‘m’ on the keyboard. In this way, the task had a non-linguistic outcome. Instructions were provided on the computer screen for the participants to read (see Figure 15).

Cronbach’s α, calculated based on the response options, was found to be acceptable (α=.66).

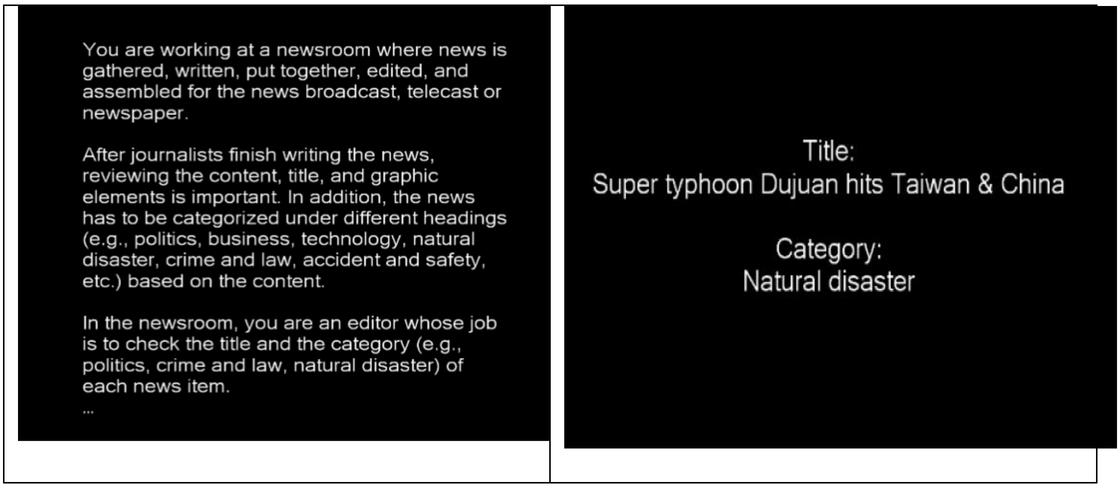


Figure 15. Experimental Treatment Task for Study 2

The news clips were downloaded from YouTube and modified using Camtasia 8.0. For the non-enhanced captions group and the enhanced captions group, captions, using Calibri 14-point font, were inserted into each news clip. Non-manipulated captions were provided for the non-enhanced captions group whereas the target constructions (present perfect and past simple) were enhanced using a different colour, i.e., yellow, with a program called Subtitle Edit for the enhanced captions group. Audio without captions was available for the no captions group (control group). While performing the tasks, participants were exposed to audio, visual stimuli, and/or textual (captions) simultaneously (see Figure 16).

Figure 16. No captions, Non-enhanced captions and Enhanced captions for Study 2

4.2.6. Materials

4.2.6.1. General Proficiency Test

An Oxford Placement test, including a listening and a grammar test, was used to assess the participants’ general proficiency in English. Each section contained 100 questions which asked students to choose the correct answer. Cronbach's alpha for the listening section was .72 and for the grammar section it was .66.

4.2.6.2. Assessment Tasks

In Study 2, an oral productive test, a written productive test and a receptive test were used to measure changes in the participants' knowledge of the target linguistic constructions. Three different versions of each test were developed, and these were counterbalanced among the participants across sessions. To contribute to the existing literature, the present study sought to further explore how the different types of assessment tasks used might mediate the effects of captions, textually non-enhanced and enhanced, on L2 grammatical knowledge development.

4.2.6.2.1. Productive Tests

Two meaning-oriented productive tests, an oral productive test and a written productive test, had the same format, differing only in modality. The participants were asked to view a series of news clips in their native language (Korean) and report them in English. Both the oral productive and written productive tests were contextualized to achieve authenticity. In the oral productive test, a situation was created for participants to tell breaking news that they saw to their friends (see Figure 17). The written productive test required the participants to post news on their Social Networking Service (SNS) in English to share it with their friends (see Figure 18). The participants' oral reports of news were audio recorded for transcription purposes (see Appendix H). The written productive test, on the other hand, asked the participants to write a brief news report in English after watching each news clip (see Appendix I).

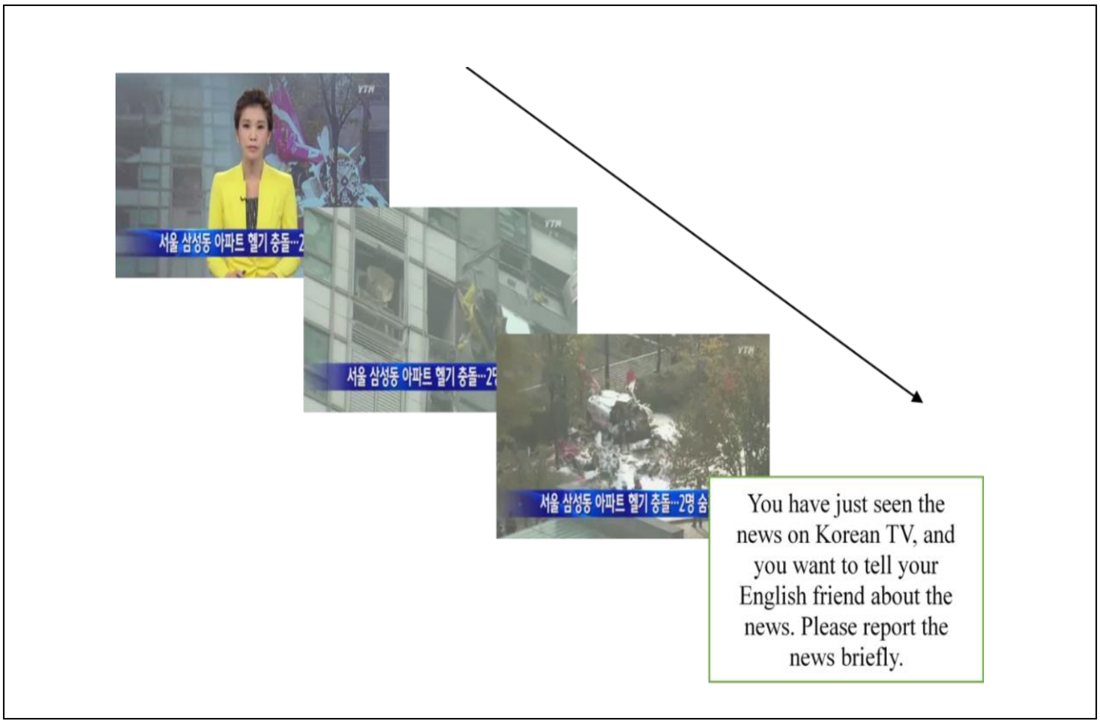


Figure 17. Oral Productive Test

Figure 18. Written Productive Test

Five news clips were included in the oral productive test and the written productive test, respectively. News on diverse topics (see Table 22) without titles was presented on a computer screen and there was no word limit for the participants' responses.

Table 22. Titles of News Clips for Productive Tests

Oral Productive Test	Set A 1. LG helicopter crashes into high-rise apartment, killing 2 pilots 2. K-pop group in fatal car accident 3. Samsung unveils new Galaxy S6 smartphone 4. Shin Hae Chul passed away after suffering a heart attack 5. Four people have been killed and scores injured in an apartment building fire in Uijeongbu Set B 1. Fire at South Korean hospital for elderly kills 21 2. Mountain Umyeon landslide 3. Wicked stepmother jailed 4. Sewol sank off South Korea's southwestern coast 5. South Korea to issue state history textbooks Set C 1. Two South Korean soldiers injured in apparent landmine explosion on boarder 2. Elderly man caught for alleged ring theft 3. A woman in Korea now nicknamed "cat mom" is murdered 4. Both North and South Korea have reached an agreement 5. Malaysia Airlines (MAS) flight MH17 was shot down over Ukraine.
Written Productive Test	Set A 1. Gun-range stabbing suspect arrested 2. South Korean teacher arrested for abusing a toddler 3. Cho, Hyun-ah is sentenced to one year in prison 4. Two dead, dozens injured in Yeongjong Bridge pileup 5. Volkswagen chief executive Martin Winterkorn resigns Set B 1. At least 10 dead after South Korean boat capsizes 2. Ventilation grate collapses at South Korea Concert 3. A singer passes away after battling cancer 4. Murder suspect arrested 5. South Korea's Seong-Jin Cho wins career-paving 17th Frederic Chopin Piano Competition Set C 1. Camping site kills 5 in Incheon 2. Elderly lady arrested in poison case 3. MERS outbreak in Korea 4. An Asiana Airlines Boeing 777 aircraft flight travelling from Seoul crash landed in San Francisco 5. Helicopter crashes in a residential area in Korea South Korea helicopter crash

4.2.6.2.2. Receptive Test

In the form-oriented receptive test, the participants were asked to complete sentences by filling in blanks using appropriate tenses of given verbs (see Appendix J). There were 10 target items and 30 distractors. The sentences used in the receptive tests were selectively chosen from various media outlets, including the BBC, CBS News, the Telegraph, the Guardian, USA Today, Forbes, the New York Times, the Washington Post, ITV and Fox News. The target items contained two blanks; one blank for the present perfect and the other for the past simple. Of the 10 target items, half of the questions required the active voice in the present perfect and the other half demanded the passive voice in the present perfect. The distractors also consisted of two blanks to be completed with appropriate verb tenses: (a) If/unless conditionals (10 items), (b) time clauses (10 items) and (c) subjunctives (10 items).

4.2.6.3. Working Memory Measures

Six different working memory measures were used to test different constructs of working memory capacity. Phonological short-term memory was assessed with a non-word span task, and a forward corsi block test was used to determine visual spatial short-term memory. The executive functions of updating, task-switching and inhibitory control were measured by an automated operation span (AOSPAN) task, a colour shape task and a stop signal task, respectively. A backward corsi block test was employed as additional measure of executive control, an updating function. The nonword span task was presented using PowerPoint and the participants' answers were recorded using a voice recorder. Other working memory measures, colour shape task, stop signal task, forward and backward corsi block tests and AOSPAN, were administered on a computer using Inquisit 4 Lab (Millisecond, 2015).

4.2.6.3.1. Nonword Repetition Span

The non-word span task in Korean (Jung, 2017) was administered to assess participants' phonological short-term memory. In this task, the participants were presented with 32 nonwords, which were developed based on the phonotactic rules of Korean, each with a length of 4 to 11 syllables. There were four sets for each syllable length. The nonwords were presented orally in a random order as for the participants to listen and recall them as correctly as possible. Approximately 10 seconds was allowed for the participants to give their responses. The participants' responses were audio recorded.

4.2.6.3.2. Forward Corsi Block Task

The forward corsi block task was used to measure visuospatial short-term memory capacity. In the forward corsi block task, 2 to 9 blocks were presented on a computer screen. The blocks were highlighted in different patterns, starting from 2 and going to 9 blocks, for the participants to click the blocks in the same order as they were highlighted. The participants had two trials for each block length (Figure 19).

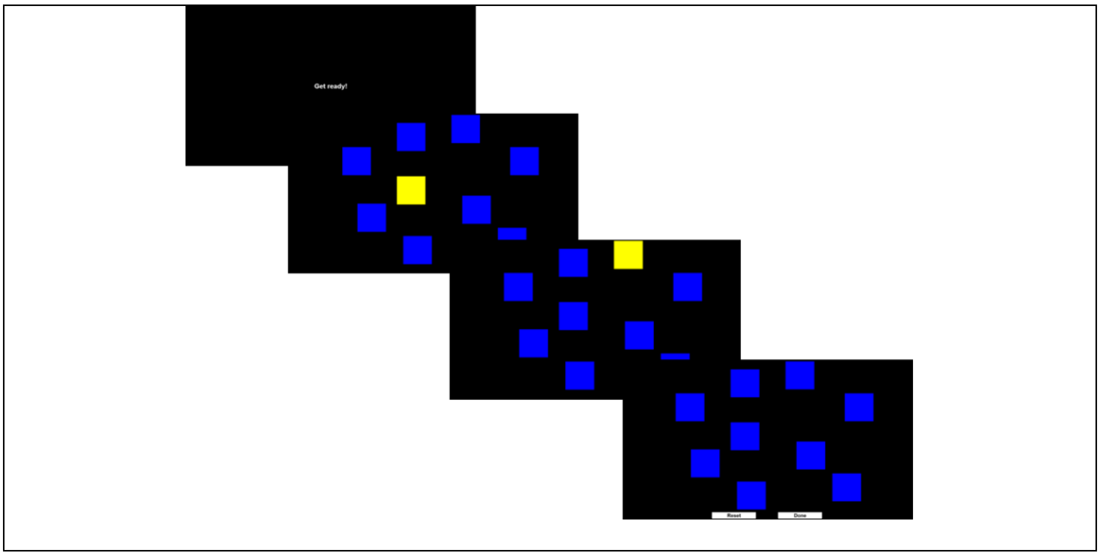


Figure 19. Corsi block Task

4.2.6.3.3. Backward Corsi block Task

The backward corsi block task was included to assess the updating function of executive control. The format of the task was same as for the forward corsi block task; however, in the backward corsi block task, the participants had to click the blocks in reverse order as they were highlighted. There were also two trials for each block length.

4.2.6.3.4. Colour Shape Task

To measure one of the executive functions, task-switching ability, a colour shape task (Miyake, Emerson, Padilla & Ahn, 2004) was employed. There were two different blocks, namely, non-switching block and switching block. In the non-switching block, there were two separate sub-blocks, either a colour or a shape block. In the colour block, the participants were asked to provide a response depending on the colour of the stimulus; when a stimulus was given on the computer screen, the participants were asked to press “A” for green and “L” for red. Similarly, the participants had to press “A” for a triangle and “L” for a circle in the shape block in which they had to respond on the basis of the shape of the stimulus (Figure 20). In switching blocks (mixed blocks), however, the participants were required to make a decision based on either a colour or a shape according to a cue letter that appeared on the screen. More specifically, the stimulus was presented randomly, with a cue letter of either ‘C’ or ‘S’ on the screen. If a stimulus with the cue letter ‘C’ appeared on the screen, the participants had to identify the colour of the stimulus (i.e., green or red), irrespective of its shape. On the other hand, if a stimulus appeared with the cue letter ‘S’, they had to make a decision about the shape (i.e., circle or triangle), irrespective of its colour (Figure 21).

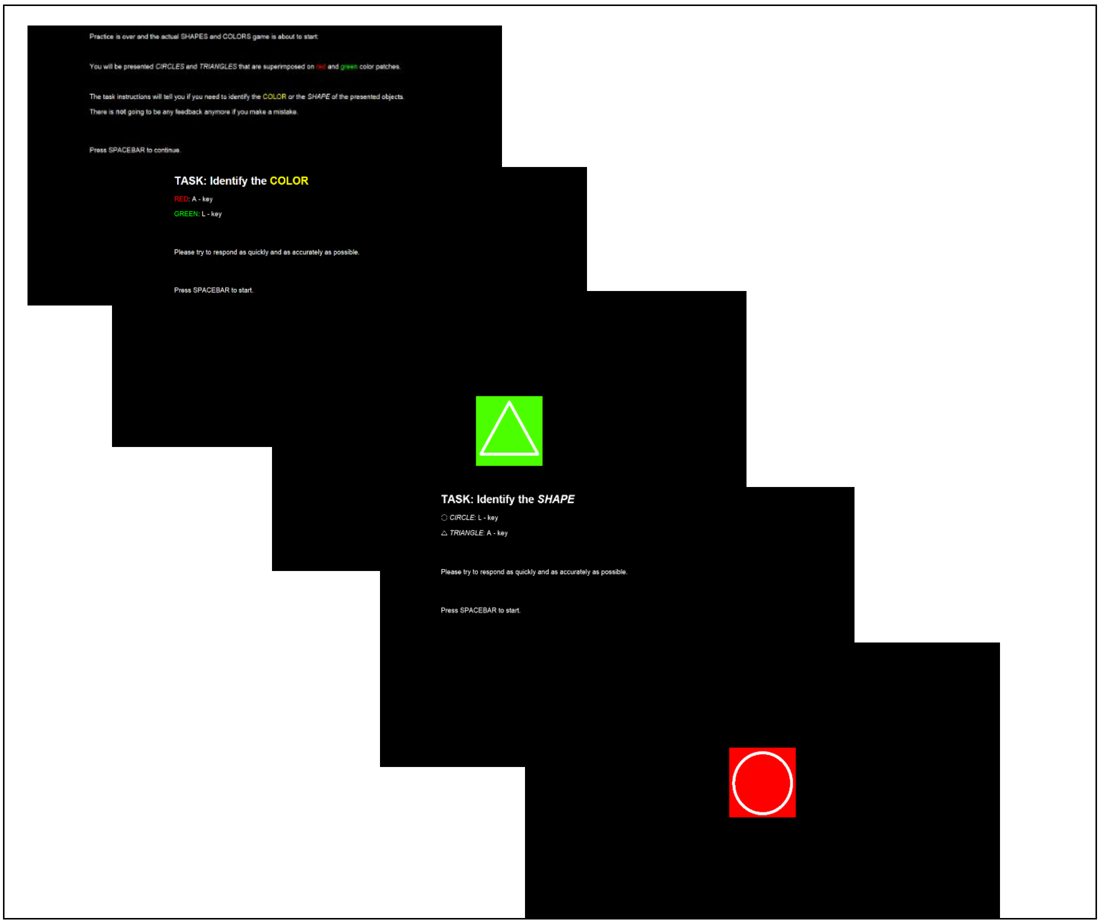


Figure 20. Colour shape Task – Non-switching block

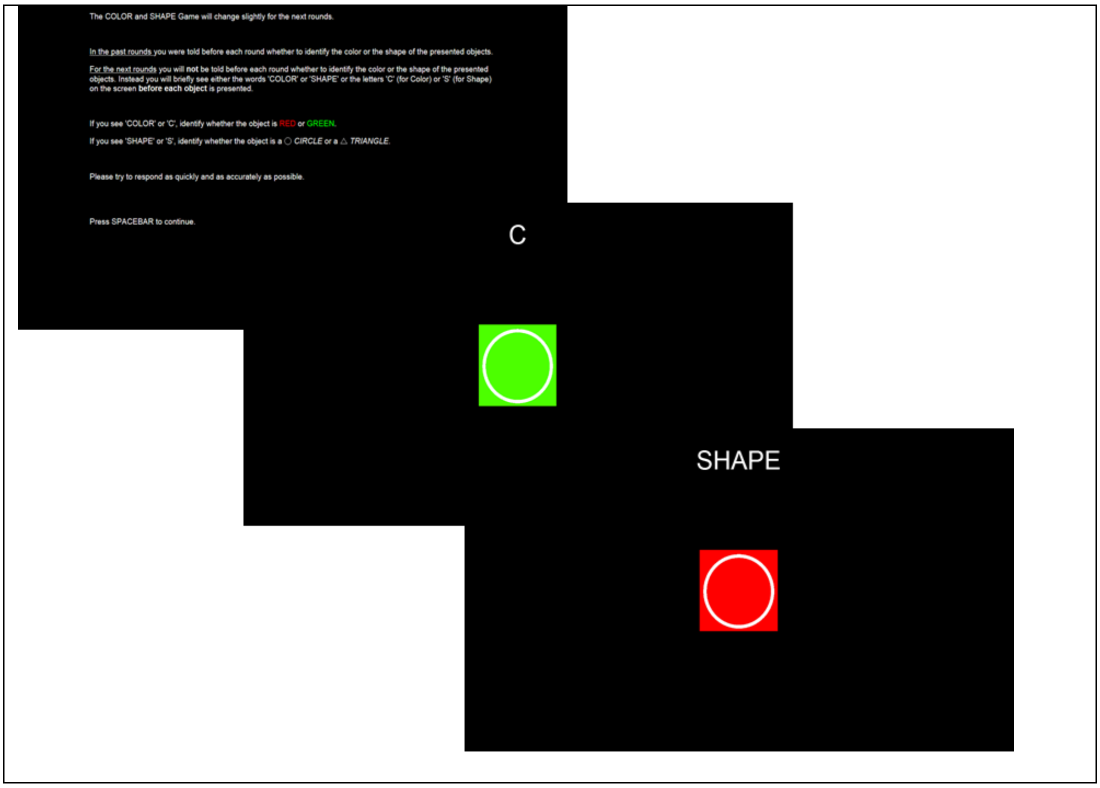


Figure 21. Colour shape Task – Switching block

4.2.6.3.5. Stop Signal Task

A stop signal task was used as a measure of inhibitory control. The participants were asked to respond to a given stimulus as quickly and correctly as they could. On the screen, an arrow stimulus pointing either left or right was displayed; the participants had to respond by pressing ‘D’ on the keyboard if the arrow pointed to the left and ‘K’ on the keyboard if the arrow pointed to the right. Some arrow stimuli were accompanied by an auditory signal (a beep). In this case, the participants were asked to withhold their response (Figure 22).

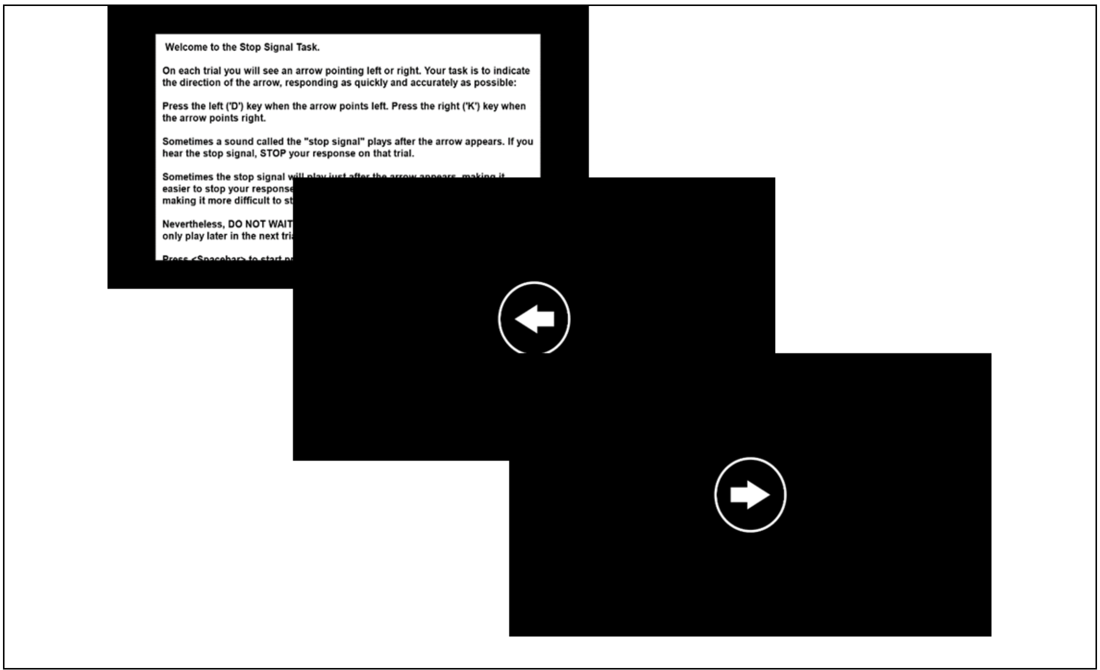


Figure 22. Stop Signal Task

4.2.6.3.6. Automated Operation Span Task (AOSPAN)

As a measure of one of the executive functions, updating, an automated version of the operation span task (Turner & Engle, 1989) was used. This task required the participants to solve a series of simple arithmetic equations and recall sets of alphabet letters. More specifically, an arithmetic equation was presented, followed by a letter (e.g., F, T, K, L, Q, P). The participants were asked to make a judgement about whether a given answer to the arithmetic equation was correct or incorrect and remember the letter. At recall, 12 possible letters in a 4 x 3 matrix were displayed for the students to select letters in the order they had appeared. The number of correctly recalled letters, ranging from 3 to 7, was defined as the set size for each participant. The task included three sets of each set size, and different set sizes were presented in a random order. While performing the task, the participants were instructed to make judgements as accurately and quickly as possible for the maths operations while remembering the letters to be recalled. The participants were provided with feedback on the accuracy of their responses after completing each equation (Figure 23). To ensure that the participants did not trade off between solving maths operations and remembering letters, they were instructed to maintain their maths accuracy at or above 85 per cent, as recommended in previous studies (e.g., Unsworth et al., 2005).

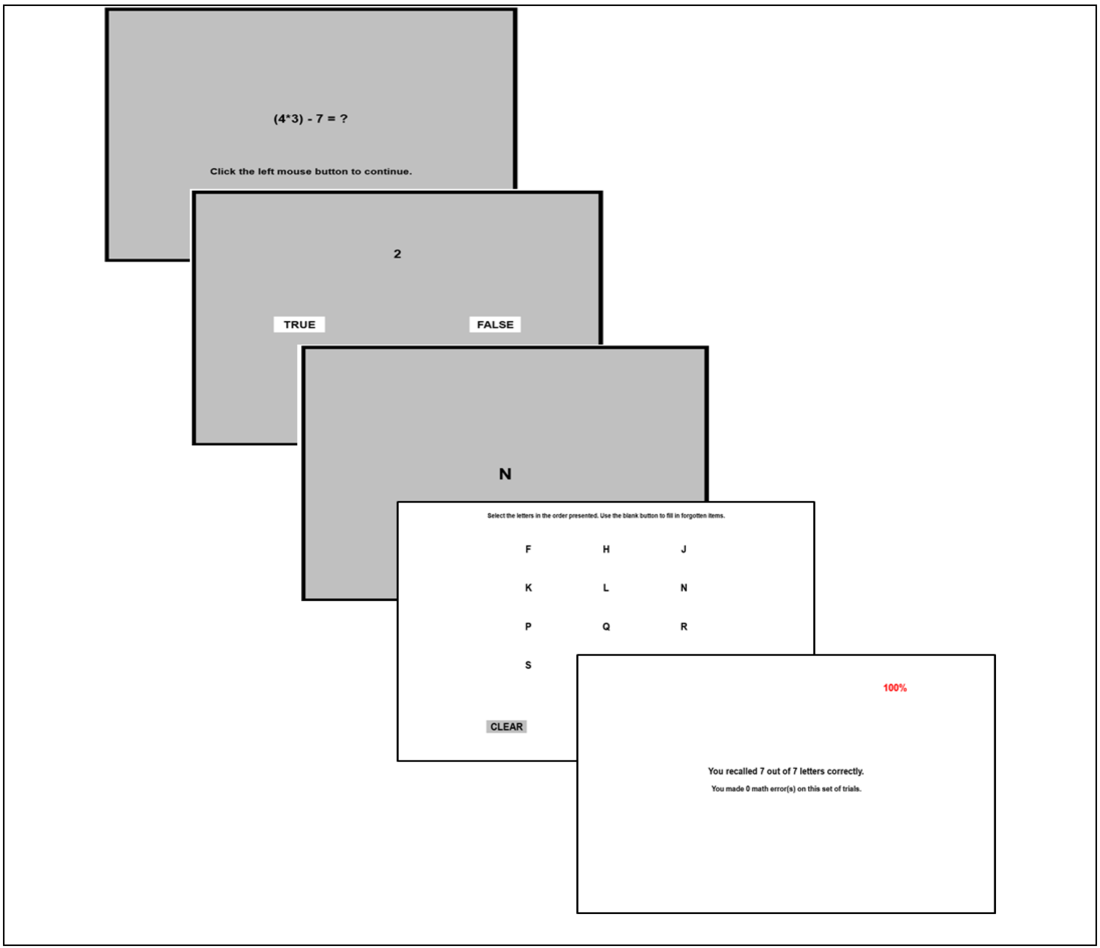


Figure 23. Automataed Operation Span Task

4.2.6.4. Background Questionnaire

A background questionnaire was developed to obtain general information about the participants, including age, major, length of time studying English, prior experience of living in an English speaking country and other languages studied (see Appendix F).

4.2.6.5. Exit Questionnaire

An exit questionnaire was administered at the end of the experiment that included questions about: (a) whether participants recognized the purpose of the study, (b) whether they focused on any particular grammatical features during the study, and (c) whether they consulted any other sources outside the study. In addition, an open question was included to obtain information about the students' perspectives regarding the usefulness of captions, textual enhancement and multimodal input-based tasks in general (see Appendix G).

4.2.7. Data Collection Procedure

The total duration of data collection was approximately five months, holding each session individually. Each participant was required to partake in four sessions, as shown in Figure 24. In the first session, a general introduction to the study was provided to the participants and their consent to take part was obtained. After completing a background questionnaire, an Oxford Placement Test was administered, which lasted for approximately 40 minutes. During the first session, the participants were also presented with a pretest in the same order as the oral productive test, a written productive test and a receptive test. The participants' responses on the oral productive test were recorded using a voice recorder, while written responses were collected for the written productive test. The duration of both the oral productive test and the written productive test was approximately between 15 and 18 minutes. The receptive test was administered as a paper-and-pencil test lasting approximately 40 minutes. The same procedure was used for an immediate posttest and a delayed posttest.

In the second session, the participants completed a series of multimodal input-based tasks, including 24 news clips, followed by an immediate posttest. While performing the tasks, an eye-tracker was used to record participants' eye gaze. The total duration of 24 treatment tasks was between 13 minutes and 15 minutes. In session 3, working memory measures were administered to the participants. The duration of each WM measure varied: (a) approximately 9 to 10 minutes for the NWS, (b) approximately 40 to 45 minutes for the CST, (b) approximately 9 to 10 minutes for the SST, (c) approximately 4 to 5 minutes for the forward corsi block task, and the backward corsi block task, respectively and (d) approximately 30 to 40 minutes for the AOSPAN. The order of the working memory tests was counterbalanced across participants. Session 4 took place a month later in which the participants were asked to complete a delayed posttest and an exit questionnaire. Each session lasted approximately from 2 hours to 3 hours.

4.2.8. Eye-tracking Procedure

Following the procedure used in Study 1, a remote eye-tracker, a Tobii X2-60 with a temporal resolution of 60 Hz, was used. The experiment was conducted individually with the eye-tracker mounted on a laptop computer with a 15-inch screen in a quiet room. The participants were seated approximately 60 cm from the computer screen. The eye-tracking system was calibrated before each set of treatment tasks, using nine points for calibrating each eye. The participants were asked to maintain the same position as much as possible while performing the tasks. The experiment was conducted using Tobii Studio 3.3.1 software (Tobii Technology, 2015).

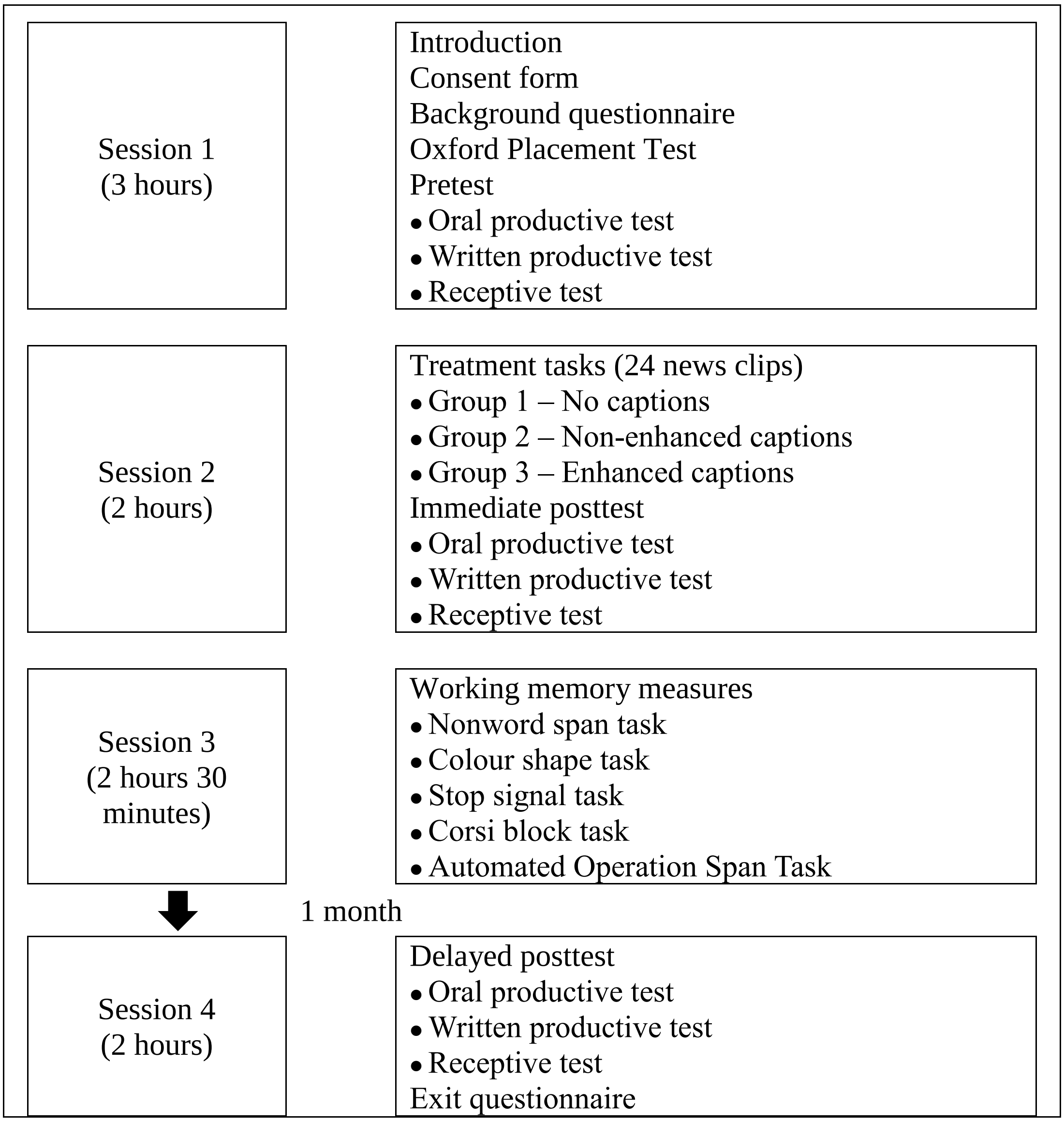


Figure 24. Data Collection Procedure for Study 2

4.2.9. Scoring and Data Analysis

4.2.9.1. Scoring

4.2.9.1.1. Oxford Placement Test

The listening section and the grammar section of the Oxford Placement Test each included 100 questions. One point was given for each correct response, resulting in a total score of 200.

4.2.9.1.2. Oral Productive Test and Written Productive Test

For both the oral productive test and the written productive test, a partial scoring procedure was used. For each obligatory context of the present perfect, the maximum score was 2 (Figure 25). Supplying the correct form of the present perfect was awarded the maximum score of 2. For cases where a partially correct form of the present perfect was used (e.g., correct use of have/has with incorrect from of past participle or incorrect use of have/has with reference to the subject), a partial score (1 point) was given. No points were allocated if different tenses, such as the present tense, were used.

2 points	Use of correct form of the present perfect tense e.g. A Belfast hospital **has carried** out five kidney transplants in a single day e.g. Two people **have been hurt** in an accident.
1 point	Use of the present perfect tense with incorrect past participle form e.g. Two people **have been hurted** in an accident.
0 point	Use of different tenses, such as the future tense or present tense. e.g. A Belfast hospital **carries** out five kidney transplants in a single day e.g. Two people **are hurt** in an accident.

Figure 25. Scoring System for Present Perfect

For the past simple tense, however, Suppliance in Obligatory Contexts was calculated to assess the accurate use of the past simple in contexts where it was required (Brown, 1973; Dulay & Burt, 1974; Pica, 1983). Different from the present perfect, which was supposed to be used once at the beginning of each news item, the number of past simple tenses used in the productive test varied for individual participants. Hence, the obligatory contexts method, using the formula below, was employed to identify how many uses of the past simple in obligatory contexts were correct.

$$\frac{\text{(Number of correct items in obligatory contexts x 2) +}\ \text{(Number of partially correct items in obligatory contexts x 1)}}{\text{(Number of total obligatory contexts x 2)}}$$

4.2.9.1.3. Receptive Test

In the receptive test, each question included two blanks, one for the present perfect tense and the other for the past simple. For both the present perfect and the past simple, a partial scoring system was used. The same criterion used for the productive test was applied for the present perfect: 2 for a correct form of the present perfect, 1 for a partially correct form of the present perfect (e.g., correct use of have/has with an incorrect form of the past participle or incorrect use of have/has with reference to the subject) and 0 for other tenses, such as the present. For the past simple, the maximum score of 2 was given for the use of a correct form of the past simple tense. For the use of the past simple with an incorrect verb form (e.g., *hurted), a partial score, 1, was given. In cases of using different tenses other than the past simple tense, such as the present tense, 0 point were given. In the test, 10 target items along with 30 distractors were included; thus, the maximum total score for the target items, excluding distractors, was 20 for both the present perfect and the past simple.

4.2.9.2. Eye movement Data

To analyse eye-movement data, the same procedure used in Study 1 was employed, but with the difference being in the interest in target constructions. In

Study 2, two types of areas of interest were selected: one for present perfect and the other for past simple forms (Figure 26). The attention paid to the target linguistic constructions was examined using four measurements – first pass reading, second pass reading, total fixation duration and number of visits (Godfroid, Boers, & Housen, 2013; Keating, 2009; Winke, 2013).

Figure 26. Areas of Interest for Study 2

4.2.9.3. Working Memory Measures

4.2.9.3.1. *Nonword Span Task*

Each of the nonword recalls was scored either correct or incorrect. Span length was determined as the maximum number of syllables that participants correctly recalled at least twice for each syllable-length, ranged from 4 to 11.

4.2.9.3.2. *Colour Shape Task*

For the colour shape task, reaction time of individual participants were trimmed to values within two standard deviations above and below the mean reading time as a preliminary step. Then, the switching cost was used to determine the participants' task switching ability by calculating differences in mean reaction time between non-switching blocks and switching blocks (e.g., Altgassen et al., 2014; Friedman et al., 2006; Gold et al., 2013; Miyake et al., 2004).

4.2.9.3.3. *Stop Signal Task*

The stop signal reaction time (SSRT) was used as an individualized measure of inhibitory control. More specifically, SSRT is an estimate of the stop signal reaction time, measured in milliseconds, which refers to the time required for an individual participant to inhibit the response and thus the strength of their inhibitory ability. Participants with slower SSRT are more likely to exhibit difficulties in inhibiting their responses, whereas participants with faster SSRT tend to show easiness in inhibiting their responses. SSRT was also trimmed to values within two standard deviations above or below the mean stop signal reaction time (Congdon et al., 2012; Enticott, Ogloff & Bradshaw, 2006).

4.2.9.3.4. *Corsi Block Task*

For both the forward and backward corsi block tasks, total score, which refers to the number of correctly repeated sequence until the test ends, was employed. The total score system was considered more reliable as it "takes into account the performance on both trials of an equal length and thus is more reliable than the block span alone" (Kessels et al., 2000, p. 254).

4.2.9.3.5. *AOSPAN*

To measure performance on AOSPAN, two different scoring systems have been used in previous research: absolute AOSPAN score and total score. The absolute AOSPAN score only represents sets with all letters recalled correctly. If an individual correctly recalled 3 letters in a size of 3, 4 letters in a size of 4, and 3 in a set size of 5, his or her Ospan score would be 7 (3+4+0). The total score reflects the total number of letters recalled in their correct positions within a particular string. In this study, the total scoring system was used as an index of AOSPAN, with a maximum score of 75, since some researchers have suggested that 'total performance score' better reflects working memory span, which includes continuous variables (e.g., Friedman & Miyake, 2005, Miyake, 2001; Miyake, Emerson, & Friedman, 1999).

4.2.10. Statistical Analyses

For Study 2, SPSS 24.0 (Statistical Package for the Social Sciences) for Windows was used to calculate descriptive statistics and linear mixed-effects models, with the help of the statistical package R (R development core team, 2016), were employed to compute inferential statistics.

4.2.10.1. Descriptive Statistics

Descriptive statistics were computed for the participants' performance on the three assessment measures, six working memory measures and eye-movement data.

4.2.10.2. Preliminary Analyses

Before conducting the main analyses, preliminary analyses were undertaken to ensure the reliability of the tests and the validity of the results. The internal consistency reliability of the Oxford Placement Test, the receptive test and the experimental treatment task was examined using Cronbach's alpha. For the productive tests -- oral productive test and written productive test – inter-coder reliability was determined with Cohen's kappa.

4.2.10.3. Main Statistical Analysis

The main statistical analyses were conducted with linear mixed-effects models using the statistical package R (R development core team, 2016). As a first step, the comparability of the three groups at the outset of the study was examined using the pretest scores. Then, a series of linear mixed-effects models was constructed to explore: (a) the extent to which multimodal input-based tasks without captions versus with captions affected development in L2 grammatical knowledge, (b) the effects of textually non-enhanced captions versus enhanced captions on development in L2 grammatical knowledge, (c) the effects of textually non-enhanced captions versus enhanced captions on drawing learners' attention to the target linguistic constructions and (d) the moderating effect of working memory on the relationship between type of captioning and development in L2 grammatical knowledge. To examine the extent to which attentional allocation was related to the development in L2 grammatical knowledge and individuals' working memory capacity, Pearson correlation analyses in R were performed.

4.2.10.3.1. Mixed-effects Models

In the field of second language acquisition, parametric statistics, such as *t*-tests or ANOVA, have been widely used (Norris & Ortega, 2000). However, with the advances in statistical analysis techniques, considerable attention has been paid to the use of mixed-effects models. One of the advantages of using mixed-effects models is their capacity to generate results that are generalizable. In mixed-effects models, fixed and random effects are included (Cunnings, 2012); fixed effects are independent variables and random effects are variance that is attributed to the random selection of participants or items in the study. The fact that random variance of participants and items is taken into consideration in mixed-effects models allows one to generalize the results beyond the participants and items included in the study (Baayen, 2008; Cunnings, 2012; Gagné & Spalding, 2009; Linck & Cunnings 2015; Rogers, 2016; Winter, 2013). Mixed-effects models provide another advantage in that both continuous factors (e.g., test scores or reaction times) and categorical factors (e.g., correct and incorrect responses) can be included as fixed effects, using either linear mixed-effects models or logic mixed-effects models, respectively.

Given that the current study included the random recruitment of participants and random selection of items, taking random variances into account using mixed-effects models seemed to produce more valid data analyses. Thus, in this study, mixed-effects models were mainly used for statistical analysis with the help of R package '*lme4*' (Bates, Maechler, & Bolker, 2012). Considering that interval-scale data were collected, which included test scores, reaction times, fixation durations, total numbers of visits and scores of different working memory measures in this study, linear mixed-effects models were constructed. As linear model summaries only provide *t* statistics without *p*-values, absolute *t*-values above 2.0 were set as a criterion for testing the significance of the models (Gelman & Hill, 2007). In addition, *lmertest* was also used, when it was applicable, to obtain *p*-values.

4.2.10.3.2. Mixed-effects Models in the Current Study

The first series of mixed-effects model analyses were constructed to determine the comparability of the three groups – no captions group, non-enhanced captions group and enhanced captions group – at the outset of the study. For these analyses, the three groups' scores obtained on the pretests were included as dependent variables. To the null model which only included random effects (i.e., participant and item), group was added as a fixed effect to determine whether there were statistically significant differences among the three groups at the pretest stage.

The mixed-effects models constructed to address the research questions consisted of two independent variables, group and time, as fixed effects. As for random effects, intercepts for participants and items were included in the models. In addition, since the use of maximal models, which include a corresponding random slope for each fixed effect, is recommended (Barr, Levy, Scheepers & Tily, 2013), by-participant and by-item random slopes for the fixed effects were also added to the models: time as a random slope by participant and group as a random slope by item (Barr et al., 2013). However, if the maximal models failed to converge, the random effect that accounted for the least variance was removed until convergence was achieved (Blom, Paradis, & Sorenson Duncan, 2012; Cunnings & Sturt, 2014).

To determine (a) the extent to which multimodal input-based tasks without captions or with captions affected the development in L2 grammatical knowledge (RQ1) and (b) the extent to which textually non-enhanced versus enhanced captions affected development in L2 grammatical knowledge (RQ2), a null model was first constructed for each assessment measure. The null model included test score as a dependent variable and random effects (i.e., participant and item). To this model, fixed effects, group and time, were added. Then, the model including the fixed effects was compared to the null model with $\chi 2$ statistic. If a significant difference was revealed, the fixed effects were interpreted as having a significant relationship with the dependent variable. After identifying the fixed effects that improved the null models significantly, maximal random effects structures, including fixed effects and random effects as well as random slope for each fixed effect, were constructed (Barr et al., 2013). However, as aforementioned, if the maximal random structure models failed to converge, the random effects that accounted for the least variance were removed from the model (Cunnings & Sturt, 2014) until convergence was achieved. The linear mixed-effects model that reached convergence was used for subsequent analyses to determine whether there was an overall group by time-interaction effect. If a significant overall interaction effect was revealed, post hoc analyses were performed for pairwise comparison. The post-hoc models had the same structure, but each only included data from only two testing times for two of the groups. For all models, the effect size was calculated with the command 'r.squared GLMM' from the "*MuMin*' package. R^2 values of .06, .16 and .36 were interpreted as small, medium and large, respectively (Plonsky & Oswald, 2014).

To address RQ 3, which was concerned with the extent to which textually non-enhanced and enhanced captions in the multimodal input-based task drew learners' attention to the target linguistic constructions, a series of linear mixed-effects models was constructed, each having attention measurement (i.e., first pass reading, second pass reading, total fixation duration and total number of visits) as a dependent variable. The same procedure used for the first and second research questions was applied, differing only in the fixed effect; that is, group was the only fixed effect included in the model. As learners' attention allocated to the target linguistic constructions was measured once, time was not a variable in this analysis.

For RQ 4, Pearson correlation analyses using R were performed to examine the relationship between the amount of attention the participants allocated to the target linguistic constructions and the gains participants displayed on each assessment measure. Correlations between the four attention measurements (i.e., first pass reading, second pass reading, total fixation duration and number of visits) and gain scores on the three assessment tasks (i.e., oral productive test, written productive test and receptive test) were examined.

RQ 5 asked the extent to which individual differences in working memory capacity moderated the effects of no captions, non-enhanced and textually enhanced captions in multimodal input-based tasks on L2 development. To answer this question, the linear mixed-effects models that were previously constructed for research questions one and two were extended. To these existing linear mixed-effects models, including group and time as fixed effects, each of the working memory measures was entered as an additional fixed effect in separate analyses to examine the moderating effects of working memory. If the model displayed a three-way group by time by working memory interaction effect, the result was interpreted as indicating that working memory capacity differentially affected the extent of gains participants displayed in the various groups. The convergence issue of the maximal random-structure models was resolved following the same steps mentioned above. If a significant overall interaction effect was revealed, post hoc analyses were performed with data from two testing times of two groups to determine where differences existed among the groups. In addition, a series of correlation analyses was conducted to further examine the moderating effect of working memory on the learning gains of each group.

To identify the relationship between the amount of attention the participants allocated to the target linguistic constructions and their working memory capacity (RQ 6), Pearson correlation analysis in R package was employed to investigate the relationships between the four attention measurements (i.e., first pass reading, second pass reading, total fixation duration, and number of visits) and indices of the six working memory measures (i.e., nonword span task, forward corsi block task, backward corsi block task, colour switching task, stop signal task, and automated operation span task). To ensure the normality of distribution with regard to assumptions, skewness and kurtosis ratios were taken into consideration with values more than 2 and less than 2 as the criterion (Larson-Hall, 2016). For Pearson correlation analyses, an alpha level of $p < .05$ was set to avoid Type II error in consideration of the relatively small sample size included in this study, and r values of .25, .40 and .60 were considered to be small, medium and large, respectively (Plonsky & Oswald, 2014).

4.3. Analyses and Results

4.3.1. Preliminary Analyses

4.3.1.1. Internal Consistency Reliability

Table 23 summarizes the Cronbach's alpha values computed to identify the internal consistency reliability of the Oxford Placement test, the three versions of the receptive test, and the experimental treatment task. Cronbach's α for each test was found to be acceptable.

Table 23. Internal Consistency Reliability

	N	*M*	*SD*	*α*
Oxford Placement Test - Listening	72	89.80	5.15	.719
Oxford Placement Test - Grammar	72	88.20	4.72	.662
Receptive Test – A	72	18.74	4.42	.660
Receptive Test – B	72	24.35	6.87	.677
Receptive Test – C	72	24.17	7.03	.750
Experimental treatment task	72	20.87	2.72	.661

Oxford Placement Test – Listening: Max. = 100; Oxford Placement Test Grammar: Max. = 100; Receptive Test: Max. = 20; Experiment Treatment Task: Max. = 27

4.3.1.2. Inter-coder Agreement

The productive tests – oral productive test and written productive test – were coded and scored by the researcher and a second coder. The second rater, who has a

PhD degree in Applied Linguistics, scored 30 per cent of both the oral productive test and the written productive test. Intercoder agreement was determined by calculating Cohen's kappa. The results of the inter-coder agreement analyses are shown in Table 24. The inter-coder agreement for the oral productive test was .98 with a kappa of .96, and 1.00 for the written productive test with a kappa of 1.00, indicating a high level of agreement between the two raters beyond chance.

Table 24. Inter-coder Agreement on Oral and Written Productive Test

	Inter-coder Agreement	Cohens' kappa
Oral Productive Test	.98	.96
Written Productive Test	1.00	1.00

4.3.1.3. Comparability of the Groups

Table 25 and Table 26 summarize the descriptive statistics for each group's performance on the pretests – oral productive test, written productive test and receptive test – for the present perfect and past simple items, respectively.

Table 25. Descriptive Statistics for the Pretest – Present Perfect

					95% CI	
		N	*M*	*SD*	Lower	Upper
Oral productive test	No captions	24	1.833	2.277	.872	2.795
	Non-enhanced captions	24	1.667	2.014	.816	2.517
	Enhanced Captions	24	1.708	1.756	.967	2.450
Written productive test	No captions	24	1.583	2.106	.695	2.472
	Non-enhanced captions	24	1.917	2.669	.790	3.044
	Enhanced captions	24	1.667	2.200	.738	2.596
Receptive test	No captions	24	2.417	3.537	.923	3.910
	Non-enhanced captions	24	2.833	2.899	1.609	4.058
	Enhanced captions	24	2.792	3.230	1.428	4.156

Oral productive max. score = 10; Written productive max. score = 10; Receptive max. score = 20

Table 26. Descriptive Statistics for the Pretest – Past Simple

		N	*M*	*SD*	95% CI Lower	95% CI Upper
Oral productive test	No captions	24	4.724	.368	4.568	4.879
	Non-enhanced captions	24	4.627	.411	4.453	4.801
	Enhanced captions	24	4.601	.753	4.283	4.919
Written productive test	No captions	24	4.762	.456	4.570	4.954
	Non-enhanced captions	24	4.601	1.031	4.165	5.036
	Enhanced captions	24	4.309	1.396	3.719	4.898
Receptive test	No captions	24	15.625	2.428	14.600	16.650
	Non-enhanced captions	24	16.333	2.408	15.317	17.350
	Enhanced captions	24	16.208	2.395	15.197	17.220

Oral productive max. score = 5; Written productive max. score = 5; Receptive max. score = 20

The linear mixed-effects model constructed to investigate whether there were significant differences among the three groups at the time of the pretest found no differences for either the present perfect items (Table 27) or the past simple items (Table 28). Thus, the groups were considered comparable with regard to the scores on the three pretests at the outset of the study.

Table 27. Results for the Linear Mixed-effects Model Examining Performance on Three Pretests – Present Perfect

		Fixed effects				Random effects	
						by participant	*by Item*
		Estimate	*SE*	*t*	r^2	*SD*	*SD*
Oral productive	Intercept	.361	.128	2.809		.239	.067
	Group	−.004	.058	−.072	<.01		
	Formula: OralPP ~ Group + (1|Participant) + (Group|Item)						
Written productive	Intercept	.328	.150	2.177		.365	.091
	Group	.008	.067	.124	<.01		
	Formula: WrittenPP ~ Group + (1|Participant) + (Group|Item)						
Receptive	Intercept	.230	.102	2.254		.265	.064
	Group	.019	.046	.406	<.01		
	Formula: ReceptivePP ~ Group + (1|Participant) + (Group|Item)						

Table 28. Results for the Linear Mixed-effects Model Examining Performance on Three Pretests – Past Simple

		Fixed effects				Random effects	
						by participant	*by Item*
		Estimate	*SE*	*t*	*r²*	*SD*	*SD*
Oral productive	Intercept	.955	.041	23.289		.067	.053
	Group	−.021	.015	−.797	<.01		
	Formula: OralPS ~ Group + (1|Participant) + (Group|Item)						
Written productive	Intercept	.974	.057	17.080		.159	.032
	Group	−.024	.025	-.958	.01		
	Formula: WrittenPS ~ Group + (1|Participant) + (Group|Item)						
Receptive	Intercept	1.548	.087	17.804		.072	.139
	Group	.029	.035	.835	<.01		
	Formula: ReceptivePS ~ Group + (1|Participant) + (Group|Item)						

4.3.2. Effects of Multimodal Input-based Tasks without Captions, with Non-enhanced Captions or with Enhanced Captions on Development in L2 Grammatical Knowledge (RQs1–2)

4.3.2.1. Oral Productive Test

4.3.2.1.1. Present Perfect

As shown in the descriptive statistics (Table 29), all three groups exhibited pretest-immediate posttest gains on the oral productive test. The enhanced captions group attained the highest gain scores, followed by the non-enhanced captions group. On a delayed posttest, further gains were observed for the enhanced captions group, whereas both the no captions and the non-enhanced captions groups showed some reduction.

Table 30 summarizes the results for the linear mixed-effects model examining the participants' development in their use of the present perfect on an oral productive test. There was a statistically significant interaction effect between the two fixed effects, with a relatively small effect size.

Given that a group by time interaction effect was revealed, post-hoc models, having the same structure, were constructed to further compare the pretest-immediate posttest and pretest-delayed posttest gains displayed by the two groups. To examine the effects of captions on promoting L2 grammatical knowledge, the performance of the no captions group and the non-enhanced captions group from pretest to immediate posttest and from pretest to delayed posttest was first compared (RQ1).

Table 29. Descriptive Statistics for Oral Productive Test – Present Perfect

		M	Mean Gain	*SD*	95% CI Lower	95% CI Upper
No captions (*N* = 24)	Pretest	1.833	–	2.277	.872	2.795
	Immediate posttest	3.083	1.250	3.175	1.743	4.424
	Delayed posttest	2.250	.417	2.382	1.244	3.256
Non-enhanced captions (*N* = 24)	Pretest	1.667	–	2.014	.816	2.517
	Immediate posttest	3.750	2.083	3.404	2.313	5.187
	Delayed posttest	3.000	1.333	3.230	1.636	4.364
Enhanced captions (*N* = 24)	Pretest	1.708	–	1.756	.9667	2.450
	Immediate posttest	6.875	5.167	2.593	5.780	7.970
	Delayed posttest	8.000	6.292	3.752	3.535	5.298

Max. score = 10

Table 30. Results for the Linear Mixed-effects Model Examining Performance on Oral Productive Test – Present Perfect

		Fixed effects					Random effects by participant	Random effects by item
		Estimate	*SE*	*t*	*p*	*r²*	*SD*	*SD*
Oral productive	Intercept	.722	.208	3.472	<.001***		.464	.022
	Group	−.274	.096	−2.849	<.001***			
	Time	−.317	.077	−4.116	<.001***			
	Group*Time	.294	.036	8.235	<.001***	.04		

*Formula: OralPP ~ Group*Time + (1|Participant) + (1|Item)*

*** $p < .001$, ** $p < .01$, * $p < .05$

As shown in Table 31, for the present perfect, the results revealed no significant interaction between time and group. That is, the non-enhanced captions group did not achieve significantly greater gain scores than the no captions group on the oral productive test.

Table 31. Results for Post hoc Contrasts for No Captions Group and Non-enhanced-Captions Group on Oral Productive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Pretest ~ Immediate posttest	Intercept	.317	.393	.805	.422		.385	.000
	Group	−.200	.249	−.804	.422			
	Time	.083	.223	.374	.708			
	Group*Time	.167	.141	1.183	.237	.003		
Pretest ~ Delayed posttest	Intercept	.450	.289	1.556	.121		.352	.000
	Group	−.125	.183	−.684	.495			
	Time	−.050	.107	−.465	.642			
	Group*Time	.092	.068	1.348	.178	.03		

*Formula: OralPP ~ Group * Time + (1|Participant) + (1|Item)*

However, when the performance of the no captions group was compared against that of the enhanced captions group, there was a significant difference between the groups (Table 32); the pretest-posttest and pretest-delayed posttest gains score of the enhanced captions group were higher than those of the no captions group, with a relatively small effect size. Compared to no captions, only textually enhanced captions exhibited effectiveness in promoting the use of the present perfect, as measured by the oral productive test.

Table 32. Results for Post hoc Contrasts for No Captions Group and Enhanced Captions Group on Oral Productive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Pretest ~ Immediate posttest	Intercept	.521	.275	1.895	.059		.343	.000
	Group	−.404	.123	−3.288	$<.01^{**}$			
	Time	−.142	.159	−.890	.374			
	Group*Time	.392	.071	5.503	$<.001^{***}$	.14		
Pretest ~ Delayed posttest	Intercept	.631	.192	3.278	$<.01^{**}$		.283	.000
	Group	−.306	.086	−3.556	$<.001^{***}$			
	Time	−.252	.076	−3.326	$<.001^{***}$			
	Group*Time	.294	.034	8.667	$<.001^{***}$	.09		

*Formula: OralPP ~ Group * Time + (1|Participant) + (1|Item)*

*** $p < .001$, ** $p < .01$, * $p < .05$

To further examine the effects of textually enhanced captions against non-enhanced captions regarding developing L2 grammatical knowledge, the performance of the non-enhanced captions group and the enhanced captions group was compared using post hoc contrasts with pretest-immediate posttest and pretest-delayed posttest gain scores (RQ2). The analyses yielded significant time-group interaction effects for both pretest-immediate posttest and pretest-delayed posttest gains, with a small effect size (Table 33). These results mean that the enhanced captions group showed greater development in their use of the present perfect than the non-enhanced captions group on the oral productive test.

Table 33. Results for Post hoc Contrasts for Non-enhanced Captions Group and Enhanced Captions Group on Oral Productive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Pretest ~ Immediate posttest	Intercept	.633	.616	1.028	.305		.320	.000
	Group	−.442	.242	−1.827	.068			
	Time	−.567	.360	−1.574	.116			
	Group*Time	.533	.141	3.776	<.001***	.02		
Pretest ~ Delayed posttest	Intercept	1.175	.450	2.613	<.001**		.307	.076
	Group	−.487	.176	−2.772	<.001***			
	Time	−.858	.173	−4.995	<.001***			
	Group*Time	.496	.068	7.297	<.001***	.07		

*Formula: OralPP ~ Group*Time + (1|Participant) + (1|Item)*

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

4.3.2.1.2. Past Simple

Table 34 displays descriptive statistics for the participants' performance on the past simple. Considering the maximum score for the past simple on the oral productive test, which was 5, all three groups' pretest scores were considerably high and, consequently, their gains were more likely to be small. The mean gain scores of the non-enhanced captions group and the enhanced captions group obtained were in a similar range, from .131 to .178, whereas the no captions group exhibited the lowest gains.

Table 34. Descriptive Statistics for Oral Productive Test – Past Simple

			Mean		95% CI	
		M	Gain	*SD*	Lower	Upper
No captions	Pretest	4.724	--	.368	4.568	4.879
(*N* = 24)	Immediate posttest	4.766	.042	.437	4.582	4.951
	Delayed posttest	4.638	−.086	.431	4.456	4.821
Non-enhanced captions	Pretest	4.627		.411	4.453	4.801
(*N* = 24)	Immediate posttest	4.798	.171	.435	4.615	4.982
	Delayed posttest	4.778	.151	.347	4.632	4.924
Enhanced captions	Pretest	4.601	--	.753	4.283	4.919
(*N* = 24)	Immediate posttest	4.732	.131	.443	4.545	4.919
	Delayed posttest	4.779	.178	.358	4.268	4.930

Max = 5

The linear mixed-effects model, as summarized in Table 35, revealed that there was no significant interaction between the fixed effects (i.e., group and time) in terms of the participants' performance on the oral productive test. The results could thus be interpreted as indicating that none of the three groups showed a significant change in their use of the past simple from oral productive pretest to posttests.

Table 35. Results for the Linear Mixed-effects Model Examining Performance on Oral Productive Test – Past Simple

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Oral	Intercept	.974	.048	20.407	<.001***		.064	.017
Productive	Group	−.022	.022	−1.009	.315			.001
	Time	−.004	.020	−.219	.826		.001	
	Group*Time	.003	.009	.301	.763	.001		

*Formula: OralPS ~ Group*Time + (Time|Participant) + (Group|Item)*

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

4.3.2.2. Written Productive Test

4.3.2.2.1. Present Perfect

The extent to which multimodal input-based tasks without or with non-enhanced and textually enhanced captions influenced development in the use of L2 grammatical knowledge was also examined by a written productive test. As presented in Table 36, the mean gain score of the enhanced captions group was considerably higher than that of the non-enhanced captions group, which in turn was higher than that of the no captions group from pretest to immediate posttest. On the delayed posttest, however, all three groups displayed slight losses in comparison to the immediate posttest.

Table 36. Descriptive Statistics for Written Productive Test – Present Perfect

		M	Mean Gain	*SD*	95% CI Lower	95% CI Upper
No captions (*N* = 24)	Pretest	1.583	--	2.104	.695	2.472
	Immediate posttest	1.625	.042	1.715	.901	2.349
	Delayed posttest	1.580	−.003	1.767	.8371	2.330
Non-enhanced captions (*N* = 24)	Pretest	1.917	--	2.669	.790	3.044
	Immediate posttest	4.500	2.583	3.551	3.001	5.999
	Delayed posttest	4.250	2.333	3.791	2.649	5.851
Enhanced captions (*N* = 24)	Pretest	1.667	--	2.200	.738	2.596
	Immediate posttest	8.000	6.333	2.703	6.859	9.141
	Delayed posttest	6.500	4.833	3.550	5.001	7.999

Max Score = 10

A linear mixed-effects model including group and time as fixed effects was constructed to examine the overall interaction effect (Table 37). A maximal model was first constructed with intercepts for participants and items as well as by-participant (i.e., time) and by-item (i.e., group) random slopes; however, the random effect that accounted for the least variance (i.e., group) was removed to achieve convergence. The results revealed a significant interaction between the fixed effects of group and time, with a small effect size.

Table 37. Results for the Linear Mixed-effects Model Examining Performance on Written Productive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	*r²*	*SD*	*SD*
Written productive	Intercept	.433	.198	2.185	.032*		.378	.547
	Group	−.104	.091	−1.143	.257			
	Time	−.244	.100	−2.450	.017*		.218	
	Group*Time	.242	.046	5.232	<.001***	.04		

*Formula: WrittenPP ~ Group*Time + (Time|Participant) + (1|Item)*

*** $p < .001$, ** $p < .01$, * $p < .05$

Post hoc contrasts were conducted to examine whether the gains of the no captions group and the non-enhanced captions group statistically differed on the written productive test in relation to the present perfect (RQ1). According to Table 38, there was a significant difference between the mean gain scores of the two groups from pretest to immediate posttest and from pretest to delayed posttest, both with relatively small effect sizes.

Table 38. Results for Post hoc Contrasts for No Captions Group and Non-enhanced Captions Group on Written Productive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	*r²*	*SD*	*SD*
Pretest ~ Immediate posttest	Intercept	.783	.395	1.981	.053		.505	.047
	Group	−.458	.250	−1.835	.073			
	Time	−.533	.249	−2.140	.038*		.317	
	Group*Time	.525	.158	3.331	<.01**	.04		
Pretest ~ Delayed posttest	Intercept	.483	.293	1.651	.105		.414	.000
	Group	−.167	.185	−.900	.373			
	Time	−.233	.129	−1.809	.077		.179	
	Group*Time	.233	.081	2.860	<.01**	.03		

*Formula: WrittenPP ~ Group * Time + (Time|Participant) + (1|Item)*

*** $p < .001$, ** $p < .01$, * $p < .05$

Similar results were obtained for a comparison of performance between the no captions group and the enhanced captions group (Table 39). The differences in the pretest-posttest and pretest-delayed posttest gains were statistically significant, with small to medium effect sizes. The higher gains obtained by the non-enhanced captions group and the enhanced captions group indicated that the provision of captions, regardless of enhancement, had a positive effect on participants' use of the present perfect when this was measured by a written productive test.

Table 39. Results for Post hoc Contrasts for No Captions Group and Enhanced Captions Group on Written Productive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	*r²*	*SD*	*SD*
Pretest ~ Immediate posttest	Intercept	.954	.278	3.426	<.01**		.55	.09
	Group	−.629	.123	−5.110	<.001***			
	Time	−.646	.179	−3.780	<.001***			
	Group*Time	.637	.076	8.343	<.001***	.13		
Pretest ~ Delayed posttest	Intercept	.550	.211	2.601	.012*		.439	.026
	Group	−.233	.094	−2.471	.017*			
	Time	−.242	.103	−2.339	.023*			
	Group*Time	.242	.046	5.231	<.001***	.11		

*Formula: WrittenPP ~ Group * Time + (Time|Participant) + (1|Item)*

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

Further analysis with the post hoc models, including the gains of the non-enhanced captions group and the enhanced captions group, was conducted to examine to what extent textually enhanced captions affected the development of L2 grammatical knowledge (RQ2). Significant interaction effects across all testing times were found, with small effect sizes, as shown in Table 40.

Table 40. Results for Post hoc Contrasts for Non-enhanced Captions Group and Enhanced Captions Group on Written Productive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	*r²*	*SD*	*SD*
Pretest ~ Immediate posttest	Intercept	1.467	.663	2.210	.032*		.587	.097
	Group	−.800	.260	−3.081	<.01**			
	Time	−.983	.444	−2.213	.032*		.423	
	Group*Time	.750	.174	4.303	<.001***	.06		
Pretest ~ Delayed posttest	Intercept	.750	.510	1.469	.015		.495	.011
	Group	−.300	.199	−1.505	.139			
	Time	−.267	.265	−1.008	.319		.083	
	Group*Time	.250	.104	2.409	.020*	.02		

*Formula: WrittenPP ~ Group*Time + (Time|Participant) + (1|Item)*

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

This finding indicated that the gains of the enhanced captions group were significantly greater compared with the non-enhanced captions group. That is, the provision of textually enhanced captions was more effective than non-enhanced captions in promoting the participants' use of the present perfect, as measured by a written productive test.

4.3.2.2.2. Past Simple

Table 41 summarizes the descriptive statistics for the groups' performance on the written productive test in terms of the past simple. Similar to the oral productive test, all three groups obtained high scores on the pretest and exhibited relatively small gains across the tests. From pretest to immediate posttest, the enhanced captions group obtained the highest gains, followed by the non-enhanced captions group; the no captions group's gain was substantially low compared to the other groups. On a delayed posttest, however, a slight decrease in gains was observed for all three groups.

Table 41. Descriptive Statistics for Written Productive Test – Past Simple

			Mean		95% CI	
		M	Gain	*SD*	Lower	Upper
No captions (N = 24)	Pretest	4.762	--	.456	4.570	4.954
	Immediate posttest	4.769	.007	.394	4.603	4.936
	Delayed posttest	4.502	−.260	.968	4.093	4.911
Non-enhanced captions (N = 24)	Pretest	4.601	--	1.031	4.165	5.036
	Immediate posttest	4.778	.177	.348	4.631	4.925
	Delayed posttest	4.558	−.043	.827	4.209	4.907
Enhanced captions (N = 24)	Pretest	4.309	--	1.396	3.719	4.898
	Immediate posttest	4.645	.336	.541	4.417	4.874
	Delayed posttest	4.256	−.053	1.020	3.825	4.687

Max. score = 5

The results for the linear mixed-effects model showed nonsignificant interaction effects between the fixed effects (time and group) for the past simple items (Table 42). This can be interpreted as indicating that the presence of captions, irrespective of whether they were enhanced or not, did not have a statistically significant effect on developing the learners' use of the past simple tense when this was assessed with a written productive test.

Table 42. Results for the Linear Mixed-effects Model Examining Performance on Written Productive Test – Past Simple

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Written productive	Intercept	1.035	.079	13.075	<.001***		.208	.007
	Group	−.048	.037	−1.307	.196			.009
	Time	−.032	.033	−.974	.333		.083	
	Group*Time	.010	.015	.669	.506	.042		

*Formula: WrittenPS ~ Group*Time + (Time|Participant) + (Group|Item)*

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

4.3.2.3. Receptive Test

4.3.2.3.1. Present Perfect

Table 43 shows descriptive statistics for the participants’ performance on the present perfect when this was measured by a receptive test. The enhanced captions group obtained the highest mean gain score, followed by the non-enhanced group. The no captions groups exhibited considerably lower gains compared to these two groups. On a delayed posttest, only the no captions group maintained its gains while the non-enhanced captions group and the enhanced captions group exhibited slight reductions in their gains.

A linear mixed-effects model was constructed to identify an overall interaction effect between fixed effects, group and time. As recommended by Barr et al. (2013), a maximal model, which consisted of both random intercepts (i.e., participants and items) and a corresponding random slope for each fixed effect, was first constructed; however, it was restructured by removing the random effect that accounted for the least variance in the data (i.e., group) to achieve convergence. As shown in Table 44, a significant overall group-by-time interaction emerged with a medium effect size.

Table 43. Descriptive Statistics for Receptive Test – Present Perfect

			Mean		95% CI	
		M	Gain	*SD*	Lower	Upper
No captions (*N* = 24)	Pretest	2.417	--	3.537	.923	3.910
	Immediate posttest	2.750	.333	2.786	1.574	3.926
	Delayed posttest	2.792	.375	2.919	1.559	4.024
Non-enhanced captions (*N* = 24)	Pretest	2.833	--	2.899	1.609	4.058
	Immediate posttest	6.083	3.250	4.671	4.111	8.056
	Delayed posttest	5.208	2.375	5.267	2.984	7.432
Enhanced captions (*N* = 24)	Pretest	2.792	--	3.230	1.428	4.156
	Immediate posttest	13.208	10.416	4.539	11.291	15.125
	Delayed posttest	12.875	10.083	4.494	10.977	14.773

Max. score = 20

Table 44. Results for the Linear Mixed-effects Model Examining Performance on Receptive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Receptive	Intercept	.412	.124	3.314	$<.001^{**}$		.201	.082
	Group	−.137	.056	−2.431	$.017^{*}$			
	Time	−.272	.061	−4.465	$<.001^{***}$		.118	
	Group*Time	.243	.028	8.621	$<.001^{***}$	.12		

*Formula: ReceptivePP ~ Group*Time + (Time|Participant) + (1|Item)*

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

The presence of an overall interaction effect called for further analyses to compare the gains of the no captions group with the non-enhanced captions groups and the enhanced captions group and explore whether the provision of captions affected the participants' use of the target constructions (RQ1). As Tables 45 and 46 show, the results of post hoc contrasts indicated that there were significant group-by-time interaction effects with small effect sizes from pretest to posttest and from pretest to delayed posttest for both comparisons. In other words, the participants benefited from captions, regardless of textual enhancement, in developing their knowledge of the present perfect, when this was measured by a receptive test.

Table 45. Results for Post hoc Contrasts for No Captions Group and Non-enhanced Captions Group on Receptive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	*r²*	*SD*	*SD*
Pretest ~ Immediate posttest	Intercept	.643	.225	2.861	<.01**		.230	.077
	Group	−.434	.141	−3.071	<.01**			
	Time	−.351	.137	−2.566	.014*		.122	
	Group*Time	.384	.086	4.438	<.001***	.03		
Pretest ~ Delayed posttest	Intercept	.418	.177	2.363	.022*		.253	.078
	Group	−.195	.111	−1.764	.084			
	Time	−.127	.075	−1.687	.098		.102	
	Group*Time	.146	.047	3.066	<.01**	.011		

*Formula: ReceptivePP ~ Group * Time + (Time|Participant) + (1|Item)*

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

Table 46. Results for Post hoc Contrasts for No Captions Group and Enhanced Captions Group on Receptive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	*r²*	*SD*	*SD*
Pretest ~ Immediate posttest	Intercept	.599	.198	3.028	<.01**		.409	.054
	Group	−.390	.088	−4.425	<.001***			
	Time	−.423	.102	−4.143	<.001***			
	Group*Time	.456	.046	9.985	<.001***	.05		
Pretest ~ Delayed posttest	Intercept	.375	.165	2.275	.027*		.383	.092
	Group	−.152	.072	−2.097	.041*			
	Time	−.200	.055	−3.608	<.001***			
	Group*Time	.219	.025	8.827	<.001***	.06		

*Formula: ReceptivePP ~ Group * Time + (Time|Participant) + (1|Item)*

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

Further post hoc analyses were also carried out to compare the effects of textually enhanced captions with non-enhanced captions on the development of L2 grammatical knowledge (RQ2). Following the procedure used previously, post hoc models were constructed to compare the pretest-posttest and pretest-delayed posttest gains of the two groups (i.e., non-enhanced captions group and enhanced captions group). The results showed that there was a significant but small size difference in the gains between the two groups (Table 47), indicating that textually enhanced captions resulted in promoting the participants' L2 grammatical knowledge (i.e., present perfect).

Table 47. Results for Post hoc Contrasts for Non-enhanced Captions Group and Enhanced Captions Group on Receptive Test – Present Perfect

		Estimate	*SE*	*t*	*p*	r^2	*by participant SD*	*by Item SD*
Pretest ~ Immediate posttest	Intercept	.467	.455	1.026	.310		.380	.027
	Group	−.346	.178	−1.939	.058			
	Time	−.642	.295	−2.175	.035*			
	Group*Time	.529	.116	4.573	<.001***	.05		
Pretest ~ Delayed posttest	Intercept	.244	.346	.704	.485		.335	.078
	Group	−.108	.135	−.800	.428			
	Time	−.419	.157	−2.673	.010*			
	Group*Time	.292	.061	4.747	<.001***	.06		

*Formula: ReceptivePP ~ Group*Time + (Time|Participant) + (1|Item)*

*** $p < .001$, ** $p < .01$, * $p < .05$

4.3.2.3.2. Past Simple

Descriptive statistics for the receptive test in terms of the past simple are presented in Table 48. Unlike the present perfect, all three groups exhibited slight but similar gains from pretest to immediate posttest. The gains made by the three groups were maintained until the delayed posttest.

Table 48. Descriptive Statistics for Receptive Test – Past simple

		M	Mean Gain	*SD*	95% CI Lower	95% CI Upper
No captions (*N* = 24)	Pretest	15.625	--	2.428	14.600	16.650
	Immediate posttest	16.708	1.083	1.922	15.897	17.520
	Delayed posttest	16.750	1.125	2.090	15.867	17.633
Non-enhanced captions (*N* = 24)	Pretest	16.333	--	2.408	15.317	17.350
	Immediate posttest	17.250	.917	2.326	15.268	18.232
	Delayed posttest	17.458	1.125	1.955	16.633	18.284
Enhanced captions (*N* = 24)	Pretest	16.208	--	2.395	15.197	17.220
	Immediate posttest	17.042	.834	2.349	16.050	18.034
	Delayed posttest	17.417	1.209	2.165	16.502	18.331

A linear mixed-effects model was used to examine possible differences among the three groups' performance in relation to the past simple in a receptive test. A maximal model, including random effects and random slopes, was first constructed, as suggested by Barr et al. (2013); however, it was reconstructed by removing the random effect that accounted for the least variance, i.e., group, to resolve the convergence issue. The results show no statistically insignificant difference in the gains of the groups (Table 49). That is, neither non-enhanced captions nor the enhanced captions had a positive effect on developing L2 learners' grammatical knowledge, particularly in relation to the past simple item, when it was measured by a receptive test.

Table 49. Results for the Linear Mixed-effects Model Examining Performance on Receptive Test – Past Simple

		Fixed effects					Random effects	
							by participant	*by item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Receptive	Intercept	1.517	.120	12.590	<.001***		.167	.135
	Group	.021	.052	.418	.676			
	Time	.053	.046	1.147	.252		.007	
	Group*Time	.002	.021	.104	.917	.000		

*Formula: ReceptivePS ~ Group*Time + (Time|Participant) + (1|Item)*

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

4.3.3. Effects of Textually Non-enhanced Captions and Enhanced Captions on Allocation of Attention (RQ3)

To examine the effects of non-enhanced versus textually enhanced captions on drawing learners' attention to the target linguistic constructions (RQ3), the eye-gaze behaviour of the non-enhanced captions group and the enhanced captions group was compared. For the analyses, four different measurements were used as evidence of the amount of attention the participants allocated to the target linguistic constructions: (a) first pass reading, (b) second pass reading, (c) total fixation duration, and (d) total number of visits. On the basis of descriptive statistics for the measurements summarized in Table 50, which shows the mean for the individual present perfect item, the enhanced captions group fixated for longer and more frequently on the present perfect while performing the treatment tasks.

Table 50. Descriptive statistics for Attention Measurements – Present Perfect

					95% CI	
		N	*M*	*SD*	*Lower*	*Upper*
First pass reading	Non-enhanced captions	24	131.35	62.55	104.93	157.76
	Enhanced captions	24	174.74	51.91	152.82	196.66
Second pass reading	Non-enhanced captions	24	89.78	76.33	57.55	122.02
	Enhanced captions	24	270.04	81.62	235.58	304.51
Total fixation	Non-enhanced captions	24	221.13	132.77	165.07	277.20
	Enhanced captions	24	444.78	126.04	391.56	498.01
Number of visits	Non-enhanced captions	24	1.62	.68	1.33	1.91
	Enhanced captions	24	2.20	.47	2.01	2.40

Linear mixed-effects models were constructed using attention measurements as the dependent variables. In each model, group was included as a fixed effect and participant and item were specified as crossed random effects to determine whether there was a significant difference between the non-enhanced captions group and the enhanced captions group in the amount of attention allocated to the target linguistic constructions. As suggested by Barr et al. (2013), a maximal model, which included group as a random slope by item, was used for subsequent analyses. According to Table 51, the results showed that there were significant differences between the two groups in terms of all four eye-movement indices. The effect sizes for these differences were in the small (first pass reading time and visit counts) and medium range (second pass reading and total fixation duration). This means that textually enhanced captions were more effective than non-enhanced captions in drawing learners' attention to the target linguistic construction while they were engaged in a multimodal input-based task.

Table 51. Results for the Linear Mixed-effects Models Examining Attention Allocated to Target Linguistic Construction - Present Perfect

		Fixed effects					Random effects	
							by participant	*by item*
		Estimate	*SE*	*t*	*p*	*r²*	*SD*	*SD*
First pass reading	Intercept	44.560	43.81	1.017	.324		54.360	55.750
	Group	43.400	17.410	2.493	.016*	.04		25.770
	Formula: PP1st ~ Group + (1|Participant) + (Group|Item)							
Second pass reading	Intercept	−277.40	61.26	−4.528	<.001***		76.350	85.340
	Group	182.08	24.86	7.325	<.001***	.29		45.75
	Formula: PP2nd ~ Group + (1|Participant) + (Group|Item)							
Total fixation	Intercept	−270.73	60.82	−4.451	<.001***		75.55	87.28
	Group	180.26	24.68	7.304	<.001***	.22		46.14
	Formula: PPTotal ~ Group + (1|Participant) + (Group|Item)							
Number of visits	Intercept	.448	.461	.970	.336		.562	.806
	Group	.585	.179	3.271	.002**	.08		.285
	Formula: PPVC ~ Group + (1|Participant) + (Group|Item)							

*** $p < .001$, ** $p < .01$, * $p < .05$

With regard to the past simple item, descriptive statistics, including the mean for the individual past simple item, are summarized in Table 52. Overall, the enhanced captions group exhibited longer fixation duration and a greater number of visits to the target linguistic construction, i.e. the past simple.

Table 52. Descriptive statistics for Attention Measurements – Past Simple

					95% CI	
		N	*M*	*SD*	*Lower*	*Upper*
First pass reading	Non-enhanced captions	24	236.74	205.17	150.11	323.38
	Enhanced captions	24	354.36	198.82	270.40	438.31
Second pass reading	Non-enhanced captions	24	109.17	92.11	70.27	148.06
	Enhanced captions	24	197.71	141.04	138.16	257.27
Total fixation	Non-enhanced captions	24	345.91	290.58	223.21	468.62
	Enhanced captions	24	552.07	331.00	412.30	691.84
Number of Visits	Non-enhanced captions	24	2.83	1.76	2.09	3.57
	Enhanced captions	24	3.95	1.73	2.86	3.92

Linear mixed-effects models were constructed to analyse the amount of attention the two groups allocated to the past simple. In the models, each of the measurements was included as a dependent variable with group as a fixed effect. In addition, participant and item were specified as random effects along with group as a by-item random slope. However, the model with a maximal structure failed to converge; thus, a new model was constructed by removing the random effect that accounted for the least variance (i.e., group by item). As presented in Table 53, the enhanced captions group's second reading duration, total reading duration and total number of visits were found to be significantly different from those of the non-enhanced captions group, with a relatively small effect size. The results can be interpreted as suggesting that a textually enhanced target linguistic construction, i.e., the past simple, yielded an increase in the time of rereading and of overall fixating, as well as the total number of visits.

Table 53. Results for the Linear Mixed-effects Models Examining Attention Allocated to Target Linguistic Construction – Past Simple

		Fixed effects					Random effects	
							by participant	*by item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
First pass reading	Intercept	1.519	148.80	.010	.992		197.33	28.54
	Group	117.613	59.700	49.840	.054	.03		62.55
Formula: PS1st ~ Group + (1|Participant) + (Group|Item)								
Second pass reading	Intercept	−67.93	87.75	−.774	.443		114.70	18.38
	Group	88.55	34.81	2.544	$.014^{*}$	.05		26.42
Formula: PS2nd ~ Group + (1|Participant) + (Group|Item)								
Total fixation	Intercept	−66.41	229.28	−.290	.773		305.98	25.30
	Group	206.16	91.33	2.257	$.028^{*}$	.05		78.51
Formula: PSTotal ~ Group + (1|Participant) + (Group|Item)								
Number of visits	Intercept	.597	1.284	.465	.644		1.690	.214
	Group	1.16	.513	2.176	$.034^{*}$	.04		.487
Formula: PSVC ~ Group + (1|Participant) + (Group|Item)								

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

4.3.4. Relationship between Attention and L2 Development (RQ4)

For RQ 4, which addressed a potential relationship between the amount of attention the participants allocated to the target linguistic constructions and their development in L2 grammatical knowledge, Pearson correlational analyses were conducted using R. The relationships between the four measurements of attention (i.e., first pass reading, second pass reading, total fixation duration and total number of visits) and the gain scores of the three assessment measures – oral productive test, written productive test, receptive test – were examined for each group, the non-enhanced-captions group (Table 54) and the enhanced captions group (Table 55). For the non-enhanced captions group, second pass reading, total fixation duration and number of visits had medium size correlations only with gains in a written productive test. In other words, participants who spent more time rereading, fixated for longer on and visited more frequently the target linguistic structure (i.e., the present perfect) under the non-enhanced condition and exhibited higher gains only on a written productive test.

Table 54. Correlations between Attention Measurements and Non-enhanced Captions Group's Performance on the Tests– Present Perfect

		Oral Productive		Written Productive		Receptive	
		Pretest – Immediate	Pretest – Delayed	Pretest – Immediate	Pretest – Delayed	Pretest – Immediate	Pretest – Delayed
First pass reading	*r*	.31	.41	.36	.38	.39	.39
	p	(.14)	(.05)	(.09)	(.06)	(.06)	(.06)
Second pass reading	*r*	.26	.31	$.49^{*}$	$.56^{**}$	.29	.28
	p	(.21)	(.15)	(.01)	(.00)	(.16)	(.18)
Total fixation	*r*	.30	.37	$.45^{*}$	$.50^{*}$	.35	.35
	p	(.16)	(.08)	(.03)	(.01)	(.09)	(.10)
Number of visits	*r*	.38	.30	$.71^{***}$	$.73^{***}$	.32	.39
	P	(.06)	(.16)	(.00)	(.00)	(.12)	(.06)

N = 48 $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

The correlational analyses, however, yielded a number of significant medium to large relationships between the gain scores of the enhanced captions group and the eye-movement measures (Table 55). All four eye-movement measurements significantly correlated to a medium to large degree with pretest-immediate posttest gains as well as with pretest-delayed posttest gains on an oral productive test. Relatively large correlations were also identified between four eye-movement measures and gains from a written productive pretest to a written immediate posttest.

As for the receptive test, the total number of visits to the present perfect had a medium-size positive correlation with the gain score on a receptive pretest to a receptive immediate posttest. Overall, the results indicated that the participants who fixated for longer and more frequently on the enhanced target linguistic construction (i.e., the present perfect) obtained higher gains on an oral productive test and a written productive test.

Table 55. Correlations between Attention Measurements and Enhanced Captions Group's Performance on the Tests– Present Perfect

		Oral Productive		Written Productive		Receptive	
		Pretest – Immediate	Pretest – Delayed	Pretest – Immediate	Pretest – Delayed	Pretest – Immediate	Pretest – Delayed
First pass	r	$.58^{**}$	$.55^{**}$	$.81^{***}$	.31	.36	.24
reading	p	(.00)	(.00)	(.00)	(.14)	(.09)	(.26)
Second pass	r	$.72^{***}$	$.59^{**}$	$.74^{***}$	.37	.37	.29
reading	p	(.00)	(.00)	(.00)	(.07)	(.07)	(.16)
Total	r	$.70^{***}$	$.61^{**}$	$.81^{***}$	.37	.39	.29
Fixation	p	(.00)	(.00)	(.00)	(.07)	(.06)	(.17)
Number of	r	$.66^{***}$	$.63^{**}$	$.67^{***}$	.30	$.46^{*}$	.28
visits	p	(.00)	(.00)	(.00)	(.16)	(.02)	(.18)

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

However, as for the past simple, none of the attention measurements had significant relationships with the gain scores on the three assessment measures for either group, as shown in Table 56 and in Table 57.

Table 56. Correlations between Attention Measurements and Non-enhanced Captions Group's Performance on the Tests– Past Simple

		Oral Productive		Written Productive		Receptive	
		Pretest – Immediate	Pretest – Delayed	Pretest – Immediate	Pretest – Delayed	Pretest – Immediate	Pretest – Delayed
First pass	r	−.27	−.03	.02	−.03	.14	−.03
reading	p	(.20)	(.88)	(.94)	(.90)	(.50)	(.90)
Second pass	r	−.23	.01	−.13	−.11	.12	−.01
reading	p	(.27)	(.94)	(.54)	(.62)	(.56)	(.94)
Total	r	−.26	−.02	−.03	−.05	.14	−.02
fixation	p	(.21)	(.93)	(.89)	(.81)	(.51)	(.91)
Number of	r	−.34	−.04	−.05	−.01	.50	.02
visits	p	(.11)	(.84)	(.80)	(.96)	(.14)	(.91)

$N = 48$ $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

Table 57. Correlations between Attention Measurements and Enhanced Captions Group's Performance on the Tests– Past Simple

		Oral Productive		Written Productive		Receptive	
		Pretest – Immediate	Pretest – Delayed	Pretest – Immediate	Pretest – Delayed	Pretest – Immediate	Pretest – Delayed
First pass reading	*r*	.42	.13	.03	−.02	.01	.19
	p	(.04)	(.53)	(.88)	(.93)	(.96)	(.37)
Second pass reading	*r*	.39	.12	.20	.15	.06	.10
	p	(.06)	(.57)	(.35)	(.49)	(.78)	(.63)
Total fixation	*r*	.42	.13	.10	.05	.03	.16
	p	(.04)	(.54)	(.63)	(.81)	(.88)	(.46)
Number of visits	*r*	.39	.15	.11	.11	.07	.14
	p	(.06)	(.49)	(.61)	(.60)	(.73)	(.51)

$N = 48$ $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

4.3.5. Moderating Effects of Working Memory in Attention Allocation and L2 Grammatical Knowledge Development (RQ5)

Research question 5 asked the extent to which individual differences in working memory capacity moderated the effects of textually non-enhanced and enhanced captions included in multimodal input-based tasks on development in L2 grammatical knowledge. The participants' working memory was assessed using six different working memory measures: (a) nonword span task, (b) stop signal task, (c) colour shape task, (d) forward corsi block task, (e) backward corsi block task, and (f) automated operation span task. The descriptive statistics for each measure are summarized in Table 58.

Table 58. Descriptive Statistics for Working Memory Measures

		N	*M*	*SD*	95% CI *Lower*	95% CI *Upper*
NWS	No captions	24	8.417	1.863	7.630	9.203
	Non-enhanced captions	24	8.333	1.857	7.549	9.118
	Enhanced captions	24	8.958	1.654	8.260	9.657
SST (ms)	No captions	24	299.650	74.032	268.388	330.911
	Non-enhanced captions	24	280.497	64.216	253.380	307.613
	Enhanced captions	24	291.030	88.980	253.458	328.602
CST (ms)	No captions	24	410.148	203.339	324.286	496.011
	Non-enhanced captions	24	476.061	259.250	366.589	585.533
	Enhanced captions	24	371.425	185.051	293.285	449.566
Forward Corsi block	No captions	24	62.625	15.446	56.094	69.156
	Non-enhanced captions	24	58.750	18.997	50.728	66.772
	Enhanced captions	24	63.458	16.741	56.389	70.527
Backward Corsi block	No captions	24	57.000	17.134	49.765	64.235
	Non-enhanced captions	24	55.000	15.094	49.127	61.874
	Enhanced captions	24	61.417	13.577	55.683	67.150
AOSPAN	No captions	24	53.125	10.481	51.217	61.533
	Non-enhanced captions	24	56.375	12.024	49.637	58.946
	Enhanced captions	24	54.292	11.185	51.969	57.226

NWS max. total score = 11; Corsi block max. total score = 88; AOSAPN max. total score = 75

As a first step, correlations among the working memory measures were determined to assess construct validity. The results of the correlation analyses, for all three groups as well as for each group, are summarized in Table 59. While no significant correlations emerged among most of the working memory measures, a medium-size positive correlation between the nonword span task and AOSPAN was observed for the three groups combined (r = .40, p < .01), the non-enhanced captions group (r = .54, p < .05) and the enhanced captions group (r = .48, p < .05). There was also a large-size positive correlation between the forward corsi block and backward corsi block tasks: total group: r = .69, p < .01; no captions group: r = .71, p < .01; non-enhanced captions: r = 80, p < .01; enhanced captions: r = .54, p < .05. On the basis of this finding, it can be hypothesized that the stop signal task, the colour shape task and AOSPAN tap into a distinctive function of the central executive.

Table 59. Correlations among Working Memory Measures for the Three Groups

		NWS	SST	CST	Forward Corsi block	Backward Corsi block	AOSPAN
Total (N = 72)							
SST	*r*	−.26					
	p	(.05)					
CST	*r*	−.08	.16				
	p	(.50)	(.17)				
Forward Corsi block	*r*	−.05	−.20	.05			
	p	(.68)	(.09)	(.67)			
Backward Corsi block	*r*	.03	−.21	−.01	$.69^{**}$		
	p	(.77)	(.08)	(.93)	(.00)		
AOSPAN	*r*	$.40^{**}$	−.13	.10	−.00	.03	
	p	(.00)	(.29)	(.40)	(.99)	(.77)	
No Captions Group (N = 24)							
SST	*r*	−.52					
	p	(.06)					
CST	*r*	−.21	.47				
	p	(.32)	(.20)				
Forward Corsi block	*r*	.26	−.37	.13			
	p	(.21)	(.08)	(.55)			
Backward Corsi block	*r*	.22	−.28	.04	$.71^{**}$		
	p	(.30)	(.19)	(.84)	(.00)		
AOSPAN	*r*	.20	−.07	−.02	−.06	−.16	
	p	(.35)	(.73)	(.93)	(.79)	(.44)	
Non-enhanced Captions Group (N = 24)							
SST	*r*	−.21					
	p	(.33)					
CST	*r*	−.08	.26				
	p	(.70)	(.23)				
Forward Corsi block	*r*	−.19	−.17	.05			
	p	(.38)	(.44)	(.81)			
Backward Corsi block	*r*	−.10	−.23	−.05	$.80^{**}$		
	p	(.63)	(.27)	(.80)	(.00)		
AOSPAN	*r*	$.54^{*}$	−.18	.12	.13	.32	
	p	(.01)	(.39)	(.59)	(.54)	(.13)	
Enhanced Captions Group (N = 24)							
SST	*r*	−.08					
	p	(.69)					
CST	*r*	.18	−.13				
	p	(.40)	(.53)				
Forward Corsi block	*r*	−.25	−.15	.06			
	p	(.24)	(.49)	(.77)			
Backward Corsi block	*r*	−.15	−.15	.10	$.54^{**}$		
	p	(.49)	(.47)	(.62)	(.01)		
AOSPAN	*r*	$.48^{*}$	−.09	.15	−.08	−.05	
	p	(.01)	(.66)	(.48)	(.72)	(.83)	

The moderating effects of working memory capacity were examined using a series of linear mixed-effects models. The models that were originally constructed for the first and second research questions, which addressed the extent to which no captions, with non-enhanced or enhanced captions affected L2 grammatical knowledge development, were used as the null models. These models included group and time as fixed effects and participant and item as random effects. To these models, each working memory measurement was entered as an additional fixed effect. Each of these models was first constructed as a maximal random effects structure; that is, the models consisted of random effects (group and item) and random slopes for each fixed effect. Those models that failed to converge were restructured by removing the random effects that accounted for the least variance and used for subsequent analyses. Tables 60 and 61 show the results for linear mixed-effects models determining the extent to which learners' working memory moderated the effects of textually non-enhanced captions and enhanced captions on gains in knowledge of the present perfect.

Of the various measures, a significant 3-way interaction effect was observed between (a) group, time and forward corsi block when the participants' use of the present perfect was measured by an oral productive test, with a very small effect size and (b) group, time and AOSPAN when the participants' use of the present perfect was measured by a receptive test, with a small effect size.

Table 60. Results for the Linear Mixed-effects Models Examining Moderating Effect of Working Memory on Gains on Productive Tests – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Oral Productive Test								
NWS	Intercept	−.487	.944	−.517	.606		.343	.006
	Group:Time:NWS	.033	.020	1.664	.096+	.07		
*Formula: OralPP ~ Group*Time*NWS + (1|Participant) + (1|Item)*								
SST	Intercept	1.428	.850	1.680	.094		.407	.000
	Group:Time:SST	.000	.000	−.183	.855	.01		
*Formula: OralPP ~ Group*Time*SST + (1|Participant) + (1|Item)*								
CST	Intercept	1.152	.472	2.441	.015*		.411	.000
	Group:Time:CST	−.000	.000	−.396	.692	.01		
*Formula: OralPP ~ Group*Time*CST + (1|Participant) + (1|Item)*								
Cosiblock forward	Intercept	2.626	.854	3.075	<.01**		.409	.000
	Group:Time:CF	.004	.002	−2.035	.04*	.00		
*Formula: OralPP ~ Group*Time*CF + (1|Participant) + (1|Item)*								
Corsiblock backward	Intercept	2.137	.782	2.734	<.01**		.413	.000
	Group:Time:CB	−.003	.002	−1.221	.222	.01		
*Formula: OralPP ~ Group*Time*CB + (1|Participant) + (1|Item)*								
AOSPAN	Intercept	.608	1.036	.587	.558		.372	.000
	Group:Time:AOS	.003	.003	1.090	.276	.05		
*Formula: OralPP ~ Group*Time*AOSPAN + (1|Participant) + (1|Item)*								
Written Productive Test								
NWS	Intercept	−1.00	.959	1.048	.298		.375	.055
	Group:Time:NWS	.033	.026	1.233	.222	.00		
*Formula: WrittenPP ~ Group*Time*NWS + (Time|Participant) + (1|Item)*								
SST	Intercept	.909	.814	1.117	.268		.381	.055
	Group:Time:SST	.000	.000	.016	.987	.01		
*Formula: WrittenPP ~ Group*Time*SST+ (Time|Participant) + (1|Item)*								
CST	Intercept	.114	.449	.253	.801		.378	.055
	Group:Time:CST	.000	.000	.936	.352	.00		
*Formula: WrittenPP ~ Group*Time*CST + (Time|Participant) + (1|Item)*								
Corsiblock forward	Intercept	.388	.650	.597	.553		.393	.055
	Group:Time:CF	.001	.002	.255	.800	.00		
*Formula: WrittenPP ~ Group*Time*CF + (Time|Participant) + (1|Item)*								
Corsiblock backward	Intercept	.458	.752	.608	.545		.392	.055
	Group:Time:CB	.002	.003	.887	.378	.00		
*Formula: WrittenPP ~ Group*Time*CB + (Time|Participant) + (1|Item)*								
AOSPAN	Intercept	.098	1.033	.095	.924		.387	.055
	Group:Time:AOS	.002	.004	.406	.686	.01		
*Formula: WrittenPP ~ Group*Time*AOSPAN + (Time|Participant) + (1|Item)*								

Table 61. Results for the Linear Mixed-effects Models Examining Moderating Effect of Working Memory on Gains on Receptive Test – Present Perfect

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	*r²*	*SD*	*SD*
Receptive Test								
NWS	Intercept	−.154	.687	−.224	.823		.306	.083
	Group:Time:NWS	.22	.013	1.728	.084+	.02		
*Formula: ReceptivePP ~ Group*Time*NWS + (1|Participant) + (1|Item)*								
SST	Intercept	.567	.604	.939	.349		.326	.082
	Group:Time:SST	−.000	.000	−1.230	.219	.02		
*Formula: ReceptivePP ~ Group*Time*SST + (1|Participant) + (1|Item)*								
CST	Intercept	.482	.334	1.456	.147		.325	.082
	Group:Time:CST	−.000	.000	−.940	.347	.01		
*Formula: ReceptivePP ~ Group*Time*CST + (1|Participant) + (1|Item)*								
Corsiblock forward	Intercept	1.598	.588	2.716	<.01**		.310	.082
	Group:Time:CF	.000	.001	.028	.978	.02		
*Formula: ReceptivePP ~ Group*Time*CF + (1|Participant) + (1|Item)*								
Corsiblock backward	Intercept	1.293	.542	2,386	.018*		.313	.082
	Group:Time:CB	.002	.001	1.423	.155	.02		
*Formula: ReceptivePP ~ Group*Time*CB + (1|Participant) + (1|Item)*								
AOSPAN	Intercept	.413	.710	.582	.561		.279	.082
	Group:Time:AOS	.004	.002	2.217	.027*	.07		
*Formula: ReceptivePP ~ Group*Time*AOSPAN + (1|Participant) + (1|Item)*								

Note. AOS = AOSPAN; $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

Given an interaction effect reported among group, time and forward corsi block, a post hoc analysis, including data from only two of the groups, was conducted to determine where the difference existed. First, a moderating effect of the forward corsi block task was observed, with a medium effect size, between the no captions group and the textually enhanced captions group when the use of present perfect was measured by an oral productive test, as shown in Table 62. A further post hoc test was performed with data from only two testing times of the two groups. An examination of the two groups – no captions group and enhanced captions group – showed that the visuospatial short-term memory function of working memory mediated the gain scores from pretest to immediate posttest and from pretest to delayed posttest, with medium effect sizes (Table 63).

Table 62. Results for Post hoc Contrasts Examining the Moderating Effects of a Forward Corsi Block Task on the Relationship between Captions, Textually Non-enhanced or Enhanced, and L2 Development – Oral Productive Test

		Fixed effects				Random effects	
						by participant	*by Item*
		Estimate	*SE*	*t*	*r²*	*SD*	*SD*
Between Groups							
No captions	Intercept	2.906	1.239	2.344		.184	.429
Non-enhanced captions	Group:Time:CF	−.008	.004	−1.923	.04		
Non-enhanced captions	Intercept	2.141	1.709	1.253		.388	.028
Enhanced captions	Group:Time:CF	−.001	.004	−.361	.25		
No captions	Intercept	2.715	.870	3.120		.363	.000
Enhanced captions	Group:Time:CF	−.005	.002	−2.093	.24		
*Formula: OralPP ~ Group*Time*CF+ (1|Participant) + (1|Item)*							

Table 63. Results for Post hoc Contrasts Examining the Moderating Effects of a Forward Corsi Block Task on the Relationship between No Captions and Textually Enhanced Captions and L2 Development – Oral Productive Test

		Fixed effects				Random effects	
						by participant	*by Item*
		Estimate	*SE*	*t*	*r²*	*SD*	*SD*
Between Times							
Pretest-	Intercept	3.822	1.157	3.305		.340	.000
Immediate	Group:Time:CF	−.010	.004	−2.348	.21		
Pretest-	Intercept	2.993	.797	3.757		.276	.012
Delayed	Group:Time:CF	−.004	.002	−2.164	.33		
Posttest-	Intercept	.496	.174	.285		.457	.140
Delayed	Group:Time:CF	.001	.004	.311	.25		
*Formula: OralPP ~ Group*Time*CF + (1|Participant) + (1|Item)*							

To explore how the forward corsi block scores moderated the participants' gain scores, a series of correlational analyses were performed. They examined whether there were correlations between the gain scores of each group and the forward corsi block scores, as shown in Table 64. For the no captions group, the results showed that visuospatial short-term memory significantly correlated to a medium degree with the gain scores on an oral productive test, from both pretest to immediate posttest and pretest to delayed posttest. However, no significant links were observed between visuospatial short-term memory and gains on an oral productive test for the enhanced captions group.

Table 64. Correlation analysis between a Forward Corsi Block Task and Gain Scores

		Pretest – Immediate Test	Pretest – Delayed Test
No Captions Group			
Forward Corsi Block	*r*	.45*	.50*
	p	(.03)	(.01)
Enhanced Captions Group			
Forward Corsi Block	*r*	−.17	−.19
	p	(.41)	(.38)

** $p < .01$, * $p < .05$

Regarding AOSPAN, post hoc contrasts, including data from only two groups at a time, were also examined to determine where differences lay among the groups. The analyses revealed that the updating function of working memory, as measured by an automated operation span task (AOSPAN), moderated the effects of enhanced captions in comparison with no captions, with relatively large effect sizes (Table 65).

Table 65. Results for Post hoc Contrasts Examining the Moderating Effects of an Automated Operation Span Task on the Relationship between Captions, Textually Non-enhanced or Enhanced, and L2 Development – Receptive Test

		Fixed effects				Random effects	
						by participant	*by Item*
		Estimate	*SE*	*t*	*r²*	*SD*	*SD*
Between Groups							
No captions	Intercept	.901	.960	.938		.259	.079
Non-enhanced captions	Group:Time:AOS	.007	.004	1.783	.09		
Non-enhanced captions	Intercept	−.256	1.594	−.161		.253	.089
Enhanced captions	Group:Time:AOS	.002	.004	.510	.25		
No captions	Intercept	.613	.729	.841		.274	.684
Enhanced captions	Group:Time:AOS	.004	.002	2.121	.30		

*Formula: RecptivePP ~ Group*Time*AOSPAN+(1|Participant) + (1|Item)*

Another series of post hoc tests were conducted, this time including only two groups (i.e., no captions group and enhanced captions group) at two testing times at a time. As shown in Table 66, the updating function of working memory was found to mediate the gain scores from receptive pretest to posttest and from receptive pretest to delayed posttest, with effect sizes falling in the small to medium range.

Table 66. Results for Post hoc Contrasts Examining the Moderating Effects of an Automated Operation Span Task on the Relationship between No Captions and Textually Enhanced Captions and L2 Development – Receptive Test

		Fixed effects				Random effects	
						by participant	*by Item*
		Estimate	*SE*	*t*	r^2	*SD*	*SD*
Between Times							
Pretest-	Intercept	.946	.892	1.060		.279	.059
Immediate	Group:Time:AOS	.008	.004	2.077	.32		
Pretest-	Intercept	.690	.711	.970		.269	.094
Delayed	Group:Time: AOS	.004	.002	2.265	.30		
Posttest-	Intercept	−.048	1.440	−.033		.266	.099
Delayed	Group:Time: AOS	.000	.004	.202	.34		

*Formula: ReceptivePP ~Group*Time*AOSPAN+ (1|Participant) + (1|Item)*

As shown in Tables 67, follow-up correlation analyses revealed that there was a strong medium-size correlation between AOSPAN and the pretest-posttest and pretest-delayed posttest gain scores of the enhanced captions group. This means that, in the enhanced captions group, learners who had greater updating ability obtained larger gains on a receptive test.

Table 67. Correlation analysis between an Automated Operation Span Task and Gain Scores

		Pretest – Immediate Test	Pretest – Delayed Test
No Captions Group			
AOSPAN	*r*	.37	.24
	p	(.07)	(.27)
Enhanced Captions Group			
AOSPAN	*r*	$.64^{**}$	$.60^{**}$
	p	(.00)	(.00)

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

For the past simple, however, none of the working memory measurements had significant moderating effects on L2 grammatical knowledge development as shown in Tables 68 and 69.

Table 68. Results for the Linear Mixed-effects Models Examining Moderating Effect of Working Memory on Gains on Productive Tests – Past Simple

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Oral Productive Test								
NWS	Intercept	1.045	.222	4.714	<.001***		.072	.020
	Group:Time:NWS	−.003	.005	−.058	.954	.01		
*Formula: OralPS ~ Group*Time*NWS + (Time|Participant) + (1|Item)*								
SST	Intercept	1.018	.185	5.491	<.001***		.070	.019
	Group:Time:SST	.000	.000	−.228	.820	.01		
*Formula: OralPS ~ Group*Time*SST + (Time|Participant) + (1|Item)*								
CST	Intercept	.949	.103	9.207	<.001***		.071	.017
	Group:Time:CST	.000	.000	.585	.559	.00		
*Formula: OralPS ~ Group*Time*CST + (Time|Participant) + (1|Item)*								
CosiBlock forward	Intercept	.943	.188	4.993	<.001***		.074	.017
	Group:Time:CF	.000	.001	.526	.600	.00		
*Formula: OralPS ~ Group*Time*CF + (Time|Participant) + (1|Item)*								
Corsiblock backward	Intercept	1.150	.168	6.849	<.001***		.067	.017
	Group:Time:CB	.001	.001	−1.154	.250	.01		
*Formula: OralPS ~ Group*Time*CB+ (Time|Participant) + (1|Item)*								
AOSPAN	Intercept	1.051	.237	4.433	<.001***		.074	.017
	Group:Time:AOS	.001	.001	−.613	.540	.00		
*Formula: OralPS ~ Group*Time*AOSPAN + (Time|Participant) + (1|Item)*								
Written Productive Test								
NWS	Intercept	.894	.391	2.284	.025*		.212	.007
	Group:Time:NWS	.010	.009	1.080	.284	.02		
*Formula: WrittenPS ~ Group*Time*NWS + (Time|Participant) + (1|Item)*								
SST	Intercept	1.108	.335	3.302	<.01**		.217	.007
	Group:Time:SST	.000	.000	−.225	.823	.01		
*Formula: WrittenPS ~ Group*Time*SST + (Time|Participant) + (1|Item)*								
CST	Intercept	.986	.186	5.292	<.01***		.218	.007
	Group:Time:CST	.000	.000	−.238	.813	.01		
*Formula: WrittenPS ~ Group*Time*CST + (Time|Participant) + (1|Item)*								
CorsiBlock forward	Intercept	1.189	.336	3.538	<.001***		.216	.007
	Group:Time:CF	−.000	.001	−.457	.649	.01		
*Formula: WrittenPS ~ Group*Time*CF + (Time|Participant) + (1|Item)*								
Corsiblock backward	Intercept	1.069	.306	3.489	<.001***		.217	.007
	Group:Time:CB	.000	.001	.132	.896	.01		
*Formula: WrittenPS ~ Group*Time*CB + (Time|Participant) + (1|Item)*								
AOSPAN	Intercept	1.114	.428	2.603	.011*		.220	.007
	Group:Time:AOS	.000	.001	.250	.803	.00		
*Formula: WrittenPS ~ Group*Time*AOSPAN + (Time|Participant) + (1|Item)*								

Table 69. Results for the Linear Mixed-effects Models Examining Moderating Effect of Working Memory on Gains on Receptive Test – Past Simple

		Fixed effects					Random effects	
							by participant	*by Item*
		Estimate	*SE*	*t*	*p*	r^2	*SD*	*SD*
Receptive test								
NWS	Intercept	.868	.543	1.600	.110		.154	.135
	Group:Time:NWS	.015	.012	1.253	.210	.00		
*Formula: ReceptivePS ~ Group*Time*NWS + (1|Participant) + (1|Item)*								
SST	Intercept	2.119	.454	4.669	<.001***		.145	.135
	Group:Time:SST	.000	.000	-.800	.424	.01		
*Formula: ReceptivePS ~ Group*Time*SST + (1|Participant) + (1|Item)*								
CST	Intercept	1.595	.257	6.213	<.001***		.155	.135
	Group:Time:CST	.000	.000	-.587	.557	.00		
*Formula: ReceptivePS ~ Group*Time*CST + (1|Participant) + (1|Item)*								
CosiBlock forward	Intercept	1.516	.461	3.287	<.01**		.154	.135
	Group:Time:CF	.000	.001	.276	.782	.00		
*Formula: ReceptivePS ~ Group*Time*CF + (1|Participant) + (1|Item)*								
Corsiblock backward	Intercept	1.852	.418	4.426	<.001***		.150	.135
	Group:Time:CB	.000	.001	.248	.804	.01		
*Formula: ReceptivePS ~ Group*Time*CB + (1|Participant) + (1|Item)*								
AOSPAN	Intercept	1.799	.577	3.117	<.01**		.152	.135
	Group:Time:AOS	.000	.002	.216	.829	.00		
*Formula: ReceptivePS ~ Group*Time*AOSPAN + (1|Participant) + (1|Item)*								

N = 72 $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

4.3.6. Relationship between Attention and Working Memory (RQ6)

The relationship between the learners' attention allocated to the target linguistic constructions and their working memory capacity for each group (RQ6) was examined using Pearson correlation analyses in R. As shown in Table 70, no significant correlation was observed between the attention allocated to the target linguistic construction, the present perfect, and working memory capacity for the non-enhanced captions group. For the enhanced captions group, however, three out of four eye-movement indices, including second pass reading, total fixation duration, and number of visits, positively correlated to a medium degree with AOSPAN. In addition, second pass reading also had a medium size positive correlation with NWS (Table 71). The results indicate that learners who were better at updating information and had greater phonological short-term memory were more likely to spend more time rereading the perceptually salient target linguistic construction (i.e., the present perfect).

Table 70. Correlations between Attention Measurements and Working Memory Capacity of Non-enhanced Captions Group – Present Perfect

		First pass reading	Second Pass reading	Total fixation duration	Total number of visits
NWS	*r*	.17	.11	.15	.10
	p	(.42)	(.59)	(.49)	(.66)
SST	*r*	−.27	−.34	−.32	−.30
	p	(.20)	(.10)	(.12)	(.15)
CST	*r*	−.15	−.08	−.11	−.29
	p	(.48)	(.72)	(.59)	(.17)
Corsiblock-forward	*r*	.24	.21	.23	−.12
	p	(.26)	(.33)	(.28)	(.58)
Corsiblock-backward	*r*	.36	.21	.29	−.00
	p	(.09)	(.33)	(.17)	(.99)
AOSPAN	*r*	.39	.25	.32	.24
	p	(.06)	(.24)	(.12)	(.26)

Table 71. Correlations between Attention Measurements and Working Memory Capacity of Enhanced Captions Group – Present Perfect

		First pass reading	Second Pass reading	Total fixation duration	Total number of visits
NWS	*r*	.28	.43*	.39	.31
	p	(.18)	(.04)	(.06)	(.13)
SST	*r*	−.16	−.16	−.17	−.10
	p	(.45)	(.45)	(.42)	(.63)
CST	*r*	−.16	.01	−.06	−.11
	p	(.44)	(.98)	(.76)	(.61)
Corsiblock-forward	*r*	−.01	.11	.07	.01
	p	(.98)	(.60)	(.74)	(.94)
Corsiblock-backward	*r*	.27	.19	.24	.35
	p	(.20)	(.37)	(.27)	(.10)
AOSPAN	*r*	.33	.42*	.41*	.41*
	p	(.11)	(.04)	(.04)	(.04)

** $p < .01$, * $p < .05$

As for the past simple, significant, negative medium-size correlations were found between the nonword span task and all four eye-movement measurements as well as between the AOSPAN and two eye-movement measurements (second pass reading and total fixation duration) for the non-enhanced group (Table 72). In addition, the corsi block forward task was found to have a medium-size correlation with second pass reading.

Table 72. Correlations between Attention Measurements and Working Memory Capacity of Non-enhanced Captions Group – Past Simple

		First pass reading	Second Pass reading	Total fixation duration	Total number of visits
NWS	*r*	$-.50^{*}$	$-.66^{***}$	$-.56^{***}$	$-.53^{**}$
	p	(.01)	(.00)	(.00)	(.01)
SST	*r*	.07	−.03	.04	−.09
	p	(.76)	(.89)	(.86)	(.66)
CST	*r*	−.34	−.29	−.34	−.34
	p	(.10)	(.17)	(.11)	(.10)
Corsiblock-forward	*r*	.22	$.41^{*}$	.28	.38
	p	(.30)	(.05)	(.18)	(.07)
Corsiblock-backward	*r*	.16	.23	.19	.26
	p	(.45)	(.28)	(.38)	(.21)
AOSPAN	*r*	−.39	$-.43^{*}$	$-.41^{*}$	−.22
	p	(.06)	(.03)	(.04)	(.29)

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

For the enhanced captions group, no statistically significant correlation between attention measurements and working memory capacity emerged, as displayed in Table 73.

Table 73. Correlations between Attention Measurements and Working Memory Capacity of Enhanced Captions Group – Past Simple

		First pass reading	Second Pass reading	Total fixation duration	Total number of visits
NWS	*r*	.07	.15	.11	.24
	p	(.75)	(.47)	(.62)	(.26)
SST	*r*	−.18	−.24	−.21	−.11
	p	(.41)	(.27)	(.33)	(.60)
CST	*r*	−.20	−.14	−.18	−.17
	p	(.36)	(.52)	(.41)	(.43)
Corsiblock-forward	*r*	.05	.03	.04	−.16
	p	(.81)	(.89)	(.84)	(.45)
Corsiblock-backward	*r*	.26	.19	.24	.10
	p	(.22)	(.38)	(.27)	(.63)
AOSPAN	*r*	.09	−.02	.05	.16
	p	(.66)	(.93)	(.82)	(.46)

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

4.3.7. Summary of Results

RQ 1 To what extent do multimodal input-based tasks without captions versus with captions affect development in L2 grammatical knowledge?

Compared to no captions, non-enhanced captions had a positive effect on development in the use of the present perfect; there were significant differences in the performances of the two groups on the written productive test and the receptive test, with small effect sizes. Regarding the past simple, however, no significant group-by-time interaction effect was revealed for any of the three assessment measures. That is, the performance on both productive tests (oral and written test) and a receptive test was not significantly different between the two groups in relation to the past simple.

RQ2 To what extent do textually non-enhanced versus enhanced captions in multimodal input-based tasks affect development in L2 grammatical knowledge?

The results indicate that the provision of textually-enhanced captions had a greater positive effect than no captions and non-enhanced captions on participants' use of the present perfect. The enhanced captions group outperformed both the non-enhanced captions group and the no captions group on all three assessment tasks (i.e., oral productive test, written productive test and receptive test). The effect sizes reported for the differences in performance fell in the small to medium range. However, no significant interaction effect was revealed for the past simple, regardless of the type of assessment task, indicating no significant difference in the performance of the three groups with respect to the past simple.

RQ3 To what extent do textually non-enhanced versus enhanced captions in multimodal input-based tasks draw learners' attention to the target linguistic construction?

For all eye-movement measures (i.e., first pass reading, second pass reading, total fixation duration, and total number of visits), two groups –non-enhanced captions group and enhanced captions group – exhibited significant differences in the amount of attention allocated to the target linguistic constructions, with small effect sizes, for the present perfect. Overall, the enhanced captions group fixated for longer and visited the present perfect constructions more than the non-enhanced captions group. Similar results were revealed for the other target linguistic construction, the past simple. The enhanced input drew learners' attention more, resulting in an increase in second reading duration, total reading duration, and total number of visits to the past simple item. The effect sizes for these differences were in the small range. In sum, the effectiveness of textual enhancement in drawing learners' attention to the target linguistic constructions was evidenced.

RQ4 To what extent does learner attention allocated to the target linguistic construction relate to development in L2 grammatical knowledge? Is this relationship influenced by whether learners are exposed to non-enhanced or textually enhanced captions?

A series of Pearson correlations was performed between eye-tracking measures and gain scores on three assessment tasks to address the relationship between attention and L2 development. For the non-enhanced captions group, medium positive correlations emerged between gains on a written productive test and three eye-tracking indices (second pass reading, total fixation duration, and number of visits) computed for the present perfect. The results indicated that the participants who reread, fixated for longer and visited non-enhanced present perfect constructions more frequently achieved higher gains on a written productive test.

For the enhanced captions group, more significant correlations were identified between eye-tracking indices and gain scores on the tests. The oral productive gain scores, from pretest to immediate posttest and pretest to delayed posttest, were found to have medium to large positive correlations with all four eye-tracking indices. Significant large correlations were also found between the gain scores on a written productive test, from pretest to immediate posttest, and four eye-tracking indices. The pretest-posttest gains on the receptive test had a significant medium correlation with the number of visits to the targeted present perfect. The results can thus be interpreted as suggesting that more attention allocated to textually enhanced captions resulted in higher gains on oral and written productive tests. Taken together, compared to the non-enhanced captions group, participants who paid more attention to textually enhanced captions showed more development in their use of the present perfect, reflected in stronger significant correlations between more developmental and attentional measures. As for the past simple, however, no significant relationships were identified between the eye-tracking indices and the gain scores on the three assessment measures for either group.

RQ5 To what extent do individual differences in working memory capacity moderate the effects of captions, non-enhanced or textually enhanced, in multimodal input-based tasks on L2 development?

To address the possible moderating effect of working memory, six different measures were used in an attempt to investigate whether the different functions of working memory had differential effects on L2 development under two captioning conditions. Of the various measures, the results revealed that a forward corsi block task and AOSPAN had a moderating, medium-size effect on the relationship between textual enhancement in captions and L2 development. Follow-up correlation analyses revealed that visuospatial short-term memory positively correlated with gain scores on the oral productive test for the no captions group. The effect size was in the medium range. This result indicated that learners who were not provided with captions performed better on an oral productive test for the present perfect if they had greater visuospatial short-term memory. On the other hand, the updating function, operationalized as AOSPAN, had large-size positive correlations with gain scores on a receptive test for the enhanced captions group. That is, learners who had greater updating ability exhibited higher gains in receptive scores when they were exposed to textually enhanced captions.

RQ6 To what extent does learners' attention allocated to the target linguistic construction correlate with their working memory capacity? Is this relationship influenced by whether learners are exposed to non-enhanced or enhanced captions?

To identify the possible correlations between the attention allocated to the target linguistic constructions and individual differences in working memory, a series of correlation analyses was computed. For the non-enhanced captions group, no significant correlations emerged between attention allocated to the present perfect and working memory. However, for the enhanced captions group, medium positive correlations were identified between AOSPAN and three eye-tracking indices (second pass reading, total fixation duration and number of visits). In addition, NWS was also found to have a medium-size positive correlation with second pass reading.

The results can be interpreted as suggesting that learners who had greater updating ability and larger phonological short-term memory were more likely to engage in reanalysis of the input including the target linguistic construction (i.e., the present perfect) when it was highlighted in the input.

Somewhat opposite results were revealed for the past simple, with significant correlations being reported between the attention drawn to the past simple and working memory capacity only for the non-enhanced captions group. More specifically, NWS had medium negative correlations with first pass reading, second pass reading, total fixation duration and total number of visits to the target linguistic construction, i.e., the past simple. AOSPAN was also found to correlate negatively to a medium degree with second pass reading time and total fixation duration. That is, learners who had greater phonological short-term memory and better updating ability were more likely to spend less time rereading the target linguistic construction (i.e., the past simple) when it was not visually enhanced. However, the significant medium-size positive correlation observed between the forward corsi block task and second pass reading suggests that learners with greater visuospatial short-term memory spent longer reanalysing the non-enhanced target linguistic construction. For the enhanced captions group, on the other hand, no significant relationship between the amount of attention allocated to the past simple and working memory was observed.

4.4. Interim Discussion

To extend Study 1, this study further examined the extent to which textually enhanced captions rendered the target linguistic constructions more salient to learners so as to result in drawing their attention to and promoting their development of L2 grammatical knowledge. Another goal of the present investigation was to shed more light on whether individual differences in working memory mediated learners' allocation of attentional resources to input and development in the use of L2 grammatical knowledge. Study 2 differed from Study 1 in that it involved: (a) the recruitment of more proficient learners of English, (b) the inclusion of a no captions group as a control group, (c) the administration of a delayed posttest, (d) the use of both form-oriented and meaning-oriented tests, (e) the employment of linear mixed-effects models for statistical analysis, and (f) the use of six different working memory measures.

4.4.1. Effects of Multimodal Input-based Tasks with or without Captions on L2 Grammatical Knowledge Development

In this study, learners benefited from the presence of captions in developing their use of L2 grammatical knowledge. In line with the accumulated evidence of previous studies reporting the positive effects of captions on L2 listening comprehension and L2 vocabulary learning, this study provided further evidence that the provision of captions is beneficial for promoting L2 grammatical knowledge acquisition. Thus, the findings lend support to the predictions of the Dual coding Theory (Paivio, 1986) and the Cognitive Theory of Multimedia Learning (Mayer, 2001); the simultaneous presentation of input through multiple modes (i.e., aural, visual and textual) in captioned videos seems to have facilitated information processing, rendered a greater depth of processing, and as a result, led to better retention for later use (Montero Perez et al., 2013a; Winke et al., 2010).

4.4.2. Effects of Textually Non-enhanced versus Enhanced Captions in Multimodal Input-based Tasks on Development of L2 Grammatical Knowledge

The textual enhancement included in the multimodal input-based tasks promoted development in the use of the present perfect. The results of this study are in line with the findings from Study 1, in which textually enhanced captions facilitated development in the knowledge of pronominal anaphoric references. However, previous reading research, in which learners were exposed to textually enhanced input while they were engaged in reading, has reported somewhat different results (Loewen & Inceoglu, 2016; Winke, 2013). Overall, the findings indicated a negligible effect for textual enhancement in the development of L2 grammatical knowledge development. A meta-analytic review of reading-based research conducted by Lee and Huang (2008) reported only a marginal effect of textual enhancement on L2 grammar learning.

As in Study 1, learners' previous knowledge of the target linguistic forms could have contributed to the positive effect of textual enhancement in this study (Lee, 2007). Participants in the study were Korean university students, who were at higher proficiency levels of English, comparable to C1 band and above according to the Common European Framework for Reference. Given the fact that students in South Korea learn the English tense and aspect system in middle school, following a national curriculum developed by the Korean Ministry of Education, the participants in this study are likely to have been exposed to explicit instruction in the basic uses of the present perfect, such as describing an action in the past with a result now, an experience and an uncompleted action at the time of speaking. Learners' explicit knowledge of the present perfect, 'past with present relevance', might have served as an initial impetus for learning a particular use of the present perfect in reporting news.

Another factor that might have contributed to the positive effects of enhancement in the present study has to do with the salience of the target linguistic construction (Gass, Spinner & Behney, 2017). According to Goldschneider and DeKeyser (2001), perceptual salience "refers to how easy it is to hear or perceive a given structure" (p. 47). In morphosyntactic studies, some researchers have argued that such perceptual salience of a linguistic property might play an important role in acquiring target features (N. Ellis, 2017; Gass, Spinner, & Behney, 2017; Goldschneider & DeKeyser, 2001). N. Ellis (2017) explains that if "[s]alient items or features are attended [to], [they] are more likely to be perceived and are more likely to enter into subsequent cognitive processing and learning" (p. 21). Building on this assumption, the relatively high physical salience of the target linguistic construction might have contributed to the positive effects of textual enhancement in the present study.

However, textual enhancement did not have a significant effect on developmental change in the use of the past simple. This was probably due to a ceiling effect observed at the pretest stage. Considering the maximum total scores, which were 5 for the oral and written productive tests and 20 for the receptive test, the participants obtained considerably high mean scores, with scores ranging from 4.56 to 4.76 on the productive tests (oral and written tests) and 15.62 to 16.33 on the receptive test. This was not an unexpected finding given the proficiency level of the participants.

With regard to the use of these two target linguistic constructions, another interesting finding was that the participants did not show a tendency to overgeneralize from use of the present perfect to the past simple. Particularly for the enhanced captions group, it was reasonable to assume that an increase in the salience of the present perfect might lead to overuse (Han et al., 2008). However, the scores obtained with enhanced captions for the past simple items on all three tests (i.e., oral productive test, written productive test and receptive test) showed no decrease in scores from pretest to immediate posttest and, further, to a delayed posttest. In other words, the participants did not use the present perfect in contexts in which the past simple was obligatory. One possible explanation for this finding might be the number of occurrences of the target linguistic constructions in the multimodal input-based tasks. In this study, 27 tasks were developed using news clips, in which the participants were exposed to the target linguistic constructions. Consistency was maintained in each news clip, in that the present perfect was only used to announce the news and the past simple was used in subsequent parts to give more details. Thus, the fact that the participants were repeatedly exposed to the use of the present perfect in the same context (i.e., to announce the news) could have made it easier for them to recognise this context for use of the target linguistic construction.

In explaining the results for textual enhancement, an additional point worth noting is that this study adopted what Leow (2009) refers to as a conflated approach. According to Leow (2009), studies on input enhancement should be categorized into "'non-conflated' (i.e., a pure comparison between an enhanced versus an unenhanced experimental group) as opposed to 'conflated' (i.e., input enhancement formed part of a combination of two or more independent variables)" (p. 18). That is, whether input enhancement was the only variable or whether it was investigated along with one or more other variables is important to consider when interpreting the results. Given that the beneficial effect of captions alone on second language acquisition has been well documented, it could be reasonably assumed that the use of textual enhancement with captions may have triggered greater depth of processing of the highlighted target construction than using textual enhancement alone, thereby resulting in more developmental benefits (Leow & Martin, 2007).

4.4.3. Effects of Textually Enhanced Captions on Allocation of Attention

To address whether textually enhanced captions rendered the target linguistic forms more salient to the learners so as to draw their attention, eye-tracking was used as a concurrent measure to assess the amount of attention allocated to the target linguistic constructions in the input. Learners' eye movements were analysed using four eye-tracking indices: first pass reading, second pass reading, total fixation duration and total number of visits. The results showed significant differences between the non-enhanced captions group and the enhanced captions group for both target linguistic constructions (i.e., present perfect and past simple), with a medium effect size. That is, textual enhancement was successful in directing learners' attention to the target linguistic constructions. This finding is consistent with the results reported in previous studies that investigated the effectiveness of textual enhancement in reading texts (Issa et al., 2015; Simard & Foucambert, 2013; Winke, 2013). In Winke's study, for instance, learners who were presented with textually enhanced forms spent more time rereading forms and, as a result, fixated on forms longer overall. Simard and Foucambert also found that textual enhancement led participants to pay more attention to forms, as indicated in total reading time, first fixation duration and regression duration. Another study that reported a positive effect of textual enhancement on drawing learners' attention to target linguistic constructions is Issa et al., which showed that participants who were exposed to textually enhanced input exhibited a decrease in one of the attention measures, i.e., skipping rate, on target forms.

Among previous eye-tracking studies investigating textual enhancement in the context of reading, the one most comparable to the current research is Winke's study. Both studies reported that perceptually more salient target linguistic constructions were more likely to capture learners' attention. However, in Winke's study, the positive effects of textually enhanced input were only evidenced in late measures, which is similar to the results of Study 1. Winke argued that input enhancement may not be effective in increasing the initial processing of target constructions. However, somewhat different results were revealed in Study 2 here: the effectiveness of textual enhancement was reflected in both early (i.e., first pass reading) and late measures (i.e., second pass reading and total fixation duration). This difference between Study 1 and Study 2 might relate to the nature of the target linguistic construction. The target linguistic construction in Study 1 was anaphoric reference, which describes the "relation between two linguistic elements, in which the interpretation of one (called an anaphora) is in some way determined by the interpretation of the other (called the antecedent)" (Huang, 2005, p. 231). That is, anaphoric references involve the processing of two related items, which might have led the learners to reanalyse the enhanced antecedent in relation to the following grammatical form, i.e., pronouns. The target linguistic construction in Study 2, on the other hand, was the present perfect (have/has + past participle), which was presented first to announce the news. Consequently, as soon as learners encountered the present perfect, they might have attended to this construction, especially when it was highlighted. This could have resulted in an increase in both initial and subsequent processing. This possible explanation requires further empirical confirmation.

Another noteworthy finding of Study 2 was that textual enhancement increased second pass reading and total fixation duration and the total number of visits to the past simple. As discussed before, learners achieved relatively high past tense scores on all three tests across testing times. The presentation of the past simple alone, without enhancement, might have remained unattended to by learners since they had a relatively good knowledge of the use of the past simple, and thus deeper processing might not have been necessary. However, the fact that the past simple was highlighted could have led the participants to re-examine perceived differences between the present perfect and the past simple.

Although these findings are intriguing, it is also important to note that not all studies have reported positive effects of textual enhancement on learner attention. Unlike this research, Loewen and Inceoglu (2016), for instance, found that highlighted grammatical forms in the text did not facilitate learners' attention to target forms during input processing. Similar results were reported in a study conducted by Indrarathne and Kormos (2017). One possible reason for these mixed findings across studies concerns the eye-tracking measures used. More specifically, Winke (2013) measured the amount of attention allocated to target linguistic constructions using total fixation time, number of visits, first pass reading time, and rereading time. Simard and Foucambert (2013), on the other hand, used total reading time, first fixation duration, and skipping rate on target forms, whereas two different indices were employed in Indrarathne and Kormos's (2017) study, including total fixation duration and the difference between observed total fixation duration and expected total fixation duration on areas of interest. Loewen and Inceoglu (2016) analysed learners' eye gaze using the number of fixations on each target item, the amount of time spent on each targeted item and the duration of the first fixation. As can be seen, the lack of congruence in the use of eye-tracking measurements across studies makes it difficult to draw a firm conclusion about the effects of textual enhancement on learners' allocation of attention.

An additional factor that might have contributed to the mixed findings concerns the difference in the types of typographical cues used in studies. In Winke's (2013) study and Simard and Foucambert's (2013) study, forms were enhanced with underlining, whereas Indrarathne and Kormos (2017) used bold facing as a means to highlight target forms. In this study, along with a study by Isaa et al. (2015), a different colour was employed as a type of typographical cue. Such inconsistencies in ways to enhance target constructions across studies could also have contributed to contradictory findings.

4.4.4. Relationship between Attention and L2 Development

The question of whether the amount of attention allocated to target linguistic constructions is related to learning target linguistic forms has been of interest to many researchers in the field of SLA (e.g., Godfroid et al., 2013; Izumi, 2002; Leow, 2001; Pellicer-Sánchez, 2016; Winke, 2013). This study has expanded on this line of research by investigating the relationships between attention, operationalized as eye-tracking indices, and gains on three different tests (i.e., oral productive test, written productive test, and receptive test). A number of significant relationships were evidenced between the amount of attention allocated to target linguistic constructions and L2 development. For the enhanced captions group, several significant medium to large relationships emerged between the eye-tracking indices and the gain scores on the oral and written productive tests. These results indicate that those learners who paid more attention to enhanced forms showed greater development in productive tests. Likewise, some of previous studies on this link found an association between increased attention to target forms and learning gains in vocabulary knowledge (e.g., Choi, 2017; Godfroid et al., 2013; Pellicer-Sánchez, 2016). In Pellicer-Sánchez's (2016) study, for instance, the total time participants spent reading target items was significantly related to their ability to recall the meanings of words. Similar result was also reported in Choi's (2017) study in that increased attention allocated to the textually enhanced collocations resulted in more learning.

However, some opposite results were reported in other studies, some of which, like the present study, focused on the effectiveness of textual enhancement in promoting L2 grammatical knowledge and used eye-tracking to measure the attention paid to target forms (e.g., Issa et al., 2015; Winke, 2013). In these studies, no significant link was observed between the attention allocated to target linguistic constructions and development in the use of L2 grammatical knowledge. In Study 1, the same result was also obtained: no relationship emerged between eye-movement indices and GJT gain scores for the enhanced captions group. However, in Study 2, eye-tracking indices were significantly related to gains on oral and written productive tests, but not to the gains on the receptive test, for the enhanced captions group. The contradictory findings elicited in these studies may be explained by the types of assessments that were employed to gauge development in learners' grammatical knowledge. The oral and written productive tests included in Study 2 were more likely to be meaning-oriented, in that participants were asked to watch the news in their native language (Korean) and report it in the target language (English). On the other hand, the receptive test, administered in a format of fill-in-the-blanks, could be characterized as a form-oriented measure, like the GJTs in Study 1. The fact that meaning-oriented productive tests were contextualized in the same way as the experimental treatment tasks, in which participants were exposed to target linguistic constructions, could have helped the learners use the constructions they had attended to while performing the tasks, thus reflecting the principle of transfer-appropriate processing (Lightbown, 2008). This principle claims that recall is more likely if a learning task and a retention task are comparable. However, many textual enhancement studies (e.g., Issa et al., 2015; Lee, 2007; Winke, 2013), including Study 1, have used form-oriented measurements (e.g., grammaticality judgement test, form correction task, sentence production) to assess development in the use of L2 grammatical knowledge (Han et al., 2008) that were not aligned with the meaning-oriented nature of the treatment. An interesting avenue of inquiry for future research is to explore whether the nature of the assessment can indeed influence the results of studies of textual enhancement.

It is also worthwhile to note that significant relationships emerged between the amount of time allocated to target linguistic forms (i.e., present perfect) and learners' gains on a written productive test for the non-enhanced captions group. For those learners presented with non-enhanced target linguistic constructions, medium positive correlations were observed between three eye-tracking indices (second pass reading, total fixation duration, and number of visits) and gains on a written productive test. Although fewer correlations were identified in comparison to the enhanced captions group, this finding indicates that those who spent more time rereading and fixating on unenhanced target constructions were, overall, more likely to achieve higher learning gains. The fact that students only showed a significant link on the written productive test could also be interpreted as providing further support for the principle of transfer-appropriate processing, as this test was the closest in nature to the treatment.

4.4.5. Moderating Effects of Working Memory in Attention Allocation and L2 Grammatical Knowledge Development

This study has examined the extent to which individual differences in working memory capacity moderated the effects of captions, textually non-enhanced or enhanced, included in multimodal input-based tasks on L2 development. To achieve this aim, different sub-constructs of working memory were measured using six measures. A non-word span task was used to assess phonological short-term memory, and a forward corsi block test was used to determine visual spatial short-term memory. The executive functions of updating, task-switching and inhibitory control were measured by an automated operation span (AOSPAN) task, a colour shape task, and a stop signal task, respectively. A backward corsi block test was employed as an additional measure of executive control, i.e., updating function.

The results of the linear mixed-effects models showed that visuospatial short-term memory, operationalized as a forward corsi block test, had a moderating effect on the relative gains of the no captions group and the enhanced captions group when use of the present perfect was measured by an oral productive test. A follow-up correlation analysis revealed a medium-size positive correlation between visuospatial short-term memory and gain scores on the oral productive test for the no captions group, but no significant correlation emerged for the enhanced captions group. That is, when learners were exposed to only aural and visual input, without captions, learners who had greater visual-spatial short-term memory performed better on an oral productive test for the present perfect. This result was somewhat unexpected since the visuospatial short-term memory, in general, was presumed to play an insignificant role in processing linguistic input. One possible explanation,however, may be given based on Wickens' (2007) the multiple-resource framework. According to Wickens, there are multiple cognitive resource pools differing along three dichotomous dimensions: processing stages (i.e., perception vs. response), modality (i.e., auditory perception with vocal response vs. visual perception with manual response), and codes of processing (i.e., verbal vs. spatial). The model also posits that interference between two tasks is more likely to occur if “any two tasks share common levels along more dimensions” (p. 187). Following this line of logic, the fact that both the processing of the visual stimuli and the processing of the aural stimuli required the participants to carry out perception (as opposed to response) tasks might have interfered with information processing. However, those with better visuospatial short-term memory in the no captions group were probably more successful in handling this competition given their superior ability to process visual information. This interpretation is tentative, more research is needed to explore this relationship.

In addition, the updating function of the central executive, measured by AOSPAN, had a mediating effect on the relative gains displayed by the no captions group and the enhanced captions group, with a medium effect size, when developmental change in the L2 target construction was assessed by a receptive test. A follow-up correlation analysis showed that the updating function positively correlated to a large extent with the gain scores on a receptive test for the enhanced captions group, but not for the no captions group. This result could be interpreted as indicating that learners who had greater updating ability did better on a receptive test for the present perfect when they were presented with highlighted target linguistic constructions. According to Miyake et al. (2000), the updating function is responsible for monitoring and manipulating information with reference to the task learners are engaged in; this function includes the ability to replace old and irrelevant information with new and more relevant information. The results of this study may indicate that participants who were better at updating their existing grammatical knowledge with newer information provided through multimodal input-based tasks benefited more from textually enhanced input and, consequently, obtained higher gains on a receptive test for the present perfect. Similarly, in Indrarathne and Kormos' (2018) study, working memory abilities were found to have an association with gains in receptive knowledge of the target construction, both in explicit (textual enhancement + instruction and textual enhancement + instruction + metalinguistic explanation) and implicit instructional conditions (input flood and textual enhancement). However, this study included a composite working memory of four measures (forward digit test, keep track task, plus minus task, and stroop task), so no more direct comparisons can be made.

Another finding that is worthy of discussion is that working memory had a negligible effect on the productive use of grammatical knowledge in this study. Indrarathne and Kormos reported a similar finding, in that working memory ability was associated with gains in receptive knowledge of the target construction in all input conditions, both explicit and implicit, whereas a weaker relationship was observed between gains in productive knowledge under their implicit learning condition, which is most similar to the condition in the present research.

For the past simple, no moderating effects of working memory on the effects of captions, non-enhanced or textually enhanced, in multimodal input-based tasks were observed. This could be due to the fact that learners' gains from pretest to immediate posttest and further to a delayed posttest in relation to the past simple were relatively small, leading to little variance in the scores.

4.4.6. Relationship between Working Memory and Attention

Considering that one of the central executive's functions is to regulate attention while performing complex cognitive tasks (Mackey et al., 2010), it can thus be reasonably assumed that learners' working memory may influence the extent to which learners allocate attention to target linguistic constructions. To address this hypothetical link, the extent to which the amount of attention allocated to target linguistic constructions correlated with working memory capacity was explored in this study. To accomplish this goal, correlation analyses between eye-tracking indices and working memory measures were computed.

For the non-enhanced captions group, no significant links between learners' working memory and the amount of attention paid to the present perfect were identified. However, for the enhanced captions group, medium-size positive correlations were observed between ASOPAN scores and three eye-tracking indices, namely, second pass reading, total fixation duration, and number of visits, to the present perfect. Given that the updating function of the central executive involves "monitoring and coding incoming information for relevance to the task at hand and then appropriately revising the item held in working memory" (Miyake et al., 2000, p. 57), it is logical that learners who were better at monitoring and updating old information with new information were more likely to spend longer reanalysing textually enhanced the target linguistic construction (i.e., the present perfect), as reflected in late measures of eye fixations (Roberts & Siyanova-Chanturia, 2013).

In addition, phonological short-term memory, operationalized as NWS, also had a medium-size positive correlation with second pass reading time for the present perfect for the enhanced captions group. This implies that learners who were better at storing and manipulating verbal information spent longer reanalysing the target linguistic construction, probably because they were more likely to notice new linguistic information.

Interestingly, with regard to the past simple, medium negative correlations were observed between NWS and all eye-tracking indices: first pass reading, second pass reading, total fixation duration, and total number of visits. AOSPAN was also negatively related to second reading with a medium size correlation coefficient. This finding indicated that learners who had greater phonological short-term memory and who were better at updating information did not engage in the process of reanalysing the textually non-enhanced target linguistic construction (i.e., the, past simple). One plausible explanation for this finding is that these learners could store and process more verbal information in their short-term memory and update information in working memory more effectively while receiving input through multiple modes; thus, it might not have been necessary for them to reanalyse the target linguistic construction.

CHAPTER 5

OVERALL DISCUSSION AND CONCLUSION

5.1. Overall Discussion

The primary goal of this thesis was to examine the extent to which textual enhancement included in multimodal input-based tasks can draw learners' attention to and, further, promote the use of L2 grammatical knowledge. Previous research on textual enhancement has been conducted in the context of reading, that is, the potential of textual enhancement in promoting learning has been explored in a single modality. Little research has attempted to investigate the effects of textual enhancement in captions, that is, when learners are exposed to textual enhancement during multimodal activities. Against this background, the aim of this study was to examine whether textual enhancement in captions influenced learners' allocation of attentional resources and helped promote the use of L2 grammatical knowledge. To achieve these aims, two empirical studies were conducted.

The primary purpose of Study 1 was to compare the capacity of two types of captions, textually non-enhanced and enhanced, to draw learners' attention to and foster second language development in anaphoric references. A pretest-posttet experimental design was employed with 48 Korean learners of L2 English, who were randomly assigned into a captions group (*n* = 24) and an enhanced captions group (*n* = 24). Both groups were asked to complete a series of treatment tasks, but for the enhanced captions group, the components of pronominal anaphoric reference were boldfaced in the treatment task input. Learner attention to anaphora antecedents and personal pronouns was assessed with eye-movement indices, and written and oral grammaticality judgement tests were used to measure learning gains. Overall, the results showed that textual enhancement succeeded in directing learners' attention to anaphora antecedents and led to increased gains in receptive knowledge of pronominal anaphoric references. Furthermore, significant links between attention and L2 development were observed for the unenhanced captions group. Based on the findings and motivated by the limitations of Study 1, a follow-up study, Study 2, was conducted to further examine the extent to which increased visual salience of the target linguistic construction, achieved through textual enhancement, could promote learners' attention to and subsequent development in second language (L2) grammar.

In Study 2, participants at a higher proficiency level were included. Accordingly, different target linguistic constructions were selected, which were use of the present perfect versus the past simple. As for the research design, a control group, which was not provided with captions, was included in order to obtain baseline data. A delayed posttest was also administered in an attempt to determine the retention of any effects of textual enhancement in captions on the use of L2 grammatical knowledge. In addition, the experimental treatment task was modified, using news clips, to increase authenticity. Given that many language learners watch news, movies and/or dramas to learn a second and foreign language, tasks incorporating real news clips were considered to have more face validity with students. In the captions, unlike in Study 1, which used bold fonts to highlight the target linguistic constructions, a different colour (yellow) was used to increase the visual salience of the constructions. As post-instructional measures, both form-oriented (i.e., fill in the blanks) and meaning-oriented measures (i.e., oral and written productive tests) were used. Given that individual differences in working memory might account for variations in second language acquisition (Juffs & Harrington, 2011), an additional aim was to determine the extent to which working memory mediated the effects of non-enhanced captions and textually enhanced captions on allocating attention to and promoting the use of the target linguistic constructions. Finally, a different statistical approach was adopted, namely, linear mixed-effects modelling.

The novelty of the studies lay in the potential of visually salient input to draw learners' attention being examined while input was presented through multiple modalities (aural, visual and/or textual), and the effects of different types of captioning on grammatical rather than lexical knowledge being investigated. With the help of eye-tracking methodology, an attempt was made to expand on existing knowledge by addressing the relationship between attention and L2 learning (e.g., Godfroid, Boers, & Housen, 2013). Another novel feature of this research was that each function of the central executive (i.e., shifting, inhibition and updating) was measured separately and, further, the role of various functions in allocating attentional resources and L2 grammar learning was also explored.

5.1.1. Effects of Multimodal Input-based Tasks with or without Captions on L2 Grammatical Knowledge Development

One of the primary purposes of Study 2 was to examine the extent to which multimodal input-based tasks without captions or with captions affected development in L2 grammatical knowledge. The results of Study 2 indicated that the provision of captions had a positive effect on learning the target construction. The captions group outperformed the no captions group on a written productive test and a receptive test. This finding is consistent with previous research, which found positive effects for captioning on L2 listening comprehension (e.g., Danan, 2004; Garza, 1991; Rodgers & Webb, 2017, Winke, Gass, & Sydorenko, 2010) and L2 vocabulary learning (e.g., Bird & Williams, 2992; Chai & Erlam, 2008; Danan, 1992; Markham, 1999; Sydorenko, 2010; Winke, Gass, & Sydorenko, 2010). This study has shown that captions can also benefit development in the use of L2 grammatical knowledge.

5.1.2. Effects of Textually Non-enhanced versus Enhanced captions on Development of L2 Grammatical Knowledge

On the basis of the results of Study 1 and Study 2, it appears that learners at both lower and higher proficiency levels are likely to benefit from the provision of textually enhanced captions in terms of development in L2 grammatical knowledge. In Study 1, the participants were characterized as less proficient users of English, comparable to bands B1–B2 according to the Common European Framework for Reference, and the target linguistic construction was the use of pronominal anaphoric references. In Study 2, more proficient users of English, falling into C1 band and above, were included and the target construction was use of the present perfect versus the past simple when delivering news. Regardless of the proficiency level of the learners and the type of target linguistic construction, textually enhanced input in captions led to greater gains in grammatical knowledge than exposure to unenhanced captions.

As discussed in Chapters 3 and 4, the positive results obtained here, relative to the mixed results emerged from previous studies of textual enhancement in the context of reading, might be attributed to factors such as (a) learners' prior knowledge of target linguistic constructions, (b) learners' proficiency level, (c) the salience of linguistic properties and (d) the fact that textual enhancement was included in the captions. Regarding the role of learners' previous knowledge of the target linguistic constructions, the participants in these studies were likely to have explicit knowledge of the target linguistic constructions (pronominal anaphoric references and the present perfect) at the onset of the study. This might have increased the likelihood that they would benefit from an implicit focus-on-form technique (Jourdenais et al., 1998). The proficiency level of the participants also needs to be taken into account when interpreting the positive effects of textual enhancement on the learning gains observed in the two studies. Given the fact that participants' proficiency level fell into B1-B2 bands in Study 1 and C1 band and above in Study 2 according to the Common European Framework for Reference, it might have been relatively easy for participants to notice the target linguistic constructions (i.e., pronominal anaphoric resolutions and the use of present perfect), as these grammatical constructions were perceptually salient in the input and the processing of the listening text was probably less demanding for these learners than it would have been for participants at lower levels of proficiency.

The degree of physical saliency of the linguistic constructions may also have contributed to the positive effects of textual enhancement on observed gains. For instance, in Study 2, the target linguistic construction, the present perfect, is made up of two adjacent components 'have/has + participle', which can be considered perceptually more salient. Anaphoric references in Study 1 are also salient features, in that they involve two full words rather than inflections.

Another factor that should be noted concerns the joint or "conflated" (Leow, 2009) use of textual enhancement and captions in this study. It has been argued that drawing a distinction between the two sub-strands of research, namely "non-conflated" and "conflated" studies, is important when interpreting the results of textual enhancement studies (Leow, 2009). The effects of textual enhancement might be more significant and pronounced when it is used in conjunction with other variables (e.g., feedback, explicit instruction or homework) in comparison to when it is used in isolation (Leow, 2009). Following this logic, in this research, the combination of textual enhancement with another attention-getting tool (i.e., captioning) may have triggered greater depth of processing and thereby contributed to better learning gains.

5.1.3. Effects of Non-enhanced Captions and Textually Enhanced Captions on Allocation of Attention

Overall, textual input enhancement provided in the multimodal input-based tasks had a positive impact on drawing learners' attention to the target linguistic constructions, as reflected in eye-gaze indices. Slight differences, however, were observed between Studies 1 and 2 regarding the degree to which the various eye-tracking indices changed as a function of enhancement. In Study 1, when presented with enhanced constructions, participants' second pass reading and total fixation duration were found to be significantly longer compared to when unenhanced constructions were provided. These two eye-tracking indices, second pass reading duration and total fixation duration, are categorized as late measures of eye-movements (Rayner, 1998, 2009). They are known "to be sensitive to later processes associated with comprehension of a text, such as information reanalysis, discourse integration, and recovery from processing difficulties" (Roberts & Siyanova-Chanturia, 2013, p. 217). Thus, the results of Study 1 suggested that textual enhancement included in multimodal input-based tasks was able to trigger more reflection and re-analysis of the target antecedents, but not initial processing of them. In Study 2, however, textual enhancement led to increased values for all eye-tracking indices, both an early measure (first pass reading) and late measures (second pass reading and total fixation duration). The differences observed between Study 1 and Study 2 in terms of eye-tracking results might be explained by the nature of the target linguistic constructions included in each study. The target linguistic construction of Study 1 was anaphoric references, which entail the processing of two related items; thus, learners might have spent more time re-analysing enhanced antecedents in relation to subsequent grammatical constructions (pronouns). However, in Study 2, the target linguistic construction (i.e., the present perfect) could be processed independently, making less reference to other forms.

It should also be noted that different types of typographical cues to highlight the target linguistic constructions were used in Study 1 and Study 2. Anaphoric reference constructions were enhanced using a boldfaced font in Study 1, whereas the font colour of the present perfect and the past simple was changed, using yellow, in Study 2. A review of previous eye-tracking studies, which examined the effects of textual enhancement on drawing learners' attention, highlighted that there are variations in the ways target linguistic constructions have been enhanced across studies. For instance, the target construction was underlined and coloured in red in Winke's (2013) study. In Issa et al.'s (2015) study and Loewen and Inceoglu's (2016) study, a different colour was used to highlight the target construction, whereas bolding was employed in Indrarathne and Kormos (2017). Some researchers (LaBrozzi, 2016; Simard, 2009) have argued that this variation in typographical cues might determine, at least partly, the degree to which textual enhancement influences L2 acquisition. Thus, it appears plausible that types of typographical cues can also have an impact on attentional allocation. To the best of my knowledge, the relationship between type of typographical cue and amount of attention allocated to target linguistic constructions has not been investigated. Further studies on this issue are therefore warranted to achieve a better understanding of the potential of visual salience in drawing learners' attention to target linguistic constructions.

An additional point that deserves attention is the different nature of visual stimuli presented in Study 1 and Study 2. The multimodal input-based tasks were developed using static visual input (i.e., pictures) in Study 1, whereas real news clips were included in Study 2. This could have influenced the learners' processing of the captions and further, might have contributed to the difference in the amount of attention allocated in the target linguistic constructions. That is to say, in Study 1, three static pictures were presented on the computer screen and the participants were asked to choose the correct one based on their understanding. Compared to the dynamic visual input (i.e. news clips) used in Study 2, the processing of static pictures fixed at the top of the computer screen might have consumed fewer attentional resources. This could explain why participants in Study 1 showed increased rereading and total reading duration of the target linguistic constructions. Probably they had more attentional resources left for processing linguistic information. Another possibility is that the fact that the pictures were static might have allowed participants to revisit the captions after viewing each picture before they had decided upon their answer.

5.1.4. Relationship between Attention and L2 Development

The relationship between attention, operationalised in terms of eye-tracking measures, and development in the use of target linguistic constructions was addressed in both Study 1 and Study 2. However, the results obtained in the two studies were found to be somewhat inconsistent. In Study 1, for the textually enhanced captions group, no significant links were observed between the amount of attention learners allocated to the target linguistic constructions and their learning gains, as measured by oral and written grammaticality judgement tests. In other words, increased attention paid to textually enhanced target constructions was not related to development in the use of an L2 grammatical construction when knowledge of grammar was assessed by a form-oriented measure. However, in Study 2, correlational analyses yielded a number of significant medium to large relationships between the eye-tracking indices and gain scores on oral and written productive tests, but consistent with the findings of Study 1, significant correlations were not evidenced between eye-tracking indices and gain scores on a form-oriented measure (i.e., a receptive test).

Taken together, in neither Study 1 nor Study 2, there was an association observed between the amount of attention allocated to target linguistic constructions and gains on form-oriented measures, the grammatical judgement test in Study 1 or the receptive fill-in-the-blanks test in Study 2. It is possible that the learners, in general, exhibited more improvement on form-oriented tests leading to less variation among their scores, thus the amount of attention allocated to target linguistic constructions did not influence gains.

Another possible explanation is that learners who paid more attention to the enhanced target linguistic constructions provided in multimodal input-based tasks performed better on meaning-oriented productive tests (but not form-oriented tests), since they were contextualized in the same way as the treatment tasks. As discussed before, this finding lends support to transfer-appropriate processing (Blaxton, 1989; Morris, Bransford, & Franks, 1977), which claims that compatibility between a learning task and a retrieval task may result in learners having better recall. In many textual enhancement studies, however, only form-oriented measurements have been used (Han et al., 2008); consequently, further studies employing a variety of post-instructional outcome measures, ranging from form-oriented to meaning-oriented, are needed.

As discussed previously (Chapter 3), somewhat mixed results emerged for the relationships between the amount of attention allocated to the target linguistic constructions and learning gains in Study 1 and Study 2. This can be interpreted as suggesting that longer fixations might not necessarily indicate that leaners were engaged in elaborate processing (Montero Perez et al., 2015). Alternatively, the increased amount of attention allocated to the target linguistic constructions may reflect that the learners encountered difficulty in processing the input. That is, they might have fixated longer on the target grammatical construction because it was difficult for them to process. Thus, the results appear to imply that the longer fixations on the target linguistic construction can reflect both greater cognitive effort to acquire novel L2 knowledge or greater effort invested in resolving a processing problem.

5.1.5. Moderating Effects of Working Memory in Attention Allocation and L2 Grammatical Knowledge Development

Premised on the assumption that "individual differences in memory and attentional capacity both affect the extent of noticing, thereby directly influencing SLA" (Robinson, 1995, p. 283), an examination of the extent to which working memory mediated the effects of textual enhancement in captions on development in the use of L2 grammatical knowledge was undertaken in Study 2. The novelty of this study lay in the fact that an attempt was made to measure different functions of working memory and assess their distinctive role in L2 attentional allocation and acquisition. To achieve this aim, six different working memory measures were administered. A non-word span task was employed to determine the capacity of phonological short-term memory. A forward corsi block task and a backward corsi block task were used to identify visuospatial short-term memory and the updating function of the central executive, respectively. In addition, three different functions of the central executive were measured with an automated operation task (AOSPAN, updating function), a colour shape task (switching function) and a stop signal task (inhibition function). A series of correlation analyses in R revealed that visuospatial short-term memory mediated gains in use of the present perfect as measured by an oral productive test, but only for learners who performed the treatment tasks without captions.

In fact, the role played by visuospatial short-term working memory in written language production has been addressed in previous studies (Kellogg, 1996; Kellogg, Olive, & Piolat, 2007). There is also one recent study which reported the potential role of visuospatial short-term memory in second language learning. Révész, Marije and Lee (2017) explored how phonological short-term memory, visuospatial short-term memory and each function of the central executive (i.e., updating, shifting and inhibition) influenced writing behaviour and text qualities. Somewhat similar to the findings of this study, visuospatial short-term memory along with phonological short-term memory and the switching function were found to have relationships with some of the measures of text quality or writing behaviour. However, in general, the relationship between visuospatial short-term memory and second language learning seems to be little researched while extensive SLA studies have focused on the role of phonological short-term memory and the central executive function. The results of this study along with previous studies (Kellogg, 1996; Révész et al., 2017), may raise the question of whether the potential of visuospatial short-term memory, along with phonological short-term memory and complex working memory, need to be taken into consideration in language learning and processing, particularly if input is to be provided through multiple modalities including a visual mode.

Furthermore, in the present study, the updating function of the central executive, operationalized as AOSPAN, was also found to have a moderating effect on gains on a receptive test for the learners when they were exposed to textually enhanced captions. Learners with higher AOSPAN scores were found to benefit more from an implicit focus on form (i.e., textual enhancement), which consequently led them to obtain higher gains on a receptive test of the present perfect. This was probably due to the fact that learners who had a greater updating function were better at replacing old and irrelevant information with new and more relevant information. Further studies including diverse working memory measures are needed to shed more light on the role of WM in learning from multimodal input.

5.1.6. Relationship between Working Memory and Attention

Given the theoretical assumption that learners' working memory (WM) may play a role in their allocation of attention (Robinson, Mackey, Gass, & Schmidt, 2012), possible associations between working memory measures and attention measures were examined in Study 2. More specifically, correlations were computed between eye-tracking indices (i.e., first pass reading, second pass reading, total fixation duration, and number of visits) and learners' performance on each working memory measure. For the enhanced captions group, two working memory measures were found to have significant positive relationships with eye-tracking indices, namely, NWS and AOSPAN. More specifically, NWS had a medium-size positive correlation with second pass reading while AOSPAN was positively related to second pass reading and total fixation. NWS and AOSPAN were used to measure phonological short-term memory and the updating function of the central executive, respectively; the results thus indicated that when textually enhanced input was provided, learners who had greater phonological short-term memory and who were better at updating information appeared to spend more time reanalysing visually enhanced target linguistic constructions.

However, somewhat different results were obtained for the non-enhanced captions group. NWS had significant but negative correlations with all eye-tracking indices (i.e., first pass reading, second pass reading, total fixation duration and total number of visits). AOSPAN also had a negative relation with second pass reading times. These results can be interpreted as meaning that learners who had greater phonological loop capacity and greater ability to update information appeared to spend less time reanalysing target linguistic constructions when they were not textually enhanced in the input. In other words, unenhanced target linguistic constructions were not reprocessed by those learners who could store and process more verbal information in their short-term memory and update information in working memory more effectively while receiving input through multiple modes.

An alternative interpretation of these findings is that textual enhancement made the target construction more perceptually salient and thereby might have attracted more attention from learners with higher working memory. Conversely, unenhanced target constructions did not draw the attention of learners who had greater working memory. These learners probably allocated their attentional resources to other, more salient, parts of the input, resulting in a negative relationship between attentional allocation and working memory for the non-enhanced captions group. Further research is needed to obtain a more fine-tuned understanding of these links.

5.2. Implications

5.2.1. Theoretical Implications

The importance of attention in second language acquisition has been addressed by many SLA researchers (e.g., Robinson 1995, 2003; Schmidt, 1990, 1993b, 2001) in recognition of the notion that not all input that is available is processed by learners (Corder, 1967). That is, learners' attention is necessary for input to be processed further and become intake and thus promote learning (Schmidt, 1990. 1993b, 2001). In line with this theoretical assumption, one goal of the current research was to shed more light on the association between the amount of attention allocated to target linguistic constructions and development in L2 grammatical knowledge. The results, similar to those of Godfroid and Uggen (2013) and Indrarathne and Kormos (2017), provide some evidence for the claim that the amount of attention L2 learners pay to grammatical constructions is linked to the extent of grammatical development they display.

The overall findings of Study 1 and Study 2 also provide evidence for the principle of transfer-appropriate processing, which argues that learners better recall what they have learned if the encoding and retrieval processes are similar (Blaxton, 1989; Lightbown, 2008; Morris, Bransford, & Franks, 1977). The combined results of Studies 1 and 2 showed a closer relationship between the amount of attention allocated to target linguistic constructions and learning gains for the enhanced group when the treatment and assessment tasks were more similar in nature. More specifically, participants did better on meaning-oriented assessment tasks (i.e., productive tests) than on form-oriented assessment tasks (i.e., grammaticality judgement test and receptive test), and meaning-oriented assessments were more aligned with the treatment.

The findings of this study also support the tenets of both the Dual coding Theory (Paivio, 1986, 2007) and the Cognitive Theory of Multimedia Learning (Mayer, 2001, 2002). These theoretical models assume that the provision of input through multiple modalities (aural and textual) can stimulate both verbal and imagery systems, which can in turn facilitate information processing and further improve learning outcomes. In this study, compared to the no captions group, learners provided with captions exhibited greater gains in the use of L2 target linguistic constructions in Study 2. Thus, the findings suggest, as predicted by Dual coding Theory and the Cognitive Theory of Multimedia Learning, that learners benefit more from input provided through multiple modes.

The results further confirmed the theoretical assumption that learners' working memory is related to how effectively they allocate their attentional resources (Robinson, Mackey, Gass, & Schmidt, 2012). In this research significant relationships between individual differences in working memory and the allocation of attentional resources were identified. In particular, learners who had greater phonological short-term memory and who were better at updating information devoted more time to re-analysing textually enhanced target linguistic constructions. An additional theoretical contribution made by this thesis concerns the possible distinctive role of three central executive functions, namely, shifting, inhibition and updating in second language acquisition. According to Miyake et al. (2000), these central executive functions may not be a single process, they may rather be separable; thus, an attempt was made in this thesis to measure each function and to examine its relation to the amount of attention directed to the target linguistic constructions and to learning gains. A series of correlation analyses showed no significant correlations among most of the working memory measures except for a link between a nonword span task and AOSPAN and between forward and backward corsi block tasks. Thus, these findings support the notion that the three functions of the central executive (updating, switching and inhibition) are indeed separable. Furthermore, as noted above, only one of the central executive functions (i.e., updating) was found to play a role in attention allocation and development in L2 grammatical knowledge. This result can also be interpreted as initial evidence for a theoretical assumption that central executive functions may be separable and, further, may have different roles in learners' use of attentional resources and second language acquisition.

5.2.2. Methodological Implications

The present thesis has a number of methodological implications. First, the findings underscore the importance of using a variety of eye-tracking measurements when investigating the relationship between attentional allocation and L2 development, including both early and late measures. Early measures capture initial processing, whereas late measures are an indicative of reanalysis or processing difficulties (Roberts & Siyanova-Chanturia, 2013). Early measures (e.g., first pass reading) and late measures (e.g., second pass reading) were found to yield different results in the present study, demonstrating that it is helpful to incorporate both early and late measures in instructed SLA research to obtain a fuller picture of the cognitive process triggered by instructional interventions.

This research also offers a key methodological implication regarding the use of assessment tasks. The results obtained from Study 1 and Study 2 suggest that the type of outcome measure may play a critical role in whether learners can retrieve what they have encoded during the learning phase. In this research, a close relationship was revealed between the amount of attention allocated to perceptually salient target linguistic constructions and the learning gains displayed on meaning-oriented tests (i.e., productive tests), but not on form-oriented tests (i.e., a grammatical judgement test and a receptive test). Thus, the administration of both form-oriented and meaning-oriented measures is recommended for further studies in order to gain a fuller understanding of the relationship between attentional allocation and development in L2 grammatical knowledge.

An additional methodological implication concerns the use of different measures to determine distinctive functions of working memory. In previous studies, two common components of working memory that have been investigated are the phonological loop and the central executive; accordingly, a nonword span task or a digit span task have commonly been used to measure phonological short-term memory, with a reading span task or an operation task being employed to assess the central executive. However, motivated by the argument that functions of the central executive are separable (Miyake et al., 2000), Study 2 included measures to assess various functions of the central executive, i.e., switching, inhibition and updating. As distinct results were found for different components of WM, future studies investigating the central executive are also advised to include measures tapping into its various subcomponents.

Finally, a methodological contribution made by this study is the advanced statistical analysis, linear mixed-effects modelling, which was used to analyse the data. Differing from parametric statistics (e.g., *t*-test or ANOVA), mixed-effects models take participants and items into account as random variance (Baayen, 2008; Cunnings, 2012; Gagné & Spalding, 2009; Linck & Cunnings 2015; Winter, 2013). This can help to address the “language-as-fixed-effect fallacy” (Clark, 1973), that is, the difficulty of generalizing results beyond the participants and linguistic items included in a given study (Linck & Cunnings, 2015). Thus, although the use of mixed effects modelling is gaining ground in SLA, empirical studies that include randomly selected participants and items as independent variables should routinely use linear mixed-effects models to produce more valid and generalizable results.

5.2.3. Pedagogical Implications

One pedagogical implication of this research is that highlighting target grammatical constructions in captions can draw learners’ attention and thereby promote L2 learning. Contrary to previous claims that textual enhancement has only marginal effects on promoting L2 grammatical knowledge (Lee & Huang, 2008), the results of this research suggest the textual enhancement is valuable and useful as an instructional intervention for language learners, particularly for those who already have prior knowledge of target linguistic constructions.

Another important pedagogical implication is that multimedia input-based tasks, using captioned videos, can be successfully used for instructional purposes. This is a key finding, given that multimedia materials (e.g., podcasts, DVDs and YouTube) are becoming increasingly available and used by many L2 learners in both formal and informal L2 settings. The findings of the present study also provide evidence for the pedagogical value of using captions in promoting L2 grammatical knowledge development, thereby extending previous research which has reported positive effects of captions for developing L2 listening comprehension and vocabulary knowledge.

5.3. Limitations and Future Directions

In interpreting the findings of this research, it is also important to take into account the limitations of this study. One methodological weakness lies in the fact that the eye-tracking measures were not triangulated with verbal protocol comments. The combination of eye-gaze and verbal protocol data would have allowed to tap into not only the amount of attention participants paid to enhanced features but also into the level of processing in which they engaged when encountering highlighted anaphoric and tense-aspect constructions. As a result, the interpretations of the findings made here could have been less tentative. In future research, the incorporation of stimulated recall protocols would appear suitable to gain insights into conscious operations that may be induced by textually enhanced captions. Unlike think-alouds, which require verbalisation, this procedure does not interfere with the online processing of aural input, although it is potentially affected by memory decay (Gass & Mackey, 2016).

Another limitation concerns the choice of target constructions. The target grammatical constructions, the pronominal anaphoric references for Study 1 and the use of the present perfect for Study 2, might have been somewhat easy for the participants, which could have affected the amount of attention paid to them during the treatment, participants' performance on the posttests, and furthermore, the relationship between attentional allocation and learning gains. In the preparation stage, great caution was exercised when selecting the target linguistic constructions. First, I intended to choose features that L2 learners, in general, have difficulties with when learning English as a second language. Second, my aim was to select features that are demanding for Korean learners due to differences between the grammatical systems of Korean and English. In addition, the outcome measures included a considerable number of distractors so as to divert learners' attention from the target constructions and thereby prevent them from easily identifying the focus of the assessments. However, since the participants had learned the rules of the target linguistic constructions through explicit explanation, it might still have been easy for them to notice the target constructions, especially when these were visually salient in the input. Other target linguistic constructions that are less salient in the input may yield differential effects for textual enhancement in future studies.

As for Study 2, another consideration has to do with the frequency of use of the present perfect and the past tense in news clips in present day English overall and across various dialects. Although one of the usages of the present perfect is to announce recent news (McCawley, 1971), and the news items used in the present study came from various dialects including British and American English, there are differences in the frequency of use of the present perfect across dialects. For example, the present perfect is more commonly used in British English than in American English (Quirk, Greenbaum, Leech, & Svartvik, 1985; Swan, 2005). Considering the fact that English learners in South Korea are more commonly exposed to American English, an investigation of a target linguistic construction to which learners are less likely to be exposed in real-life contexts could be a limitation of this study. Hence, further textual enhancement studies could focus on other uses of the present perfect and/or be conducted in other contexts.

An additional limitation concerns the content of the Information Sheet which was used to explain the purpose and procedures of the study to the participants at the outset of the study. The fact that the specific aim of the study and the target linguistic constructions were specified in the Information Sheet could have undermined the primary purpose of the study, which was to examine the effects of textual enhancement, a type of implicit focus on form instruction, on learners' attentional allocation and subsequent development. However, given that the majority of the participants appeared to listen to the researcher's brief oral explanation, rather than read the Information Sheet before providing consent, it is unlikely that participants had become aware of the target linguistic constructions prior to the experiment. In fact, the comments on the exit questionnaire indicated that the majority of participants had gained no awareness of the target linguistic constructions during the experiment. Nevertheless, this remains a weakness of the research that needs to be considered when interpreting the findings.

A further limitation of the study concerns the design of the multimodal input-based tasks, particularly the nature of the captions. In an effort to ensure ecological validity, the captions were inserted at the bottom of the screen, using a 14-point font, similar to how captions are presented in real-life settings. However, this led to some limitations from the perspective of construct validity. Analyses of eye-movements require a selection of areas of interest (i.e., target linguistic constructions) to determine fixation duration or the numbers of visits to areas of interest. To ensure that a learner's eye fixation is located within selected areas of interest, that is on target linguistic constructions, an enlarged font with wider line spacing is preferable to increase eye-movement data accuracy. However, in the current studies, a normal font size with regular line spacing was used to achieve ecological validity. The debate about whether it is better to improve ecological validity or increase the accuracy of eye-movement data seems to be ongoing (Spinner, Gass, & Behney, 2013) and remains to be resolved.

A further consideration has to do with the areas of interests defined for the personal pronouns in Study 1. Pronouns, such as ‘he’ and ‘she’, are categorized as function words and relatively short in terms of length. Thus, there is a possibility that they have been skipped by the participants while reading the captions. This, in turn, might have influenced the results for number and duration of eye-fixations at the pronouns. One way to overcome this in further studies is to include the preceding word in the area of interest.

Another apparent limitation concerns the eye-tracker system used in this thesis to collect the eye-movement data. A remote eye-tracker was used, which is considered more ecologically valid than head-and-chin-rest eye-trackers because learners' movements are not constrained. However, due to the fact that learners can move their head and change position while performing tasks, there is a high probability that eye gaze may not be captured as precisely as when the treatment is administered under more controlled conditions. An additional issue that needs to be raised has to do with the resolution of the eye-tracker. In this study, Tobii X2-60 mobile eye tracker with a temporal resolution of 60 Hz was used to record participants’ eye movements while performing the treatment tasks. In general, the accuracy of eye-movement data collected with an eye-tracker using a low-temporal resolution could be questioned. However, since this study focused only on fixation analysis, the use of an eye tracker with 60 Hz temporal solution is arguably acceptable. According to Raney, Campbell and Bovee (2014, p. 2), “the average temporal error will be approximately half the duration of the time between samples." Thus, a sampling rate of 60 Hz will result in an error of about 8 msec on average. As argued by Raney et al., while an 8 msec error might be too large to examine saccade durations, it is not too large to investigate fixation durations. Nevertheless, an eye-tracker with higher resolution would have provided more accurate data in terms of fixations as well.

Other weaknesses of the study include its focus on a limited number of grammatical features and particular task types. Replication of this research with constructions that are of lower physical salience and communicative value would be especially warranted, since such linguistic targets are less prone to attracting attention and thus being acquired by L2 learners from exposure to input alone (e.g., Long & Robinson, 1998). Future studies are also needed to examine whether the results found here can transfer to different tasks. For instance, in this thesis, learners were exposed to target linguistic constructions mainly through input-based tasks in which they were not required to produce any output. It could be worthwhile examining whether learners would also benefit from textual enhancement in developing L2 grammatical knowledge through tasks that also entail some kind of production, such as speaking or writing.

Despite its obvious limitations, the present research extends the line of research on the effectiveness of increasing the visual salience of target constructions in captions to draw learner attention to target linguistic constructions and, further, to promote L2 grammatical knowledge. Future studies on various grammatical features are needed in order to elucidate further the pedagogical value of textual enhancement in multimodal input-based tasks for facilitating the acquisition of L2 grammatical knowledge.

REFERENCES

Adams, A.M., & Guillot, K. (2008). Working memory and writing in bilingual students. *International Journal of Applied Linguistics, 156,* 13–28.

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259–302). Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Centre.

Alptekin, C., & Erçetin, G. (2009). Assessing the relationship of working memory to L2 reading: Does the nature of comprehension process and reading span task make a difference? *System*, *37*, 627–639.

Alptekin, C., & 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*, 206–219.

Alptekin, C., & Erçetin, G. (2015). Eye movements in reading span tasks to working memory functions and second language reading. *Eurasian Journal of Applied Linguistics, 1,* 35–56.

Alsadoon, R., & Heift, T. (2015). Textual input enhancement for vowel blindness: A study with Arabic ESL learners. *Modern Language Journal, 99*, 57–79.

Altgassen, M., Vetter, N. C., Phillips, L. H., Akgün, C., & Kliegel, M. (2014). Theory of mind and switching predict prospective memory performance in adolescents. *Journal of experimental child psychology, 127,* 163–175.

An, S. (2008). The function of third-person '*Ku, Kunye*'. *Korean Linguistics, 38,* 145–164.

Ardila, A. (2003). Language representation and working memory with bilinguals. *Journal of Communication Disorders, 36,* 233–240.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge, UK: Cambridge University Press.

Bachman, L., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language test*s. Oxford: Oxford University Press.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D. (1992). Working memory. *Science, 255*, 556–559.

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A*, 5–28.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Science*, *4*, 417–423.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews: Neuroscience, 4*, 929–939.

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–90). New York: Academic Press.

Baltova, I. (1999). Multisensory language teaching in a multidimensional curriculum: The use of authentic bimodal video in core French. *The Canadian Modern Language Review, 56*, 31–48.

Barcroft, J. (2003). Distinctiveness and bidirectional effects in input enhancement for vocabulary learning. *Applied Language Learning, 13*, 133–159.

Bardovi-Harlig, K. (2001). Another piece of the puzzle: The emergence of the present perfect. *Language learning, 5*, 215–264.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278.

Başaran, H. F., & Köse, G. D. (2013). The effects of captioning on EFL learners' listening comprehension. *Procedia – Social and Behavioral Sciences, 70,* 702–708.

Bates, D. M., Maechler, M., & Bolker, B. (2012*). lme4: Linear mixed-effects models using S4 classes*. R package version 0.999999-0.

Bax, S., & Weir, C. (2012). Investigating learners' cognitive reading processes during a computer-based CAE reading test. *University of Cambridge ESOL Examinations Research Notes, 47,* 3–14

Bensalem, E. (2016). Effects of captioning video clips on vocabulary learning among ESL learners. *International Journal of Humanities and Cultural Studies, 3,* 381–395.

Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Linguistics, 23,* 509–533.

Bisson, M. J., van Heuven W. J. B., Conklin, K., & Tunney, R. J. (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics, 35*, 399–418.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 657-688.

Blom, E., Paradis, J., & Sorenson Duncan, T. (2012). Effects of input properties, vocabulary size, and L1 on the development of third person singular –s in child L2 English. *Language Learning, 62,* 965–994.

Borras, I., & Lafayette, R. C. (1994). Effect of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Language Journal*, *78*, 61–75.

Bowles, M. A. (2003). The effects of textual input enhancement on language learning: An online/offline study of fourth-semester Spanish students. In P. Kempchinsky & C. E. Piñeros (Eds.), *Theory, practice, and acquisition: Papers from the 6th Hispanic linguistics symposium and the 5th conference on the acquisition of Spanish and Portuguese* (pp. 395–411). Somerville, MA: Cascadilla Press.

Bowles, M. A., & Leow, R. P. (2005). Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation. *Studies in Second Language Acquisition*, *27,* 415–440.

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics, 36*, 1–22.

Brown, R. (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.

Brunfaut, T., & McCray, G. (2015). Looking into test-takers' cognitive processes whilst completing reading tasks: A mixed-method eye-tracking and stimulated recall study. *ARAGs Research Reports Online*, AR/2015/001.

Bygate, M., Skehan, P., & Swain, M. (2001). *Researching pedagogic tasks: Second language learning, teaching and testing*. New York: Longman.

Candlin, C. (1987). Towards task-based language learning. In C. Candlin & D. Murphy (Eds.). *Language learning tasks* (pp. 5–22). Englewood Cliffs, NJ: Prentice Hall.

Cantor, J., & Engle, R. (1993). Working memory capacity as long-term activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1101–1114.

Chai, J., & Erlam, R. (2008). The effect and the influence of the use of video and captions on second language learning. *New Zealand Studies in Applied Linguistics, 14*, 25–44.

Chang, A. C-S. (2009). Gains to L2 listeners from reading while listening vs. listening only in comprehension short stories. *System, 37*, 652–663.

Chang, S-J. (1996). *Korean*. Amsterdam/Philadelphia: John Benjamins.

Chapelle, C. A. (2003). *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology*. Amsterdam, Netherlands: John Benjamins.

Chaudron, C. (1985). Intake: On models and methods for discovering learners' processing of input. *Studies in Second Language Acquisition, 7*, 1–14.

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology, 32*, 867–873.

Cho, M., & Reinders, R. (2013). The effects of aural input enhancement on L2 acquisition. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka, (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 133–148). Honolulu: University of Hawai'i, National Foreign Language Resource Center.

Choi, S. (2017). Processing and learning of enhanced English collocations: An eye movement study. *Language Teaching Research*, *21*, 403-426.

Choi, U-H. (2007). A study on person in Korean. *The Journal of Linguistic Science, 40*, 33–52.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behaviour*, *12*, 335–359.

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 341–372). Oxford, UK: Elsevier.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). New York: Academic Press.

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

Congdon, E., Mumford, J. A., Cohen, J. R., Galvan, A., Canli, T., & Poldrack, R. A. (2012). Measurement and reliability of response inhibition. *Frontiers in psychology, 3,* 1–10.

Conklin, K., & Pellicer–Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research, 32,* 453–467.

Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-Tracking: A Guide for Applied Linguistics Research*. Cambridge University Press.

Corder, S. (1967). The significance of learners' errors. *International Review of Applied linguistics, 5,* 161–170.

Corsi, P. M. (1972). *Human memory and the medical temporal region of the brain.* Unpublished doctoral dissertation. McGill University, Montreal, Canada.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*, 163–191.

Cowan, N. (1993). Activation attention, and short-term memory. *Memory & Cognition*, *21*, 162–167.

Cowan, R. (2008). *The teacher's grammar of English*. Cambridge: Cambridge University Press.

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*, 369–382.

Cunnings, I., & Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language, 75*, 117–139.

Danan, M. (1992). Reversed subtitling and dual coding theory: New directions for foreign language instruction. *Language Learning, 42*, 497–527.

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta, 49*, 67–77.

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research, 2*0, 445–464.

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

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25*, 1–18.

Dekeyser, R, M. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press.

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern spans: A tool for unwelding visuo-spatial memory. *Neuropsychologia*, *37*, 1189–1199.

Denhovska, N., Serratrice, L., & Payne, J. (2016). Acquisition of second language grammar under incidental learning conditions: The role of frequency and working memory. *Language Learning, 66*, 159–190.

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativiation. *Studies in Second Language Acquisition, 13*, 431–469.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). New York: Cambridge University Press.

Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 256–310). New York: Basil Blackwell.

Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197–261). New York: Cambridge University Press.

Dulay, H., & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning, 26*, 321–351.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30*, 257–303.

Dussias, P. E. (2010). Uses of eye-tracking data in second language sentence processing research. *Annual Review of Applied Linguistics, 30*, 149–166.

Dussias, P. E., Valdés Kroff, J. R., Tamargo, R. E. G., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition, 35,* 353–387.

d'Ydewalle, G., & De Bruycker, W. (2007). Eye movement of children and adults while reading television subtitles. *European Psychologist, 12*, 196–205.

Eastwood, J. (1994). *Oxford guide to English grammar*. Oxford: Oxford University Press.

Ellis, N. (2017). Salience in usage-based SLA. In S. Gass, P. Spinner & J. Behney (Eds.). *Salience in Second Language Acquisition* (pp. 21-40). New York: Routledge

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (2003). *Task-based language teaching and learning*. Oxford: Oxford University Press.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.

Ellis, R. (2009). Task-based language teaching: sorting out the misunderstandings. *International Journal of Applied Linguistics, 19,* 221–246.

Ellis, R. (2013). Task-based language teaching: Responding to the critics. *University of Sydney Papers in TESOL, 8*, 1–27.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. New York: Routledge.

Engle, R. W., Cantor, J., & Carullo, J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18*, 972–992.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999a). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). New York: Cambridge University Press.

Enticott, P. G., Ogloff, J. R., & Bradshaw, J. L. (2006). Associations between laboratory measures of executive inhibitory control and self-reported impulsivity. *Personality and Individual Differences, 41*, 285–294.

Erlam, R. (2003). Evaluating the relative effectiveness of structured-input and output-based instruction in foreign language learning. *Studies in Second Language Acquisition, 25*, 559–582.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research, 9*, 147–172.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics, 27*, 464–491.

Felser, C., Roberts, L., Gross, R., & Marinis, T. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24*, 453–489.

Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: a cross-modal priming study. *Second Language Research, 23*, 9–36.

Foote, R. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics, 32*, 187–220.

Fortkamp, M. B. M. (1998). Measures of working memory capacity and L2 oral fluency. *Ilha do Desterro, 35*, 201–238.

Foster, P. (2009). Task-based language learning research: Expecting too much or too little? *International Journal of Applied Linguistics, 19,* 221–246.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*, 172–179.

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language, 51*, 136–158.

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior research methods, 37,* 581–590.

Frumuselu, A. D., De Maeyer, S., Donche, V., & Plana, M. D. M. G. C. (2015). Television series inside the EFL classroom: Bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education*, *32*, 107–117.

Gabriele, A. (2009). Transfer and transition in the SLA of aspect. *Studies in Second Language Acquisition, 31,* 371–402.

Gagné, C. L., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language, 60*, 20–35.

Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals, 24*, 239–258.

Gascoigne, C. (2006). Explicit input enhancement: Effects on target and non-target aspects of second language acquisition. *Foreign Language Annals, 59*, 551–564.

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.

Gass, S. M., & Mackey, A. (2016). *Stimulated Recall Methodology in Applied Linguistics and L2 Research*. New York: Routledge.

Gass, S. M., Spinner, P., & Behney, J. (2017). Salience in second language acquisition and related field. In S. Gass, P. Spinner & J. Behney (Eds.). *Salience in Second Language Acquisition* (pp. 1-18). New York: Routledge.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27,* 513–543.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. New York: Psychology Press.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press.

Gilabert, R., & Muñoz, C. (2010). Differences in attainment and performance in a foreign language: The role of working memory capacity. *International Journal of English Studies, 10,* 19–42.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition, 35*, 483–517.

Godfroid, A., Housen, A., & Boers, F. (2010). A procedure for testing the noticing hypothesis in the context of vocabulary acquisition. In M. Pütz & L. Sicola

(Eds). *Inside the learner's mind: Cognitive processing and second language acquisition* (pp. 169–197). Amsterdam/Philadelphia: John Benjamins.

Godfroid, A., & Uggen, M. S. (2013). Attention to irregular verbs by beginning learners of German. *Studies in Second Language Acquisition, 35*, 291–322.

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *The Journal of Neuroscience, 33,* 387–396.

Goldschneider, J., & DeKeyser, R. (2001) Explaining the 'natural order of L2 morpheme acquisition' in English: A meta-analysis of multiple determinants. *Language Learning, 51,* 27–77.

Grabe, W. (2012). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.

Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. *System, 36,* 85–93.

Guillory, H. C. (1998). The effects of key word captions to authentic French video in foreign language instruction. *CALICO Journal, 15*, 89–108.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18,* 1279–1296.

Hama, M., & Leow, R. P. (2010). Learning without awareness revisited. *Studies in Second Language Acquisition, 32*, 465–491.

Han, J., & Hong, S. (2015). The acquisition problem of English present perfect to Korean adult learners of English: L1 transfer matters. *English Language and Linguistics, 213,* 141–164.

Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics, 29*, 597–618.

Harrington, M. (1992). Working memory capacity as a constraint on L2 development. *Advances in psychology, 83*, 123–135.

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

Havik, E., Roberts, L., van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning, 59*, 73–112.

Hayati, A., & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology, 42*, 181–192.

Huang, H., & Eskey, D. (2000). The effects of closed-captioned television on the listening comprehension of intermediate English as second language students. *Educational Technology Systems, 28*, 75–96.

Huang, Y. (2005). Anaphora, cataphora, exophora, logophoricity. In K. Brown (Ed.) *Encyclopaedia of language and linguistics* (pp. 231–238). New York: Elsevier.

Im, C. G. (1993). The acquisition of English anaphora and its usage. *Studies in Modern Grammar, 2*, 157–177.

Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions: An eye-tracking study. *Studies in Second Language Acquisition, 39*, 401–430.

Indrarathne, B., & Kormos, J. (2018). The role of working memory in processing L2 input: Insights from eye-tracking. *Bilingualism: Language and Cognition, 21*, 355–374.

Issa, B. Morgan-Short, K. Villegas, B. & Raney, G. (2015). An eye-tracking study on the role of attention and its relationship with motivation. In L. Roberts, K.

McManus, N. Vanek & D. Trenkic (Eds.), *EUROSLA Yearbook 2015* (pp. 114–142). Amsterdam, Netherlands: John Benjamin.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*, 541–577.

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition. *TESOL Quarterly, 34*, 239–278.

Jackson, C. N., Dussias, P. E., & Hristova, A. (2012). Using eye-tracking to study the on-line processing of case-marking information among intermediate L2 learners of German. *International Review of Applied Linguistics, 50*, 101–133.

Jahan, A., & Kormos, J. (2015). The impact of textual enhancement on EFL learners' grammatical awareness of future plans and intentions. *International Journal of Applied Linguistics, 25*, 46–66.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, Whole No. 89.

Jiang, N. (2012). *Conducting reaction time research in second language studies*. New York: Routledge.

Johnson, J. (1992). Critical period effects in second language acquisition: The effect of written versus auditory materials on the assessment of grammatical competence. *Language Learning, 42*, 217–248.

Jourdenais, R., Ota, M., Stauffer, S., Boyson, B., & Doughty, C. (1995). Does textual enhancement promote noticing? A think aloud protocol analysis. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 183–216). Honolulu: University of Hawai'i National Foreign Language Resource Center.

Juffs, A. (2004) Representation, processing and working memory in a second language. *Transactions of the Philological Society, 102*, 199–225.

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research, 21*, 121–151.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching, 44*, 137–166.

Jung, J. (2017). *Effects of task complexity, glossing, and working memory capacity on L2 reading and L2 learning*. Unpublished dissertation, University College London Institute of Education, London, UK.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive psychology, 8*, 441–480.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Kaushanskaya, M., & Marian, V. (2007). Bilingual language processing and interference in bilingualism: Evidence from eye tracking and picture naming. *Language Learning, 57*, 119–163.

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning, 59*, 503–353.

Keating, G. D., & Jegerski, J. (2015). Experimental designs in sentence processing research. *Studies in Second Language Acquisition, 37*, 1–32.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.). *The science of writing: Theories, methods, individual differences and applications* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum.

Kellogg, R. T., Olive, T., & Piolat, A. (2007). Verbal, visual, and spatial working memory in written language production. *Acta Psychologica, 124*, 382–397.

Kessels, R. P., Van Zandvoort, M. J., Postma, A., Kappelle, L. J., & De Haan, E. H. (2000). The Corsi block-tapping task: standardization and normative data. *Applied neuropsychology, 7*, 252–258.

Kim, Y. J. (1997). The acquisition of Korea. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (pp. 409–427). Mahwah, NJ: Lawrence Erlbaum.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11*, 261–271.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. London/New York: Longman.

LaBrozzi, R. M. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research, 20*, 75–91.

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.

Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' analytic review. *Language Learning 57*, 87–118.

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*, 307–331.

Leech, G. N. (1971). *Meaning and the English verb*. London: Longman.

Leeser, M. J. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning, 57*, 229–270.

Leow, R. P. (1997). The effects of input enhancement and text length on adult L2 readers' comprehension and intake in second language acquisition. *Applied Language Learning, 8*, 151–182.

Leow, R. P. (1999). Attention, awareness, and focus on form research: A critical overview. In J. F. Lee & A. Valdman (Eds.), *Form and meaning: Multiple perspectives* (pp. 69–96). Boston: Heinle & Heinle.

Leow, R. P. (2000). A study of role of awareness in foreign language behavior: Aware versus unaware learners. *Studies in Second Language Acquisition, 22,* 557–584.

Leow. R. P. (2001). Attention, awareness, and foreign language behavior. *Language Learning, 51*, 113–155.

Leow, R. P. (2007). Input in the L2 classroom: An attentional perspective on receptive practice. In R. DeKeyser (Ed.), *Practice in second language learning: Perspectives from applied linguistics and cognitive psychology* (pp. 21–50). Cambridge, UK: Cambridge University Press.

Leow, R. P. (2009). Input enhancement and L2 grammatical development: What the research reveals. In J. Watzinger-Tharp, & S. L. Katz, (Eds), *Conceptions of L2 grammar: Theoretical approaches and their application in the L2 classroom* (pp. 16–34). Boston, MA: Heinle Publishers.

Leow, R. P. (2015). *Explicit learning in the L2 classroom*. New York: Routledge.

Leow, R. P., Egi, T., & Moreno, N. (2008). Attention to form and meaning revisited. *Language Learning*, 58, 665–695.

Leow, R. P., Egi, T., Nuevo, A. M., & Tsai, Y. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning, 13,* 1–16.

Leow, R. P., Grey, S., Marijuan, S., & Moorman, C. (2014). Concurrent data elicitation procedures, processes, and the early stages of L2 learning: A critical overview. *Second Language Research, 30*, 111–127.

Leow, R. P., & Martin, (2017). Enhancing the input to promote salience of the L2: A critical overview. In S. Gass, P. Spinner & J. Behney (Eds.), *Salience in SLA* (pp. 167–186). New York: Routledge.

Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition, 26*, 35–57.

Li, S. (2012). The effects of input-based practice on pragmatic development of requests in L2 Chinese. *Language Learning, 62*, 403–438.

Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal, 97*, 634–654.

Li, S., & Taguchi, N. (2014). The effects of practice modality on pragmatic development in L2 Chinese. *The Modern Language Journal, 98,* 794–812.

Lightbown, P. M. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Z. Han (Ed.), *Understanding second language process* (pp. 27–44). Clevedon, UK: Multilingual Matters.

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning, 65*, 185–207.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*, 861–883.

Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology, 94*, 778–784.

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language teaching, 40*, 243–249.

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 55–75). Oxford: Elsevier.

Loewen, S., & Inceoglu, S. (2016). The effectiveness of visual input enhancement on the noticing and L2 development of the Spanish past tense. *Studies in Second Language Learning and Teaching, 6*, 89–110.

Logan, G. D. (1985). On the ability to inhibit simple thoughts and actions: II. Stop-signal studies of repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 675–691.

Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.

Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4*, 126–141.

Long, M. (1985). Input and second language acquisition theory. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 377–393). Rowley, MA: Newbury House.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. deBot, C. Kramsch, & R. Ginsberg, (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: John Benjamin.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.

Long, M. (2000). Focus on form in task-based language teaching. In R. D. Lambert & E. Shohamy (Eds.), *Language policy and pedagogy: Essays in honor of A. Ronald Walton* (pp. 179–192). Philadelphia: Benjamins.

Long, M., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge: Cambridge University Press.

Lwo, L., & Lin, M. C-T. (2012). The effects of captions in teenagers' multimedia L2 learning. *ReCALL, 24*, 188–208.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied linguistics*, *27*, 405–430.

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, *60*, 501–533.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, *22*, 471–497.

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–209). Amsterdam: John Benjamins.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, *27*, 53–78.

Markham, P. (1989). The effects of captioned television videotapes on the listening comprehension of beginning, intermediate, and advanced ESL students. *Educational Technology, 29*, 38–41.

Markham, P. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals, 32*, 321–328.

Markham, P. (2001). The influence of culture-specific background knowledge and captions on second language comprehension. *Journal of Educational Technology Systems, 29*, 331–343.

Markham, P., Peter, L., & McCarthy, T. (2001). The effects of native language vs. target language captions on foreign language students' DVD video comprehension. *Foreign Language Annals, 34*, 439–445.

Markham, P., & Peter, L. (2002-2003). The influence of English language and Spanish language captions on foreign language listening/reading comprehension. *Journal of Educational Technology Systems, 31*, 331–341.

Martin, K. I., & Ellis, N. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34*, 379–413.

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

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge: Cambridge University Press.

McCawley, J. D. (1971). Tense and time reference in English. In C. J. Filmore, & T. C. Langdoen (Eds.), *Studies in linguistic semantics* (pp. 96–113). New York: Holt, Rinehart & Winston.

McLaughlin, B. (1987). *Theories of second language learning*. London: Edward Arnold.

Millisecond. (2015). Inquisit 4 Lab [Computer software]. Retrieved from http://www.millisecond.com/

Mitkov, R. (2002). *Anaphora resolution*. London: Longman.

Miyake, A. (2001). Individual differences in working memory: Introduction to the special section. *Journal of Experimental Psychology: General, 130*, 163–168.

Miyake, A., Emerson, M. J., & Friedman, N. P. (1999). Good interactions are hard to find. *Brain & Behavioral Sciences, 22*, 108–109.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J-C. (2004). Inner speech as a retrieval aid for task goals: the effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115*, 123–142.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive function: Four general conclusions. *Current Directions in Psychological Science, 21*, 8–14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language, Learning & Technology, 18*, 118–141.

Montero Perez, M., Peters, E., & Desmet, P. (2015). Enhancing vocabulary learning through captioned Video: An eye-tracking study. *The Modern Language Journal, 99*, 308–328.

Montero Perez, M., Peters, E., & Desmet, P. (2013a). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension. *ReCALL, 26*, 21–43.

Montero Perez, M., van den Noortgate, W., & Desmet, P. (2013b). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*, 720–739.

Montrul, S., & Slabakova, R. (2002). The L2 Acquisition of Morphosyntactic and Semantic Properties of the Aspectual Tenses Preterite and Imperfect. In A.T. Perez-Lerouz & J. Liceras (Eds.), *The acquisition of Spanish morphosyntax: the L1-L2 connection* (pp. 113–149). Dordrecht: Kluwer.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of verbal learning and verbal behavior, 16*, 519–533.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*, 111–121.

Mota, M. B. (2003). Working memory capacity and fluency, accuracy, complexity, and lexical density in L2 speech production. *Fragmentos, 24*, 69 –104.

Muñoz, C. (2017). The role of age and proficiency in subtitle reading. An eye-tracking study. *System, 67*, 77–86.

Murphy, V. (1997). The effect of modality on a grammaticality judgment task. *Second Language Research, 13*, 34–65.

Nagata, N. (1998a). Input vs. output practice in educational software for second language acquisition. *Language Learning & Technology, 1*, 23–40.

Nagata, N. (1998b). The relative effectiveness of production and comprehension practice in second language acquisition. *Computer Assisted Language Learning, 11*, 153–177.

Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehension input: effects of incidental word learning from context for language minority students. *Reading Research Quarterly, 27*, 95–106.

Norouzian, R., & Plonsky, L. (2017). Eta-and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research*. Doi 10.1177/0267658316684904

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–528.

Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.

O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition, 29*, 557–582.

Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society, 30*, 287–289.

Osaka, M., Osaka, N., & Groner, R. (1993). Language-independent working memory: Evidence from German and French reading span tests. *Bulletin of the Psychonomic Society, 31*, 117–118.

Overstreet, M. H. (1998). Text enhancement and content familiarity: The focus of learner attention. *Spanish Applied Linguistics*, *2*, 229–258

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical interpretation*. Mahwah, NJ: Lawrence Erbaum Associate, Inc.

Park, E. S. (2004). Constraints of implicit focus on form: Insights from a study of input enhancement. *Teachers College, Columbia University Working Papers in TESOL and Applied Linguistics. 4*, 1–30.

Park, E. S., & Nassif, L. (2014). Textual enhancement of two L2 Arabic forms: a classroom-based study. *Language Awareness, 23,* 334–352.

Pellicer-Sánchez, A. (2016). Learning L2 collocations incidentally from reading. *Language Teaching Research, 19*, 1–22.

Pica, T. (1983). Methods of Morpheme Quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition, 6*, 69–78

Pienemann, M., & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (Ed.), *Applying second language acquisition research* (pp. 45–141). Adelaide: National Curriculum Resource Center.

Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning, 64*, 878–912.

Prabhu, H. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A contemporary grammar of the English language*. London: Longman.

R Development Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rai, M. K., Loschky, L. C., Harris, R. J., Peck, N. R., & Cook, L. G. (2010). Effects of stress and working memory capacity on foreign language readers' inferential processing during comprehension. *Language Learning, 61*, 187–218.

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of Visual Experimentation, 83*, e50780.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research, 7*, 4–22.

Révész, A. (2012). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning, 62*, 93–132.

Révész, A., Marije, M., & Lee, M. (2017). Investigating IELTS Academic Writing Task 2: Relationships between cognitive writing processes, text quality, and working memory. *IELTS Research Reports Online Series, 44*.

Richards, J. C. (2005). Second thoughts on teaching listening. *Regional Language Centre Journal*, *36*, 85–92.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition, 30*, 333–357.

Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition, 3*5, 213–235.

Robinson, P. (1995). Attention, memory, and the “noticing” hypothesis. *Language Learning, 45*, 283–331.

Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed condition. *Studies in Second Language Acquisition, 19*, 223–247.

Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631–678). Oxford: Blackwell.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition, 27*, 235–268.

Robinson, P. (2011). Task-based language learning: A review of issues. *Language Learning, 61*, 1–36.

Robinson, P., Mackey, A., Gass, S., & Schmidt, R. (2012). Attention and awareness in second language acquisition. In S. Gass & A. Mackey (Eds.) *The Routledge Handbook of Second Language Acquisition* (pp. 247–267). New York: Routledge.

Rodgers, M. P. H., & Webb, S. (2017). The effects of captions on EFL learners' comprehension of English language television programs. *CALICO Journal, 32*, 20–38.

Rogers, J. R. (2016). *Developing implicit and explicit knowledge of L2 case marking under incidental learning conditions.* Unpublished dissertation, University College London Institute of Education, London, UK.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 207–231.

Rosa, E. M., & Leow, R. P. (2004). Awareness, different learning condition, and second language development. *Applied Psycholinguistics, 25*, 269–292.

Rost, M. (2002). *Teaching and researching listening*. London: Longman.

Rost, M. (2011). *Teaching and researching listening*. London: Longman.

Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, *29*, 67–100.

Sagarra, N. (2007). From CALL to face-to-face interaction: The effect of computer-delivered recasts and working memory on L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 229–248). Oxford: Oxford University Press.

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. London: Palgrave Macmillan.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129–158.

Schmidt, R. (1993a). Consciousness, learning and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 21–42). Oxford: Oxford University Press.

Schmidt, R. (1993b). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206–226.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11–26.

Schmidt, R. (1995). Consciousness and foreign language learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu, HI: University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 1–32). Cambridge: Cambridge University Press.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *"Talking to learn": Conversation in second language acquisition* (pp. 237–326). Rowley, MA: Newbury House.

Segalowitz, N. (2003). Automaticity and second languages. In C. Doughty & M. H. Long, (Eds.), *The handbook of second language acquisition* (pp. 382–408). Malden, MA: Blackwell.

Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learners. *Second Language Research*, *7*, 118–132.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA. *Studies in Second Language Acquisition, 15*, 165–179.

Shiffrin, R. (1993). Short-term memory: A brief commentary. *Memory & Cognition, 21*, 193–197.

Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research, 15*, 137–158.

Shintani, N. (2012a). Repeating input-based tasks with young beginner learners. *RELC Journal, 43*, 39–51.

Shintani, N. (2012b). Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research, 16*, 253–279.

Shintani, N. (2014). Using tasks with young beginner learners: the role of the teacher. *Innovation in Language Learning and Teaching, 8,* 279–294.

Shintani, N., & Ellis, R. (2010). The incidental acquisition of English plural –s by Japanese children in comprehension-based and production-based lessons. *Studies in Second Language Acquisition, 32*, 607–637.

Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language Learning, 63*, 296–329.

Shook, D. J. (1994). FL/L2 reading, grammatical information, and the input to intake phenomenon. *Applied Language Learning, 5*, 57–93.

Simard, D. (2009). Differential effects of textual enhancement formats on intake. *System, 37*, 124–135.

Simard, D., & Foucambert, D. (2013). Observing noticing while reading in L2. In J. M. Bergsleithner, S.N. Frota & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 207–226). Honolulu, HI: National Foreign Language Resource Center, University of Hawai`i at Mānoa.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research, 27*, 1–22.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Smith, B. (2012). Eye tracking as a measure of noticing: A study of explicit recasts in SCMC. *Language Learning & Technology, 16*, 53–81.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283*, 1657–1661.

Sohn, S-O. (1995). *Tense and Aspect in Korea*. Honolulu, Hawaii: University of Hawaii Press.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*, 263–308.

Spinner, P., Gass, S. M., & Behney, J. (2013). Ecological validity in eye-tracking. *Studies in Second Language Acquisition, 35*, 389-415.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Suvorov, R. (2014). The use of eye tracking in research on video-based second language (L2) listening assessment: A comparison of context videos and content videos. *Language Testing, 32*, 463–483.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury.

Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, *26*, 376–401.

Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 159–168). New York: Cambridge University Press.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, *14*, 50–73.

Tagarelli, K. M., Mota, M. B., & Rebuschat, P. (2011). The role of working memory in implicit and explicit language learning. *Proceedings of the Cognitive Science Society*, *33*, 2061–2066.

Takimoto, M. (2007). The effects of input-based tasks on the development of learners' pragmatic proficiency. *Applied Linguistics*, *30*, 1–25.

Tanaka, T. (2001). Comprehension and production practice in grammar instruction: Does their combined use facilitate second language acquisition? *JALT Journal*, *23*, 6–30.

Tateyama, Y., Kasper, G., Mui, L. P., Tay, H-M., & Thananart, O. (1997). Explicit and implicit teaching of pragmatic routines. In L. Bouton (Ed.), *Pragmatics and language learning, monograph series 8* (pp. 163–178). Urbana-Champaign, IL: Division of English as an International Language, University of Illinois, Urbana-Champaign.

Taylor, G. (2005). Perceived processing strategies of students watching captioned video. *Foreign Language Annals, 38*, 422–427.

Tobii Studio. (2014). *User Manual – Tobii Studio (Version 3.3.0)*. Retrieved from http://www.tobii.com/Global/Analysis/Downloads/User_Manuals_and_Guides/Tobii_X2-30_EyeTrackerUserManual_WEB.pdf

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*, 183–203.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Uggen, M. S. (2012). Reinvestigating the noticing function of output. *Language Learning, 62*, 506–540.

Unsworth, N., Heitz, R., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498–505.

Van den Noort, M. W. M. L., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist*, *11*, 289–296.

Van den Noort, M. W. M. L., Bosch, P., Haverkort, M., & Hugdahl, K. (2008). A standard computerized version of the reading span test in different languages. *European Journal of Psychological Assessment, 24*, 35–42.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language teaching, 40*, 191-210.

Vanderplank, R. (1988). The value of teletext sub-titles in language learning. *ELT journal*, *42*, 272–281.

VanPatten, B. (1985). The acquisition of ser and estar by adult learners of Spanish: A preliminary investigation of transitional stages of competence. *Hispania, 68*, 399–406.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research.* Norwood, NJ: Ablex.

VanPatten, B., & Benati, A. G. (2010). *Key terms in second language acquisition*. London, UK: Continuum.

Varol, B., & Erçetin, G. (2016). Effects of working memory and gloss type on L2 text comprehension and incidental vocabulary learning in computer-based reading. *Procedia – Social and Behavioral Sciences*, *232*, 759–768.

Verbruggen, F., Logan G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods, 40*, 479–483.

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*, 315–339.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology, 49A*, 51–79.

Wen, Z. E. (2012). Working memory and second language learning. *International Journal of Applied Linguistics*, *22*, 1–22.

Wen, Z. E. (2015). Working memory in second language acquisition and processing: The phonological/executive model. In Z. E. Wen, M. B. Mota, & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 41–62). Bristol: Multilingual Matters.

Wen, Z. E., & Skehan, P. (2011). A new perspective on foreign language aptitude research: Building and supporting case for "working memory as language aptitude." *Ilha Do Desterro: A Journal of English Language, Literatures in English and Cultural Studies, 60*, 15–43.

White, J. (1998). Getting the learners' attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 85–113). New York: Cambridge University Press.

Widdowson, H. G. (2003). *Defining issues in English language teaching*. Oxford: Oxford University Press.

Williams, J. N. (1999). Memory, attention and inductive learning. *Studies in Second Language Acquisition, 21*, 1–48.

Williams, J. N. (2011). Working memory and SLA. In S. M. Gass & A. Mackey (Eds.), *The handbook of second language acquisition* (pp. 427–441). New York, NY: Routledge.

Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning, 53*, 67–121.

Willis, J. (1996). *A Framework for Task-based Learning*. London: Longman.

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.

Winke, P. (2013). The effects of input enhancement on grammar learning and comprehension: A modified replication of Lee (2007) with eye-movement data. *Studies in Second Language Acquisition, 35*, 323–352.

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*, 65–86.

Winke, P., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal, 97*, 254–275.

Winter, B. (2013). *Linear models and linear mixed effects models in R with linguistic applications*. arXiv:1308.5499. [http://arxiv.org/pdf/1308.5499.pdf]

Wong, W. (2003). Textual enhancement and simplified input: Effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning, 13*, 17–46.

Yang, H-Y. (2014). The effect of advanced organizers and subtitles on EFL learners' listening comprehension skills. *CALICO Journal, 31*, 345–373.

Yang, J. C., & Chang, P. (2014). Captions and reduced forms instruction: The impact on EFL students' listening comprehension. *ReCALL, 26*, 44–61.

Yi, B., & Luo, S. (2013). Working memory and lexical knowledge in L2 argumentative writing. *Asian Journal of English Language Teaching, 23,* 83–102.

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors, 5*, 7–17.

Yuill, N., Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology, 80,* 351–361.

APPENDICES

List of Appendices

Appendix A: Information Sheet and Consent Form for Study 1

Institute of Education

UCL

STUDY INFORMATION SHEET:

Effect of Input Enhancement on L2 learning in Instructed Second Language Acquisition

I would like to invite you to participate in a research study that examines how the use of captions in listening influences second / foreign language learning and how language learners perceive the use of captions in learning second / foreign language. I am an MPhil/PhD student in the Department of Culture, Communication and Media at the UCL Institute of Education, University College London, interested in task-based language teaching and instructed second language acquisition.

The aim of this study is to examine how the use of captions during listening affects learning English. Specifically, the potential effects of different types of captioning (audio with captions vs. audio with highlighted captions) on learning grammatical features of English (pronoun reference) will be investigated in the current study.

If you decide to participate, I will ask you to partake in five individual sessions.

At the outset of the experiment, you will be asked to complete a brief questionnaire about yourself, and to complete a general English proficiency test and a pretest (grammatical judgment test). Approximately 40 minutes will be designated to complete the questionnaire and the pretest.

During the intervention, you will be asked to participate in three individual sessions. In each session, you will view a series of captioned video and perform the related tasks. While watching the captioned video, your eye movements will also be recorded. Each session will last approximately 10 minutes.

After completing three sessions, you will be asked to undertake a posttest consisting of a grammatical judgment test. The posttest will last approximately 20 minutes.

In total, the experiment is expected to require approximately 1hour and 30 minutes. For the purposes of the study, it is important that you complete all activities. Otherwise, I will not be able to use the data collected from you.

Any data obtained from you will be kept securely. At every stage of the project and beyond, your name will remain confidential. Your identity will be anonymised by the use of a unique identifier. The results may be presented at professional conferences and in research publications.

You are free to withdraw from the study at any time without reason and without any impact on you. If you decide to withdraw, any data collected from you will be destroyed. If you have any queries about the study, please feel free to contact me at mlee10@ioe.ac.uk, +44 07899346797 (UK), +821088273337 (Korea) or Dr. Andrea Révész at a.revesz@ioe.ac.uk, +44 (0)20 7612 5158.

I would be very grateful if you would agree to take part!

Minjin Lee, MPhil/PhD student (mlee10@ioe.ac.uk)
Department of Culture, Communication, and Media
UCL Institute of Education, University of London
20 Bedford Way, London WC1H OAL

Institute of Education

CONSENT FORM

Effect of Input Enhancement on L2 learning in Instructed Second Language Acquisition

		YES	NO
1.	I have read and had explained to me by MinJin Lee the Information Sheet relating to this project.	☐	☐
2.	I have had explained to me the purposes of the project and what will be required of me, and any questions have been answered to my satisfaction. I agree to the arrangements for my participation as described in the Information Sheet.	☐	☐
3.	I understand that my participation is entirely voluntary and that I have the right to withdraw from the project any time.	☐	☐

Name:

Signed:

Date:

Appendix B: Information Sheet and Consent Form for Study 2

Institute of Education

UCL

STUDY INFORMATION SHEET:

Input Enhancement in Multimodal Input-based Tasks: An Eye-Tracking Study

I would like to invite you to participate in a research study that examines how the use of captions in listening influences second / foreign language learning and how language learners perceive the use of captions in learning second / foreign language. I am an MPhil/PhD student in the Department of Culture, Communication and Media at the UCL Institute of Education, University College London, interested in task-based language teaching and instructed second language acquisition.

The aim of this study is to examine how the use of captions during listening affects learning English. Specifically, the potential effects of different types of captioning (audio with captions vs. audio with highlighted captions vs. audio only) on learning grammatical features of English (present perfect and past simple) will be investigated in the current study. Another aim of this study is to examine the role of individual differences in working memory in mediating learners' noticing and processing of the enhanced input.

If you decide to participate, I will ask you to partake in four individual sessions.

First, at the outset of the experiment, you will be asked to complete a brief questionnaire about yourself, a general English proficiency test and a pretest (a receptive test and a productive test). Approximately 60 minutes will be designated to complete the questionnaire and the pretest. During the intervention, you will view a series of a video (news clips) and perform the related tasks. While watching the captioned video, your eye movements will also be recorded. After viewing video clips, you will be asked to undertake a posttest consisting of both receptive and productive tests. During the third session, you will carry out some memory activities, which will require you to remember and repeat nonwords and particular visual patterns. After approximately one month, you will be invited to complete a delayed posttest, which includes both receptive and productive tests, to assess long-term effects of enhanced input in multimodal input-based tasks on L2 development.

Each session is expected to last approximately 75 minutes. For the purposes of the study, it is important that you participate all sessions. Otherwise, I will not be able to use the data collected from you.

Any data obtained from you will be kept securely. At every stage of the project and beyond, your name will remain confidential. Your identity will be anonymised by the use of a unique identifier. The results may be presented at professional conferences and in research publications.

You are free to withdraw from the study at any time without reason and without any impact on you.
If you decide to withdraw, any data collected from you will be destroyed. If you have any queries about the study, please feel free to contact me at m.lee.14@ucl.ac.uk, +821088273337 (Korea) or Dr. Andrea Révész at a.revesz@ucl.ac.uk, +44 (0)20 7612 5158.

I would be very grateful if you would agree to take part!

MinJin Lee, MPhil/PhD student (m.lee.14@ucl.ac.uk)
Department of Culture, Communication, and Media
UCL Institute of Education, University of London
20 Bedford Way, London WC1H OAL

Institute of Education

<u>CONSENT FORM</u>

Input Enhancement in Multimodal Input-based Tasks: An Eye-Tracking Study

		YES	NO
4.	I have read and had explained to me by MinJin Lee the Information Sheet relating to this project.	☐	☐
5.	I have had explained to me the purposes of the project and what will be required of me, and any questions have been answered to my satisfaction. I agree to the arrangements for my participation as described in the Information Sheet.	☐	☐
6.	I understand that my participation is entirely voluntary and that I have the right to withdraw from the project any time.	☐	☐

Name:

Signed:

Date:

Appendix C: A Sample of Multimodal Input-based Task

Group 1 (Normal Caption Group)

A ☐ B ☐ C ☐

What is the grandmother's job now?

A ☐ B ☐ C ☐

My grandmother always wanted to be a teacher when she was a little girl,

A ☐ B ☐ C ☐

but she had to leave school when she was fourteen and help her mother clean offices and shops.

A ☐ B ☐ C ☐

When she was in her thirties she went to college, but she had to work as a waitress in the evenings to pay for her studies.

A ☐ B ☐ C ☐

A few years later she finally got the job she'd always wanted and she's done it ever since.

Group 2 (Enhanced Caption Group)

A ☐ B ☐ C ☐

What is the grandmother's job now?

A ☐ B ☐ C ☐

My grandmother always wanted to be a teacher when **she** was a little girl,

A ☐ B ☐ C ☐

but she had to leave school when she was fourteen and help her mother clean offices and shops.

A ☐ B ☐ C ☐

When she was in her thirties she went to college, but she had to work as a waitress in the evenings to pay for her studies.

A ☐ B ☐ C ☐

A few years later she finally got the job she'd always wanted and she's done it ever since.

Appendix D: Grammaticality Judgement Test – Target Items

Singular – HE

1	S / G	My brother read the journal in the room. He wrote a report.
	S / UG	My brother read the journal in the room. They wrote a report.
	P / G	My brothers read the journal in the room. They wrote a report.
	P / UG	My brothers read the journal in the room. He wrote a report.

2	S / G	My brother got the access to the room. He found a present.
	S / UG	My brother got the access to the room. They found a present.
	P / G	My brothers got the access to the room. They found a present.
	P / UG	My brothers got the access to the room. He found a present.

3	S / G	My brother found an error on the form. He gave some comments.
	S / UG	My brother found an error on the form. They gave some comments.
	P / G	My brothers found an error on the form. They gave some comments.
	P / UG	My brothers found an error on the form. He gave some comments.

4	S / G	My brother used the table in the room. He drew the design.
	S / UG	My brother used the table in the room. They drew the design.
	P / G	My brothers used the table in the room. They drew the design.
	P / UG	My brothers used the table in the room. He drew the design.

5	S / G	My uncle sold a hotel on the street. He made a profit.
	S / UG	My uncle sold a hotel on the street. She made a profit.
	P / G	My uncles sold a hotel on the street. They made a profit.
	P / UG	My uncles sold a hotel on the street. She made a profit.

6	S / G	My uncle closed the account at the bank. He had some problems.
	S / UG	My uncle closed the account at the bank. She had some problems.
	P / G	My uncles closed the account at the bank. They had some problems.
	P / UG	My uncles closed the account at the bank. She had some problems.

7	S / G	My uncle bought some flowers at the store. He wrote the letter.
	S / UG	My uncle bought some flowers at the store. She wrote the letter.
	P / G	My uncles bought some flowers at the store. They wrote the letter.
	P / UG	My uncles bought some flowers at the store. She wrote the letter.

8	S / G	My uncle made a promise in the room. He kept the secret.
	S / UG	My uncle made a promise in the room. She kept the secret.
	P / G	My uncles made a promise in the room. They kept the secret.
	P / UG	My uncles made a promise in the room. She kept the secret.

Singular – SHE

9	S / G	My sister saw the market on the street. She bought some cookies.
	S / UG	My sister saw the market on the street. They bought some cookies.

	P / G	My sisters saw the market on the street. They bought some cookies.
	P / UG	My sisters saw the market on the street. She bought some cookies.

10	S / G	My sister gave same advice for the task. She made some comments.
	S / UG	My sister gave same advice for the task. They made some comments.
	P / G	My sisters gave same advice for the task. They made some comments.
	P / UG	My sisters gave same advice for the task. She made some comments.

11	S / G	My sister earned the money from the work. She built a hotel.
	S / UG	My sister earned the money from the work. They built a hotel.
	P / G	My sisters earned the money from the work. They built a hotel.
	P / UG	My sisters earned the money from the work. She built a hotel.

12	S / G	My sister drank some coffee at the bar. She met old friends.
	S / UG	My sister drank some coffee at the bar. They met old friends.
	P / G	My sisters drank some coffee at the bar. They met old friends.
	P / UG	My sisters drank some coffee at the bar. She met old friends.

13	S / G	My daughter left the report on the train. She called the office.
	S / UG	My daughter left the report on the train. He called the office.
	P / G	My daughters left the report on the train. They called the office.
	P / UG	My daughters left the report on the train. He called the office.

14	S / G	My daughter had a party at the house. She cooked the dinner.
	S / UG	My daughter had a party at the house. He cooked the dinner.
	P / G	My daughters had a party at the house. They cooked the dinner.
	P / UG	My daughters had a party at the house. He cooked the dinner.

15	S / G	My daughter rode the bicycle on the road. She found a flower.
	S / UG	My daughter rode the bicycle on the road. He found a flower.
	P / G	My daughters rode the bicycles on the road. They found a flower.
	P / UG	My daughters rode the bicycles on the road. He found a flower.

16	S / G	My daughter passed the exam for the course. She won the award.
	S / UG	My daughter passed the exam for the course. He won the award.
	P / G	My daughters passed the exam for the course. They won the award.
	P / UG	My daughters passed the exam for the course. He won the award.

Plural – HE

17	S / G	The student did a project in the room. He liked the outcome.
	S / UG	The student did a project in the room. They liked the outcome.
	P / G	The students did a project in the room. They liked the outcome.
	P / UG	The students did a project in the room. He liked the outcome.

18	S / G	The student solved the problems in an hour. He got the credit.
	S / UG	The student solved the problems in an hour. They got the credit.
	P / G	The students solved the problems in an hour. They got the credit.
	P / UG	The students solved the problems in an hour. He got the credit.

19	S / G	The student learned the language for a year. He made an effort.
	S / UG	The student learned the language for a year. They made an effort.
	P / G	The students learned the language for a year. They made an effort.
	P / UG	The students learned the language for a year. He made an effort.

20	S / G	The student caused the trouble in the town. He got a warning.
	S / UG	The student caused the trouble in the town. They got a warning.
	P / G	The students caused the trouble in the town. They got a warning.
	P / UG	The students caused the trouble in the town. He got a warning.

21	S / G	The teacher wrote a report at the desk. He had a meeting.
	S / UG	The teacher wrote a report at the desk. They had a meeting.
	P / G	The teachers wrote a report at the desk. They had a meeting.
	P / UG	The teachers wrote a report at the desk. He had a meeting.

22	S / G	The teacher sent a letter t to the bank. He asked some questions.
	S / UG	The teacher sent a letter t to the bank. They asked some questions.
	P / G	The teachers sent a letter t to the bank. They asked some questions.
	P / UG	The teacher sent a letter t to the bank. He asked some questions.

23	S / G	The teacher placed the package on the floor. He found the address.
	S / UG	The teacher placed the package on the floor. They found the address.
	P / G	The teachers placed the package on the floor. They found the address.
	P / UG	The teachers placed the package on the floor. He found the address.

24	S / G	The teacher ate some pizza in the house. He cleaned the table.
	S / UG	The teacher ate some pizza in the house. They cleaned the table.
	P / G	The teachers ate some pizza in the house. They cleaned the table.
	P / UG	The teachers ate some pizza in the house. He cleaned the table.

Plural – SHE

25	S / G	The teacher baked some cookies in the house. She used the oven.
	S / UG	The teacher baked some cookies in the house. They used the oven.
	P / G	The teachers baked some cookies in the house. They used the oven.
	P / UG	The teachers baked some cookies in the house. She used the oven.

26	S / G	The teacher viewed the website for an hour. She found the data.
	S / UG	The teacher viewed the website for an hour. They found the data.
	P / G	The teachers viewed the website for an hour. They found the data.
	P / UG	The teachers viewed the website for an hour. She found the data.

27	S / G	The teacher saved some money for one week. She bought the tickets.
	S / UG	The teacher saved some money for one week. They bought the tickets.
	P / G	The teachers saved some money for one week. They bought the tickets.
	P / UG	The teachers saved some money for one week. She bought the tickets.

28	S / G	The teacher put the notice on the board. She gave a warning.

	S / UG	The teacher put the notice on the board. They gave a warning.
	P / G	The teachers put the notice on the board. They gave a warning.
	P / UG	The teachers put the notice on the board. She gave a warning.
29	S / G	The student missed the lessons for one week. She failed the exam.
	S / UG	The student missed the lessons for one week. They failed the exam.
	P / G	The students missed the lessons for one week. They failed the exam.
	P / UG	The students missed the lessons for one week. She failed the exam.
30	S / G	The student planned the budget for one month. She asked some questions.
	S / UG	The student planned the budget for one month. They asked some questions.
	P / G	The students planned the budget for one month. They asked some questions.
	P / UG	The students planned the budget for one month. She asked some questions.
31	S / G	The student took the lessons for one year. She won the contest.
	S / UG	The student took the lessons for one year. They won the contest.
	P / G	The students took the lessons for one year. They won the contest.
	P / UG	The students took the lessons for one year. She won the contest.
32	S / G	The student drew some pictures in the room. She used white paper.
	S / UG	The student drew some pictures in the room. They used white paper.
	P / G	The students drew some pictures in the room. They used white paper.
	P / UG	The students drew some pictures in the room. She used white paper.

Appendix E: Grammaticality Judgement Test – Distractors

1) Passive

1	Correct	We were told to be ready by noon. We were quite late.
	Incorrect	We told to be ready by noon. We were quite late.
	Correct	I was told to be ready by noon. I was quite late.
	Incorrect	I told to be ready by noon. I was quite late.

2	Correct	We were woken by the alarm this morning. The sound was loud.
	Incorrect	We woken by the alarm this morning. The sound was loud.
	Correct	I was woken by the alarm this morning. The sound was loud.
	Incorrect	I woken by the alarm this morning. The sound was loud.

3	Correct	I was fired by the director last month. It was a surprise.
	Incorrect	I fired by the director last month. It was a surprise.
	Correct	We were fired by the director last month. It was a surprise.
	Incorrect	We fired by the director last month. It was a surprise.

4	Correct	I was invited to the party last night. The party was nice.
	Incorrect	I invited to the party last night. The party was nice.
	Correct	We were invited to the party last night. The party was nice.
	Incorrect	We invited to the party last night. The party was nice.

5	Correct	I was sent to the other office by the boss. I liked the office.
	Incorrect	I sent to the other office by the boss. I liked the office.
	Correct	We were sent to the other office by the boss. I liked the office.
	Incorrect	We sent to the other office by the boss. I liked the office.

6	Correct	I was accused of stealing a car last week. There was no evidence.
	Incorrect	I accused of stealing a car last week. There was no evidence.
	Correct	We were accused of stealing a car last week. There was no evidence.
	Incorrect	We accused of stealing a car last week There was no evidence.

7	Correct	I was attacked by the man on the street. The man was strange.
	Incorrect	I attacked by the man on the street. The man was strange.
	Correct	We were attacked by the man on the street. The man was strange.
	Incorrect	We attacked by the man on the street. The man was strange.

8	Correct	I was offered a new job by the leader. I was quite pleased.
	Incorrect	I offered a new job by the leader. I was quite pleased.
	Correct	You were offered a new job by the leader. You were quite pleased.
	Incorrect	You offered a new job by the leader. You were quite pleased.

9	Correct	We were kicked out of the restaurant. It was not nice.
	Incorrect	We kicked out of the restaurant. It was not nice.
	Correct	I was kicked out of the restaurant. It was not nice.
	Incorrect	I kicked out of the restaurant. It was not nice.

10	Correct	The man was arrested by the police. There was a crime.
	Incorrect	The man arrested by the police. There was a crime.
	Correct	You were arrested by the police. There was a crime.
	Incorrect	You arrested by the police. There was a crime.

11	Correct	We were paid ten dollars to do the job. We did the work.
	Incorrect	We were paying ten dollars to do the job. We did the work.
	Correct	I was paid ten dollars to do the job. I did the work.
	Incorrect	I was paying ten dollars to do the job. I did the work.

12	Correct	The baby was born in this hospital tonight. We saw the baby.
	Incorrect	The baby born in this hospital tonight. We saw the baby.
	Correct	The girl was born in this hospital tonight. We saw the baby.
	Incorrect	The girl born in this hospital tonight. We saw the baby.

13	Correct	I was given two hours to make my choice. It was quite short.
	Incorrect	I given two hours to make my choice. It was quite short.
	Correct	We were given two hours to make my choice. It was quite short.
	Incorrect	We given two hours to make my choice. It was quite short.

14	Correct	You were hired by the team last month. You were quite lucky.
	Incorrect	You hired by the team last month. You were quite lucky.
	Correct	I was hired by the team last month. I was quite lucky.
	Incorrect	I hired by the team last month. I was quite lucky.

15	Correct	I was hired by the company two months ago. I was quite happy.
	Incorrect	I hired by the company two months ago. I was quite happy.
	Correct	We were hired by the company two months ago. We were quite happy.
	Incorrect	We hired by the company two months ago. We were quite happy.

16	Correct	The people were forced to do the survey. It was very long.
	Incorrect	The people forced to do the survey. It was very long.
	Correct	We were forced to do the survey. It was very long.
	Incorrect	We told to do the survey. It was very long.

2) Verb + ing / Verb + to

17	Correct	We decided to take a bus to the hotel. The car was broken.
	Incorrect	We decided taking a bus to the hotel. The car was broken.
	Correct	I decided to take a bus to the hotel. The car was broken.
	Incorrect	I decided taking a bus to the hotel. The car was broken.

18	Correct	The child wanted to look around the store. The store was big.
	Incorrect	The child wanted looking around the store. The store was big.
	Correct	We wanted to look around the store. The store was big.
	Incorrect	We wanted looking around the store. The store was big.

19	Correct	You agreed to work late at the office. You were very kind.
	Incorrect	You agreed working late at the office. You were very kind.
	Correct	We agreed to work late at the office. We were very kind.
	Incorrect	We agreed working late at the office. We were very kind.

20	Correct	I avoid driving in the rush hour. It takes two hours.
	Incorrect	I avoid to drive in the rush hour. It takes two hours.
	Correct	We avoid driving in the rush hour. It takes two hours.
	Incorrect	We avoid to drive in the rush hour. It takes two hours.

21	Correct	The kids finished eating an hour ago. The food was good.
	Incorrect	The kids finished to eat an hour ago. The food was good.
	Correct	We finished eating an hour ago. The food was good.
	Incorrect	We finished to eat an hour ago. The food was good.

22	Correct	I expect to pass the test with a good score. The test seems easy.
	Incorrect	I expect passing the test with a good score. The test seems easy.
	Correct	We expect to pass the test with a good score. The test seems easy.
	Incorrect	We expect passing the test with a good score. The test seems easy.

23	Correct	You need to get a visa to go there. I will help you.
	Incorrect	You need getting a visa to go there. I will help you.
	Correct	I need to get a visa to go there. You will help me.
	Incorrect	I need getting a visa to go there. You will help me.

24	Correct	I suggested going to the park this morning. The weather was nice.
	Incorrect	I suggested to go to the park this morning. The weather was nice.
	Correct	We suggested going to the park this morning. The weather was nice.
	Incorrect	We suggested to go to the park this morning. The weather was nice.

25	Correct	We decided to take a taxi home tonight. It was quite cold.
	Incorrect	We decided taking a taxi home tonight. It was quite cold.
	Correct	I decided to take a taxi home tonight. It was quite cold.
	Incorrect	I decided taking a taxi home tonight. It was quite cold.

26	Correct	I intend to visit my sister next year. We will have fun.
	Incorrect	I intend visiting my sister next year. We will have fun.
	Correct	We intend to visit our sister next year. We will have fun.
	Incorrect	We intend visiting our sister next year. We will have fun.

27	Correct	We waited to buy tickets for the game. The line was long.
	Incorrect	We waited buying tickets for the game. The line was long.
	Correct	I waited to buy tickets for the game. The line was long.
	Incorrect	I waited buying tickets for the game. The line was long.

28	Correct	I avoided telling you about my plan. It was a secret.
	Incorrect	I avoided to tell you about my plan. It was a secret.
	Correct	We avoided telling you about my plan. It was a secret.
	Incorrect	We avoided to tell you about my plan. It was a secret.

29	Correct	My husband promised to help me today. I was very pleased.
	Incorrect	My husband promised helping me today. I was very pleased.
	Correct	My wife promised to help me today. I was very pleased.
	Incorrect	My wife promised helping me today. I was very pleased.

30	Correct	We hope to pass the exam next month. We need to pray.
	Incorrect	We hope passing the exam next month. We need to pray.
	Correct	I hope to pass the exam next month. I need to pray.
	Incorrect	I hope passing the exam next month. I need to pray.

31	Correct	You appeared to be quite happy last night. I was happy too.
	Incorrect	You appeared being quite happy last night. I was happy too.
	Correct	The kid appeared to be quite happy last night. I was happy too.
	Incorrect	The kid appeared being quite happy last night. I was happy too.

32	Correct	We plan to have a party in the park. It will be nice.
	Incorrect	We plan having a party in the park. It will be nice.
	Correct	I plan to have a party in the park. It will be nice.
	Incorrect	I plan having a party in the park. It will be nice.

3) **Adjective –ed / -ing**

33	Correct	We were quite surprised at the test results. The score was bad.
	Incorrect	We were quite surprising at the test results. The score was bad.
	Correct	I was quite surprised at the test results. The score was bad.
	Incorrect	I was quite surprising at the test results. The score was bad.

34	Correct	I am excited about the football match. It is on TV.
	Incorrect	I am exciting about the football mach. It is on TV.
	Correct	We are excited about the football match. It is on TV.
	Incorrect	We are exciting about the football mach. It is on TV.

35	Correct	I was impressed with the new film. It was very good.
	Incorrect	I was impressing with the new film. It was very good.
	Correct	We were impressed with the new film. It was very good.
	Incorrect	We were impressing with the new film. It was very good.

36	Correct	You look tired from working late last night. You need a rest.
	Incorrect	You look tiring from working late last night. You need a rest.
	Correct	I look tired from working late last night. I need a rest.
	Incorrect	I look tiring from working late last night. I need a rest.

37	Correct	We were shocked by the news on the radio. It was very sad.
	Incorrect	We were shocking by the news on the radio. It was very sad.
	Correct	I was shocked by the news on the radio. It was very sad.
	Incorrect	I was shocking by the news on the radio. It was very sad.

38	Correct	My boss told me a very interesting story. It was quite funny.
	Incorrect	My boss told me a very interested story. It was quite funny.
	Correct	You told me a very interesting story. It was quite funny.
	Incorrect	You told me a very interested story. It was quite funny.

39	Correct	You were really satisfied with the offer. You were quite happy.
	Incorrect	You were really satisfying with the offer. You were quite happy.
	Correct	I was really satisfied with the offer. I was quite happy.
	Incorrect	I was really satisfying with the offer. I was quite happy.

40	Correct	I was pleased to hear the news last night. I was very happy.
	Incorrect	I was pleasing to hear the news last night. I was very happy.
	Correct	We were pleased to hear the news last night. We were very happy.
	Incorrect	We were pleasing to hear the news last night. We were very happy.

41	Correct	I was interested in the new project. I worked very hard.
	Incorrect	I was interesting in the new project. I worked very hard.
	Correct	We were interested in the new project. We worked very hard.
	Incorrect	We were interesting in the new project. We worked very hard.

42	Correct	You should feel relaxed in your house. The house seems nice.
	Incorrect	You should feel relaxing in your house. The house seems nice.
	Correct	The man should feel relaxed in his house. The house seems nice.
	Incorrect	The man should feel relaxing in his house. The house seems nice.

43	Correct	I felt tired after running all that way. It was quite cold.
	Incorrect	I felt tiring after running all that way. It was quite cold.
	Correct	We felt tired after running all that way. It was quite cold.
	Incorrect	We felt tiring after running all that way. It was quite cold.

44	Correct	I got very confused with one of the rules. It was not clear.
	Incorrect	I got very confusing with one of the rules. It was not clear.
	Correct	We got very confused with one of the rules. It was not clear.
	Incorrect	We got very confusing with one of the rules. It was not clear.

45	Correct	We found the game to be very exciting. The game started today.
	Incorrect	We found the game to be very excited. The game started today.
	Correct	I found the game to be very exciting. The game started today.
	Incorrect	I found the game to be very excited. The game started today.

46	Correct	We were very pleased with the outcome. We were very happy.
	Incorrect	We were very pleasing with the outcome. We were very happy.
	Correct	I was very pleased with the outcome. I was very happy.
	Incorrect	I was very pleasing with the outcome. I was very happy.

47	Correct	We are excited about a new house. It has a garden.
	Incorrect	We are exciting about a new house. It has a garden.
	Correct	I was excited about a new house. It has a garden.
	Incorrect	I was exciting about a new house. It has a garden.

48	Correct	You are concerned about the test result. It will be fine.
	Incorrect	You are concerning about the test result. It will be fine.
	Correct	We are concerned about the test result. It will be fine.
	Incorrect	We are concerning about the test result. It will be fine.

Appendix F: Background Questionnaire

Participant # __________________

We are conducting research on second language acquisition. Please fill in the form as truthfully as possible. Information on this form is kept entirely confidential. Only your participant number appears with the information you provide. By completing the questionnaire, you consent to our use of the information.

1. Date:__________
2. Age:___________
3. Gender:_________
4. Major:______________________
5. Standardized English proficiency test score (e.g., TOEIC, TOEFL, IELTS, etc.)?____________
6. At what age did you start to learn English? ____________________
7. Do you have experience living in English speaking countries?

   Years of residence in the Country:________________ Year(s)_______ Month(s) ______

   Years of formal education in total:____________________

   Years of formal education in English speaking countries_____________
8. Estimate your level of English on a scale of 1 (beginner) to 5 (advanced)

   Speaking 1 2 3 4 5 Listening 1 2 3 4 5

   Reading 1 2 3 4 5 Writing 1 2 3 4 5
9. List any other languages you speak besides English and your month tongue________________

   Estimate your level for each:

   Language:_________________

   a. Beginner b. Intermediate c. Advanced

   Language:_________________

   a. Beginner b. Intermediate c. Advanced

Appendix G: Exit Questionnaire

Directions: Write down your responses in the provided blanks (________).

1. What do you think were the goals of this study?

____

2. Did you learn anything from this study?

____

3. Did you focus on any specific grammatical or lexical feature in this study?

____

4. Did you look up or study any forms or information outside of this study?

____

5. If you have any comments about this task, please provide them in the space below:

____

____

____

____

Appendix H: Oral Productive Test

You have just seen the news on Korean TV, and you want to tell your English friend about the news. Please report the news briefly.

1.

2.

3.

4.

5.

Appendix I: Written Productive Test

You have just seen the news on Korean TV, and you want to post the news on your SNS in English to share it with your friends. Please write the news briefly.

6.

7.

8.

9.

10.

Appendix J: Receptive Test

Set A

1.	A major earthquake just offshore ____________________ (hit) Chile, killing three people. Rescuers ____________________ (work) into the early hours assessing damage in several coastal towns that saw flooding from small tsunamis set off by the quake.
2.	In the first round of discussions, India and other SAARC nations ____________________ (propose) that a regional learning outcome assessment system ____________________ (be) put in place to address the issues of poor learning levels.
3.	Two orphaned sisters separated decades ago in South Korea ____________________ (reunite) after being hired at the same hospital in Florida. The women, now both in their 40s, ____________________ (be) stunned to learn that they were related, having not seen each other since the early 1970s.
4.	Abhorring for-profit insurance companies, the left-wingers in Congress ____________________ (insist) that the Affordable Care Act ____________________ (establish) nonprofit insurance cooperatives to compete with big players like Aetna and Humana.
5.	Nebraska's death penalty ____________________ (stay) on the books until voters ____________________ (decide) next year whether to keep it.
6.	Saudi Arabia's King Salman ____________________ (order) a safety review for the Hajj pilgrimage after at least 717 people ____________________ (die) in a stampede near the holy city of Mecca.
7.	The International Monetary Fund ____________________ (refuse) to participate in a new bailout for Greece until there ____________________ (be) an "explicit and concrete agreement" on debt relief from the country's eurozone creditors.
8.	If you ____________________ (buy) a qualifying computer from the Microsoft Store for over $599, the company ____________________ (give) you a rebate after you send in your old laptop.
9.	NPAD (New Politics Alliance for Democracy) Floor Leader Lee Jong-kul ____________________ (propose) that rival parties and the government

	____________________ (set) up a committee of politically neutral figures to work on improving history textbooks.
10.	A 14-year-old boy ____________________ (arrest) on suspicion of starting a fire at Shinewater Primary School in Eastbourne and remains in police custody. Sixty firefighters ____________________ (be) needed to put out the blaze at around 1 a.m. on Monday morning.
11.	Iranian President Hassan Rouhani publicly ____________________ (propose) that the U.S. and Iran ____________________ (swap) prisoners held in each other's jails, including the *Washington Post*'s Tehran bureau chief Jason Rezaian, whom Iran has charged with espionage.
12.	Children ____________________ (be) eligible for free childcare from three years of age until they ____________________ (start) primary school.
13.	Florida's Governor Rick Scott ____________________ (refuse) to drop his lawsuit against the federal government until the Obama administration ____________________ (make) a decision on the Legislature's proposal for $2 billion in shared spending next year for hospital payments and a raise in Medicaid rates.
14.	England's players ____________________ (share) a jackpot of almost £6 million if they ____________________ (win) the World Cup, by far the largest incentive ever for any team.
15.	A Belfast hospital ____________________ (carry) out five kidney transplants in a single day. The first transplant ____________________ (begin) at 1:30 a.m. on Sunday.
16.	The commandant of the Marine Corps ____________________ (recommend) that women ____________________ (be) excluded from competing for certain front-line combat jobs, U.S. officials said Friday, as the Corps distanced itself from the other military services that are expected to allow women to serve in battlefield posts.
17.	Operations ____________________ (continue) until the area ____________________ (be) cleared of all "terrorist concentrations."

18.	Ministers ____________________ (intervene) on planning applications for controversial fracking operations if local authorities ____________________ (fail) to act quickly enough.
19.	The report ____________________ (recommend) that police ____________________ (include) community leaders in response planning and be more open and transparent with the public about those plans.
20.	Volkswagen Australia ____________________ (write) to affected owners and arrange for free repairs as soon as ~~the~~ a solution ____________________ (be) found, which may not be available until early next year.
21.	At least 10 people ____________________ (kill) in a suicide attack on a hotel in the Someli capital Mogadishu. Gunman ____________________ (use) a vehicle packed with explosives to blast their way into the Sahafi hotel compound before storming the building.
22.	A federal audit released in March ____________________ (recommend) that Maryland ____________________ (pay) the U.S. government $28.4 million that was misallocated because the state waited too long to formally update enrollment projections and numbers with federal grant providers.
23.	British luxury designer Burberry ____________________ (show) off its latest fashion collection on Snapchat before it ____________________ (hit) the runway ~~at~~ during London Fashion Week.
24.	Swiss prosecutors ____________________ (open) a criminal investigation into Sepp Blatter, the head of world soccer body FIFA, on suspicion of criminal mismanagement and misappropriation of funds. Blatter ____________________ (be) interrogated after a meeting of FIFA's executive committee in Zurich.
25.	Generally, a pharmacy ____________________ (process) a price reduction almost as soon as it ____________________ (be) announced and carried out by the manufacturer.
26.	The committee ____________________ (recommend) that federal law enforcement ____________________ (be) directed to identify and arrest dangerous criminals who

	try to buy illegal guns, and notify and work local and state officials when there are attempts.
27.	The European Central Bank ____________________ (take) further measures if it ____________________ (see) a significant risk to the outlook for inflation, Vice President Vitor Constancio said.
28.	Two Libyans ____________________ (identify) as suspects in the ongoing investigation into the Lockerbie bombing. The bombing of Pan Am Flight 103 in December 1988 ____________________ (kill) 270 people - the deadliest terror attack ever on British soil.
29.	A consultant ____________________ (recommend) that the commission members actively ____________________ (be) recruited to serve on the San Bernardino County Museum Association's board of directors and assist in fundraising activities.
30.	Donald Trump said on Wednesday that if he ____________________ (be) elected president, he ____________________ (send) Syrian refugees arriving in the West who could be ISIL fighters in disguise back to Syria.
31.	Amtrak ____________________ (suspend) service on parts of its national network by December unless Congress ____________________ (extend) its deadline for implementing advanced safety technology.
32.	Dublin City Council ____________________ (propose) that cars ____________________ (be) banned from key streets in a radical transport plan to make the city more cycle and pedestrian friendly.
33.	A two-storey building ____________________ (collapse) in China's central Henan province killing 17 construction workers and injuring 23 others. Rescue workers ____________________ (search) overnight pulling bodies and survivors from the debris in Beiwudu in Wuyang county.
34.	David Cameron's flagship Big Society plan ____________________ (fail) unless there ____________________ (be) big changes to the way the civil service works.
35.	The stock market ____________________ (anticipate) those higher oil prices as soon as it ____________________ (see) significantly more mergers and takeunders occur in the next few quarters.

36.	A toddler who was reported missing from his New Jersey home ____________________ (find) dead in nearby woods just three hours later. Three-year-old Brendan Creato ____________________ (disappear) from his family's apartment on Cooper Street.
37.	More ____________________ (die) from cholera unless we ____________________ (secure) clean water.
38.	Foreign college students ____________________ (be) banned from working and forced to leave the UK when their course ____________________ (end).
39.	The fragile care market ____________________ (collapse) unless councils ____________________ (act) quickly.
40.	The foreign secretary believes Britons ____________________ (opt) to leave the European Union unless EU leaders ____________________ (agree) to changes and that an in/out referendum campaign could be launched in spring 2016.

Set B

1.	Family members of passengers aboard a Malaysia Airlines flight - ____________________ (express) their frustration with the airline. Families ____________________ (meet) with representatives of the airline on Tuesday afternoon to discuss an agreement on compensation and conditions for potential travel to Malaysia.
2.	A total of 165 jobs ____________________ (be) lost at Wylfa Nuclear Power Station on Anglesey when the plant ____________________ (move) from energy generation to defueling.
3.	Two men and a woman ____________________ (detain) in Malaysia in connection with last month's bomb blast in Bangkok, which killed 20. Inspectors - ____________________ (say) they were of Pakistani nationality and Malaysian nationality.

4.	If Hurricane Joaquin ____________________ (cause) another gas crisis in New Jersey, gas stations ____________________ (have) more flexibility lowering prices under a law signed by the government.
5.	Tennessee lawmakers ____________________ (set) to vote on loosening vehicle emissions testing requirements as soon as they ____________________ (return) in January.
6.	The Prime Minister ____________________ (visit) Norwich to unveil plans to create millions more apprenticeships. He ____________________ (say) they would be paid for by a "levy" on large companies that aren't investing enough into staff training.
7.	Stanley S. Hubbard, the media mogul based in Minnesota, ____________________ (suggest) to the president that he ____________________ (stay) out of social issues.
8.	According to **The Telegraph**, Kiev ____________________ (insist) that the elections ____________________ (be) held in tandem with the rest of the country on October 25, but the self-proclaimed People's Republic of Donetsk and Luhansk have set ad-hoc elections on different dates.
9.	If the lawmakers ____________________ (be) successful, Nebraska ____________________ (become) the first largely conservative state in more than 40 years to strike down the death penalty.
10.	Fifteen traders including six children ____________________ (kill) in a fire at a public market in the Philippines, police say. About 13 others ____________________ (injure) in the blaze at a market in the southern city of Zamboanga.
11.	Autonomous vehicles ____________________ (be) in wide use in as little as three years, slipping into traffic before transportation departments and disrupted industries ____________________ (have) a chance to prepare.
12.	Under Kim Jong Un, North Korea ____________________ (insist) that it ____________________ (be) accepted as a nuclear weapons state and said it has no interest in an Iran-style deal to scale back its nuclear program.

13.	If a prime-age worker voluntarily ____________________ (leave) the labor force, the prime-age employment-to-population ratio ____________________ (fall), indicating that the economy isn't using labor resources effectively.
14.	A company called Orion says it ____________________ (pay) your baggage fees if you ____________________ (use) one of their[its?] shiny, new suitcases that's plastered with an advertisement.
15.	Two convicted murderers ____________________ (escape) from New York State'[Choose either simple or cursive and be consistent.]s largest maximum security prison using power tools. Between Friday night and Saturday morning, the pair - ____________________ (place) clothes in their bed to look like they were sleeping, cut through the wall and got onto a builder walkway.
16.	The prime minister ____________________ (refuse) to make a decision on the timing of a referendum until he ____________________ (judge) the progress made by December.
17.	Kiev and its Western backers ____________________ (demand) that local elections ____________________ (take) place according to Ukrainian law as set under the peace plan, not on the rebels' terms.
18.	If GOP presidential candidate Donald Trump ____________________ (be) elected, Billionaire Barry Diller ____________________ (move) out of the country.
19.	Toronto's former city hall building ____________________ (serve) as a courthouse for an additional five years as the city council considers extending [council has decided to extend?] the province's lease until a new downtown facility ____________________ (be) built.
20.	When European authorities ____________________ (demand) that Google ____________________ (remove) links that people found unflattering from its search results last year, that was already bad enough for the public's access to information.
21.	A Harvard employee ____________________ (accuse) of stealing $80,000 from the university and spending it on Lego and Apple products. He ____________________ (use) his university-issued credit card to pay for personal items during his 17 years as a computer lab manager.

22.	The committee ____________________ (propose) that each institution ____________________ (nominate) one colleague to form a task group to report to us with alternatives and recommendations by November 15, 2015.
23.	The British people ____________________ (vote) to leave the European Union unless Brussels ____________________ (give) "substantial and irreversible" reforms to the UK, Philip Hammond has warned.
24.	Russian cosmonaut, Gennady Padalka, ____________________ (return) safely to Earth after spending 879 days in orbit. The world's most experienced space flier ____________________ (join) by two crewmates, including Denmark's first-ever astronaut.
25.	The pool at the Meadows Community Recreation Centre ____________________ (be) closed while warranty work ____________________ (be) carried out.
26.	A legislative panel ____________________ (recommend) that Democrat Rep. Frank Mautino ____________________ (serve) as the new state auditor general, who oversees how tax dollars are spent.
27.	The Food and Drug Administration on Friday ____________________ (propose) that every nutrition fact label ____________________ (include) the percentage of ~~your~~ the recommended daily intake of added sugars in a food item.
28.	An intact tomb dating to the fourth century B.C. ____________________ (discover) in Pompeii by French archaeologists. The tomb ____________________ (construct) by the Samnites, who lived in south-central Italy and fought against the Romans.
29.	Sam Mitchell ____________________ (replace) Flip Saunders as coach of the Minnesota Timberwolves while Saunders ____________________ (continue) to undergo treatment for cancer.
30.	The suspension in selling Volkswagen's diesel vehicles ____________________ (remain) until the emission issues ____________________ (be) addressed in those vehicles.

31.	Britain ____________________ (need) to build three cities the size of Birmingham by 2020 unless action ____________________ (be) taken to tackle the migrant crisis.
32.	A Boston mother says United Airlines ____________________ (suggest) that she ____________________ (use) an indoor pet relief area at Dulles International Airport to pump breast milk.
33.	Police ____________________ (arrest) a man accused of opening fire at a party, killing one and injuring four others. The arrests ____________________ (make) during a 90-day investigation of two of Stockton's most notorious gangs.
34.	South Korea's anti-North propaganda broadcasts ____________________ (continue) to be blared across the border unless Pyongyang ____________________ (apologize), the country's president has said.
35.	Bayern Munich winger Franck Ribery ____________________ (return) to action this year after he fully ____________________ (recover) from a prolonged ankle injury.
36.	Five people, including two RAF personnel, ____________________ (kill) in a helicopter crash in Afghanistan. The Puma Mk2 helicopter ____________________ (crash) as it was landing at Nato's training and support mission HQ, in Kabul.
37.	Britain ____________________ (vote) to leave the EU unless European leaders ____________________ (agree) to a "substantial package of reform" demanded by David Cameron, the foreign secretary has warned.
38.	The Gangnam District Office in southern Seoul ____________________ (demand) that Seoul Metropolitan Government (SMG) ____________________ (grant) it autonomy in a protest against the latter's development plan.
39.	The first experiments away from the International Space Station ____________________ (take) place in cislunar space – the area of space between the Earth and the Moon – before missions ____________________ (begin) venturing further afield.
40.	Iran ____________________ (refuse) to sign any agreement, unless all economic sanctions ____________________ (be) lifted on the first day of the implementation of the deal.

Set C

1.	Nigeria ____________________ (launch) a major crackdown on fake bank-account holders in a bid to reduce fraud in the banking sector. People ____________________ (ask) to enroll at their local banks to have their fingerprints taken, along with a photograph of their face by the end of Friday.
2.	The two judges sitting at the High Court in London made it clear that if no assurance of a fair trial ____________________ (be) given, they ____________________ (refuse) to hand over Roger Giese, 40, to stand trial in California, where he is charged with sexually abusing a boy under the age of 14 from 1998 until 2002.
3.	Seven Hong Kong police officers ____________________ (charge) with allegedly attacking a demonstrator during pro-democracy protests in the territory last year. Activist Ken Tsang ____________________ (be) filmed being led away in handcuffs and beaten by police for several minutes on 15 October 2014.
4.	The report ____________________ (propose) that the intelligence services ____________________ (retain) the power to collect bulk communications data on the private lives of British citizens.
5.	Pope Francis ____________________ (meet) with Fidel Castro during his upcoming visit to communist-controlled Cuba if the former president's health ____________________ (be) strong enough.
6.	A UK university professor ____________________ (die) after falling 40-50 ft (12-15 m) onto rocks on a hiking trail in the US. Professor Alexei Likhtman, 44, of the University of Reading, ____________________ (loose) his balance and fell on a section of the Appalachian Trail in Maryland.
7.	Washington ____________________ (propose) that all members of the World Trade Organization (WTO) ____________________ (undertake) voluntary commitments to reduce their market price support and input subsidy programs.
8.	The economic gap between London and Northern UK cities ____________________ (continue) to widen until Chancellor George Osborne ____________________ (provide) substantially more funding through his "Northern Powerhouse."

9.	If the Federal Reserve ____________________ (raise) interest rates, housing in San Francisco ____________________ (begin) to see a significant cooldown.
10.	A former undercover policeman ____________________ (sentence) to six and a half years in prison for stealing $700,000 of the virtual currency bitcoin. Agent Carl Force ____________________ (be) part of the Drug Enforcement Administration (DEA) investigation into the black market website Silk Road.
11.	Regulators ____________________ (insist) that Sir Mike ____________________ (remain) on the Barclays board until a new chief executive is in place in order to avoid excessive influence being wielded by Mr. McFarlane.
12.	The Chicago Public Schools' chief says the district ____________________ (have) to issue layoff notices by Thanksgiving if there ____________________ (be) no financial help from the state.
13.	The government ____________________ (announce) its plan for new history textbooks as soon as the National Assembly audit of the Park Geun-hye administration ____________________ (wrap) up, admitting that the president told the Education Ministry to develop a new textbook.
14.	A Beijing judge ____________________ (suggest) that her face ____________________ (be) examined by medical professionals to certify its authenticity.
15.	Nearly 600,000 migrants ____________________ (reach) the EU by sea so far this year, many of them travelling from Turkey to Greece before seeking to head north. Turkey ____________________ (make) a number of demands in exchange for helping to stem the flow.
16.	As soon as the contract ____________________ (be) signed, traditional members ____________________ (get) a $4,000 up-front lump sum bonus and in-progression members will get a $3,000 payment.

17.	Poland ____________________ (introduce) border controls if there ____________________ (be) any threat to national security amid the ongoing refugee crisis in Europe.
18.	Efforts to seize people traffickers' boats in the Mediterranean ____________________ (fail) unless the EU also ____________________ (fight) the criminal gangs in Europe.
19.	An Islamic committee in Kuwait ____________________ (propose) that the government ____________________ (set) up a television facility to censor foreign satellite broadcasts and redistribute the programs by cable.
20.	A prominent Islamic leader in Indonesia, Abdurrahman Wahid, ____________________ (propose) that the ousted President Suharto ____________________ (be) granted amnesty in exchange for returning any money allegedly siphoned from the state during his years in office.
21.	A man whose legal name is Santa Claus ____________________ (elect) as a city councilman in the town of North Pole. The 68-year-old, formerly known as Thomas Patrick O'Connor, ____________________ (win) a three-year term in the community of about 2,200 residents south-east of Fairbanks.
22.	Justin Trudeau's fledgling government ____________________ (have) less than a month to get up and running before the new prime minister ____________________ (take) off for the first in a whirlwind series of international summits.
23.	This ____________________ (be) his third tour of duty in Afghanistan before he ____________________ (take) up a new role as a trainer.
24.	A study of Ebola survivors in west Africa ____________________ (find) a group of women who appear to be immune to the deadly virus. The discovery ____________________ (be) made by a team of British and European scientists who are studying Ebola survivors in Guinea.
25.	The UN says it ____________________ (have) to suspend humanitarian work in Gaza unless it ____________________ (receive) fresh fuel supplies.

26.	It is hoped a deal ____________________ (be) reached before India's Prime Minister ____________________ (meet) the European Commission in February.
27.	Politicians in Guinea ____________________ (propose) that elections ____________________ (be) held by the end of 2009 to replace the army officers who took power in a coup in December.
28.	Eurotunnel services ____________________ (disrupt) for the third time in three days. They ____________________ (be) suspended for about an hour earlier because of "intruder activity on our French platforms" at Coquelles.
29.	It ____________________ (proposed) that camping management bylaws ____________________ (be) introduced to two new areas, covering many lochs in the Trossachs, much of the west side of Loch Lomond and the north-east tip of Loch Long.
30.	Afternoon temperatures ____________________ (be) slightly warmer than normal until the rain ____________________ (arrive) on Friday.
31.	Australia ____________________ (end) up behind other countries unless it ____________________ (increase) funding for university research to match competitors in Asia, Europe, and the US.
32.	Banks in Cyprus ____________________ (remain) closed until at least Thursday while talks ____________________ (continue) over controversial plans to put a levy on savers' deposits.
33.	Steinbach, one of Germany's best-known makers of traditional wooden nutcrackers, ____________________ (file) for bankruptcy. The company ____________________ (blame) high wage costs resulting from the country's minimum wage, introduced at the beginning of this year.
34.	Unless Congress quickly ____________________ (pass) a 2016 budget including adequate defense appropriations, our military ____________________ (find) it harder to live up to its responsibilities under the Constitution.

35.	Joni Mitchell's lawyer, Rebecca J. Tyne, ____________________ (recommend) that Mitchell ____________________ (remain) under a conservatorship until she fully recovers from her stroke.
36.	The Ten Commandments monument ____________________ (remove) from the state capitol grounds under cover of darkness. On Monday, using a heavy-duty crane and cutting tools, workers from a private contractor ____________________ (move) the tablet-style monument to the offices of the Council of Public Affairs.
37.	The scouting process ____________________ (continue) for players for the U17 World Cup until the registration window ____________________ (close).
38.	Unless Bulgaria immediately ____________________ (start) rearming its army, it ____________________ (have) to pay other countries for its security.
39.	Opposition parties ____________________ (demand) that the Welsh agriculture secretary ____________________ (set) the record straight after a leaked email led to fears that the beef-on-the-bone ban is to be brought back.
40.	Ms. Bouchart ____________________ (demand) that Britain ____________________ (open) its border to let them travel freely from northern France to Britain.