Attentional processing of different types of written L2 input and its relationship with learners' working memory capacity

Bimali Indrarathne

This thesis is submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Applied Linguistics
Department of Linguistics and English Language
Lancaster University

October 2016

Acknowledgement

This thesis would not have been possible without the help of a large group of people. First of all, I owe a great debt of gratitude to Professor Judit Kormos, my supervisor, for her constant guidance and encouragement during my PhD journey. She has always been there for me whenever I needed her help, PhD or otherwise. I'm very much grateful to her for all the support that she has given me in many ways.

I'm grateful to my participants at Kotelawala Defence University (KDU) in Sri Lanka for taking part in this study. I would also like to thank my colleagues, especially Wimansha, Kithsiri, Lakshani and Vasuki and the management at KDU, who not only facilitated the data collection, but also have encouraged me throughout. I would like to express my gratitude to Dr Gareth McCray for teaching me eye-tracking, Dr Kristof Savski for helping me with corpus data and Carol Bennett for proofreading the thesis.

I would also like to thank Dr Andrea Revesz from UCL, who first encouraged me to apply for a PhD and has always been helpful in my studies. My thanks go to the staff at the Department of Linguistics at Lancaster University for their constant support and especially to Marjorie for all the kindness.

I also wish to acknowledge the EAP studentship of Lancaster University which partly funded this PhD.

Finally, I would like to thank my family and friends back home who have always been there for me. Kathy and the housemates at 29, Queen Street created a second home for me in Lancaster, my friends in the UK have made my life so special here during the PhD journey. Thank you all!

Abstract

It is widely accepted that sufficient target language input is necessary for successful L2 learning. However, the type and the amount of input as well as several cognitive mechanisms play an important role in input processing. The present study investigated how different cognitive mechanisms are influential in processing different types of input.

The study was conducted with 100 undergraduate L2 learners of English in Sri Lanka. Seven examples each of the target construction (*causative had*) were embedded in three short stories that the participants read on three occasions under four different input conditions: two implicit and two explicit input conditions. The implicit conditions were input flood and textual enhancement (TE). In both explicit conditions learners were asked to pay attention to the highlighted target examples. In addition, the participants in one of the explicit input conditions were provided with an explicit explanation of the target construction. Improvement on the knowledge of the target construction was measured by a pre/post-test design, which included a sentence reconstruction (SR) and a grammaticality judgment (GJ) task. The four experimental conditions contained 20 participants each and the rest were in the control condition.

In terms of cognitive mechanisms, the study investigated how attention paid to target examples in these different input conditions influences learning gains using two eye-tracking measurements: total fixation duration (TFD) and the difference between observed and expected TFD. In addition, the involvement of working memory (WM) in facilitating the regulation of attentional resources was studied based on four WM tests that measured the capacity of the PSTM and the functions of the central executive (CE).

In order to analyse the relationship between learners' awareness of the target construction and their performance, a questionnaire was administered.

The eye-tracking data revealed that explicit input conditions could draw attention of the learners more successfully to the target examples than the implicit input conditions. In line with this finding, the pre/post-test data analysis highlighted that the learners in explicit conditions could develop both the explicit and the implicit knowledge of the target construction while the learners in one of the implicit conditions (TE) could only develop the implicit knowledge. The results also showed that those with a high WM capacity are able to direct their attention to the target constructions more successfully and consequently they could develop both explicit and implicit knowledge of the target construction. The study indicated that gains of either explicit or implicit knowledge are independent of learners being aware of the target syntactic construction. A free writing task administered as a part of the pre/post-tests highlighted that only a few learners were able to use the target construction in a meaningful communicative situation after the experiment.

Contents

List of tables

List of figures

List of appendices

List of abbreviations

AOI	Areas of Interest
BNC	British National Corpus
CE	Central Executive
CQ	Comprehension question
DS	Digit Span
EAP	English for academic purposes
EFL	English as a foreign language
ELT	English language teaching
FFI	Form focused instruction
HL	Heritage language
KT	Keep Track
MFI	Meaning focused instruction
FonF	Focus on Form
FonM	Focus on Meaning
GJ	Grammaticality judgement
L1	Fist language
L2	Second language
LTM	Long-term memory
NH	Noticing hypothesis
PM	Plus Minus
PSTM	Phonological short-term memory
RQ	Research question
SAS	Supervisory Activating System
SLA	Second language acquisition
SR	Sentence reconstruction
ST	Stroop Task
STM	Short-term memory
TL	Target language
TE	Textual enhancement
TFD	Total fixation duration
TGC	Target grammatical construction
WM	Working memory
ΔOE	Difference between observed and expected fixation duration

Chapter 1: Introduction

Second language (L2) acquisition (SLA) theories developed to date unanimously agree that the "*sine qua non*" (Gass & Mackey, 2015, p. 181) of successful L2 learning is the exposure to sufficient target language (TL) input. Nevertheless, providing sufficient input does not guarantee that learners acquire the TL successfully (Doughty & Williams, 1998) mainly because not all input is processed by learners. This leads to the argument that the type, the amount and the mode of input provided can influence L2 learning. Moreover, cognitive mechanisms, such as attention, consciousness and awareness, that are involved in how learners process input play an important role in the L2 learning process. Individuals differ in their attention regulation abilities and thus these differences also account for successful L2 learning. Due to the important place that input processing enjoys in the L2 learning process, a plethora of research in the SLA field has been devoted to understanding how input influences language development and how individual differences can explain input processing. However, within this area of research, there are still under-researched questions and unexplained and ambiguous phenomena. The study presented in this thesis attempted to investigate the processing of different types of L2 input and the involvement of several cognitive mechanisms and individual differences in input processing.

This introductory chapter of the thesis provides the background of this study. It starts with a section that explains my motivation behind the study (1.1) and Section 1.2 outlines the main theoretical concepts on which the study was based and also highlights the research gaps this study attempted to fulfil. Section 1.3 briefly explains the main aims of the study and the research design. Section 1.4 provides the layout of the thesis.

1.1 Motivation

The motivation behind this study comes from several observations that I made as a second language learner and a teacher. In my experience as an L2 learner of English and also as a teacher of English, I have observed that adult L2 learners sometimes struggle to learn certain grammatical features of the English language. Moreover, despite receiving the same input, some learners find it more difficult to acquire certain grammatical features than the other learners. There can be various reasons behind these difficulties. For example, DeKeyser (2005) lists five factors that can influence adult L2 acquisition: ‘‘the characteristics of the L2, the influence of the first language (L1), the role of age, the role of individual differences in cognitive and affective ‘aptitudes’, and the role of learning context” (p. 2). The findings of a recent study by Shiu (2011), which investigated the perception of grammatical difficulty by Taiwanese university level English as a Foreign Language (EFL) learners, are in line with the features identified by DeKeyser. Shiu’s study shows that the learners’ perceptions on grammatical difficulty depended on if “the rules to describe the formation of language features are easy or difficult to articulate” (p. ii), the learners’ L1 and L2 proficiency and their L2 grammar learning experience.

It is assumed that the difficulties adult learners have with the acquisition of certain grammatical features might be alleviated by explicit grammar explanations. In relation to this assumption, I have also noticed that most adult learners I have taught are interested in learning grammar of the L2 through explanations of language rules and conventions. Moreover, I have come across learners who would even request the teacher to give a metalinguistic explanation of language rules despite receiving plentiful practice opportunities in class with examples of TL structures. Shiu's (2011) findings shed some light on this observation. The perceptions of grammatical difficulty among the participants of her study varied according to the participants' explicit and implicit knowledge of the particular features. In error correction, error identification and error explanation tasks, which measured the explicit knowledge of the participants, the performance was accurate for the features of the TL that the participants themselves identified as less difficult. However, in the implicit knowledge measurement tasks i.e. oral picture cued elicitation tasks, the participants' performance of the same language features was not as accurate as in the explicit tasks. According to these findings, it can be assumed that learners' requests for metalinguistic explanations can be related to the type of tasks that they have to complete in class or under examination conditions. In this study, I was interested in investigating this further in order to understand how explicit and implicit input tasks assist learners in their development of grammatical knowledge.

Although different types of syllabi are used in different adult ELT programmes, nowadays, most of the ELT course books such as *Cutting Edge* (Cunningham & Moor, 2013), *Headway* (Soars & Soars, 2014) *Face to Face* (Redston & Cunningham, 2013) tend to teach grammar components through implicit techniques. For example, they provide an abundance of examples of target words/phrases/sentences without any explicit explanations of the target constructions. This can be identified as input flood, which is entirely an implicit input technique. Moreover, there are other mechanisms such as underlining or boldfacing, which can be recognised as textual enhancement (TE) techniques. TE also falls under implicit input since an explicit explanation of the target construction is not provided. Some of these course books contain explicit explanations of the target features as an appendix and it is up to the teacher to decide whether to give these explanations in lessons. As a teacher, when I used different techniques to present TL input, I realised that certain techniques are more beneficial in developing learners' syntactic knowledge of grammatical constructions. Due to the interest that I developed on more productive methods to provide TL input, I decided to investigate the effectiveness of different input techniques in this study.

I have also observed that individual learners process L2 input differently despite receiving the same input. As DeKeyser (2005) points out, individual differences can be decisive in L2 acquisition. This inspired me to study the individual differences involved in L2 input processing. Although there are several individual differences such as motivation, gender, aptitude and attitude, which can be influential in L2 learning, the focus of this study was the working memory (WM) capacity of the learners. The main reason behind this choice is the lack of studies in the field of SLA investigating the relationship between L2 input processing and its relationship with WM capacity. The few existing studies have examined the influence of the WM by measuring the capacity of the phonological short term memory (PSTM) in the phonological loop, which is only one component of the WM. Thus, this study was designed to understand the role of both PSTM capacity and central executive (CE) function of the WM in L2 input processing.

1.2 Theoretical background to the study

[1]Providing sufficient input of the TL to learners is a necessary condition for language development (R. Ellis, 1994; Gass, 1997; Krashen, 1985; VanPatten, 2002). Nonetheless, TL input alone may be insufficient for learning to take place (Doughty & Williams, 1998; Swain, 1985) because learners may not pay attention to input or might not be able to process it in a way that would facilitate their linguistic development. Thus, attention to TL input is recognised as a crucial factor for language learning (Gass, 1997). There have been several attempts in SLA literature to explain how attention assists language learning (e.g., Massaro, 1975; Schmidt, 1990; VanPatten, 1996); however, conceptualisation of attention and its associated terms such as awareness and consciousness in the SLA field seems to be vague. For example, in cognitive psychology, attention and consciousness are treated as separate phenomena (Koch & Tsuchiya, 2006; Lamme, 2003) while in some SLA theorisations, consciousness is not separated from attention (Carr & Curran, 1994; Schmidt, 1993). Furthermore, there is no agreement in SLA literature on the involvement of awareness in input processing. One argument is that attention and awareness are necessary conditions for L2 learning (Schmidt, 1990, 1993, 2010) and the other is that awareness is not a necessary condition (Tomlin & Villa, 1994; Robinson, 1995). This lack of accurate interpretation of key terminology such as attention, awareness and consciousness that are involved in input processing is one shortcoming that I have observed in the existing SLA literature.

The methodology applied so far in SLA studies to measure attention and awareness has not always been able to provide accurate information about the involvement of those

[1] Sections of this paragraph were published in Indrarathne & Kormos (2016)

cognitive processes. For example, some methods have not separated attention and awareness in collecting data. In Williams' (1999, 2001) studies, classroom observation was used to collect data on attention; however, not only attention but also other complex cognitive mechanisms may be involved when learners are in a classroom context. Thus, the findings may not reveal accurate information related to attentional processing of language input through such a methodology. A common data collection method used in attention research is think-aloud. However, this is criticised for giving an extra processing load to the participants since they have to verbalise what they are attending to while processing input during the think-aloud process. Methods such as asking participants to underline the language features that they are attending to in the input (Gass, Svetics & Lemlin, 2003) do not seem to have provided accurate information on the amount of attention paid either.

[2]There is an abundance of studies in the field of SLA that have previously investigated the various conditions under which provision of input might assist language learning (Green & Hicks, 1984; Long, 1991; Robinson, Mackey, Gass & Schmidt, 2012; Schmidt, 1995; VanPatten, 1996). Yet little is known about the attentional processing involved in learning from different types of input and how it relates to learning gains in various input conditions. Most studies to date have focused on attentional processing in conditions where participants were exposed to input flood and where the input was textually enhanced (Simard & Foucambert, 2013; Winke, 2013), but no research has previously examined how instruction to pay attention to specific aspects of the input and explicit metalinguistic explanation facilitate attentional processing.

[2] Sections of this paragraph were published in Indrarathne & Kormos (2016)

WM abilities are key individual cognitive characteristics that can potentially influence L2 learning outcomes (Juffs & Harrington, 2011). WM assists in the comprehension of L2 input, directing learners' attention to the relevant features of the input and processing and encoding this perceived input in the long-term memory (LTM). WM is part of what, in educational psychology, is called aptitude complexes (Snow, 1992), in other words individual differences that dynamically interact with the situation in which learning takes place. While the role of language learning aptitude under various learning conditions has been studied in previous studies (e.g., De Graaff, 1997; Erlam, 2005; Robinson, 2005; Yilmaz & Granena, 2015), little is known about how WM and its associated function of regulating attention influence the acquisition of previously unknown syntactic constructions in explicit and incidental learning conditions. Currently available findings in the field of SLA research and cognitive psychology with regard to the interaction of WM abilities with instructional treatment types are inconclusive.

In summary, this study has attempted to fill four main gaps in the existing research. These are: (1) the lack of accurate interpretation of three key terms i.e. attention, awareness and consciousness in the SLA literature (2) inappropriate or inaccurate data collection methods used to investigate the involvement of those key concepts in L2 input processing (3) lack of knowledge of how explicit explanation of a target grammatical feature and instructions to pay attention to that feature influence the subsequent development of knowledge of it, and (4) lack of knowledge on how WM capacity including both PSTM and the CE influences processing of input in explicit and implicit input conditions.

1.3 Aims and the design of the study

This thesis presents the findings of a study that investigated how L2 learners pay attention to a target syntactic construction in written L2 input in four different conditions: input flood, whereby frequency of the target item is increased in the input; TE, whereby the target construction is highlighted in the text; instruction to pay attention to the highlighted grammatical construction; and explicit metalinguistic explanation of the highlighted TL construction. According to Norris and Ortega (2000, p. 437), the instruction conditions that have "neither rule presentation nor directions to attend to particular forms were part of a treatment" can be treated as implicit instruction. Thus, input flood and TE conditions in this experiment can be regarded as implicit learning conditions, and an instruction to pay attention to the target feature and an explicit metalinguistic explanation as explicit conditions (Spada & Tomita, 2010).

One hundred L2 learners of English at a Sri Lankan university were exposed to a target syntactic construction (*causative had*) in these four different conditions (one control group and four experimental conditions). How learners' knowledge of the target construction changed because of the different types of input received in these conditions was measured through a pre-test and a post-test. The study also investigated how the change in knowledge under different input conditions is related to attentional processing of input by collecting data on attention with the help of eye-tracking. In addition, a post-exposure interview was used with the aim of understanding if awareness was involved during attentional processing of input. Learning TL features involve not only understanding the form and meaning but also the use of those features for communication. Therefore, it was deemed necessary to evaluate the ability of learners to use the target grammatical construction (TGC) on this occasion, so a free production task was administered as a part of the pre- and post-tests for this purpose.

The study further investigated how WM capacity of individual learners influences the way that they process input in incidental and intentional learning conditions. A novel feature of the study was that in addition to a test of PSTM capacity, which most SLA studies have used to measure WM capacity, three additional tests that assess individual differences in attention regulation controlled by the CE were also administered.

1.4 Structure of the thesis

This thesis consists of seven chapters including the introduction. In Chapter 2 a detailed discussion of theoretical concepts related to the study is provided. This chapter analyses the key concepts such as attention, awareness and consciousness in cognitive psychology as well as in SLA literature. It also provides the rationale behind the main data collection method (eye-tracking) used to collect data on attention in this study. The theoretical and empirical findings related to input processing in explicit and implicit conditions and the relationship between WM and L2 learning are also discussed in detail.

Chapter 3 of this thesis presents the pilot study of this research. The main aim of the pilot study was to investigate the plausibility of the data collection methods. This chapter discusses the background to the pilot study including the context and participants and presents a detailed description of the instruments used in data collection. The chapter also presents the findings and gives a detailed discussion of the lessons learnt in piloting the study and how the design of the main study was informed by the outcomes of the pilot. Chapter 4 then outlines the methodology of the main study. It contains a description of the context where data collection took place, the participants, the study design, instruments, data collection procedure and data analysis methods.

This study answered eight research questions categorised into three sections. In the first section, the relationship between attentional processing and explicit and implicit learning gains are discussed. The second section explores the relationship between attentional processing, learning gains and WM. In the last section, the involvement of awareness in attentional processing in relation to learning gains and different types of input is investigated. The results of these research questions are presented in Chapter 5, which starts with a detailed discussion of how data screening was performed. Then it presents the results in relation to the three sets of research questions. The discussion of the findings is included in Chapter 6. It presents the discussion of the findings under the three sets of research questions.

The concluding chapter (Chapter 7) summarises the main findings in relation to the eight research questions, discusses the theoretical and methodological contribution of the study and the pedagogical implications of the findings. This chapter also outlines the limitations of the study and future research directions.

Chapter 2: Literature review

2.1 Introduction

This study had two main aims: (1) the analysis of the attentional processing of L2 input in different instructional conditions and (2) the investigation of the relationship between attentional processing and learners' WM capacity. In order to lay the theoretical foundations of this study, it was vital to define and elucidate the key concepts, such as attention, awareness and noticing. Moreover, a detailed description of how WM capacity influences language learning was also deemed crucial.

The review of SLA literature, however, did not provide convincing clarifications of the terminology, such as attention, consciousness and awareness. Thus, Section 2.2 of this literature review discusses how the terms are defined in cognitive psychology. Section 2.3 reviews how these concepts are interpreted in SLA literature and also establishes the theoretical basis related to these cognitive processes on which the study was based.

Not only is the definition of this terminology in the SLA literature vague, but also the methodology that investigated attention, awareness and consciousness in empirical studies on L2 learning seems to have potential problems. Therefore, Section 2.4 provides a critical review of the data collection methods applied in previous empirical studies and a theoretical rationale for the selection of data collection methods in this study.

This study investigated attentional processing in different input conditions: therefore, a section of this literature review (2.5) is devoted to discussing different input techniques.

Section 2.6 outlines the theoretical background behind input processing, which is another key concept related to the design of this research. The final section includes a discussion of WM models and the relationship between WM and language learning in general and L2 learning in particular.

2.2 Attention, awareness and consciousness in cognitive psychology[3]

Although attention is a key concept in the field of SLA research as well as in cognitive psychology, there is still no existing consensus on what attention actually is (Allport, 1993; Shinn-Cunningham, 2008; Wolfe & Horowitz, 2004). Conceptualisations of attention, in cognitive psychology have been influenced by an early definition by James (1890) who identified attention as:

> the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others (pp. 403-404).

Two key phrases in this definition are particularly important for current theories of attention. The phrase “taking possession by the mind” indicates the nature of control involved in attention. Parallel terms are used in current definitions that all refer to attention as a “perceptual process” (McFarland, 2006, p.7), and the intentional control of the allocation of attention (Shiffrin, 1988; Styles, 2006). The phrase “one out of what seem several simultaneously possible objects or trains of thought” describes the selectiveness involved in attention. In more recent psychological theorisations, this has

[3] Several parts of this section of the literature review were published in Indrarathne & Kormos (2016).

been referred to as “selective response to stimuli” (McFarland, 2006, p.7) and “selecting some information for further processing” (Smith & Kosslyn, 2006, p. 28). Another feature of attention is its limited capacity, which restricts the number of stimuli one can process at a given time (Desimone & Duncan, 1995; Shiffrin, 1988; Styles, 2006).

While acknowledging that attention is not a unitary system, Chun, Golomb and Turk-Browne (2011) define it as “a core property of all perceptual and cognitive operations. Given limited capacity to process competing options, attentional mechanisms select, modulate, and sustain focus on information most relevant for behavior” (p. 73). They also draw up a taxonomy of attentional processing in which they interpret James’ terms, *objects and trains of thought* as external and internal attention. On the one hand, we can select and modulate stimuli perceived externally through our senses (e.g., specific information presented on a screen), which they call external attention. This can be a stimulus-driven bottom-up process, as well as a goal-directed top-down process. On the other hand, internal attention operates over internally generated representations such as task rules, responses, LTM and WM.

Due to the difficulty of defining what attention is, different terms related to attention are used in psychological studies and they provide a narrower focus to attention. For example, selective attention and visual attention are two such terms widely used in cognitive psychology. According to Fuster (2005), the human mind is comprised of several networks, and when someone pays attention to something, resources in the mind are allocated to activate a particular network. Selective attention involves selecting one of those networks to be attended for further processing. Several theories in psychology also support the concept of selective attention. For example, Kahneman's (1973) capacity theory notes that the human mind has a limited capacity; hence, certain amounts of information can be attended to at a given time and the selection of stimuli to be attended depends on task demand and importance. The filter model by Broadbent (1958) emphasises that the stimuli that someone receives externally are filtered in the mind, suggesting that the mind itself selects information to be attended to. Inhibition is a process that takes place in selective attention in order to stop the processing of other stimuli (Smith & Kosslyn, 2006). Although inhibition plays a role in attention, Shettleworth (1998) points out that attention does not necessarily filter out all other stimuli, but has the ability to process the most important stimuli at a given time. In line with this argument, Treisman's attenuation model (1964) suggests that unattended stimuli in the input do not completely get blocked or ignored, rather they become attenuated or in other terms, less effective.

The second term discussed in relation to attention is visual attention. Desimone and Duncan (1995) explain that "at any given time only a small amount of the information available on the retina can be processed and used" (p. 193) and that is called visual attention. The use of the term 'retina' indicates that the attention described here is related to what one actually processes through the eyes rather than what is available to be heard or what is already available in the mind. Chelazzi, Perlato, Santandrea and Della Libera (2013) who combine both visual and selective attention provide a definition that "visual selective attention is the brain function that modulates ongoing processing of retinal input in order for selected representations to gain privileged access to perceptual awareness and guide behaviour" (p. 58). Despite the still continuing debate about the exact nature and definition of attention, the consensus position in cognitive psychology is that it can be seen as a process with limited capacity, controlled by one's mind that selects certain stimuli to be processed. Thus, attention can be summarised as "a core property of all perceptual and cognitive operations. Given limited capacity to process competing options, attentional mechanisms select, modulate, and sustain focus on information most relevant for behavior" (Chun, Golomb & Turk-Browne, 2011, p. 73)

Consciousness, which in simple terms describes the understanding of one's experiences (Nagel, 1974; Max Velmans, 2009), is a concept that is inseparable from attention, yet it still needs to be distinguished from attention (Bachman, 2006; O'Regan & Noe, 2001). Koch and Tsuchiya (2006) distinguish consciousness from attention based on bottom-up and top-down control processes. Bottom-up control means paying attention to a salient stimulus among other stimuli, whereas top-down control refers to the search for specific stimuli demanded by the task (Koch & Tsuchiya, 2006). Koch and Tsuchiya argue that even certain top-down processes might not necessarily involve consciousness. For example, in processes such as priming, adaptation and visual search, "subjects can attend to a location for many seconds and yet fail to see one or more attributes of an object at that location" (p. 17). This demonstrates the potential lack of consciousness in these cognitive processes. However, according to Koch and Tsuchiya, attention with consciousness is necessary for the registration of stimuli in WM, the distinction between stimuli and the provision of a full report of the stimuli attended.

Table 2.1

A fourfold classification of conscious and unconscious percepts and behaviors

	Might not give rise to consciousness	Gives rise to consciousness
Top-down attention is not required	Formation of afterimages Rapid vision (<120 ms*) Zombie behaviors**	Pop-out in search Iconic memory Gist Animal and gender detection in dual tasks Partial reportability
Top-down attention is required	Priming Adaptation Visual search Thoughts	Working memory Detection and discrimination of unexpected and unfamiliar stimuli Full reportability

*ms=milliseconds

**Zombie behaviours: "highly trained, automatic, stereotyped and fluid visuomotor behaviors" (p. 19) such as driving home "on automatic pilot" (p. 19)
Adapted from "Attention and consciousness: two distinct brain processes" by C. Koch & N. Tsuchiya. 2006. *TRENDS in Cognitive Sciences, 11*(1), p. 17.

Awareness is another concept closely related to attention and consciousness. Lamme's (2003) cognitive psychological model is helpful in elucidating how attention can be distinguished from both consciousness and awareness (see Figure 2.1). Lamme argues that input can be perceived either consciously or unconsciously. Consciously perceived input can either be attended to or remain unattended. Thus, in his model, attention is seen as distinct from consciousness. Input that has been consciously perceived and attended to is then available for conscious reporting and can be stored in the WM. This final level in the model is regarded as awareness. There is still disagreement in the field of cognitive psychology on the question whether awareness is needed for learning to take place. On the one hand, based on pain sensitivity studies, Becker, Kleinbohl, and Holzl (2012) claim that awareness is not a necessary pre-condition for learning. On the other hand, Shanks and John (1994) argue that it is indispensible for subliminal, conditioning, artificial grammar, instrumental, and sequence learning.

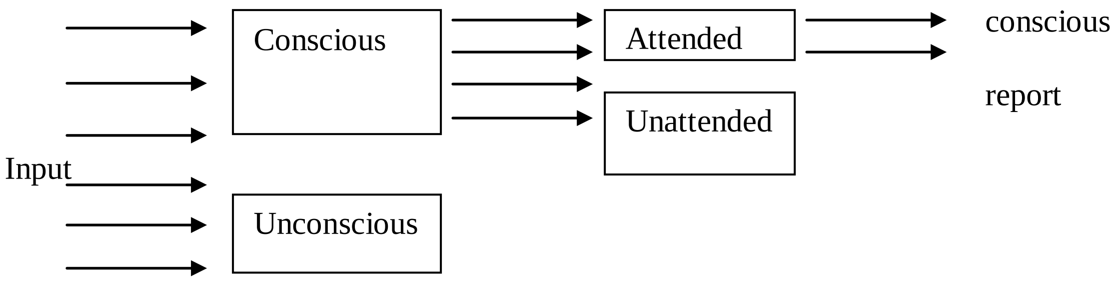


Figure 2.1. Four models of visual awareness and its relation to attention (d). Adapted from "Why visual attention and awareness are different" by V.A.F. Lamme. 2003. *TRENDS in Cognitive Sciences, 7*(1), p. 13.

Whereas Koch and Tsuchiya (2006) argue that WM functions and reporting involve consciousness, according to Lamme (2003), WM activation and reporting involves awareness. Since Koch and Tsuchiya do not discuss awareness, it is difficult to determine how it fits in their model. Their argument that input can be attended consciously or unconsciously seems to be closely related to Lamme's model of conscious/unconscious input. Therefore, it can be argued that consciously attended stimuli in the input can be registered in the WM if this attention involves awareness.

As discussed so far, attention, consciousness and awareness are treated as closely related phenomena in cognitive psychology yet they are different from each other. Kentridge (2011) who summarizes empirical evidence on the three phenomena is also of the view that they are different processes. However, there is a common agreement that consciousness/awareness is necessary to register stimuli in the WM or to provide a report on the stimuli attended. All these three phenomena have been widely used in SLA research and in the next section I will summarize how they are interpreted in SLA literature.

2.3 Attention, consciousness, awareness and noticing in second language acquisition[4]

The role of attention, awareness and consciousness is key in relation to how L2 learners process TL input. Existing studies in the field of SLA unequivocally agree that paying attention to certain features in the input is necessary for language development (e.g., Green & Hicks, 1984; Long, 1991; Schmidt, 1995; VanPatten, 1996; Robinson, Mackey, Gass & Schmidt., 2012; Leow, 2013), and also it is often claimed to be the "most important aspect of language learning" (Gass, 1997, p. 1). However, making TL input available to learners in itself may be insufficient for learning to take place (Doughty & Williams, 1998; Swain, 1985) since learners may not process everything in the input. Hence, a distinction between input and intake is made (Corder, 1967; Schmidt, 1990; VanPatten, 1996, 2004) and intake is defined as "that part of input which has actually been processed by learners and turned into knowledge of some kind" (Corder, 1967, p. 8). In other words, "intake is that part of the input that the learners notice" (Schmidt, 1990, p. 139).

Several attempts have been made to identify what in the input becomes intake and what mechanisms are involved in processing input. The Information Processing Model by Massaro, (1975) and VanPatten's (1996) Input Processing Model highlight that learners

[4] Several parts of this section of the literature review were published in Indrarathne & Kormos (2016).

are likely to process certain features in the input as a result of them paying attention to it. Schmidt (1995) too believes that learning without attention is impossible and all forms of learning, such as implicit, explicit, declarative or procedural need attention.

Nevertheless, early conceptualisations of the three key constructs: attention, awareness and consciousness, in the field of SLA are often vague and are not aligned with relevant theories in cognitive psychology. For example, Carr and Curran (1994), note that “if you are conscious of something, then you are attending to it… and if you are attending to something, then you are conscious of it” (p. 209). The second part of this definition seems to suggest that their position is that all forms of attention are conscious. Schmidt’s interpretation of consciousness does not seem to be very different from this either. Based on Baars (1988), Schmidt (1993) argues that “when people pay attention to something, they become conscious of it” (p.209). However, as discussed in Section 2.2, this is different from the interpretations provided in cognitive psychology in which attention and consciousness are identified as two distinct phenomena. Vygotsky (1962) has also identified the role of consciousness in language learning and accordingly he argues “to become conscious of a mental operation means to transfer it from the plane of action to that of language, i.e., to recreate it in the imagination so that it can be expressed in words” (p. 88). This form of consciousness i.e. the possibility to offer a report is similar to the model proposed by Koch and Tsuchiya (2006) and it may thus be considered awareness, as Lamme (2003) highlights in his model. Although attention has not been clearly defined in early SLA literature, Gass (1988) and Long (1991) identify selective attention as important in language learning. Moreover, based on capacity theories discussed in Section 2.2., Carr and Curran (1994), Green and Hicks (1984) and VanPatten (1996) claim that only a certain amount of information can be attended to at a given time. Thus there is a common agreement in SLA literature that selective attention to some input is necessary for language development.

In more recent discussions of attention and awareness in SLA, attention and awareness are treated as different phenomena. For example, Williams (2013, p. 51) highlights that “attention and awareness can be dissociated.” Based on empirical evidence available, both Leow (2013) and Williams (2013) note that awareness may or may not be necessary for second language acquisition, however, attention is a necessary condition.

As discussed in Section 2.2, in cognitive psychology, awareness is associated with WM activation and/or reporting of stimuli attended (Lamme, 2003). This concept of awareness is elaborated by Allport (1988) who suggests three possible measures to identify whether awareness is involved in learning: “a show of some behavioural or cognitive change due to the experience, a report of being aware of the experience and a description of this subjective experience” (p.51). Robinson’s (1995) discussion of awareness in SLA aligns with this description since he believes that it involves reporting the processes involved in a task. In addition, Leow (2013, p.3) highlights three ways that awareness can be identified: “some resulting behavioral or cognitive change, a meta-report of the experience but without any metalinguistic description of a targeted underlying rule, or a metalinguistic description of a targeted underlying rule.” Schmidt (2010), however, argues that “what we are aware of is what we attend to, and what we attend to determines what enters phenomenal consciousness” (p. 725). This argumentation suggests that attention is inseparable from awareness and consciousness, a position also advocated in Schmidt’s (1990) Noticing Hypothesis.

2.3.1 Noticing hypothesis

Schmidt's (1990) Noticing Hypothesis (NH) identifies three stages of consciousness i.e. consciousness as awareness, consciousness as intention and consciousness as knowledge, and then describes three levels of consciousness as awareness: perception, noticing and understanding. Green and Hicks (1984) and Chaudron (1985) also identify perception as the primary attentional process. Schmidt categorises perception under awareness and some empirical evidence for this claim is provided by Curran and Keele (1993), Marcel (1983) and Nissen and Bullemer (1987). In contrast, perception is recognised as an unconscious process in the late selection models of attention by Broadbent (1958) and Treisman (Kahneman & Treisman, 1984). Although Schmidt categorises perception under awareness, he contradicts his own categorisation by arguing that "perceptions are not necessarily conscious" (1990, p. 132). Thus, it is possible to argue that perception is an attentional process that does not involve awareness.

Noticing, according to Schmidt (1995), is "focal awareness" (p. 132) which is necessary for language learning. He further notes that noticing is "nearly 'isomorphic' with attention" (Schmidt, 1995, p. 1) and "[what is noticed is] consciously registered" (2010, p. 2). These interpretations show that Schmidt considers noticing as an attentional process that involves both consciousness and awareness. Schmidt (1990) also suggests that noticing is a "private experience, although noticing can be operationally defined as availability for verbal report, subject to certain conditions" (p. 132). According to definitions in cognitive psychology cited in Section 2.2, awareness involves the possibility of offering verbal reports, and therefore what Schmidt calls ‘noticing’ can be equated with ‘awareness’. Schmidt (1990, 1995, 2010) distinguishes between two different levels of awareness: awareness at the level of noticing and awareness at the level of understanding. Noticing can be considered as a surface level phenomenon, “the conscious registration of the occurrence of certain events” (Schmidt, 1995, p. 19), whereas understanding represents a deeper level of processing involving pattern and rule recognition. Therefore, understanding can be characterized by a learner’s ability to analyse, compare and test hypotheses of L2 data.

Although noticing involves awareness, according to Schmidt (1990, 1995, 2010), it is a lower level of awareness. He distinguishes between noticing and understanding and defines understanding as a mechanism that involves a higher level of awareness and a deeper level of input processing. In his view it also involves “recognition of general principle, rule or pattern and is related to the organization of material in LTM, to restructuring, and to system learning” (Schmidt, 1993, p. 213). For example, if learners notices verb final ‘s’ in an input text and is able to express that they have noticed some verbs in the text have a final ‘s’, that can fall under noticing. If the learners are also able to mention that the verbs in the present simple tense need a final ‘s’ when the subject is singular, then the awareness has reached the stage of understanding that goes beyond noticing (Schmidt, 2010). Schmidt also points out that understanding involves generalization of instances, metalinguistic knowledge and hypothesis formation and testing. To sum up, Schmidt (2010) believes that “noticing [is] a technical term limited to the conscious registration of attended specific instances of language, and understanding, a higher level of awareness that includes generalizations across instances” (p. 5).

Truscott (1998) and Truscott and Sharwood Smith (2011) call attention to several conceptual problems with the NH. Truscott and Sharwood Smith (2011) argue that noticing, as Schmidt defines it, is an intermediate level of noticing, because it does not involve rules, generalizations about the language, or form-meaning mapping. They also point out that NH does not provide a clear explanation of the level of awareness involved in noticing and how noticing is differentiated from either no awareness or a higher level of awareness. Truscott (1998) proposed a weaker formulation of the NH by arguing that conscious detection of grammatical constructions and underlying rules is necessary for the development of metalinguistic awareness, but not for the enhancement of grammatical competence, which he sees as a primarily implicit process.

2.3.2 Other theoretical discussions on attention, awareness and consciousness in L2 learning

In contrast with Schmidt's (1990, 1995, 2010) NH, Tomlin and Villa's (1994) model distinguishes attention from awareness. Based on Posner and Rothbart's (1992) early work, Tomlin and Villa divide attentional processes into three levels: alertness, orientation and detection. They believe that no awareness is involved in any of these processes. Nevertheless, detection, which is defined as "cognitive registration of the stimuli" (p. 190), and seen as necessary for language learning, is the closest attentional process to awareness. Thus, they suggest that learning is possible without awareness at the level of detection. What is meant by 'cognitive registration of the stimuli' remains unclear in their definition. Furthermore, their conceptualization of detection seems to be contradictory to those in cognitive psychology, where it is argued that awareness is needed for stimuli to be registered in the WM (Lamme, 2003).

In another model, Robinson (1995) defines noticing as "detection plus rehearsal in short-term memory, prior to encoding in long-term memory" (p. 269). Robinson applies a similar meaning to the term 'detection' as Tomlin and Villa (1994) and clearly distinguishes it from awareness. Detection occurs prior to noticing, and noticing is the next level of processing, which involves both detection and rehearsal. Robinson assumes that neither the detection nor the noticing stages involve awareness, and that awareness is not a necessary condition for language learning. Robinson's conceptualization of noticing is not without problems either. In his definition, he refers to STM as a passive temporary memory store (as in Hutton & Towse, 2001, p. 383) that is not responsible for processing information (as in Baddeley & Hitch, 1974). If we assume that additional cognitive processing is needed for noticing, the locus of this should be in the WM. According to Koch and Tsuchiya's (2006) and Lamme's (2003) models of attentional processing, however, for a stimulus to enter WM, awareness is necessary. Therefore, it is questionable whether detection and rehearsal without awareness could take place in WM.

In a more recent discussion, Godfroid, Boers and Housen (2013) separate awareness and attention and view them as two sides of the same coin. They do not disregard the possibility of awareness having a role to play in noticing; however, their argument is that a certain level of attention is sufficient for registering new stimuli in the LTM. This is a similar view to that of Tomlin and Villa (1994) and Robinson (1995) who argue that awareness is not a necessary condition for language learning. In line with Robinson

(1995), Godfroid et al. (2013) define noticing as “a cognitive process in which the amount of attention paid to a new language element in the input exceeds a critical threshold, which causes the language element to enter WM and become the object for further processing” (p. 493). Nevertheless, the question remains how to determine the threshold when a stimulus enters WM. Furthermore, as already pointed out above, if a stimulus is processed further in WM, it requires awareness. Following from this reasoning, noticing needs to involve some degree of awareness.

Truscott (2015), who discusses attention and its function within his MOGUL (Modular Online Growth and Use of Language) framework, also considers attention as an important process. According to this framework, the mind is comprised of certain independent units and each of them has a processor and a storage system. For example, syntax is one such unit and the syntactic processor manipulates all information on syntax while the syntactic store holds this information. The information stored here is identified as representations. Processing is “activating existing representations and combining them into new ones” (p. 83). Accordingly, attention is a phenomenon with the capacity to activate certain representations more extensively than others in this process. This seems to signify the role played by selective attention. A representation becomes conscious “if and only if its current activation level is above a given threshold” (p. 98) and if a representation becomes “exceptionally active” (p. 98) it may become an object of awareness. In language processing, according to MOGUL, awareness of input means awareness of representation of a certain language construction (e.g., a syntactic structure).

Truscott who offers an example of how language learning fits into this framework highlights that when a sentence (with a particular structure such as simple past tense) is heard by someone and if it does not become highly active, it would not reach the threshold of consciousness and thus the person would be unaware of it. As a result, it would not be further processed and the person might not be able to repeat it. However, if the same sentence becomes active in the listener's mind, it will reach the threshold of consciousness and thus the person will become aware of it. They may be able repeat it as well. This is considered as simple awareness. If it gets further processed, the person may identify the structure and this creates a new representation. This level of awareness is identified as the "awareness of the presence of a particular linguistic feature" (p. 163). According to this model, attention, consciousness and awareness are different phenomena; however attention is necessary to activate consciousness and consciousness is necessary to activate awareness. Although this model attempts to provide a comprehensive description of attentional processing, it is not able to solve the existing conceptual issues relating to attention, noticing and awareness in SLA. For example, it postulates a threshold of attention for a representation to reach consciousness (as in Godfroid et al.'s, 2013 argument) and also assumes that a representation should be "exceptionally active" (p. 98) if it is to reach the level of awareness. It is unclear what this threshold is and how to distinguish between the threshold of consciousness and the point that a representation reaches awareness. Furthermore, there is no explanation on how to measure *extreme activeness*.

2.3.3 Empirical evidence on attention, awareness and consciousness in SLA

Empirical studies in SLA aiming to analyse the effect of attention, awareness and noticing do not seem to clearly differentiate between these processes. In particular, the methodology used to collect data on attention, awareness and noticing raises concerns on whether the methods have actually measured what they are supposed to have measured. These empirical studies can be categorised into three main groups: studies that measured attention, studies that measured awareness and studies that measured noticing. Since it is not clear what exactly has been measured in most of these studies, the categorisation is solely done based on the use of the terminology in these studies.

The studies that claim to have assessed attention have used different methods to collect data. For example, two early studies aimed at analysing attention to input used audiotapes of actual classroom settings as data (Williams, 1999, 2001). The instances in which learners asked about a language feature or addressed a question about a feature to the teacher and the instances when other learners talked about a language feature were taken as occasions when attentional processing took place in these studies. Williams points out that there is a relationship between the forms that emerged during classroom episodes as a result of learners paying attention to them and learners using those forms subsequently. These two studies have not necessarily taken into consideration the complex processes of attention and thus what is measured as attention may involve attention, awareness, consciousness and even the processes in LTM. Gass, Svetics and Lemlin (2003) who studied focused and non-focused attention to input asked the experimental group (focused attention) to pay attention to and underline the language features in the input. The underlined feature was taken as evidence for attentional processing. Through the use of this method it is also difficult to determine the amount of attention paid and whether the attention paid by learners involves awareness.

N. Ellis and Saggara (2010, 2011) and R. Ellis and Shintani (2014) also attempted to analyse attention; however, they did not use a specific research tool to collect data on attention per se. Instead, they compared the performance of the participants in different learning conditions and concluded that the participants had paid more attention in the conditions on which they demonstrated some development of the target feature. Godfroid et al. (2013) used eye fixation duration in eye-tracking to measure how much attention the participants paid to TL features. Accordingly, they argued that more time spent on target features demonstrates more attention. Although they do not disregard the possible involvement of awareness in learning, they argue that a certain level of attention is sufficient for learning to take place. However, they do not provide an explanation of how they measured whether awareness was involved or not in attention data. Although eye-tracking has its own shortcomings as a method of collecting data on attention, as I will argue in Section 2.4, the use of eye-tracking, as a recent research tool, seems to be currently the best available method to measure the amount of attention paid to input.

For the assessment of awareness, Leow (1997, 2000), Rosa and O'Neill (1999) and Rosa and Leow (2004) applied concurrent verbal reports. The think-aloud process provides participants with the opportunity to offer a conscious report on the material that has entered their attentional focus, and hence it can serve as a valid tool for gaining insights into awareness. Nevertheless, asking participants to think-aloud can change or affect the cognitive processes that are implicated in task performance and as a result the reports may be inaccurate (Russo, Johnson, & Stephens, 1989; Leow & Morgan-Short, 2004; Yoshida, 2008).

The third set of empirical studies has investigated the effect of noticing in SLA. These studies have used several different methods to measure noticing. They are: think-aloud (Leow, 2001 and Radwan, 2005), stimulated recall (Mackey, 2006; Park, 2007, 2010, Bao, Egi & Han, 2011 and Uggen, 2012), immediate retrospective reports (Egi, 2004), post-exposure questionnaire (Radwan, 2005; Mackey, 2006 and Park, 2010, 2013), note taking (Izumi, 2002; Mackey, 2006 and Iwanaka & Takatsuka, 2007), underlining (Izumi & Bigelow, 2000; Park, 2007, 2010 and Uggen, 2012), and eye-tracking (Smith, 2012; Godfroid et al., 2013; Godfroid & Schmidtke, 2013, Godfroid & Uggen, 2013 and Winke, 2013). Both Leow (2001) and Radwan (2005) have used think-aloud to measure noticing in enhanced and unenhanced input conditions. Leow equates awareness to noticing in this study and thus verbal reports of noticing were considered as measures of noticing. Radwan who has attempted to measure awareness at the level of noticing and understanding used verbal reports in think-aloud as evidence of noticing and the participants' ability to verbalise the structure of the target feature as understanding. The studies that have used stimulated recall (Mackey, 2006; Park, 2007, 2010, Bao, Egi & Han, 2011 and Uggen, 2012) asked the participants to comment on what they had been doing/thinking while engaging in the task and reports that indicated awareness/noticing were taken as instances of noticing. Egi (2004) used both stimulated recall and immediate retrospective reports (participants provided reports of what they had attended to immediately after the exposure) and argued that the latter has the ability to capture instances of noticing clearly without possible memory decay. Post-exposure questionnaires have also been used to measure noticing. In Radwan's (2005) Mackey's (2006) and Park's (2010, 2013) studies, the participants were directly asked whether they had noticed/paid attention to any interesting features in the input. All these data collection methods have used verbal reports of learners of what they have paid attention to. As discussed earlier, ability to offer a report is a characteristic of being aware and thus it seems that these studies have measured awareness as noticing.

Underlining has been considered a method that would give minimum extra processing load. Izumi and Bigelow (2000) and Uggen (2012) asked their participants to underline the words/phrases that are necessary for subsequent production, while Park (2007, 2010) asked her participants to underline what they notice. The number of occasions that a feature was underlined and/or the percentage of underlining the target features were taken as measurements of noticing. This, however, does not seem to be an accurate measurement of the amount of attention paid. It cannot distinguish between attention and awareness either.

Iwanaka and Takatsuka's (2007) participants had to take notes on what they noticed while reading target sentences and the participants in Izumi's (2002) study were asked to take notes on necessary language features for a subsequent production. Mackey (2006) asked the participants to record the language features that they noticed in the form of learning journals. Thus, both underlining and note taking in these studies have been used in a similar pattern and what they have measured is again questionable. Note taking on a particular language feature and using the notes in subsequent production may involve attention and awareness as well as the functions of both STM and LTM.

Thus it is apparent that noticing, which involves awareness, has been measured by these methods as well.

As highlighted in this discussion it is still difficult to determine what attention, awareness and noticing are, and more importantly still there is a lack of appropriate methodology to investigate attention and awareness. This is mainly due to the close relationship between the two, the difficulty in distinguishing between them and the difficulty of measuring cognitive processes. This discussion has also highlighted the complexity of defining the term ‘noticing’ as some of the definitions and empirical studies have used it as a phenomenon that involves awareness and some have separated it from awareness. Further, the empirical studies discussed in this section highlighted that language learning may need both attention and awareness.

Due to the fact that there is no clear definition or empirical evidence available on what ‘noticing’ is and also it is still not possible to determine if it involves awareness or not, this study avoids using the term. My research attempts to investigate how attention and awareness are involved in different input conditions considering them as two distinct phenomena. Since cognitive psychology and most SLA studies agree that awareness involves the possibility of offering a verbal report of what is attended, this study also considered verbal reports as a measurement of awareness. As discussed, eye-tracking has some shortcomings such as the difficulty of establishing if one is paying attention to something when looking at a particular point on the computer screen. However, this method can provide more accurate quantitative data on the amount of attention paid than other methods such as underlining. It also does not give an additional processing load as in methods such as think-aloud. Thus, this study used eye-tracking to measure attention and the next section offers a justification for selecting this research tool.

2.4 Eye-tracking as a method of measuring attention[5]

The basic assumption underlying eye-tracking studies is that eye-movements are to some extent controlled and guided by the attentional system and that eye-fixation durations are indicative of ongoing cognitive processing (Liversedge, Gilchrist & Everling, 2011; Rayner & Pollatsek, 1989). According to Roberts and Siyanova-Chanturia (2013), there are different types of eye-tracking data that can be used in empirical studies. For example, *first fixation duration* i.e. the time that one spends on a particular word on the first occasion that the word is encountered is used in word-recognition research. *Gaze duration* is the time that one gazes at an Area of Interest (AOI) for the first time and *total reading time* is the total amount of time that one gazes at an AOI during reading. Apart from these basic measurements, it is also possible to use *regression path duration, rereading duration* and *second pass reading time* to measure regressions in reading and rereading. Lastly, *fixation count* measures the number of times that one visits at a particular word or an AOI.

The measures used to gain an insight into attentional processing vary in the field of cognitive psychology (Reichle, Pollatsek & Rayner, 2006). In SLA research, *total reading time* has been utilised as one of the main measures of attention, and it has been named in different ways in different studies: 'total time' (Godfroid et al., 2013; Godfroid & Uggen, 2013; Issa, Morgan-Short, Villegas, & Raney, 2015), 'total fixation

[5] Several parts of this section of the literature review were published in Indrarathne & Kormos (2016).

duration' (TFD) (Winke, Gass, & Sydorenko, 2013), 'eye fixation duration' (Smith, 2010, 2012), 'total fixation time' (Winke, 2013) and 'dwell time' (Nezami, 2012). [In this thesis, I use the term TFD to refer to total reading time since it is the commonly used one]. Apart from TFD, Smith (2012), Smith and Renaud (2013) and Winke (2013) have also used *total number of eye fixations* i.e. fixation count/number of visits as a measurement of attention/noticing, while Smith (2012) and Winke and Lim (2015) have also applied heat maps to explain noticing patterns.

What has been measured through TFD in SLA research also varies because of the different terminology used in those studies. For example, Smith (2012), Godfroid et al. (2013), Godfroid and Schmidtke (2013), Godfroid and Uggen (2013) have used TFD as a measurement of noticing while Godfroid and Uggen (2013) used it as a measurement of attention. In Godfroid et al.'s (2013) study they seemed to have assumed that awareness was not involved in the eye-tracking measurements that they recorded. However, they do not discuss how they distinguished between attention and awareness within eye-tracking data. A study by Godfroid and Schmidtke (2013) used both eye fixation duration and verbal reports and its results suggest that "attention included awareness" (p. 183) indicating that eye-tracking data can be a measurement of attention that includes awareness. Thus, it is still unclear if eye-tracking data can be treated as a measurement of attention only or attention that involves awareness.

Eye-tracking has been considered a better method of measuring attentional processing since, unlike think-aloud, it does not provide an extra processing load. However, it is not free of potential problems either. It is important to note that not all eye-movements are under attentional control and one can attend to a specific visual stimulus without necessarily fixating on it (Hunt & Kingstone, 2003; Juan, Shorter-Jacobi, Schall, & Sperling, 2004). Moreover, other processes such as awareness may be involved while one pays attention to items on a computer screen, but eye-tracking cannot distinguish between those processes involved in attentional processing. While fixation duration is strongly associated with the locus of visual attention in reading, it is also influenced by the lexical characteristics of the text such as word frequency, length and predictability (Schotter, Angele & Rayner, 2012). TFD and total number of eye fixations on a target word/phrase might tell us for how long and how many times the participants viewed the target item, but it might not accurately reflect attentional processing (Reichle, Pollatsek & Rayner, 2006). If the definition of attention is revisited, which includes the selective processing of incoming stimuli (see e.g., McFarland, 2006), it becomes apparent that the existing eye-tracking measures do not provide an indication of whether one has selected the target word/phrase from the many words/phrases in the input available to be attended.

In order to overcome this problem, Smith (2010) considered TFD beyond 500 milliseconds (ms) as an indication of noticing (attention); however, how this threshold has been derived is not clear. Some other eye-tracking studies used control and experimental conditions and treated the control condition as the base line in measuring attention. In other words, the difference of TFD between the experimental condition/s and the control condition on the target items has been considered as noticing/attention. For example, Winke (2013) used an enhanced and an unenhanced input condition in order to investigate both learning gains and noticing patterns of the English passive forms. She used four eye-tracking measurements to analyse noticing: total fixation time, number of visits, first pass reading time and re-reading time. The results were reported as a comparison between the two input conditions. Comparative measurement is certainly a more accurate procedure for establishing extra attentional processing load than the use of total fixation time and the total number of eye-fixations alone.

In Godfroid et al.'s (2013) and Godfroid and Schmidtke's (2013) studies, the threshold for noticing is established by comparing processing times in control and experimental conditions. Noticing is assumed to have taken place if participants spend longer time processing stimuli in the experimental than in the control condition. This rests on the assumption that the control and the experimental conditions only differ from each other with regard to the presence or absence of noticing, and participants otherwise engage in comparable cognitive operations. In Godfroid et al.'s (2013) study the control group read texts that contained existing words, whereas the texts in the experimental groups contained pseudo words. Cross-group comparisons of pseudo- and real-word reading, however, might also reflect differences resulting from the different characteristics of the groups as well as the different nature of reading involved in reading existing vs. non-existing words (Coltheart, Curtis, Atkins & Haller, 1993). Therefore, processing time differences in the two conditions might not only result from different attentional processing mechanisms but also from different cognitive operations. The use of pseudo-words is also criticized for floor effects, i.e. the difficulty of creating pseudo-words to match the decoding ability of learners (in particular pseudo-words with more than one syllable) (Rathvon, 2004).

In both Godfroid et al.'s (2013) and Godfroid and Schmidtke's (2013) studies, the researchers assumed that all words except pseudo-words were known to learners, based on a previous analysis done with a different group of learners at the same level of language proficiency. However, this previous analysis revealed that the learners' knowledge of all words was 'fairly good' (Godfroid et al., 2013, p. 495), which does not rule out the possibility that there could have been participants in the control condition in the main study who did not know the target words. Moreover, it is also possible that the participants in the preceding analysis, who were asked to highlight the individual words, the meaning of which they were unsure of, may have guessed the meaning of words from the surrounding text and as a result might have failed to mention that they did not know the meaning.

In another study, Godfroid and Uggen (2013), who analysed if learners of German pay attention to the stem vowel changes of verbs as opposed to the regular form of those verbs, used a slightly different eye-tracking measurement of attention. They subtracted the fixation time on the regular verb from the fixation time on the irregular verb (which appeared on the same screen) and took the difference as the amount of attention paid to the irregular verbs. This seems to be a more accurate measurement than the TFD to measure attention since it provides data on the extra amount of time spent by the participants on the target items. Although comparative measurements are a more accurate way to measure attention to input, the studies reviewed so far have ignored the fact that the control group (usually also unenhanced) may also pay attention to the target items (or other items that that they select), and their attentional processing is ignored by using their data as the baseline.

Considering the shortcomings of the eye-tracking measures that have been previously applied in SLA studies, this study used two eye-tracking measurements to investigate the amount of attention paid to the target items: TFD on Areas of Interest (AOIs) and the difference between observed TFD and expected TFD on AOIs (ΔOE). As described above, the first of these measures, TFD, has been widely applied in previous studies. The second measure, which has not been used in SLA studies before, was developed to take into consideration participants' reading speed in determining extra attentional processing load. In other words, the expected TFD was the average reading time that the participants may spend on each word/phrase without selecting particular words/phrases to pay attention to. Observed time was the actual time that the participants spent on target constructions. If a participant spent more time on a word/phrase than the expected time (difference between the observed fixation duration and expected fixation duration), it was considered as a measure of additional attentional processing. Although the proposed measurement is new to SLA, a similar measurement has been used in some psychological studies (Ferreira & Clifton, 1986; Wilson & Garnsey 2009; Trueswell, Tanenhaus & Garnsey, 1994) (for more detail on the actual calculation of this measure see Section 3.7.1).

I have so far discussed three key concepts related to L2 input processing: attention, awareness and consciousness and how they are measured in SLA research. It is worth noting that these studies have used different input conditions in order to investigate how attention, awareness and consciousness influence L2 acquisition in these different conditions. The present study also included several explicit and implicit input conditions. Therefore, the next section of this literature review will look into the theoretical background of explicit and implicit input techniques in SLA.

2.5 The effect of explicit and implicit input techniques on L2 learning outcomes[6]

As mentioned before, the study presented in this thesis primarily investigated the role of attention in L2 input processing. In order to understand how learners attend to input in different input conditions, the study used four different input conditions. These were: unenhanced input, enhanced input, specific instructions to pay attention to input and explicit explanation of the target construction. While the two former conditions can be categorised as implicit learning conditions, the latter two can be identified as explicit conditions. Thus, it is necessary to review the concepts of explicit vs implicit learning, explicit vs. implicit knowledge and explicit vs. implicit instruction along with other related concepts such as intentional, incidental and statistical learning. This section of the literature review discusses these key constructs and reviews empirical research related to implicit and explicit learning conditions.

Consciousness in learning is associated with the explicit and implicit nature of learning. According to N. Ellis (2011), the knowledge gained through exposure cannot be described i.e. one would not be able to explain the underlying structures of an acquired language. He identifies this knowledge as implicit knowledge and the process of its acquisition as implicit learning. According to R. Ellis (2009), this type of implicit learning does not make demands on attentional resources. Moreover, DeKeyser (1995) notes that implicit learning does not involve awareness of what is being learnt. In summary, as Reber (1989) defines, implicit learning is "the process by which knowledge about the rule governed complexities of the stimulus environment is acquired independently of conscious attempts to do so" (p. 219). In contrast, explicit

[6] Several parts of this section of the literature review were published in Indrarathne & Kormos (2016).

learning involves learning the rules and vocabulary of the TL and the knowledge gained through this type of learning is said to be conscious (N. Ellis, 2011). According to Rebuschat and Williams (2012a), explicit learning requires learners to look for patterns. R. Ellis (2009) points out that explicit learning makes heavy demands on WM because it involves the memorisation of patterns. The knowledge gained through this process is considered explicit knowledge and learners are aware of what they know (Williams, 2009). In other words, they are able to explain the underlying structures of what they have learnt.

The distinction between implicit and explicit learning and knowledge based on *awareness* and *consciousness* can be discussed in relation to how they have been interpreted in cognitive psychology (Section 2.2). Lamme's (2003) model shows how awareness involves the ability to offer a verbal report of the attended stimuli, and Koch and Tsuchiya's (2006) model emphasizes that a verbal report is available only for consciously attended stimuli. Thus, learners who go through an implicit learning process may not be able to offer a verbal report on what is being learnt because implicit learning processes and implicit knowledge are said to be unconscious. In contrast, those who have undergone explicit learning may be able to offer a verbal report of the learning process and the explicit knowledge that they have gained because both are considered conscious. In other words, unlike implicit learning and knowledge, explicit learning and knowledge involve awareness.

A distinction between implicit and explicit instruction has also been made by R. Ellis (2009), who identifies implicit instruction as "directed at enabling learners to infer rules without awareness" (p. 16). According to R. Ellis, this type of instruction is provided to learners with the aim of giving them examples of rules or patterns. However, learners usually make fewer attempts to learn them and as a result, the focus of the input is directed to meaning. In contrast, in explicit instruction, "learners are encouraged to develop metalinguistic awareness of the rule" either inductively or deductively (R. Ellis, 2009, p. 17). VanPatten (2011) defines both explicit learning and teaching as "any attempts by either learners or teachers to manipulate learning from the outside" (p.10).

It is still debated in the SLA literature whether explicit learning only leads to explicit knowledge and implicit learning only results in implicit knowledge. MacWhinney (1997) points out that the outcome of explicit instruction might not be explicit knowledge or implicit instruction does not necessarily result in implicit knowledge. He further notes that there is a possibility that a learner could identify a target rule through implicit instruction and learners might not learn the target rule through explicit instruction. This argument can be extended to implicit and explicit learning too. For example, a learner who goes through an explicit learning process may not necessarily be able to provide a verbal report of awareness, but may demonstrate the ability to use the target construction. Thus, the knowledge gained could be implicit. VanPatten (2011) provides an example which demonstrates the difficulty of distinguishing between explicit/implicit learning mechanisms: if someone wants to learn a computer programme, they would use a manual to try out applications and after several hours of attempts, may be able to use some of the functions in the programme. If regular attempts are made, the performance improves. VanPatten argues that this process involves features of both explicit and implicit learning mechanisms. An external manipulation in the form of reading the manual with an intention to learn is involved in this process, thus it is explicit learning, even though, the person may not be able to provide a verbal report of the underlying mechanisms in the programme. This, however, does not rule out the possibility of using the programme accurately. R. Ellis' (2009) argument that explicit knowledge can become implicit knowledge when proficiency increases may also shed light on the difficulty of distinguishing between the knowledge that one could gain through explicit/implicit learning.

The effect of implicit learning in adult L2 acquisition seems to be limited, and explicit learning associated with consciousness is treated as a necessary condition to achieve accuracy in the L2 (N. Ellis, 2011). Extending this argument further, N. Ellis (2015) notes that conscious processing of input results in acquiring "novel explicit cross-modal form-meaning association" (p. 3) which helps to establish processing routines. These routines are then available to be used in subsequent implicit learning processes. Hulstijn (2015) emphasises that explicit knowledge including metalinguistic knowledge is necessary in acquiring 'higher' level lexis and grammar. Accordingly, achieving a higher level of proficiency can take an extremely long period and might require a large amount of input if metalinguistic information is absent in the input. Hulstijn (2015) further argues that both explicit and implicit knowledge can be acquired when explicit learning takes place. VanPatten and Rothman (2015) however believe that learners do not learn surface rules in the input, but they implicitly learn the morpho-phonological forms of meaning and functions in the input.

Incidental and intentional learning are two other terms used alongside explicit and implicit learning. Hulstijn (2003) identifies intentional learning as memorising vocabulary and grammar rules deliberately. In contrast, incidental learning involves "picking up" (p. 349) vocabulary and grammar through meaning-based communicative tasks. Rebuschat and Williams (2012a) identify explicit learning as intentional learning and implicit learning as incidental learning. Williams (2009) believes that the terms implicit/incidental and explicit/intentional only have a methodological difference, but in fact they mean the same. Williams assumes that the term incidental learning refers to experimental conditions in which learners are not informed of what they are learning. This can also be applied outside laboratory contexts; learners can learn language structures without having an intention to learn them. In Williams' definition, explicit learning involves intention to learn.

One other term associated with implicit learning is statistical learning. Rebuschat and Williams (2012b) define it as "our ability to make use of statistical information in the environment to bootstrap language acquisition" (p. 1). Statistical learning is assumed to be involved in both child and adult language learning through the manipulation of statistical information available in the environment in any form (e.g., auditory, visual etc.). In other words, a significant amount of statistical information is available in the input and learners have the possibility to process it (McMurray & Hollich, 2009). Misyak and Christiansen (2012) claim that both statistical learning and implicit learning involve sequential learning and thus they name it statistical-sequential learning. Perruchet and Pacton (2006) argue that both statistical learning and implicit learning share the same process and they advocate the use of the term 'implicit statistical learning' coined by Conway and Christiansen (2006). Statistical learning may also be influenced by individual differences such as WM capacity (Misyak & Christiansen, 2012).

2.5.1 Empirical studies that have investigated the L2 learning outcomes in explicit and implicit input techniques

Although it is still an unresolved question whether explicit knowledge can turn into implicit knowledge, and if yes, how (R. Ellis & Shintani, 2014), there seems to be evidence for the beneficial effects of explicit instruction in L2 learning. In Norris and Ortega's (2000), Spada and Tomita's (2010) and Goo, Granena, Yilmaz and Novella's (2015) meta-analyses of previous studies, explicit instruction was found to result in substantially higher improvement in various post-test scores than implicit instruction, in which no explicit explanation of the grammatical construction or instruction to pay attention to specific constructions in the input was provided. In line with the findings of these meta-analyses, studies where implicit and explicit Focus on Form (FonF) techniques were jointly investigated also show the advantage of explicit instruction (e.g., Robinson, 1997; Radwan, 2005; Rosa & O'Neill, 1999; Tode, 2007).

One notable issue in analysing explicit and implicit learning mechanisms in SLA research is the type of measurements used to distinguish them. Rebuschat (2013) discusses three such measurements: retrospective verbal reports, direct and indirect tests and subjective measures. The empirical studies discussed in this section have used one or more of these methods. Bowles (2011) used a battery of tests proposed by R. Ellis (2005) to measure implicit and explicit knowledge in a study that compared the performance of adult L2 learners of Spanish. Three groups: Spanish L1 speakers, Spanish L2 learners and Spanish Heritage Language (HL) learners (who were Spanish English bilinguals, but used English as the dominant language and had little exposure to formal instructions of Spanish) took five tests, namely: oral imitation, oral narration, timed grammaticality judgement (GJ), untimed GJ and metalinguistic knowledge test. The first three tests were considered to measure implicit knowledge due to the time pressure under which the learners had to solve the tasks and the intuitive knowledge of language that they demanded, and the rest were supposed to measure explicit knowledge. The results indicate that the L1 group outperformed both L2 and HL groups in all tests; however, the L2 group scored nearly the same as the L1 group in the metalinguistic knowledge test. The HL group performed better than the L2 group in all tests except in the metalinguistic knowledge test. Based on the results, Bowles validates the use of these tests to measure explicit and implicit knowledge. This argument is based on the initial assumption that the L1 speakers had implicit knowledge and the L2 learners had explicit knowledge of the 17 target grammatical structures used in the study. Since the HL group had received very little formal instruction, they were also assumed to have relied on their implicit knowledge. However, the results indicate that the L1 group performed better than the L2 group in the metalinguistic knowledge test and untimed GJ test, which were supposed to measure explicit knowledge. Therefore, it is difficult to conclude that the L1 group only had implicit knowledge if the tests had measured what they were supposed to have measured.

In another study, Rebuschat and Williams (2012a) conducted two experiments based on a semi artificial language. In the first one, the experimental group was exposed to a training task in an incidental learning condition. In the training phase, they had to listen and judge the semantic plausibility of the sentences they heard, which consisted of word order patterns of the semi artificial language. They were not informed about the target construction. Then they were assessed on a timed GJ test. The control group only took the GJ test. The results reveal that the scores of the control and experimental groups were not significantly different, which suggests that incidental exposure did not affect the performance of the experimental group. However, based on the results of the confidence judgement of the GJ test items, Rebuschat and Williams report that the 'aware' group outperformed the 'unaware' group. From this they conclude that in the incidental learning condition learners gained conscious knowledge. In the second experiment, the experimental group judged the plausibility of items and participated in an elicited imitation task (they were asked to repeat the sentences that they heard). The results revealed that they performed significantly better than the control group. Thus, Rebuschat and Williams claim that language development is possible through incidental learning. Although the experimental group outperformed the control group in the second experiment, they received more input in the second experiment. This highlights that more exposure may be necessary for incidental learning to take place.

In another study Morgan-Short, Sanz, Steinhauer, and Ullman (2010) exposed adults to noun-article and noun-adjective gender agreement of an artificial language in explicit (metalinguistic explanations with meaning) and implicit conditions (meaning only). The participants judged agreement violations in an event-related potential assessment task, which included agreement violation items and accurate items. The event related brain potentials were measured in this task. The study revealed that the performance of groups did not differ from each other. Batterink and Neville (2013) also compared the performance of L2 learners of French who were grouped into implicit (meaning based reading of stories) and explicit (grammar explanation before meaning based reading of stories) learning conditions. In the comprehension question (CQ) task, both groups performed equally well. In the GJ task, the implicit group showed a lower level of performance compared to the explicit group.

Guo et al. (2011) in their study of acquisition of L2 semantic prosody used two experimental conditions and a control condition. In the incidental condition, the participants were presented with sentences and asked to repeat them aloud, thinking of the meaning. In the intentional condition, the same sentences were presented and the participants had to work out the rule in the underlined section of the sentence. All three groups then took a post-test, which showed that both explicit and implicit knowledge development took place in the incidental condition; however, explicit knowledge gains in the intentional condition were higher than in the incidental condition.

In a recent study by Rogers, Revesz and Rebuschat (2015), the participants were exposed to case marking in artificial grammar in incidental conditions. The findings highlight that participants were able to acquire the target structure in this condition although the learning gains were small. In another study, Kachinske, Osthus, Solovyera and Long (2015) used three input conditions: incidental, intentional and control to provide input on word order of artificial grammar. Participants in both incidental and intentional conditions demonstrated significant learning gains in a GJ task; however, only the participants in the intentional condition were able to learn more complex word order rules.

2.6 Input Processing[7]

How learners process input may depend not only on the implicit/explicit nature of the input, but also on the other factors such as how input is presented, salience of the target

[7] Several parts of this section of the literature review were published in Indrarathne & Kormos (2016).

construction and the nature of the form-meaning relationship of the target construction. Therefore, theoretical concepts related to input processing are discussed in this section.

2.6.1 Focus on Form (FonF)

Form Focused Instruction (FFI) and Meaning Focused Instruction (MFI) are two approaches that have been used in second language teaching. In the former, the focus is on the formal features of grammatical constructions and in the latter, it is primarily on the meaning conveyed by grammatical constructions (Nassaji & Fotos, 2007). There has been a debate over the best approach. Swain (1985) claims that engaging in communicative tasks is not sufficient for language learning and emphasizes the necessity of drawing learners' attention to form. Norris and Ortega's (2000), Spada and Tomita's (2010) and Goo et al.'s (2015) meta-analyses also indicate that FFI could bring more positive results than MFI.

Long's (1991) distinction between FonF and Focus on FormS (FonFS) has been the most widely discussed approach to FFI. Long (1991) identifies FonFS as the traditional type of grammar teaching in which grammatical forms of the TL are explicitly taught based on a pre-designed linguistic syllabus. Moreover, in the FonFS approach, the classroom activities are not focused on the meaning that they convey, but on the forms of language taught. In the Focus on Meaning approach, the whole focus of classroom activities is on meaning and the language forms that are used to convey this meaning are paid less attention to (Burgess & Etherington, 2002). Considering the shortcomings of these two approaches, Long (1991) proposes FonF as a more effective approach in which learners' attention is drawn to grammatical features in meaning focused input. In other words, learners are exposed to comprehensible input in meaning based communicative tasks and their attention is drawn to linguistic elements in the input. Thus, the focus is on both meaning and form. Long and Robinson (1998) also highlight that in FonF, there is an occasional shift of attention from meaning to form initiated either by the teacher or learner/s. Basturkmen, Loewen and Ellis (2004) are of the view that this shift provides a time-out for learners to notice language forms in the input. Doughty (2001) suggests that the most important feature of FonF approach is that it allows learners to simultaneously attend to "form, meaning and use during one cognitive event" (p. 211). Doughty and Williams (1998) further state that learners are at an advantage in the FonF approach since meaning provides additional support to understand form.

R. Ellis, Basturkmen and Loewen (2004) identify two types of FonF activities: planned and incidental. In planned FonF, the target structures are identified earlier and the communicative activities are designed including the target structures. Nassaji (1999) calls this FonF through *design*. In contrast, in incidental FonF, tasks are not designed based on a particular grammatical structure. The tasks may elicit several linguistic structures and while engaging in the task, either the learners will pay attention to one or many structures that occur or the teacher may draw the attention of learners to these. Nassaji (1999) describes this as FonF through the *process* of teaching.

Doughty and Williams (1998) and R. Ellis, Basturkmen and Loewen (2004) divide pedagogical FonF activities into two main types. One is reactive FonF activities which focus on the production of language i.e. treatment of learner errors. The other is preemptive FonF activities based on language input. In these activities either learners pay attention to or the teacher draws learners' attention to a feature/features in the input.

Two other terms that are related to the FonF approach are proposed by Spada and Lightbown (2008). They discuss the effect of *isolated* FFI and *integrated* FFI and highlight that both may have positive impacts in different contexts. In *isolated* FFI, input is provided through lessons which are neither content based nor include communicative interaction. They are also presented in isolation i.e. such tasks are not linked to the other activities in the class. In *Integrated* instruction TL forms are presented within meaning based content focus tasks. Spada and Lightbown highlight that the former type of lessons can be useful for learners to understand the influence of L1 on the TL and the latter may help to improve fluency and automaticity of the TL.

Not only the FonF approach, but another approach called emergent pedagogy (Thornbury, 2005) also highlights the importance of providing meaning based input for language learning. In the next section, I will outline the main theoretical concepts behind emergent pedagogy.

2.6.2 Emergent pedagogy

Thornbury (2005) believes that grammar should be treated as a process not as a product. In order to discuss the difference between 'process' and 'product', he uses the analogy of making an omelette and describes how the process of making an omelette involves all the necessary ingredients and performing the tasks: beating up the egg, mixing ingredients and finally frying the egg to have the end product. Thus, Thornbury

(2005) argues that in a language, the end product is its ‘grammar’ and in traditional teaching methods this end product is taught. In other words, teaching grammar is equal to teachers breaking up the omelette into pieces and asking learners to make the omelette. According to Thornbury, teaching the process or ‘grammaring’ is important, but not teaching ‘grammar’.

Thornbury (2005) further argues that grammar learning is a non-linear complex process and it involves self-organization and emergence. At the early stage of language learning, grammar may not matter; learners might be able to communicate only with lexis in their immediate contexts. However, the need for grammar increases under four conditions. They are, when there is less context, (for example, if speakers are in the same context they will understand each other with less use of grammar); when there is a wider knowledge between the communicators; if the social gap between the communicators is wide; and when the formality of the context increases. Thus, Thornbury believes that grammar teaching should be based on activities that provide natural language input related to the above contexts. In such activities, grammatical structures should be used in such a way that form and meaning are connected i.e. learners should be able to identify the necessity of structures or why grammar ‘matters’ (Thornbury, 2005, p. 38) or, in other words, why a particular structure is necessary to express a particular meaning. Thornbury believes in the importance of promoting noticing or pattern detection by providing such input with salient examples through techniques such as input flooding. This consciousness raising is aimed at providing data that could become intake; however, immediate production of the target structures by learners is not expected in such tasks. Thornbury calls this technique emergent pedagogy since grammar is taught through grammatical constructions that emerge in the input.

Both FonF and emergent approaches argue that learners should be provided with more natural language input to facilitate the identification of both form and meaning of the TL constructions. Both approaches also highlight the possibility of creating situations in the classroom where natural language input emerges, and also the necessity of providing pre-planned input. The vital aspect is providing sufficient input that promotes identifying form-meaning relationships.

2.6.3 Salience

Providing natural language input may also depend on the type of grammatical constructions that are to be taught in class. DeKeyser (2005, p. 3) points out that there are at least three factors that determine grammatical difficulty. They are the complexity of form, meaning, and of the form-meaning relationship. He further points out that it is the transparency of the form-meaning relationship or the possibility of grasping it that actually determines the difficulty of acquisition. This transparency is determined by "the degree of importance of a linguistic form for the meaning it expresses" (DeKeyser, 2005, p. 3). For example, "certain morphemes are the one and only clue to the meaning they express; others are largely or completely redundant, because they mark grammatical agreement with meanings whose primary representations are elsewhere in the sentence or the discourse" (DeKeyser, 2005, p. 3).

DeKeyser (2005, p. 8) highlights three factors that determine the transparency of the form-meaning relationship: redundancy, optionality, and/or opacity. If a grammatical form is not necessary to express meaning because the meaning given by that form is expressed in one or more other ways in the same sentence, the form in question is considered redundant. In addition, if redundancy is linked to novelty and abstractness, the understanding of form-meaning relationships becomes even more difficult. Optionality is the possibility of not using certain forms to express meaning: e.g., null subjects in Spanish or Italian (p. 8). This becomes problematic when the optional forms are present in the input for the same meaning. Opacity is referred to as the weak correlation between form and meaning. "When a morpheme has different allomorphs, and at the same time it is homophonous with other grammatical morphemes, then the correlation between form and meaning becomes very hard to detect" (p. 9). According to DeKeyser, the best example to illustrate this is the word final –*s* in English. It is used in the plural nouns and third person singular present simple verbs, which can create confusion to learners. Thus, it seems that the grammatical structures with clear form-meaning mapping are more likely to be identified by learners.

Apart from the form-meaning relationship of language structures, there are other factors that determine what features in the input learners can identify. N. Ellis (2006) points out that learners pay more attention to certain stimuli even though many other stimuli are available in the environment. Both N. Ellis (2006) and Sharwood Smith (1993) maintain that the salience of input features may determine whether they receive the learners' attention or not. Schmidt (1990) is of the view that lack of perceptual salience or 'noticeability' in the input and the lack of processing ability of learners may hinder noticing language features in input.

Thus, it is apparent that salience is prominent in determining what features in the input learners identify. Goldschneider and DeKeyser (2005) highlight five factors that determine the degree of salience: the number of phones in a functor, semantic complexity, morphophonological regularity, syntactic category and frequency. Perceptual salience in the input can be increased by including functors with more phones, which contain a vowel in the surface form, for example, present progressive -*ing*; plural –*s*, possessive -*'s*, articles *a, an, the,* 3rd person singular present –*s* and regular past –*ed* as functors. The second factor, semantic complexity "is a measure of how many meanings are expressed by a particular form" (p. 50) and semantically less complex forms are considered to be more salient. The third, morphophonological regularity is the "degree to which the functors are (or are not) affected by their phonological environment" (p. 51) and regular phonological functors are assumed to be more salient than the functors with a number of phonological alternations. The next factor, syntactic category refers to which syntactic items are more salient and it is considered that lexical items are more salient than functional ones. The last factor is frequency and if an item occurs frequently in the input, it is predicted to be more salient. DeKeyser (2005) stresses that when relevant input on such form-meaning relationships is not available to learners, acquisition becomes extremely difficult and thus he considers frequency as a decisive factor that determines the difficulty of the form-meaning relationship. If the transparency of this relationship is clear, less input may be sufficient; however, more frequent input of the less transparent relationships may be required for acquisition to take place. Therefore, it is apparent that frequency of occurrence plays a vital role in drawing the attention of learners to specific features in the input.

Input salience can be increased externally with the help of explicit FonF techniques (Doughty & Williams, 1998) such as discussion of target forms, metalinguistic explanation, negative evidence via overt error correction and processing instructions. R. Ellis (1999) recognises this type of input as *processing instruction* in which learners' attention is drawn to the target structure through specially designed input. According to him, the manipulation of input in these ways is intentional. Input salience can also be increased implicitly by means of input flood and TE (Gascoigne, 2006; Nassaji and Fotos, 2011) which R. Ellis (1999) identifies as *enriched input*. In enriched input, learners are provided with input full of examples of the target structure in meaning focused activities. Therefore, identification of the target structure is incidental. In this research I investigated the effect of two implicit FonF techniques: input flood and TE, and two explicit means of consciousness raising: instruction to pay attention to the target feature in the input and explicit metalinguistic explanation on attentional processing.

Nassaji and Fotos (2011, p. 38) define input enhancement as "the process through which the salience of input is enhanced", and note that "such modifications of input make grammatical forms more noticeable and subsequently learnable" (p. 36). However, Sharwood Smith (1993) notes that it is difficult to predict the consequences of such manipulation of input on learners. Jourdenais, Ota, Stauffer, Boyson and Doughty (1995), Izumi (2002) and Leow, Egi, Nuevo and Tsai (2003) agree that input enhancement draws learners' attention to features in the input and they assume that this may facilitate learning. Furthermore, all of them agree that salience of input is crucial to draw the attention of learners to the target features in the input. Salience of certain features in input can be increased through external enhancement i.e. "through external operations carried out on input" (Nassaji & Fotos, 2011, p. 37). Nassaji and Fotos (2011) further state that internal enhancement is also possible and in this case learners themselves identify features in the input on their own as a result of their internal cognitive processes. For example, a learner may pay more attention to content words to understand meaning of sentences and this choice is based on internal cognitive processes.

One of the techniques that can potentially draw learners' attention to particular linguistic constructions in the input is input flood. According to Nassaji and Fotos (2011), in this technique, several examples of the TL structure are included in the input text (either aural or written) to provide a flood of input to learners. This technique is based on the assumption that frequency increases the perceptual salience of target features (Gass & Selinker, 2008; N. Ellis, 2002). Wong (2005) believes that this artificial increase of the target structure in input aids noticing and then subsequent acquisition. The findings with regards to the potential role of input flood in enhancing learners' L2 competence are mixed. Jourdenais et al. (1995) used both input flood and input enhancement tasks and reported that learners demonstrated more gains in the input flood tasks than in the input enhancement tasks. Reinders and Ellis (2009) report similar post-test gains in both the input enhancement and input flood groups in their study of English negative adverbs. Trahey and White (1993) found that input flood had a significant effect on learning word order in English. Hernandez's (2008) study on Spanish discourse markers, however, indicated that input flood combined with explicit instruction is more effective than input flood alone. Izumi (2002, 2003) also argues that input flood does not necessarily have a positive effect on acquiring target forms, while

Jahan and Kormos (2015) did not detect a significant effect of input flood compared to input enhancement.

Another commonly applied input enhancement technique involves making TL constructions visually salient by means of highlighting, underlining and making words/phrases bold. Similar to the studies investigating the effectiveness of input flood, the findings of previous research on TE are contradictory. Shook (1994), Jourdenais et al. (1995), and Alanen (1995) found TE to have a positive impact on L2 learners' grammatical development. Jahan and Kormos' (2015) results revealed that exposure to textually enhanced input facilitated development in the controlled use of the targeted future tense construction that learners had pre-existing knowledge of. No gains were obtained, however, in the construction that learners showed no previous familiarity with. The results of White (1998), Leow (2001), Izumi (2002), Wong (2003), Simard (2009) Overstreet (1998) and Leow, et al. (2003) showed no effect of TE. Overall, Lee and Huang's (2008) meta-analysis concluded that TE has a very small-sized effect, d = .22, on the acquisition of grammatical constructions in L2 learning. Two recent studies by Winke (2013) and Simard and Foucambert (2013) indicated increased attentional processing of input under enhanced conditions when attentional processing was operationalized as total fixation time, whereas in Issa et al.'s research (2015) TE had no effect on total fixation time. Neither Simard and Foucambert (2013) nor Issa et al. (2015) found a significant relationship between total fixation time and gain scores of the tests of targeted grammatical constructions.

2.6.4 Combined input conditions

The present study was aimed at analysing how learners pay attention to both enhanced/unenhanced input i.e. input enhancement and input flood as well as how attention changes when receiving explicit instructions. Moreover, explicit explanation of rules as opposed to the implicit learning conditions was also investigated. Therefore, below I summarize some of the SLA studies that have incorporated two or more of these features within the same study design.

Robinson (1997), in a study that used monosyllabic and disyllabic verbs as the target construction, provided input by means of 55 sentences in four different input conditions. The first group of participants was told that the task was to read and remember the position of words, the second group was told that the exercise was to understand the meaning, the third received the same explanation as the second group with the target forms textually enhanced. The fourth group received an explicit explanation of the rule and they were asked to remember the rules to be used when reading sentences. The results showed that the last group outperformed all other conditions while the third group with the target forms textually enhanced performed better than the first and second groups.

Rosa and O'Neill (1999), in a study of Spanish contrary-to-fact conditional sentences, used four experimental groups and a control group. The first group was given an explanation of the rule and then was instructed to work on a jigsaw and form a rule based on the sentences while solving the puzzle. Finally they were asked to write down the rule. The second group received the same materials, but they were not asked to look for a rule while solving the puzzle. Instead they were asked to memorize the sentences. The third group did not receive an explanation; they only read a short text that did not contain the target structure. When solving the puzzle, they were asked to look for a rule based on the sentences. The fourth group also read the same short text as the third, but they were not asked to find a rule. Instead they had to memorize the sentences. The last group, which was the control condition, read the reading passage and completed the puzzle without having to find a rule or memorize sentences. A significant improvement was observed only between the first group (who received the explicit explanation) and the fourth group and between the second group and the fourth group.

Radwan (2005) used four groups in a study that investigated the learning of English dative words contained in a short story. One group received the input text with textually enhanced target words. The second group was given the same input with a one page explanation of the target rule. The third group read the text without enhancement and the fourth was the control group. The rule oriented group showed a significant improvement in the post-test. In another study by Tode (2007) the first experimental group received an explicit explanation of the target construction (English copula *be*). The second group only memorized the sentences containing the target construction without focusing on the rule. The third group was the control condition. Tode reports that only the explicit explanation group showed a significant improvement in the post-test.

A common finding of these studies is that explicit rule explanation had a positive impact whenever such a condition was included. Further to this, participants in TE only conditions without explicit instructions or explicit explanations did not seem to have shown improvement that was different from unenhanced conditions (for example in Radwan, 2005). Moreover, providing explicit explanation of the target construction or instructions to pay attention to the target constructions seems to have played a positive role in all experiments discussed above.

2.7 Working memory

Although the primary aim of this study was to investigate the role of attentional processing in different input conditions, the study also examined the relationship between attentional processing of input and the WM capacity of the participants. As pointed out above, WM is closely related to attentional processing and consciousness. In particular, if one attends to a stimulus with conscious awareness, such stimuli have the possibility to enter the WM (Koch & Tsuchiya, 2006; Lamme, 2003). Therefore the investigation of the association of WM and attentional processing can yield new insights into how L2 learners attend to and process input while learning an additional language. In the following paragraphs I review the models of WM and discuss the empirical studies that investigated the role of WM in L2 learning.

2.7.1 Baddeley and Hitch's model

Atkinson and Shiffrin (1968), in an early discussion of human memory, identify two components of memory: STM and LTM. According to this model of memory, a set of information available through sensory systems is selected through selective attention and is transferred to STM for further possessing. The selected information in the STM cannot be kept for long periods of time; therefore rehearsal is necessary until information is transferred to the LTM. Otherwise, memory traces decay within a short period of time. In this model, STM is conceptualized as a unitary system which is responsible not only for the temporary storage of information but has additional functions such as retrieving stored information, processing information in learning, and reasoning. In their criticism of this model, Baddeley and Hitch (1974) argue that the unitary nature of STM would mean that if a person has a lower STM capacity, they would also have problems in other cognitive processes such as reasoning and learning. Based on the results of several dual-task experiments, they found that the storage function works separately from other cognitive functions. Therefore, they proposed a non-unitary model in which STM and WM are separated. Miyake and Shah (1999), who comment on this distinction, identify STM as a storage-oriented structure and WM as a process oriented construct.

As a result of identifying that STM and WM function independently, Baddeley and Hitch (1974) proposed a multi-component model for WM which contains three main components: a CE and two slave systems called the phonological loop and visuospatial sketchpad. The CE acts as the general attentional controller, the phonological loop is responsible for processing information related to speech, and the visuospatial sketchpad handles information related to visuospatial imagery. In this model, the CE is considered as a processor of all information that is not processed by the phonological loop and visuospatial sketchpad. Baddeley (1996) explains that the CE is responsible for mechanisms such as coordinating the subsidiary memory system, switching attention, controlling encoding, retrieval strategies and manipulating information in the phonological loop and visuospatial sketchpad. Some other functions are also assumed to be controlled by the CE, for example, switching attention between multiple tasks

(Rogers & Monsell, 1995), inhibition of irrelevant information (Roberts, Hager & Heron, 1994), monitoring and updating the content of the WM (Van der Linden, Bredart & Beerten, 1994), temporal tagging and contextual coding of incoming information (Jonides & Smith, 1997) and planning and sequencing of intended actions (Ward & Allport, 1997). The phonological loop is a limited capacity component of the WM, which is comprised of a temporary phonological store (phonological short term memory, referred to as PSTM) and an articulatory rehearsal process (Baddeley, 2000). The auditory memory traces are stored and rehearsed in the articulatory process. If not rehearsed, the traces decay within a short period of time. The visuospatial sketchpad also has a limited capacity and is responsible for processing visual and spatial information such as colour, location and shape (Baddeley, 2003). Similar to the functions of the phonological loop, the visuospatial sketchpad temporarily stores and processes this information (Baddeley, 2000).

Baddeley (2000), who reviewed the three-component model of WM he himself proposed with Hitch in 1974, identified several of its shortcomings. Since CE was considered to be only a processing mechanism and both the phonological loop and visuospatial sketchpad were assumed to have limited storage capacity, this model did not identify a place where other complex information can be stored temporarily. It did not explain either how information extracted from LTM is processed or where it is stored while being processed. In addition, it did not offer an explanation of how chunking could occur, how conscious awareness is related to WM and how information from the two slave systems is integrated (Baddeley, 2003). Thus, Baddeley (2000) proposed a fourth component to the original model called the episodic buffer. This also has a limited storage capacity and its main function is to combine information from different sections of the WM. Episodic buffer is also controlled by the CE and is accessible to conscious awareness (Baddeley, 2003). A recent review of this model by Baddeley (2015) suggests that the phonological loop feeds the buffer with information related to speech, language (lip read and signed information) and non-verbal sounds, whereas the visuospatial sketchpad provides information related to spatial location, colour, shape, touch and kinesthesis.

The early model (Baddeley & Hitch, 1974) did not clearly highlight the functions of the CE either. Based on Norman and Shallice's (1986) model of attentional control, Baddeley (2003) later attempted to provide a more comprehensive discussion of the functions of the CE. Accordingly, two main sections were identified. One is responsible for implicitly controlling habitual behaviour based on the cues provided by the environment. The other is recognised as the Supervisory Activating System (SAS) that functions when the routine system is inadequate. The first is an automatic process that does not involve awareness whereas SAS is the attentional process that does. This postulates that the CE per se is not a unitary system either. Quoting results of several studies, Miyake et al. (2000) point out that there is no correlation reported between different tasks controlled by the CE. However, according to Miyake et al., these studies have several methodological problems. One of the shortcomings is that the tasks used to measure the CE have strong relationships with the functions of the two slave systems and thus the independence of the measurements is questionable. Another issue is that when the tasks are repeatedly used to measure the same function, their effectiveness is affected. The third is that most tests that have been used to measure the CE lack sufficient validity.

Due to these shortcomings in the existing literature, Miyake et al. (2000) used a series of tests to assess three different functions of the CE with the aim of investigating the relationship between them. The functions that they measured were: shifting between tasks or mental sets (switching), updating and monitoring of WM representations, and inhibition of prepotent responses. Shifting refers to switching back and forth between multiple tasks/mental sets/operations. The second process 'updating and monitoring' involves identifying and coding the relevant information for the task at hand from the incoming data, revising what is already in the WM and updating it with appropriate information. Inhibition is the ability to "deliberately inhibit dominant, automatic or prepotent responses when necessary" (p. 57). The results of their study indicated a moderate correlation between the three functions suggesting that they work independently. Moreover, the results highlighted that these functions made a different contribution to performance in complex executive tasks such as random number generation, operation span, and dual tasking. Thus, the authors stress the importance of identifying the unitary as well as diverse functions of the CE.

2.7.2 Other models

Although Baddeley and Hitch's model is the most discussed and widely used model of WM, a few others have also been proposed. Cowan's (1988) embedded-processes model, which has more unitary features than Baddeley and Hitch's, explains that memory has three components: LTM is the largest where information is stored. Within the LTM, there is a subset of information that is ready to be activated. Within that subset, there is another subset of information that is in the current focus of attention or conscious awareness state. According to this model, activation of information is necessary because it is not possible to attend to inactive information. The main argument behind the proposal of Cowan's unitary model is the fact that Baddeley and Hitch's model, and its later additions, does not sufficiently explain how information is forwarded to the relevant slave system, or discuss the possibility of a stimulus having multiple features, such as visually related and verbal information. Thus, Cowan's (2015) model assumes that different parts of the LTM get activated by one stimulus. However, the possibility of activating information automatically without conscious attention negates the validity of the claim that one stimulus is needed to activate different parts of the LTM (Miyake & Shah, 1999).

According to Engle, Laughlin, Tuholski and Conway's (1999) model, WM consists of three components: (1) LTM traces that are active above a certain threshold, (2) the processes that control this activation and (3) controlled attention. WM capacity is identified as the capacity of "controlled, sustained attention in the face of interference or distraction" (p. 104). Lovett, Reder and Lebiere (1999) identify WM as a limited attentional resource that is focused on a particular goal. When the attention is on a particular goal, it increases the possibility of accessing the goal-relevant knowledge (p. 135) and individual differences play an important role in the access process. Several other models such as those of Kieras, Meyer, Mueller and Seymour (1999), Young and Lewis (1999), Schneider (1999) and Ericsson and Delaney (1999) also discuss the role of WM. Some commonalities in them are that all these models identify WM as a limited capacity system related to LTM. They also recognise that WM consists of several processes and mechanisms and thus it is not a unitary system.

2.7.3 Different approaches to measuring WM

The effect of WM in language learning is discussed in two parallel lines by British (such as Baddeley and Gathercole) and Canadian (such as Just, Daneman and Carpenter) researchers. While the British research tends to base its explanations largely on Baddeley and Hitch's (1974) model, the Canadian research tradition provides a slightly different view to what is proposed by Baddely and Hitch in their earlier model as well as its improvements.

For example, Gathercole and Baddeley (1993, p. 3) identify the phonological loop as the main section of the WM that influences language learning; however, the visuospatial sketchpad also plays a key role in it. The CE, the coordinator of both these slave systems, and the episodic buffer, the temporary memory store of the coordinated information, also assist in language learning. In contrast, according to Daneman and Carpenter (1980), Just and Carpenter (1992) and Daneman and Merikle (1996), the WM does not have a separate system such as the phonological loop that is specific for language processing, but it involves a set of processes and resources. Language comprehension is a blend of both processing and storage functions of the WM. Each element that enters the WM to be processed (e.g., words/ phrases/ sentences) has a particular activation level. When the activation level goes beyond the threshold that item is in the WM or in other words it gets processed. As in Baddeley and Hitch's (1974) model, they also assume that the WM has a limited capacity and as a result there is a trade-off effect. For example, if an item requires a heavy processing capacity load, the other items that enter the WM during this time of processing may not be maintained and thus they would decay. Therefore, Daneman and Carpenter (1980), Just and

Carpenter (1992) and Daneman and Merikle (1996) argue that those who are better in language comprehension have a larger WM capacity.

There is also a difference between how WM is measured in the British and Canadian research tradition. Canadian researchers tend to use reading span as a measurement because they argue that the traditional digit span and word span tests are unable to assess the processing function of WM, instead they measure the storage capacity. Although it is debatable that a reading span test is able to identify all mechanisms in the WM, listening span and speaking span (Daneman & Green, 1986) tests were also introduced in the Canadian research tradition later on. In contrast, British research uses a series of tests to measure different functions of WM. For example, digit span and non-word span tests are widely used to measure the PSTM capacity while a battery of tests is used to identify CE functions (see Miyake et al., 2000).

2.7.4 WM and L2 learning

WM is assumed to influence different L2 learning functions and here I use those that Gathercole and Baddeley (1993) believe to be affected by WM as the basis of the discussion of empirical findings. They are: vocabulary acquisition, learning to read, speech production, general language proficiency, and grammar learning.

According to Gathercole and Baddeley (1993) the phonological loop plays a crucial role in L2 vocabulary acquisition. For example, functions such as word recognition, relating familiar words to unfamiliar words and identifying phonological similarity and word length are influenced by PSTM. Service (1992), Papagno and Vallar (1995), Cheung (1996), O'Brien, Segalowitz, Collentine and Freed (2006), Akamatsu (2008), de Abreu, Pascale and Gathercole (2012) and Martin and Ellis (2012) are some of the researchers who have studied the relationship between L2 vocabulary learning and WM. Apart from Akamatsu's (2008) research, all studies mentioned here have found a significant relationship between WM and vocabulary learning in L2.

Learning to read also seems to be linked to phonological awareness and phonological STM because poor readers tend to have issues in both these aspects (Gathercole & Baddeley, 1993). L2 reading in relation to WM has been studied by Daneman and Carpenter (1980), Harrington and Sawyer (1992), Dede, Caplan, Kemtes, and Waters (2004), Walter (2004), Leeser (2007), Swanson, Orosco and Lussier (2015) and Kim, Christianson and Packard (2015). All these studies except that of Leeser have indicated a significant relationship between WM and reading comprehension i.e. higher WM ability leads to better performance in reading comprehension. Leeser examined how both WM and topic familiarity influences reading comprehension and identified that WM plays a role in reading comprehension only when the readers need to use their previous knowledge of the topic for comprehension.

Gathercole and Baddeley (1993) also highlight the possible strong relationship between WM and speech production although sufficient research has not been conducted on this. Accordingly, they predict that speech production of adults is not directly related to the PSTM capacity, but to the CE functions. The CE is believed to influence the planning of semantic content of what is being said. This has been studied by Payne and Whitney (2002), O'Brien et al. (2006) and Mojavezi and Ahmadian (2014). These three studies have used different WM measurements: non-word span (O'Brien et al., 2006), reading span and non-word repetition (Payne & Whitney, 2002) and listening span (Mojavezi & Ahamadian, 2014). The first two have identified a significant relationship between speech production and WM but Payne and Whitney report limited correlation between oral proficiency and a non-word repetition test.

Another series of empirical research studies have examined the relationship between WM and general language proficiency. For example, Isaki and Plante (1997), Segalowitz and Frenkiel-Fishman (2005), Kormos and Sáfár (2008), Vejnovic, Milin, and Zdravković (2010), Mackey and Sasch (2011) and Biedron and Szczepaniak (2012), who investigated the relationship between general L2 proficiency and WM, report a significant relationship between general L2 proficiency and WM.

Overall, the meta-analysis of 79 studies by Linck, Osthus, Koeth and Buntin (2013) report a positive relationship of WM with L2 processing and proficiency outcomes. The focus of the present study is on the relationship between L2 grammar learning and learners' WM capacity and thus the next section of this review discusses the theoretical background to L2 grammar learning and its relationship with WM.

2.7.5 WM and L2 grammar learning

Drawing upon the findings of studies on the language learning difficulties faced by dyslexic learners, who are generally characterized by smaller WM capacity (Jeffries & Everatt, 2004), N. Ellis (1996) and N. Ellis and Sinclair (1996) argue that WM plays an important role in the processes of acquiring L2 grammatical knowledge and skills. This role has been investigated in a large number of studies which, however, have applied very different measures to assess the functioning of WM. This makes the comparison of the findings and drawing conclusions based on them difficult.

N. Ellis and Sinclair's (1996) study was the first to provide indirect evidence for the role of WM in relation to grammar learning as it demonstrated that articulatory rehearsal in WM facilitates the acquisition of explicit grammatical knowledge. Williams and Lovatt (2003) compared L2 learners' ability to remember determiner-noun combinations in Italian with their scores in an immediate serial recall test. Their results indicated a moderately strong relationship between success in the grammar learning task and WM. Martin and Ellis (2012) also reported a moderately strong correlation between a non-word repetition test and the participants' ability to produce artificial foreign language forms that they were taught, and to generalise those forms to new utterances. In Santamaria and Sunderman's (2015) study, participants with a high reading span were also found to score highly in both immediate and delayed post-tests targeting the knowledge of French direct objects. Apart from these, Ahmadian (2015) observed a significant relationship between a listening span test and L2 self-repair behaviour among adult L1 Farsi learners and O'Brien et al. (2006) report a significant relationship between a non-word span test and L2 Spanish grammatical competence measured in a narrative task. In contrast, Grey, Williams and Rebuschat (2015) did not detect a significant relationship between two non-word repetition tasks that measured WM capacity and the results of an acceptability judgment and a picture matching task on verb-final word order and case marking of a semi-artificial language that the participants in their study were exposed to. A series of studies was conducted by Juffs (2005, 2006) which aimed to uncover the association between L2 sentence processing and WM capacity, as measured by a reading span task. He also found no significant relationship between WM capacity and reading speed.

Several studies also investigated the relationship between corrective feedback given in L2 learning and WM capacity. For example, Li's (2015) study on the effectiveness of recasts in learning Chinese and WM capacity measured by a listening span test shows that high proficient learners take advantage of their WM when processing recasts. In Revesz's (2012) study, the participants were provided with recasts on English past progressive construction. In the recast group, the participants who had high digit and non-word spans performed significantly better in an oral test while those who had high reading span achieved more gains in written tests. Sagarra and Abbuhl (2013) also report that the participants with higher reading span in the recast group were able to perform significantly better than the lower reading span participants in a post-test that tested the Spanish noun–adjective gender or number agreement. In another study that provided recasts on Spanish subjunctives in a story re-tell task by Baralt (2015), however, no significant relationship between feedback efficacy and WM capacity, measure by three span tasks, emerged.

WM is a key factor in the regulation of attentional processing of L2 input, but evidence for this assumption to date is mostly available from studies that have investigated the role of WM capacity in learners' uptake of interactional feedback. Mackey, Philp, Egi, Fujii and Tatsumi (2002) studied the relationship between WM capacity and noticing corrective feedback in recasts, that is in utterances that repeated what the learner said with their error corrected. They found that participants with higher WM capacity (measured by English non-word recall and listening span tests) noticed more items in the corrective recasts, and these participants also show development in question formation between the pre and post-tests based on the feedback they received. Two other studies by Trofimovich, Ammar and Gatbonton (2007), who used a letter-number sequencing test to measure MW capacity, and Saggara (2007), who used a reading span test, examined learners' processing of recasts. They also provide some evidence that WM capacity influences the noticing of recasts.

WM was traditionally assumed to be influential in explicit learning, but not in implicit learning (e.g., Reber, Walkenfeld & Hernstadt, 1991). For example, Unsworth and Engle (2005) who studied the effect of WM (measured by an operation span test) on intentional and incidental learning conditions argue that it plays a role in intentional learning tasks which require conscious attentional control, but not in incidental learning processes which are the by-products of other cognitive operations and less prone to the influence of individual differences. In contrast, recent work by Hassin, Bargh, Engell and McCulloch (2009) and Soto and Silvanto (2014), who used specially designed WM tests for their studies, indicate that WM operations might be involved in implicit learning processes. The research up to date of the effect of WM on language learning, however, provides contradictory findings. For example, Robinson (2005) and Santamaria and Sunderman (2015) found that PSTM was influential in developing explicit knowledge in explicit learning conditions whereas Tagarelli, Borges Mota and Rebuschat (2015, 2011) report no relationship between WM and implicit knowledge development under an explicit learning condition. Most studies which investigated the role of PSTM in developing explicit knowledge in implicit learning conditions report significant relationship between the two (Ellis & Sinclair, 1996; Li, 2015; Mackey et al., 2002; Martin & Ellis, 2012; Saggara & Abbuhl, 2013; Saggara, 2007; Trofimovich et al. 2007; Williams & Lovatt 2003) except in Baralt's (2015) and Grey et al.'s (2015) studies. In the case of developing implicit knowledge in implicit learning conditions, Ahmadian (2015), Ellis and Sinclair (1996), and Saggara and Abbhul (2013) found significant influence of PSTM and implicit knowledge development. However, Baralt (2015), Grey at al. (2015), Robinson (2005) and Tagarelli et al. (2015, 2011) did not find a significant relationship between the two. (See Table 2.2 for a summary of studies that investigated the relationship between implicit/explicit input/knowledge and WM).

Table 2.2

Empirical studies on implicit/explicit instruction/knowledge and WM capacity

Study	Working memory components assessed	Learning conditions	Outcome measures	Learning target	Relationship with WM components
Ahmadian (2015)	PSTM	implicit	implicit	L2 self-repair behaviour	significant
Baralt (2015)	PSTM	implicit	explicit implicit	Recasts on Spanish past subjunctive	non-significant non-significant
Ellis & Sinclair (1996)	PSTM	implicit	explicit implicit	Sound change in noun-phrases	significant significant
Grey, Williams, & Rebuschat (2015)	PSTM	implicit	explicit implicit	verb-final word order and case marking	non-significant non-significant
Li (2015)	PSTM	implicit	explicit implicit	effectiveness of recasts on Chinese classifiers	implicit- significant among high proficient learners
Mackey, Philp, Egi, Fujii and Tatsumi (2002)	PSTM	implicit	explicit implicit	noticing corrective feedback in recasts	significant
Martin & Ellis (2012)	PSTM	implicit	explicit	Determiner-noun combinations	significant
Robinson (2005)	PSTM	explicit implicit	explicit implicit	3 grammatical rules in Samoan	significant non-significant
Sagarra and Abbuhl (2013)	PSTM	implicit	explicit implicit	Recasts on Spanish noun–adjective gender or number agreement	significant significant
Saggara (2007)	PSTM	implicit	explicit	processing of recasts	significant
Santamaria and Sunderman (2015)	PSTM	explicit	explicit	French direct objects	significant
Tagarelli, Borges Mota and Rebuschat (2011, 2015)	PSTM Letter-	explicit implicit	implicit implicit	Semi-artificial language three specific verb	non-significant non-significant

	number ordering			placement rules	
Trofimovich, Ammar and Gatbonton (2007)	Letter-number sequencing test	implicit	explicit	processing of recasts	significant
Williams & Lovatt (2003)	PSTM	implicit	explicit	Determiner-noun combinations	significant

Given the contradictory findings and the different type of WM measures applied in the field of SLA research, I found it important to examine what role WM capacity plays in L2 learners' attentional processing of a novel TGC, and how this cognitive ability contributes to learning gains under different instructional conditions. Therefore, in this study, I applied three different tests of CE and a test assessing the PSTM capacity. With the administration of four different assessment tools, I aimed to overcome the shortcoming of most previous studies that used only one instrument. The inclusion of tests of the CE also allowed me to examine its potentially important role in the regulation of attentional processing in combination with the PSTM capacity. Moreover, up to date, little is known about how WM abilities affect the learning of grammar in a language that learners have prior knowledge of since most of the existing studies focused on artificial grammar learning. Thus, this study investigated the role of WM in learning a novel TGC of a language that learners have prior knowledge of.

2.8 Summary

As pointed out in cognitive psychology, attention, consciousness and awareness are three different phenomena. Attention is identified as a limited capacity that selects a stimulus from several available stimuli for further processing. Awareness may or may not involve attention. However, for an attended stimulus to get registered in the WM or to offer a report on an attended stimulus, awareness is believed to be necessary. Awareness thus occurs due to paying conscious attention. There is no agreement in cognitive psychology whether awareness is necessary for learning since empirical studies have had contradictory results.

In the SLA literature, attention, consciousness and awareness are vaguely defined due to the lack of appropriate methodology to investigate these phenomena separately. However, one line of research claims that attention with awareness is necessary for L2 learning (this is recognised as 'noticing' by Schmidt, 1990), while the others argue that a certain level of attention is sufficient for L2 learning (Tomlin & Villa, 1994; Robinson, 1995; Godfroid et al., 2013). Furthermore, empirical evidence sheds little light on understanding the role of attention and awareness in L2 learning because of the shortcomings of the methodology used to investigate the role of these two cognitive processes. For example, methods such as think-aloud, stimulated recall, classroom observation and interviews, which have been commonly used in attention/awareness studies, have potential problems such as their inability to distinguish between attention and awareness, and giving an extra processing load (think-aloud).

Although eye-tracking, a novel method used to investigate attention in L2 processing has several shortcomings, it is the best method available to date to measure attention quantitatively without causing an extra processing load. Moreover, it records attention at the same time as one is exposed to a stimulus, and therefore the data can be more accurate than the data provided by stimulus recall or post-exposure interview. However, it is necessary to attest that eye-tracking is unable to distinguish between attention and awareness. Furthermore, it only records the movement of eyes and does not take into account that one could look at the computer screen without attending to a particular stimulus. The eye-tracking measurements used so far in attention related studies also seem to be inadequate to discuss the additional processing time spent on target features while attending to input.

As this literature review highlighted, there is a substantial body of research that has explored the effect of explicit and implicit FonF techniques on language learning outcomes. According to the outcomes of these studies, the FonF approach seems to be more useful for L2 learners since it helps them identify form-meaning relationships. In addition, empirical evidence shows that enhanced FonF conditions promote attention to target features in the input, which subsequently develop the knowledge of those target features. Although there is no agreement in the SLA literature if explicit instruction leads into explicit knowledge only or implicit instruction leads into implicit knowledge only, there is evidence that explicit instruction such as explaining the rules of target constructions or asking learners to pay attention to target features is more beneficial for L2 learning than implicit instruction such as TE and input flood. Yet little is known about how successful these techniques are at directing learners' attention to the target linguistic features and how the amount of attention paid to the target constructions is associated with learning gains under different instructional conditions. Recent advances in eye-tracking technology allow us to investigate these questions in a more accurate and reliable manner than was done in previous studies, which relied on self-report or other indirect measures of attentional processing.

There have been several attempts to define WM and its functions. Nevertheless, Baddeley and Hitch's (1974) model and its extensions (Baddeley, 2000, 2003) are most widely used in research probably because of the clarity and explicitness in them in explaining the cognitive processes related to WM. Theoretically, it is believed that WM plays a role in L2 development and this has been validated by the results of several empirical studies that investigated the relationship between WM and L2 learning. While paying attention to measuring PSTM capacity, most of these studies have ignored the CE functions that can be crucial in attentional processing. In particular, none of the SLA studies that have investigated WM functions in L2 input processing have used tests to measure different functions of the CE.

The main theoretical gap related to attentional processing of L2 input that I have identified through this literature review is the vagueness of terminology used in SLA literature (such as attention and awareness). This has subsequently led to using data collection methods that do not measure what they are assumed to measure in empirical studies related to attention and awareness. Hence, in this study I separated attention from awareness and treated them as two different phenomena. Moreover, two different data collection methods were used to collect data on attention (eye-tracking) and awareness (post-exposure interview). A methodological issue that was observed is the inadequate information provided by existing eye-tracking measurements on attention. Therefore, I developed a new eye-tracking measurement (ΔOE) in this study. Another novelty of the study was combining four different input conditions: two explicit (explicit explanation and instructions to pay attention to the target construction) and two implicit (TE and input flood) within the same study design to understand what type of input would yield better results in drawing learners' attention to target structures and subsequent language development. Finally, the study investigated the relationship between attentional processing and WM capacity by focusing on both PSTM and CE functions using a battery of tests.

Chapter 3: Pilot study

3.1 Introduction

As discussed in 2.8, the study used several data collection methods: eye-tracking, post-exposure interview and WM tests. Since measuring attention and awareness has been problematic in SLA research, it was necessary to pilot the data collection methods applied in this study in order to see whether they could provide fruitful data. Moreover, to my knowledge, there has not been any SLA study on attentional processing of input in which four WM tests were used within the same study design. Thus, piloting the WM tests was also necessary to gain information about their psychometric characteristics. The methodology in the main study was developed based on the insights gained in the pilot study. Therefore, this chapter is devoted to discussing the pilot study in detail.

I will start with a description of the research context (3.2) and research questions (3.3) of the pilot study and then discuss the participants (3.4), instruments (3.5), data collection methods (3.6) and the data analysis methods (3.7). I also outline the results (3.8) and discuss the main findings (3.9). The final section of this chapter (3.10) highlights the lessons learnt and how those insights have influenced the decisions taken in developing the methodology of the main study.

3.2 Research context

Since the main study aimed at analysing how L2 learners process TL input, it was necessary to conduct the pilot study with the participation of the L2 learners. The pilot study was carried out at Lancaster University, and thus it was a challenge to find L2 learning/teaching contexts within the university. The focus of most of the English language related programmes run by the university is English for Academic Purposes (EAP), not English language teaching. The best options available to find suitable participants for the study were the four-week pre-sessional EAP course run by the Department of Linguistics and English Language and the foundation-year course run by the Lancaster University International Study Centre. The pre-sessional students are those who have already been selected for degree courses at the university but have not reached the required standard in one or more of the components of the IELTS [8] examination. As a result, they need to take the four-week EAP course to improve their academic English language proficiency. This programme receives both undergraduate and postgraduate students. The foundation year students are those who expect to start their undergraduate courses at the university; however, they have not reached the required standard of subject knowledge to fulfil the university entry requirements. Therefore, they undertake a foundation year course to improve their subject knowledge to enter undergraduate courses at the university.

There were two main reasons behind selecting participants for the pilot study from these two programmes. The first is that the majority of the course participants of these programmes are L2 learners of English. The pre-sessional EAP programme does not receive any native English speakers and the majority of the course participants are from East Asian countries such as China, Hong Kong and Taiwan. These students have been learning English as an L2 for a number of years when they enter the programme and they are mostly at the B1/B2 level of the Common European Framework of Reference

[8] IELTS: International English Language Testing System. This is a language test that measures the proficiency of those who aim to study/work in English speaking countries.

(CEFR, Council of Europe, 2011). The student population of the foundation year programme is slightly different from the four week EAP course student population because although the majority of these students are also L2 learners of English from countries such as China, Hong Kong, India and the Middle East, it does include some students whose first language is English. Due to this, the proficiency level of the L2 learners of this programme varies, but majority of them are usually at the B1/B2 level of the CEFR. The second reason for choosing participants from these programmes was that these students have less exposure to studying in an English language context because mostly they join these programmes immediately after completing their high school education in their home countries. Unlike these students, the other students on undergraduate courses or masters courses at the university may have more exposure to living and studying in an English language environment and their processing of L2 input may be affected by this exposure. Thus, if gathered from such a population, data might not have provided useful insights to L2 teaching contexts. Considering these two factors, it was decided to recruit participants for the pilot study from the above mentioned programmes.

3.3 Research questions

The pilot study was aimed at trialling all data collection methods which were going to be used in the main study. Thus, it attempted to answer the following three research questions.

i. Can attention to L2 grammatical construction be reliably investigated by means of eye-tracking?

ii. Can L2 learners at a lower intermediate/intermediate level of proficiency develop awareness as a result of paying attention to unenhanced examples of grammatical structures in input?
iii. Is the WM capacity of the participants related to the amount of attention paid to input, if so which components of WM are most strongly linked to it?

3.4 Participants

Seventeen participants took part in the data collection; however, two had to be excluded from the analysis due to the poor quality of their eye-tracking data (please see 3.7.1 for an explanation). The 15 participants whose data were included in the study were first language Chinese speakers who were either foundation year students or undergraduate pre-sessional EAP course participants at Lancaster University. None of these participants had any prior knowledge of the target constructions i.e. they scored 0 in the pre-test SR target items. Any correct answers given to the GJ items were thus considered to be due to guessing. The participants were 18-21 years of age ($M = 19.24$, $SD = 1.01$) and 13 of them were female and 2 male. Their IELTS scores were between 5.0 and 6.0 i.e. they were at a lower intermediate/intermediate level of English language proficiency. All of them had previous experience in learning English for more than 6 years. The two males were bilingual speakers of Mandarin and Cantonese and all others were monolingual Mandarin speakers. The only other language they spoke was English. At the time they took part in the data collection, they had lived in the UK for less than one month.

Consent for participation was obtained by means of a form (Appendix A) that the participants signed and handed in after reading an information sheet provided to them at the beginning of the data collection process. The information sheet described the main aims of the study, gave the reasons for approaching them as participants, outlined the benefits of participation and the options for withdrawal.

3.5 Instruments

3.5.1 Background questionnaire

A background questionnaire (Appendix B) was used to obtain data related to age, gender, IELTS score, and the first language of the participants. It also included questions on the other languages they spoke and their language learning experience. The responses obtained are summarised in Section 3.4 under participants.

3.5.2 Input text

I wrote the input text used in the study (see Appendix C for the text and Table 3.2 for the examples of the target structures in the text). This is because I could not find a text that contained both target structures in combination as well as a text that suits the participants' proficiency level in terms of vocabulary in the text. The aim of the study was to analyse attentional processing of target syntactic constructions; nevertheless, I did not want to provide input in separate sentences. The reason behind this is that it is best to provide natural language input where form and meaning are connected (Thornbury, 2005) because then learners might understand why a given structure is necessary to convey a particular meaning. This is also the position maintained in the FonF approach. Although form-meaning mapping can be created in isolated sentences, such sentences do not act as a holistic sample of natural language. A short story however serves the purpose of providing natural language input better.

Based on the assumption that ‘salience’ is one of the features that promotes attention (Caroll, 2006; N. Ellis, 2006; Schmidt, 1990; Sharwood Smith, 1993), and input flood can increase salience (Gass & Selinker, 2008; N. Ellis, 2002), several examples of the target structures i.e. 9 examples of *causative had* (e.g., I had my hair cut...) and 8 examples of past perfect (e.g., He had bought paint ...) were included in the text.

Causative:

had + article/determiner + N + past participle

e.g., I *had the roof repaired*

Past Perfect:

had + past participle + N

e.g., I *had painted the house*

The two target structures were selected considering two factors: structures that the prospective participants were assumed not to have learnt before and structures that differ in word order which were considered suitable for an eye-tracking study (Holmqvist et al., 2011, p.69). It was predicted that since the participants of this study were at the lower intermediate/intermediate proficiency level, they would not have had exposure to the target structures, in particular, to *causative had* (This structure is usually included in C1 or C2 level English language text books – e.g., *English File*, *Cutting Edge*).

The input text was also checked for complexity by Vocabprofile (Heatley, Nation, & Coxhead, 2002) to verify whether the vocabulary in the text fell within participants' proficiency level. Based on the results of the Vocabprofile analysis, 87.58% of words contained in the text were K1 words i.e. they belonged to the most frequent 1000 words in English, 6.54% of words were K2 words i.e. they belonged to the second most frequent 1000 words in English, 0.98% of words belonged to the Academic Word List and the rest (4.90%) were off list words i.e. words that do not belong to the other three categories. Most of the off-list words in the text were proper nouns. Since most words in the text belonged to the K1 category, it was assumed that the text was within the lexical competence of the participants.

3.5.3 Pre and post-tests

The pre-test (Appendix D – target items highlighted), which was conducted to verify whether the participants were familiar with the target structures, consisted of a Sentence Reconstruction (SR) task and a GJ task. In GJ tasks, learners are expected to provide a metalinguistic judgement (R. Ellis, 1991). In the type of GJ task that I used, the participants had to distinguish between well-formed and ill-formed sentences (correct/incorrect). R. Ellis (1991, p.162) notes that this type of GJ task is 'purely intuitive' and learners usually base their answers on implicit knowledge, in particular, if the task is administered with a time pressure.

GJ tasks are believed to isolate structures from communicative functions of the language (Schutze, 1996). Thus, another task that is able to measure the explicit knowledge of the learners in a more communicative functional base was also necessary. As a result, an SR task was also included in which the participants had to reconstruct a sentence keeping the meaning of the prompt sentence the same. Among the 20 SR items, 4 were *causative had*, 4 were past perfect and the others were distracter items. In these SR items, the participants had to rewrite the given sentences starting with the word/s provided so that the new sentences meant exactly the same as the first sentence. One example of a *causative had* SR item is:

> My sister paid someone to paint her house.
>
> My sister had ....

An example of a past perfect SR item is:

> She read more than 100 pages before going to bed.
>
> By the time she ....

As Task 2, 50 GJ items were given, of which 9 were *causative had* (5 correct, 4 incorrect items), 9 were past perfect (5 correct, 4 incorrect items) and the others were distracter items. In these items, the participants had to decide whether the sentences were grammatically correct or incorrect. The following was a target GJ item in the test.

	Correct	Incorrect
She had cut her hair by her sister.		

All the questions related to the two target structures in the pre-test were given to 10 native speakers and their answers were analysed to revise the items.

An identical post-test was not used in this pilot study since its main purpose was to trial the instruments. However, the participants had to answer 7 open-ended CQs (Appendix E) in the post-exposure test. The aim of this task was to examine whether awareness was involved when the participants paid attention to target constructions (target in this task was only *causative had*). There were 3 target questions (Questions 1, 6 and 8) along with other CQs used as distracter items. One of the target questions was:

> Q: Who painted James' house?
>
> A: …………………………..

In order to provide the correct answer, the participants had to understand the form-meaning relationship of the target structure.

3.5.4 WM tests

Since it was considered that measuring WM capacity in terms of PSTM and the three processes of the CE (updating, shifting and inhibition) would provide detailed insights into how the functioning of WM is related to conscious attention in processing L2 input, four different WM tests were used in this study. The four tests were selected based on previous research in the field of cognitive psychology. The Digit Span test (either forward or backward) has been used by Briscoe and Rankin (2009), Pickering and Gathercole (2001), Gathercole, Willis, Baddeley and Emslie (1994) and Shahabi, Abad and Colom (2014) for the assessment of the capacity of the PSTM. Updating is often measured by Keep Track tasks (Zheng et al., 2012; Shahabi, Abad & Colom, 2014; Tamnes et al., 2013; Wilhelm, Hildebrandt & Oberauer, 2013) and Plus Minus tasks have been applied to assess attention-shifting capacity (Hull et. al., 2008; Puric & Pavlovic, 2012; St Clair-Thompson & Gathercole, 2006; Tamnes et al, 2010). The Stroop task has been used widely in psychology as a tool to measure inhibition (Zheng et al, 2012; Shahabi et al., 2014).

The most common test that has been used to measure PSTM capacity is a Digit Span test, and thus a forward Digit Span test (Appendix F) was used in this pilot study. It had 9 sets of numbers starting with three-digit spans and going up to eleven-digit spans. Each span contained 2 options (e.g., Span 1: First option 3, 8, 6 Second option 6, 1, 3). The test also contained a set of practice tasks with two three-digit spans. The longest span that the participants could remember without a mistake was taken as the actual digit span.

Three tests were designed to measure the different functions of the CE. Based on the description that Miyake et al. (2000) provide, a Plus Minus task (Appendix G) was designed to measure attention shifting. The participants were given 3 lists of 2 digit numbers (30 in total in each list). In the first list, they were asked to add 3 to each number and write the answer, in the second list they were asked to subtract 3 from each number and write the answer and in the third list they had to shift between adding and subtracting 3 from alternate numbers. The participants were asked to complete each list as quickly as possible paying attention to the accuracy of answers. The time spent on the completion of each list was measured by a stopwatch. The shifting cost between addition and subtraction was "calculated as the difference between the time to complete the alternating list and the average of the times to complete the addition and subtraction lists" (Miyake et al., 2000, p. 65).

A Keep Track task (Appendix H) was used in this study to measure updating. It was also designed based on the description offered by Miyake et al. (2000). In this task, words from 6 different categories: animals, colours, countries, distances, metals and relatives were used. Each category contained three words (for example, animal words were cat, horse, cow). First, the participants were shown all categories and all words that would be used in the experiment. Then they were told that they would see words that belong to one category presented in a serial random order and they had to remember the last word of the category. Then they had to write down all the last words of the categories presented to them in each trial. For example, in the first practice trial, the words presented were Iran, Canada, Japan, iron, gold, silver, uncle, sister, daughter and at the end of this series, the participants were expected to write down Japan (last word of the country category), silver (last word of the metals category) and daughter (last word of the relatives category). When they were presented the words, the target category (e.g., country) was given at the bottom of the screen. The total number of words that the participants had to recall was 27 in 6 different series (3 trials with words from four categories and 3 trials with words from five categories) excluding the two practice trials with words from three different categories. The proportion of correctly recalled words was taken as the measurement of updating.

An internet based Stroop task (cognitivefun.net-test2) was used to measure inhibition. In this task, the participants were shown names of colours (e.g., red, blue, purple) and the letters on the screen also appeared in a particular colour. They had to press the first letter of the colour of the letters that they could see, not the first letter of the colour mentioned in the word (e.g., if the word that appeared on the screen was ‘green’, but if the colour of letters in the word was red, they had to press ‘r’ on the key board). The measurement taken was ‘interference time’ i.e. the difference between the time taken to press the keys when the word matched the colour of the letters and when the word did not match the colour of the letters. The programme itself provided the interference time.

The language in which the WM tests should be administered is a debated issue in WM literature. Osaka and Osaka (1992) who trialled a Japanese and an English version of a reading span test found very high correlation between the two (.84), which indicates that this test is language independent. Another study by Osaka, Osaka and Groner (1993) with French-German bilinguals also reports a high correlation (.85) between the L1 and L2 reading span tests. While two other studies also found high correlation between L1 and L2 WM tests (Hummel, 1998 & Miyake & Friedman, 1998) two others report moderate correlation (Berquist, 1997; Harrington & Sawyer, 1992). In addition, Gass and Lee (2011) note the difficulty of administering WM tests in the participants' L1. Taking into consideration that WM tests can be language independent at least to a certain extent and the practical difficulty of finding L1 tests, I decided to use the English versions of the Keep Track and the Stroop tasks. Linck et al. (2013), who comment on the language of the WM tests, highlight that it is important to consider the proficiency level of the participants if tests are administered in L2. The Stroop task selected for this study contained words for colours (e.g., red, blue, green, purple, yellow, orange), which were K1/K2 words (Heatley, Nation, & Coxhead, 2002) and 15 words out of 18 used in the Keep Track task were also K1/K2 words (Heatley, Nation,

& Coxhead, 2002). The rest of the words in the Keep Track task were off list words (country names) also familiar to this sample of participants. Thus, it was assumed that the language in these tests was within the participants' current proficiency level.

3.6 Data collection

3.6.1 Research design

The scope of the pilot study was narrower with a limited number of participants. It used only one type of input i.e. input flood. Further, it piloted the four WM tests and a pre and a post-test. Figure 3.1 illustrates the research design of the pilot study. The data was collected in two sessions. In the first session, the participants were informed of the research design and consent was obtained to use their data in the study. Then they filled in a background questionnaire and took a pre-test and the four WM tests. In the second session, i.e. 2 weeks after the first, the participants individually read the input text on an eye-tracker. They were not provided with any explicit instruction on the target structures. Then they read the text again on paper and answered the CQs in the post-test.

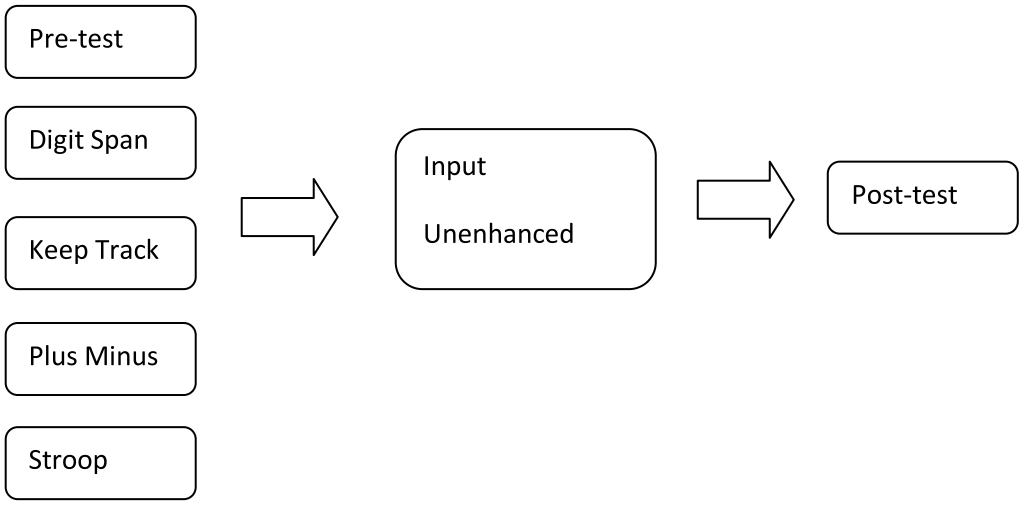


Figure 3.1. Research design of the pilot study

3.6.2 Eye-tracking

In this study, the participants read the input text on a Tobii TX 300 eye-tracker. Before the experiment, the input text was transformed into eye-tracking slides (Appendix I and see Figure 3.2 below for a sample), 7 in total. The instruction slide, (Appendix I, slide 1), was also read by the participants in addition to the verbal instructions that they received. Before deciding the font size and spacing of the input text, different font sizes and spacing were piloted and it was realized that data analysis was easier with a larger font size and sufficient spacing between the lines. This was mainly due to the necessity of marking AOIs for data analysis. First, the slides were prepared on PowerPoint using 24-point, double-spaced Calibri font. Although Calibri is not a fixed-width font, it is considered suitable for screen display (Ericson, 2013). This font size and the spacing provided sufficient room to mark AOIs. Four to six lines of text were included in each slide. This was based on the number of complete sentences that could be included on slide. Only the complete sentences were included in order to prevent the participants moving back to a previous slide to read the first half of a sentence. Depending on where the target phrases occurred in the text, it was not always possible to ensure that all slides contained the same number of target examples. This ranged from 2 to 6 examples on a slide.

Then the slides were transferred to the Tobii eye-tracking software. According to Holmqvist et al. (2011, p. 69) learners are more likely to pay attention to the centre of the screen on a monitor than to the edges; therefore, the text was placed in the middle of the slide. In addition, the brightness of the screen in the slides was made similar to the brightness of the calibration screen in order to maintain the data quality. The AOIs (examples of the target structures) were identified before designing the slides and all the words belonging to a particular example (e.g., I had my hair cut) were placed in the same line in order to facilitate measuring fixation duration on AOIs (see Figure 3.2 below).

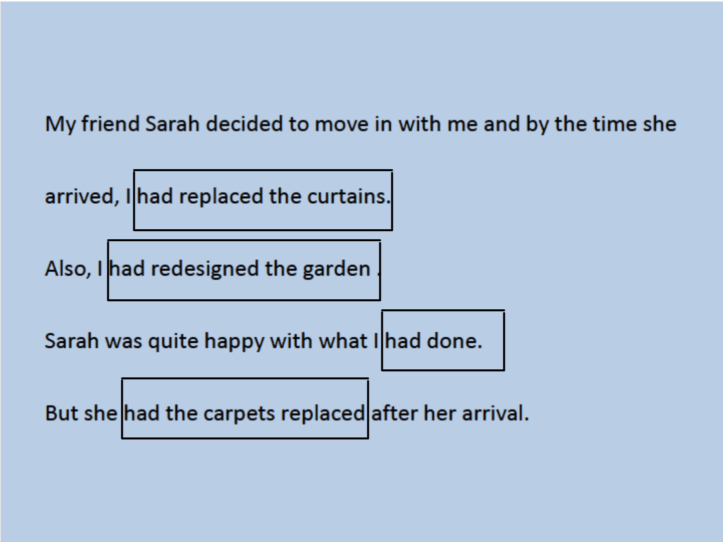


Figure 3.2. An eye-tracking slide (AOIs are highlighted)

Before the eye-tracking phase of the study began, the calibration was done on all 9 points of the screen. Then I gave verbal instructions to the participants. No time limit was set for reading. First, the participants read the instructions slide and then they continued reading the other slides. The participants were seated approximately 80cm away from the computer screen.

3.6.3 Pre and post-tests

The pre- and post-tests were paper based tests. The participants took the pre-test as a group. They were given Task 1 first and the maximum time given was 20 minutes. Then, they received Task 2. The maximum time allocated to this task was 40 minutes. After reading the input text on the eye tracker, they received the paper based post CQ task. They had 10 minutes to answer the questions. Instructions for the tests were printed on the paper (Appendix D).

3.6.4 WM tests

The participants took the four WM tests individually. In the Digit Span test, first they read the instructions slide on the PowerPoint presentation (Appendix F, slide 2), and I also gave a verbal explanation. Then they completed the first two practice spans on paper. I monitored the process and offered another explanation to the participants who did not understand the process in the two trials. After the trials, they took the test.

In the Keep Track task a similar method was applied using a PowerPoint presentation. The participants read the instructions slide (Appendix H, slides 2-4) and then listened to the verbal instructions. Next, they completed the two trials and the answers to the trials were shown on the screen. Those who could not complete the trials successfully received another explanation. At the end of the PowerPoint presentation, the answer sheet was collected.

The Plus Minus task was a paper based test. First, the participants read the instructions on the paper (Appendix G) and also listened to the verbal instructions. Then they were asked to complete the first column adding 3 to the numbers given, and the time to complete the column was measured by a stop watch. The same process was repeated for columns 2 and 3.

In the Stroop task, the participants received verbal instructions based on the web page (see Figure 3.3). Then, they had to engage in the task by pressing the correct key on the computer key board for 30 trials. At the end of this, the answers were saved under a pseudo name for each participant on the website itself. Then the PDF documents of the results were downloaded.

Figure 3.3. Screen shot of the online Stroop task homepage

3.7 Data analysis

3.7.1 Eye-tracking data

[9]As mentioned in 2.4, two main eye-tracking measurements were used in this study to investigate the amount of attention paid to the target items; TFD and the difference between observed TFD and expected TFD (this measurement is symbolised as ΔOE).

[9] Some parts of this paragraph were published in Indrarathne & Kormos (2016).

The second measurement, which has not been used in SLA studies before, was taken considering the possibility that participants may spend extra reading time (pay more attention to) on certain language features that they notice compared to other text presented in the slide while processing the input. In cognitive psychology research, Ferreira and Clifton (1986) used the number of characters including character spaces and punctuation marks of a region (i.e. AOI) to calculate the amount of reading time spent on the region. Taking insights from this study, Wilson and Garnsey (2009) calculated the expected reading time and actual reading time spent on the target items. However, Trueswell, Tanenhaus and Garnsey (1994) argue that this measurement is not effective when calculating reading time on smaller regions.

In order to investigate which measurement would be more plausible for this study, I used the number of syllables in AOIs (see table 3.1), number of letters in AOIs and the area of the AOI in square centimetres as a proportion of the number of syllables on the whole page, number of letters on the whole page and the total text area of the whole page in square centimetres. (These measurements did not include the edges of the slides, i.e., I used the text and only about 1 cm space surrounding the text). For example, the number of syllables in causative item 3 (had the roof repaired) was 5. The number of syllables on the page where this item occurred was 86. The expected TFD on the item – had the roof repaired – was calculated based on the proportion of number of syllables that the AOI had in relation to the number of syllables on the whole page (5/86 X TFD on the whole page). Then the difference between the expected TFD and the observed TFD was calculated. The same process was used when the measurement of the number of letters was taken (e.g., number of letters in causative item 3 was 18 and the total number of letters on the page that this item occurred was 267). Total area in centimetres was measured by a ruler on paper copies (the proportion of square centimetres between the AOI and the whole text area on a slide).

Table 3.1

Number of syllables in target examples

Item No.	Example	No. of syllables
Structure 1.1	had it painted	4
Structure 1.2	had the chimney replaced	6
Structure 1.3	had the roof repaired (1)	5
Structure 1.4	had the carpets replaced	6
Structure 1.5	had my hair cut	4
Structure 1.6	had the materials delivered	9
Structure 1.7	had the roof repaired (2)	5
Structure 1.8	had the carpets changed	5
Structure 1.9	had all the DIY activities done	11
Structure 2.1	had replaced	3
Structure 2.2	had changed	2
Structure 2.3	had redesigned	4
Structure 2.4	had done	2
Structure 2.5	had bought	2
Structure 2.6	had removed	3
Structure 2.7	had ordered	3
Structure 2.8	had painted	3

The data was first collated in an Excel spreadsheet and later transferred into SPSS 21 software for statistical analysis. The following variables were calculated for data analysis.

First, the TFD (raw) of each participant on each AOI was transferred from the Tobii software to the Excel spreadsheet. Then the mean TFD for each participant for all AOIs was calculated. This calculation was done for each text separately. The data was included in the SPSS analysis as TFD-P.

$$\text{Mean TFD of participants (TFD-P)} = \frac{\text{TFD of a participant for all AOIs}}{\text{Number of AOIs}}$$

The mean TFD for each AOI was also computed. The data was included in the graphical representations of the results as TFD-AOI.

$$\text{Mean TFD of an AOI} = \frac{\text{TFD of all participants of the AOI}}{\text{Number of participants}}$$

For the second eye-tracking measurement, first the TFD for each participant for the whole page (raw) was transferred from Tobii software to the Excel spreadsheet. Then expected fixation duration was calculated for each AOI.[10]

$$\text{EFD of an AOI} = \frac{\text{no. of syllables of the AOI}}{\text{no. of syllables of the whole page}} \text{ X TFD-whole page}$$

Then the difference between observed and expected TFD was calculated again for each participant for each AOI.

[10] Example: On a slide that contained a text with 30 syllables if a participant spent 2400 milliseconds on reading the slide, the expected fixation duration for the target construction that consisted of 10 syllables was 800 milliseconds (10/30 x 2400). If the participant's actual fixation duration was 900 milliseconds, then the ΔOE value was 100 milliseconds.

Difference = TFD raw – EFD raw

Finally, the mean value of the difference for each participant for all AOIs was calculated and used in the SPSS analysis (∆OE-P).

$$\text{Mean difference (∆OE-P)} = \frac{\text{TFD raw – EFD raw for a participant for all AOIs}}{\text{Number of AOIs}}$$

The mean difference for each AOI was also computed. The data was included in the graphical representations of the results as ∆OE-AOI.

$$\text{Mean difference of an AOI} = \frac{\text{TFD raw – EFD raw of all participants of the AOI}}{\text{Number of participants}}$$

Eye-tracking data of two participants in this study were not usable because either the eye- tracker was unable to track the eye movement of such participants or data was of poor quality. Such data was excluded from the analysis.

3.7.2 Pre and post-tests

In the pre-test, 1 mark was allocated to a correct answer and 0 to an incorrect answer in both SR and GJ items. In SR items, if the form of the target structure was written correctly (e.g., *had* a letter *delivered)*, a mark was awarded and spelling mistakes were ignored. For example, if the word *dilivered* was written, a mark was awarded; however, *had… deliver* was marked incorrect. Half marks were not given for these items. After entering the marks for each SR and GJ item separately, the total score of the SR and GJ items was computed. In the post-test, 1 mark was allocated to each correct answer.

3.7.3 WM test data

In the Digit Span test, the longest span that a participant could remember without a mistake was taken as the span. For example, if a participant could provide the correct answer for the first three digit span (3, 8, 6), they could move to the second three digit span (6, 1, 3). If they provided the correct answer, then they were allowed to move onto the first four digit span (3, 5, 1, 7). If they made a mistake in the first four digit span, 3 was taken as the actual digit span of that participant.

In the plus-minus test, the time spent on the completion of each list was measured by a stop watch. The shifting cost between addition and subtraction was "calculated as the difference between the time to complete the alternating list and the average of the times to complete the addition and subtraction lists" (Miyake et al., 2000, p. 65).

Shifting cost = time to complete the alternating list – (time to complete the addition list + time to complete the subtraction list)/2

The total number of words that the participants had to recall in the Keep Track task was 27. The proportion of correctly recalled words was taken as the measure of updating.

$$\text{Keep Track score} = \frac{\text{Number of words recalled correctly}}{27} \times 100$$

In the Stroop test, the measure taken was interference time (see 3.6.4 for a description). The average interference time in milliseconds for each participant calculated and displayed by the website itself was taken as the Stroop measure.

3.8 Results

3.8.1 Pre-test

According to Table 3.2, the descriptive statistics of SR items demonstrate that *causative had* items had a higher mean value (*M* = .67) than the past perfect tense items (*M* = 1.00). The maximum possible score for both structures was 4. In the GJ items, the maximum possible score was 9 and the past perfect items had a higher mean value (*M* = 6.20) than the *causative had* items (*M* = 5.13). The scores were normally distributed in all item types except in SR items of past perfect tense. The results indicate that participants had some prior knowledge of both structures.

Table 3.2

Descriptive statistics of the pre-test results (N = 15)

	SR - Past Perfect	SR - Causative	GJ - Past Perfect	GJ- Causative
Maximum possible score	4	4	9	9
M	.67	1.00	6.20	5.13
SD	.90	1.31	1.52	1.81
Skewness/ Std. Error of Skewness	2.51	1.14	-0.18	0.33
Kurtosis/Std. Error of Kurtosis	1.47	1.01	0.78	1.02

3.8.2 WM tests

According to the descriptive statistics of the four WM tests given in Table 3.3, the mean value of the Digit Span test was (*M* = 7.67). The normative score of forward digit span for Italians aged 20-90 is reported to be 5.84 (Monaco, Costa, Caltagirone, & Carlesimo, 2013, p. 752), Hungarians aged 16-20 is 6.32 (Revesz, 2012, p. 119) and Chinese whose mean age is 35.4 is 6.8 (Ostrosky-Solís & Lozano, 2006, p. 338). This demonstrates that most of the participants in this pilot study had a relatively high capacity of PSTM. However, no similar comparative data is available for the CE tests; therefore, it is not possible to comment on the performance by the participants in the CE tasks.

Table 3.3

Descriptive statistics of the WM test results (N = 15)

	Digit Span	Plus Minus	Keep Track	Stroop
Maximum possible score	11	none	27	none
M	7.67	11.98	.68	1374.03
SD	1.17	7.53	.16	230.24
Skewness/ Std. Error of Skewness	1.78	1.81	-1.28	1.67
Kurtosis/Std. Error of Kurtosis	0.98	1.23	1.44	1.04

Table 3.4 illustrates the Spearman rank order correlation between the four WM tests. None of the correlations were statistically significant. These results suggest that the four WM tests measure different components and different functions of the WM.

Table 3.4

Correlation between WM test results (N = 15)

		Plus Minus	Keep Track	Stroop
Digit Span	Spearman rho	.417	.156	-.081
	p	.122	.578	.950
Plus Minus	Spearman rho		.266	-.048
	p		.338	.864
Keep Track	Spearman rho			-.447
	p			.095

3.8.3 Eye-tracking data

As can be seen in Table 3.5 and on Figure 3.4, TFD-AOI on each item increased with time in *causative had* examples. However, TFD-AOI in items 3 (had the roof repaired), 6 (had the roof repaired) and 9 (had the DIY activities done) demonstrated a decrease compared to the other items. In past perfect items, TFD-AOI seemed to fluctuate: item number 4 (had done) demonstrated a sudden decrease, while item number 7 (had ordered) also showed a decrease.

Table 3.5

TFD-AOI in milliseconds for all items

Structure	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9
Causative	960	1228	954	1353	1509	1186	1334	1595	1464
Past Perfect	725	886	1359	407	677	1614	805	1112	-

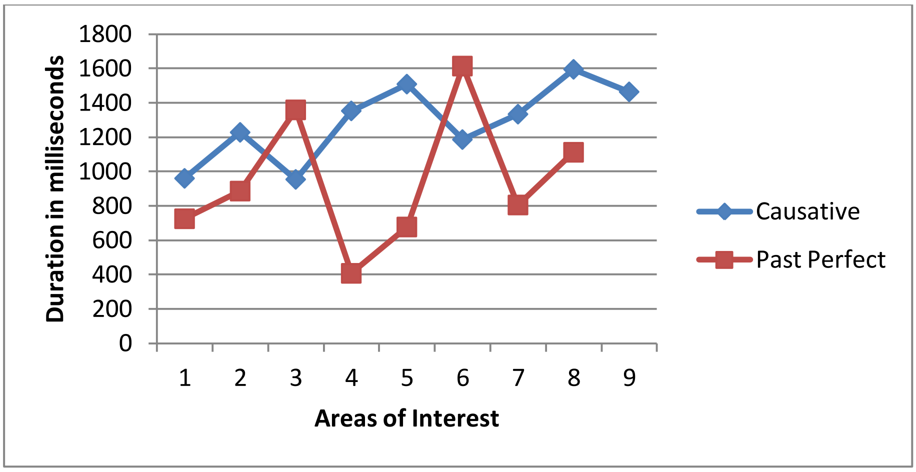


Figure 3.4. TFD-AOI for each example of the target structures (in milliseconds)

As discussed in 3.7.1, I used three measurements to calculate the ΔOE (number of syllables in AOIs, number of letters in AOIs and the actual length of AOI on the screen in centimetres as a ratio). According to Table 3.6 and on Figure 3.5, the measurement of the number of syllables indicates an increase in ΔOE-AOI on AOIs in both structures. According to Table 3.6, only past perfect item 4 (had done) and causative item 6 (had the roof repaired) and 9 (had the DIY activities done) demonstrate minus values i.e. the participants had not spent even the expected reading time gazing at these items. Causative item 9 that demonstrates a clear decrease in ΔOE-AOI has the highest number of syllables (11 syllables). Moreover, past perfect item 4 (had done) shows a decrease in ΔOE-AOI and has the lowest number of syllables (2 syllables). The measurements taken based on the number of letters and the actual length of AOIs showed a random pattern of eye-fixations. Therefore the letter and length-based calculations do not seem to yield any accurate insight into the attentional processing of the input.

Table 3.6

ΔOE-AOI (in milliseconds) based on the number of syllables in AOIs

Structure	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9
Causative	23	11	5	0.3	43	-71	7	32	-61
Past Perfect	11	12	32	-3	10	60	4	33	-

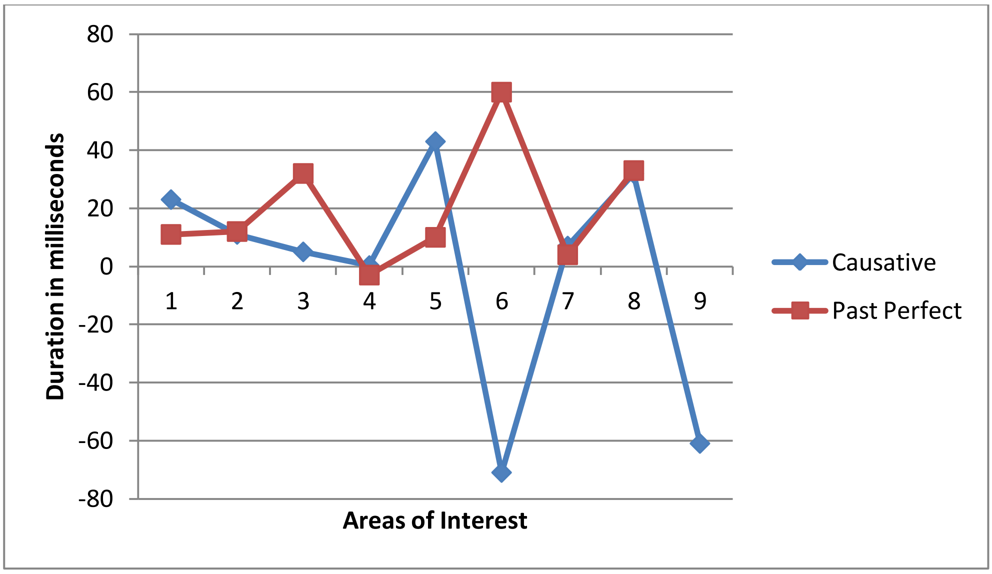


Figure 3.5. ΔOE-AOI based on the number of syllables in AOIs (in milliseconds)

Pearson chi-square test was run for all target items individually in order to measure whether there was a statistical significance in the increase of observed TFD-AOI compared to the expected TFD-AOI. Causative item 5, $X^2(1, N = 13) = 72.000$, $p = .024$, and item 8, $X^2(1, N = 13) = 97.000$, $p = .040$, showed a statistically significant difference from the expected value while past perfect item 3, $X^2(1, N = 13) = 44.057$, $p = .033$, item 6, $X^2(1, N = 13) = 76.570$, $p = .031$, and item 8, $X^2(1, N = 13) = 27.000$, $p = .021$, also showed a significant difference.

3.8.4 The relationship between eye-tracking and WM test data

According to Spearman rho correlation results presented in Table 3.7, none of the WM tests showed a strong correlation with the TFD-P of the two target structures. However, ∆OE-P of causative items and the Keep Track data indicated a positive correlation which is significant.

Table 3.7

Spearman correlation between eye-tracking data and WM test data (n = 13)

		1	2	3	4
Digit Span	Spearman rho	-.076	.383	.231	.560
	p	.787	.159	.726	.546
Plus Minus	Spearman rho	-.143	.179	-.126	-.456
	p	.612	.524	.113	-.791
Keep Track	Spearman rho	-.061	-.104	.502*	-.287
	p	.829	.712	.001	.087
Stroop	Spearman rho	-.004	.148	-.298	.190
	p	.990	.598	.078	.256

Note. 1 = TFD-P of *causative had* items, 2 = TFD-P of Past Perfect items, 3 = ∆OE-P of *causative had* items, 4 = ∆OE-P of Past Perfect items.
*Correlation significant at 0.05 level

3.8.5 Correlation between the pre-test results and eye-tracking data

As can be seen in Table 3.8, the pre-test score for the causative construction had a significant correlation with TFD-P, $r(13) = .551$, $n = 13$, $p = .033$, and ∆OE-P, $r(13) = .665$, $n = 13$, $p = .026$. However, the pre-test results of the past perfect items did not correlate with the TFD-P or ∆OE-P values.

Table 3.8

Correlation between pre-test results and eye-tracking data (n = 13)

		1	2	3	4
Pre-test Causative	Spearman rho	.551*		.665*	
	p	.033		.026	
Pre-test Past Perfect	Spearman rho		.160		.477
	p		.570		.621

Note. 1 = TFD-P of *causative had* items, 2 = TFD-P of Past Perfect items, 3 = ∆OE-P of *causative had* items, 2 = ∆OE-P of Past Perfect items.
*Correlation significant at 0.05 level

3.9 Discussion

i. *Can attention to L2 grammatical construction be reliably investigated by means of eye-tracking?*

The first research question of the pilot study attempted to establish whether eye-tracking could reliably investigate attention paid to L2 grammatical constructions in an input text. Eye-tracking data collected in this study provided some useful information related to input processing. For example, increased TFD-AOI and TFD-P with more exposure indicated possible attentional processing. The new measurement, ∆OE (∆OE-AOI and ∆OE-P), clearly indicated that the participants spent extra reading time on most of the target items which suggests potential attentional processing. The data analysis also revealed that two of the three measurements used to calculate ∆OE: namely number of letters in AOIs and length of AOI in square centimetres as a proportion of number of letters on the whole page and the total area of the whole page in square centimetres, did not provide accurate data that could provide an insight into attentional processing. However, the number of syllables in AOIs as a proportion of the number of syllables on the whole page was able to show extra reading time spent on target examples. Alderson's (2000) claim that "readers typically identify the sound of words as part of the process of identifying their meaning" (p. 14) may have played a role here since the participants in the experiment had to read a passage for meaning. Furthermore, Trueswell, Tanenhaus and Garnsey (1994) argue that letter-based standardizations are inaccurate. In English there is no one-to-one correspondence between letters and sounds and often letter combinations rather than individual letters denote phonemes. This results in the fact that beyond initial stages of learning to read, not letters but larger grain sized units such as onsets and rimes and ultimately syllables constitute the basic units of word-level decoding in English (Ziegler & Goswami, 2005). Empirical evidence also suggests that children's awareness of syllables is a strong predictor of their later reading ability (McBride-Chang & Kail, 2002; Morais, 2003). Thus, it can be concluded that ∆OE based on the number of syllables has the potential to provide insightful quantitative data on attentional processing.

ii. Can L2 learners at lower intermediate/intermediate level of proficiency develop awareness as a result of paying attention to unenhanced examples of grammatical structures in input?

The second research question that the pilot study attempted to answer was whether L2 learners at lower intermediate/intermediate level of proficiency are able to develop awareness as a result of paying attention to unenhanced examples of grammatical structures in input. There was an increase of the TFD-AOI with the number of exposures and ∆OE-AOI in all target examples except for two items. This showed that learners are likely to pay more attention to language features that frequently occur in the input even in unenhanced conditions as in this experiment. However, eye-tracking data could not provide any evidence of development of awareness of the target construction as a result of paying attention to the target examples. The post-test was mainly aimed at assessing the ability of the participants to develop awareness of the form-meaning mapping of the *causative had* construction as a result of paying attention to input. None of the participants provided correct answers for any of the target questions. This could be an indication that they had not developed awareness of the target items as a result of attention paid to target items. If I revisit the definition of awareness, the learners should be able to provide a report of what they have attended to or the attended stimuli should be registered in the WM if awareness is involved when paying attention. The inability to give accurate answers to the CQs does not provide clear evidence that awareness was not involved on this occasion because the questions did not ask the participants to provide a report of what they had attended to. It was also impossible to measure if the attended stimuli got registered in their WM. What the CQs had measured was if the participants had understood the meaning of the target construction. Thus, it can be concluded that CQs were not a reliable measurement of awareness.

iii. Is the WM capacity of the participants related to the amount of attention paid to input, if so which components of WM are most strongly linked to it?

The third research question investigated whether WM capacity of the participants was related to the attention paid to input and if so which components of WM are mostly related to it. The eye-tracking data and WM test data did not show any significant correlations except between Keep Track task and ∆OE-P of *causative had* items. Holmqvist et al. (2011) point out that there is no guarantee that WM registration occurs even though participants look at the objects on the screen of the eye-tracker. Therefore, it is possible that the participants did not pay conscious attention to target features on this occasion even though TFD and ∆OE increased with time. The small sample size could also be a reason for not showing any significant relationship between attention and WM.

3.10 Lessons learnt

The pilot study methodology was insufficient to provide a detailed discussion of the attentional processing of the input. As discussed in Chapter 2, attention with or without consciousness can develop L2 learning. Moreover, it is possible for learners to develop knowledge of the form without identifying the form-meaning relationship. This was not measured in the pilot study because it did not have an identical pre/post-test design. The CQs used in the post-test were targeted at gaining insights into whether the participants were able to create-form meaning mapping rather than investigating gains in the form. This confirmed the necessity of an identical pre/post-test design for the main study.

The findings also indicated that eye-tracking could offer an insight into attentional processing; however, it is unable to give information on whether awareness is involved. As a result, the main study used eye-tracking as a measurement of attention only. Furthermore, the main study used the number of syllables in AOIs as a proportion of the number of syllables on the whole page to calculate ∆OE since the pilot study indicated that this measurement was more accurate than the letter- or length-based variable. The definition of awareness discussed in Chapter 2 indicates that a report of the stimuli attended can be an indicator of awareness. Therefore, a post-exposure interview was included in the main study to identify the existence of possible awareness.

Another conclusion made based on the pilot study results was that using two structures that have different word orders may have confused the participants, or given them an extra processing load. Therefore, the focus of the main study was only on one target structure, *causative had*. The choice of this construction was motivated by: the assumption that the selected learner sample would have no or very little pre-existing knowledge of it; its meaning can be inferred from the context and there is a one-to-one form to meaning mapping; and it also forms an easily identifiable AOI for eye-tracking research. Further to these, causative construction in Sinhala (the L1 of the participants in the main study) is different to that of in English because Sinhalese is an SOV language and causative in Sinhala is marked by bound suffixes.

Providing a similar duration to complete each item in the GJ task was difficult in the written mode because the test takers have the possibility of revisiting a previous item. This minimises the accuracy of GJ task as a measurement of implicit knowledge because as R. Ellis (1991, 2015) points out, GJ tasks can involve explicit knowledge if not performed under time pressure. As such, the mode of input in the pre and post-tests was also reviewed in the main study instrument planning. In the pilot study pre-test, the participants took a paper-based written test (for both SR and GJ items). However, with the aim of changing the input modes, the main study pre and post-tests had writing (SR) and listening (GJ) items (Section 4.6.3 provides a detailed description of the revisions).

Moreover, it has been understood that the pre- and post-test SR and GJ items should be controlled for length to obtain accurate data. Although SR items could measure the form-meaning mapping of the target construction, neither the SR nor the GJ items could provide insights into the ability of learners to use the target structure in a meaningful context. Thus, a free writing task was also included in the main study pre- and post-tests.

The pilot study results did not indicate any relationship between attentional processing and the WM capacity of the participants. It is possible that in this input flood condition, learners did not engage in conscious attentional processing, and therefore, WM capacity was not influential. Another explanation might be that the low sample size did not result in significant findings. Even though WM tests did not show a significant relationship with attentional processing on this occasion, it does not exclude the possibility that WM may play a role in attention to input. Consequently, it was deemed necessary to investigate the relationship between WM and attention further. Therefore, the four tests that were used in the pilot study were also administered in the main study. For the same reasons listed in Section 3.5.4 and considering the context of where main study data collection was carried out (please see 4.6.5 for the discussion), I decided to use the English versions of the Keep Track and the Stroop tests.

Items 1.6 (had the materials delivered), 1.9 (had all the DIY activities done) and 2.4 (had done) showed a decrease in both TFD-AOI and ∆OE-AOI. The decrease in item 1.6 could be due to the fact that the structure had been used earlier; therefore, the participants might have paid less attention to it when it occurred the second time. Item 1.9 contained the acronym DIY which could be the reason why it was paid less attention to, in particular, if that was an unfamiliar term for the participants. The decrease of TFD-AOI in item 2.4 with 2 syllables and item 1.9 with 11 syllables is a possible indication that items with more syllables and items with fewer syllables are less likely to receive attention. Therefore, the target examples used in the main study had 4-8 syllables and none of the examples were repeated.

Based on the lessons learnt from the pilot study, the methodology of the main study was developed. Chapter 4 outlines the methodology applied in the main study.

Chapter 4: Main study design

4.1 Introduction

The pilot study discussed in Chapter 3 provided several methodological insights into how the methodology of the main study should be designed. For example, the data collection methods needed to be amended, the pre- and post-tests had to be redesigned and input text had to be revised. This chapter explains the changes made.

The chapter begins with the research questions answered in this study (4.2) and provides a detailed description of the context where the data collection took place (4.3). It then explains the background of the participants (4.4), ethical procedure (4.5), study instruments (4.6), data collection process (4.7) and data analysis methods (4.8). If a similar process to that of the pilot study was applied in any section, a description is not provided. Instead a reference to the relevant section in Chapter 3 is given.

4.2 Research questions of the main study

This study attempted to answer several research questions categorised into three sets. The first set of questions investigates the role of attention in different input conditions, the second set analyses how WM is related to input processing. The third set discusses the role of awareness in input processing. The research questions are:

Section 1[11]

RQ 1: How does knowledge of the form and meaning of the target syntactic construction *causative had* as measured by a production and a comprehension task change under explicit and implicit instructional conditions in the case of Sri Lankan pre-intermediate/ intermediate level English language learners?

RQ 2: How does attention paid to examples of the *causative had* construction in the input texts differ under explicit and implicit instructional conditions?

RQ 3: How is the change in knowledge of the target construction under explicit and implicit instructional conditions related to the attention paid to examples of the *causative had* construction in the input texts?

Section 2

RQ 4: How is the functioning of WM including both PSTM capacity and CE functions related to the change of knowledge of the TGC *causative had* in different input conditions?

RQ 5: How is the functioning of the WM related to attention paid to the target items in different input conditions?

Section 3

RQ 6: Is there an improvement in the CQ scores in different input conditions between the three exposures?

RQ 7: Is the improvement in the CQ task related to the eye fixation measures, the knowledge of the TGC *causative had* (SR and GJ tasks) and WM?

[11] These RQs were published in Indrarathne & Kormos (2016).

RQ 8: Is there a difference between the change in the knowledge of the TGC *causative had* (SR, GJ and comprehension tasks) between the participants who reported awareness and those who reported no awareness? Is the awareness reported in different input conditions significantly different to each other?

4.3 Context[12]

The main study data collection was done in a university in Sri Lanka. The higher education institutes in Sri Lanka fall under three main categories. The first is the 15+ state universities under the direct authority of the University Grants Commission (UGC). The second category is the 4 other public universities under the authority of different ministries but not under the direct authority of the UGC. However, the UGC plays an advisory role in their administration and quality control. The third category comprises of several other private and non-profit higher educational institutions, which offer degree level courses, but are not recognised as universities. The data of this study was collected in one of the four universities mentioned above under the second category.

This university is run by the Ministry of Defence, primarily because it admits military students and was a military training academy in the 1980s and 1990s. However, the structure of the university underwent changes over the time and now it admits both military and civilian students to several faculties: Engineering, Medicine, Allied Health Sciences, Law, Management Social Sciences and Humanities, Computing and Built Environment and Spatial Sciences. All undergraduate courses in all universities under

[12] Some parts of this section were published in Indrarathne & Kormos (2016).

the direct authority of the UGC in Sri Lanka are offered free of charge. However, civilian students in the said university follow undergraduate degree courses on a fee-levying basis. This is because this university does not receive funds from the UGC since it is not under the direct authority of the UGC. The military students are not required to pay the course fee since they are funded by the Ministry of Defence.

There is a limited number of universities in Sri Lanka and thus they can only admit ca. 25,000 students each year. However, about 400,000 students take the university entrance General Certificate of Education Advanced Level (GCE A/L) examination every year and nearly 50% of them qualify to enter university. As a result, there is huge competition since the students who score higher marks in each district can enter a state university where there are no course fees. Due to this vacuum of opportunities, many private institutions attract students who do not qualify to enter a university. The university where data collection was done also attracts such students.

The student population of the university is comprised of those from all parts of the country and almost all of them belong to first language Sinhala or Tamil communities. The medium of instruction in the majority of state run schools in the country is either Sinhala or Tamil and about 10% of these schools have English medium classes for students who prefer studying in the English medium. However, the said university receives the majority of its students from Sinhala medium schools. In such schools English is taught as a compulsory subject from Grade 1. Therefore, these students have learnt English language for nearly 13 years at school by the time they enter the university.

The university administers both undergraduate and postgraduate degree courses through the medium of English. In addition, English language is taught as a credit bearing course at the university for all undergraduate students. The English language curriculum is comprised of four parts taught in the first and second years of the undergraduate degree course. It focuses on both EAP and English Language. EAP related components such as summarising, paraphrasing, structuring essays and academic presentation skills are spread within the four parts of the curriculum. The English language component covers listening, speaking, reading and writing skills along with vocabulary and grammar. EAP related components of the curriculum are taught based on the materials related to the specific disciplines, and language components are taught based on general English materials. All materials were developed in-house and the British Council in Sri Lanka assisted in the materials development process. During this process, CEFR was taken as a basis for the English language materials and further insights from popular English language course books such as *Cutting Edge*, *English File* and *Headway* were taken into consideration. Since most of the students that the university receive belong to either B1 or B2 levels, the materials are aimed at B1 level to C1 level.

Usually there are 20-25 students in an English class at the university and the teachers are encouraged to use interactive methods in class. The materials encourage a lot of pair and group work promoting language use. Many audio and video materials are also used in class in order to improve listening skills. Students also engage in group presentations, discussions, and projects to improve their English language skills. Thus, the curriculum and the materials promote the use of different English language teaching methods such as communicative language teaching and task based language teaching blended together. Moreover, content and language integrated learning method is used when teaching the EAP related components of the curriculum.

4.4 Participants[13]

The participants in this study were 100 first language Sinhala speakers who were first year undergraduate students from the Bachelor of Commerce degree programme of this university. By the time they participated in the study, they had had spent 5 months at the university. The participants were 19-21 years of age (*M = 20.04, SD = 0.81*) and there were 29 females and 71 males. Since the university runs English medium degree courses only, students need to take an English language admission test during the admission process. The minimum requirement for admission to the university is B1 level of CEFR. The participants in this study were either at the B1 or early B2 levels based on the scores of the admission test.

All the participants in the main study had been participating in English language lessons since they joined the university. Further, all the participants had experience of learning English for more than 10 years. Therefore, a pre-test that contained items of the target structure was administered to 107 participants. One participant was excluded from the study since she demonstrated substantial existing knowledge of the target structure (5/6 in the SR task and 8/10 in the GJ task). Two others were excluded because they provided accurate answers to two SR items (2/6) in the pre-test. Among 100 participants who took part in the study, 97 scored 0 in the SR task. The others 3 scored 1 in the item with an irregular past participle (*put*). Thus, it was assumed that

[13] Some parts of this section were published in Indrarathne & Kormos (2016).

they did not have knowledge of the target construction, but the correct answer was provided to this item by chance. Any accurate answers to GJ task were taken being due to chance. The maximum score obtained in the GJ task was 7/10 by 3 participants. Three others were excluded because they mentioned in the pre-experiment background questionnaire that they had language learning disorders. The other student withdrew his participation due to a personal reason. None of the 100 selected participants were bilingual speakers of languages and the other additional language that they spoke was English.

4.5 Ethical procedure

All the participants were given an ethics form (Appendix J) and a description of the study written according to the guidelines provided by the Linguistics and English Language Department of Lancaster University. I also explained the purpose and the procedures of the study and the withdrawal process to all participants. Following this, the signed ethics forms were collected from the participants.

4.6 Instruments

4.6.1 Background questionnaire

As a first step, the participants filled in a background questionnaire (Appendix K) containing 12 questions. The first six were aimed at collecting basic information such as age and gender. Questions 8 and 9 were related to their language learning experience and questions 7, 10 and 11 were about their language speaking background. Question 12 asked the participants to indicate if they had any language learning disorders. The answers given to these questions were considered in the initial filtering process of the participants.

4.6.2 Input texts[14][15]

Three input texts were used in the eye-tracking sessions. I wrote all three texts (Appendix L). This is because I could not find texts that fulfil all the requirements of this research (e.g., number of examples, vocabulary within the participants' proficiency level and similar length). Then the texts were also checked for complexity by Vocabprofile (Heatley, Nation, & Coxhead, 2002) to verify whether the vocabulary fell within participants' proficiency level. Table 4.1 summarises the results of the analysis of the complexity of vocabulary. All three texts contained more K1 words than K2 words. Text 1 contained a total of 93.56% K1 and K2 words, Text 2 a total of 91.74% of K1 and K2 words and Text 3 a total of 91.84% of K1 and K2 words. Most of the off-list words in the texts were proper nouns. The few words that belonged to the academic word list were kept due to the fact that they were necessary to maintain the coherence of the story in the texts. Since most words in the texts belonged to the K1 category, it was assumed that the texts were within the lexical competence of the participants.

---

[14] Some parts of this section were published in Indrarathne & Kormos (2016).
[15] Included in the IRIS data base

Table 4.1

Complexity of vocabulary in the input texts

	Text 1	Text 2	Text 3
Total words	230	227	230
K1 words	87.12	80.87	82.83
K2 words	6.44	10.87	9.01
Academic Word List words	0.43	0.87	2.58
Off list	6.01	7.39	5.58

The target structure was *causative had*.

had + article/determiner + N + (V) Past Participle

had the tools delivered

The texts contained 21 examples of the target structure and they were taken from the British National Corpus (BNC). Slight alterations were made in nouns (e.g., had the *tools* delivered instead of had the *letters* delivered) to make the examples compatible with the overall story line. The examples of the target structures used in the pilot study contained different number of syllables ranging from two to ten. As previously mentioned, the results indicated that the examples with fewer syllables and many syllables received less attention. Thus, the examples used in the main study contained 4-8 syllables. Table 4.2 includes examples used in each text, BNC frequencies and the number of syllables.

Table 4.2

Number of syllables and BNC frequencies of target examples

Text	Item No.	Item	No. of syllables	BNC frequency of construction *(instances per million words)*	BNC frequency of main verb *(instances per million words)*
1	1.1	had the walls painted	5	0.02	43.72
	1.2	had the roof repaired	5	0.04	18.32
	1.3	had the curtains replaced	6	0.01	110.87
	1.4	had the tools delivered	6	0.01	64.46
	1.5	had the carpets changed	5	0.02	270.35
	1.6	had the garden cleaned	5	0.01	39.2
	1.7	had a door made	4	0.13	2138.73
2	2.1	had her blood tested	5	0.001	70.09
	2.2	had the reports sent	5	0.02	245.9
	2.3	had her hair cut	4	0.02	182.8
	2.4	had a letter written	6	0.04	399.02
	2.5	had her photograph taken	7	0.07	1765.87
	2.6	had sugar added	5	0.01	273.55
	2.7	had the bus stopped	4	0.02	245.81
3	3.1	had the car left	4	0.001	627.77
	3.2	had the car checked	4	0.03	97.72
	3.3	had the camera fixed	6	0.01	45.7
	3.4	had the letters printed	6	0.06	34.99
	3.5	had his photographs developed	8	0.01	237.72
	3.6	had his breakfast put	5	0.13	688.55
	3.7	had his shoes mended	5	0.001	6.09

Even though the reading of the texts on the eye-tracker was part of an experiment, it was necessary to give participants a purpose for reading. Therefore, each reading text contained four multiple choice reading CQs for the participants to answer. Two CQs out of the four for each text were based on the general understanding of the text and the other two were aimed at measuring whether the participants could make the form-meaning mapping of the target structure. For example, in Text 1 the first CQ was,

Who repaired the roof of James' house?

a. James
b. Sarah
c. Peter
d. Someone else

In order to provide the correct answer, the learners needed to understand the form-meaning mapping. For example, in the text, it was mentioned that "James moved to a new house six years ago …. He noticed water on the floor in some places, so he had the roof repaired because he knew he couldn't do that himself."

4.6.3 Pre-test and post-test[16][17]

In the pre-test of the main study (Appendix M – target items are highlighted), three tasks were used. Similar to the pilot study, SR items (20 in total) and GJ items (40 in total) were included as the first two tasks and in addition, a free production (writing) task was included.

[16] Some parts of this section were published in Indrarathne & Kormos (2016).
[17] Included in the IRIS data base

One SR item was,

Sara got someone to print invitation cards for her party.

Sara had ....

One GJ item was,

	Correct	Incorrect
My dad had his lunch delivered to his office yesterday.		

There were 6 target SR items and 10 target GJ items (5 correct and 5 incorrect) in the pre-test. Among the five incorrect GJ items, two contained an *–ed* morpheme omission error (e.g., My sister had her new house decorate before the weekend); three word order errors and one instance of the use of the wrong suffix *–ing*.

In order to minimize the effect of the length of sentences, all target SR items and 8 target GJ items contained 10 words each. The remaining two target GJ items contained 11 words each. The test included two writing tasks (SR items and production task) and a listening task (GJ items). The participants took the pre-test as a group and they wrote answers to the SR items on the test paper. Then the answer sheet (Appendix N) of the GJ items was handed out and the recording was played once for the participants to put a tick or a cross on the relevant column (correct/incorrect). In the recording, I read out the 40 sentences with a 5 second break between sentences. After this aural version of the GJ task, the participants received Task 3, a free production task (Appendix M).

The post-test (Appendix O – target items are highlighted) had a similar design and the free production task was the same in the pre- and post-test phases. However, the proper nouns and some other nouns in the SR and GJ items were changed in the post-test to distract the participants from remembering the answers that they provided in the pre-test.

All the sentences (SR and GJ items) used in the pre and post-tests were taken from the BNC; however, the nouns were altered and phrases were added to the sentences in order to maintain the length of sentences. The pre-test was piloted with 6 students who were taking a pre-sessional English language course at Lancaster University and 37 students at the same Sri Lankan university (none of them were participants in the main study). These six participants were at B2 level of proficiency and they were unable to provide correct answers to any of the SR items. The 37 Sri Lankan learners were at B1/B2 level of proficiency and none of them provided accurate answers to the SR items. Thus, it was predicted that learners at B1/B2 levels may not have pre-existing knowledge of the target structure.

4.6.4 Post eye-tracking interview

The post eye-tracking interview contained 8 questions (Appendix P). Questions 6 (What did you notice in the texts?) and 7 (Did you notice any particular grammatical structure in the texts?) were directly aimed at investigating whether the participants showed awareness of the target items. Question 8 (If your answer to question 7 is ‘yes’ can you please write down that structure?) was included to identify whether the participants gained any awareness of the target structure and if so whether they had registered the form of the target structure in their memory. These questions were predicted to give information on if awareness was involved when paying attention. Question 5 was included in the interview in order to analyse whether the participants had faced any difficulties reading on the computer. Questions 3 and 4 provided insights into the difficulty level of the input texts whereas questions 1 and 2 focused on the overall impression of the participants of the reading tasks. The interview answers were taken into account in the data analysis.

4.6.5 WM tests

The four WM tests discussed in 3.5.4 were used in the main study in order to collect data on the WM capacity of the participants. No changes were made to the designs of the WM tests used in the pilot study since no methodological issues emerged when these tests were administered when piloting. The four tests used were Digit Span, Keep Track, Plus Minus and Stroop.

As mentioned in Section 3.5.4, both the Digit Span and the Plus Minus tasks allow learners to perform these tasks in their L1 or any other language they select. The two tests in which the participants used English were the Stroop and the Keep Track tasks. These tests were only available in a standardized English version, but they both used very frequent vocabulary that participants were assumed to be familiar with. Administering these two tests in the learners' L2 might not have created an interference with participants' level of proficiency for a number of reasons. First of all, the learners were at a similar level of proficiency and the English words in these two tests belong to K1 and K2 levels i.e. within the proficiency level of the participants. Second, the university, where the main study data collection was carried out, expects the students to use only English within the institutional context even in informal interactions. Finally, due to the lack of resources at school level in the country, most of the learners at this university had not frequently used computers before entering the university. Therefore, they are more familiar with reading English on computer screen than their L1 and these learners frequently use computers at the university. Considering all these reasons, it was assumed that the administration of the Stroop and the Keep Track tasks in L2 would not significantly impact the results. However, the participants received instructions for the four tests in both L1 and L2 (see Appendices E, F and G for English versions).

4.7 Data collection process

4.7.1 Research design[18]

The data collection process contained several steps. In the first session, on Wednesday of Week 1 of the study, 107 participants took the pre-test, filled in the background questionnaire and signed the ethics forms. After a filtering process based on the pre-test answers and the background questionnaire answers, 100 participants were selected for the study. Then they were randomly grouped into five groups of twenty participants each: Groups A, B, C, D and E. Group E was the control group and the rest were experimental groups.

Table 4.3 highlights the sessions that each group participated in. The participants in the experimental groups individually met with me three times between the pre-test and post-test. Groups A and B met with me on Monday, Wednesday and Friday for the individual eye-tracking sessions in Week 2 and Groups C and D on Tuesday, Thursday and Saturday of the same week for the eye-tracking sessions. Immediately after the

[18] Some parts of this section were published in Indrarathne & Kormos (2016).

third eye-tracking session (i.e. Friday for Groups A and B and Saturday for Groups C and D), the participants took the post-test and answered the oral interview questions. The post-test and the oral interview were conducted by an English language lecturer at the university. The control group took the post-test on the Friday of Week 2. In Week 3, participants in Groups A, B, C and D met with me individually and took the four WM tests.

Table 4.3

Sessions for each group

Group	Pre-test	Eye-tracking 1	Eye-tracking 2	Eye-tracking 3	Post-test	WM Tests
A	√	√	√	√	√	√
B	√	√	√	√	√	√
C	√	√	√	√	√	√
D	√	√	√	√	√	√
E	√	X	X	X	√	X

4.7.2 Experimental groups[19]

Figure 4.1 illustrates the experimental design of the study. The four experimental groups were provided with input on the target structure in three sessions in four different ways. All groups had to read three short stories (one in each session) that contained seven examples of the target structure of *causative had*. Group D received unenhanced input i.e. the examples of the target structure were not highlighted in the text. They were not informed that there was a specific target structure in the texts.

[19] Some parts of this section were published in Indrarathne & Kormos (2016).

Group C received enhanced input i.e. all examples of the target structure were highlighted in bold in the texts. The participants in Group C were not informed that they needed to pay attention to the highlighted phrases. Group A also received enhanced input and at the beginning of each eye-tracking session they were asked to pay attention to the highlighted phrases. The verbal instruction that they received was: "There are some phrases highlighted in bold in the story. Please pay attention to them." Group B received similar input as Group A in the first eye-tracking session. Immediately before the second eye-tracking session, the participants in Group B were shown a PowerPoint presentation (see 4.7.3 and the PowerPoint slides in Appendix Q) that explained the form and the meaning of the target structure. Thereafter, they were exposed to the second input i.e. the enhanced text with the instruction asking them to pay attention to the highlighted phrases. Input session 3 for Group B was similar to input session 1.

In addition to asking the participants in groups A and B to pay attention to the highlighted phrases, participants in all groups were told that they were going to read a short story and at the end of the story they needed to answer four comprehension questions. ("You are going to read a short-story on the computer. You can spend any amount of time on a slide. When you finish reading one slide, please go to the next slide. But, you can't go back to a previous slide. At the end of the story, you have to answer four questions about the story. You can read the questions and possible answers on the slides. Then you have to tell me which answer is correct.")

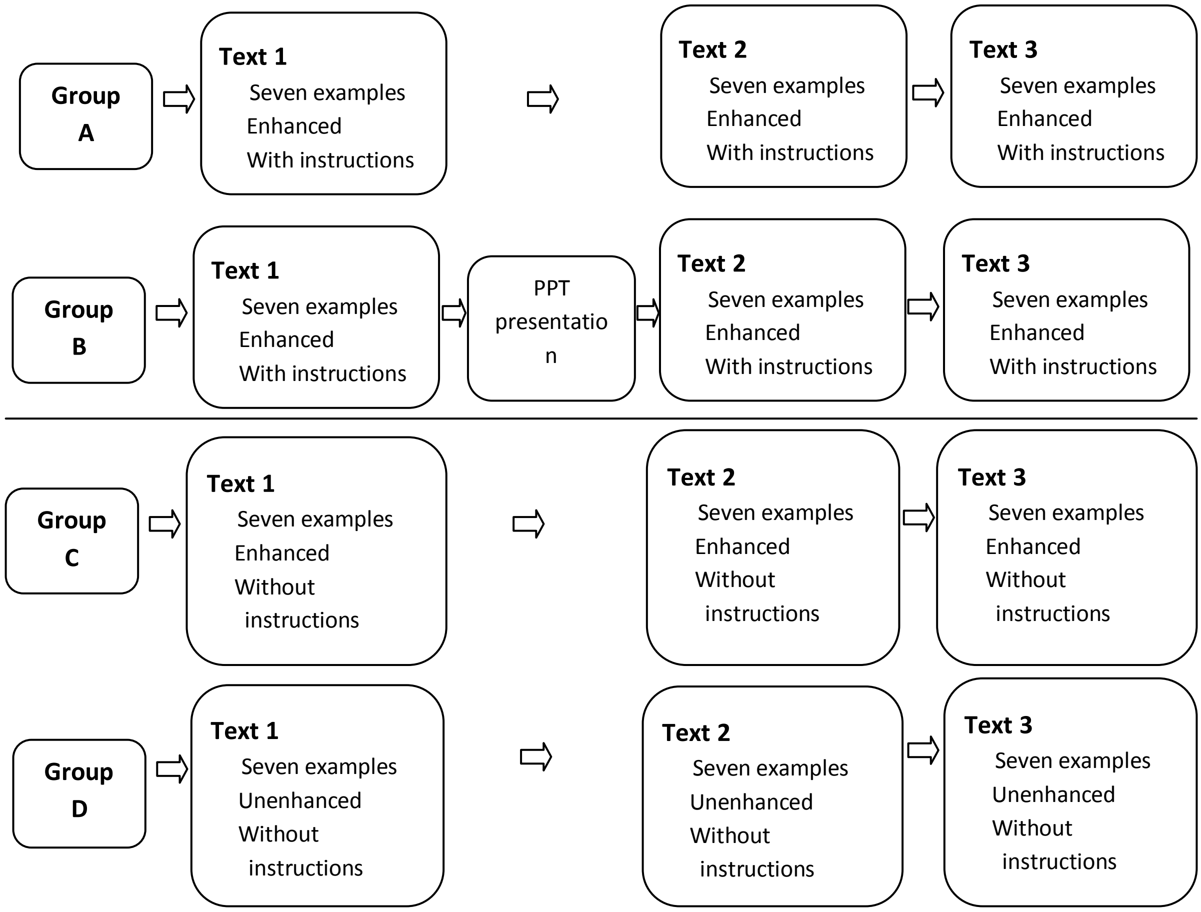


Figure 4.1. Experimental design

4.7.3 Explicit explanation provided to Group B[20]

Group B received a metalinguistic explanation in the form of a PowerPoint between first and second exposures. The slideshow contained 6 slides including the title slide. One example from Text 1 was used in slide 2 to show the form of the structure. Three other examples from Text 1 were used in slide 3 as more examples. Two of these four examples were presented in the following slide to explain the meaning of the construction. The participants went through the PowerPoint individually and no any other assistance such as verbal explanation was offered. Although providing

[20] Some parts of this section were published in Indrarathne & Kormos (2016).

metalinguistic explanation to one experimental group can be criticised for giving extra input (Andringa, De Glopper & Hacquebord, 2011), the explanation provided in this study was not considered as having a significant impact on the amount of input provided to Group B because the examples used were the ones that the group had already seen in exposure 1.

4.7.4 Eye-tracking[21]

Since the data collection was done in Sri Lanka, a Tobii X2-60 portable eye tracker was used. It was fixed to a laptop computer and placed in a quiet room. At the beginning of the first session, the participants were informed about the functioning of an eye-tracker and a trial slide was used in order to provide them with an orientation to eye-tracking. At the beginning of each reading session, a 9 point calibration was performed. The participants were seated approximately 67 cm from the computer screen.

The three input texts were put into PowerPoint slides (Appendix R-W). As discussed in 3.6.2, slides were prepared on PowerPoint using 24-point, double-spaced Calibri. All the words belonging to a particular example of a target structure (AOIs- e.g., had my hair cut) were placed in the same line in order to facilitate measuring fixation duration on AOIs (see Figure 4.2 for an example of the unenhanced input slides and Figure 4.3 for an example of the enhanced input slides).

[21] Some parts of this section were published in Indrarathne & Kormos (2016).

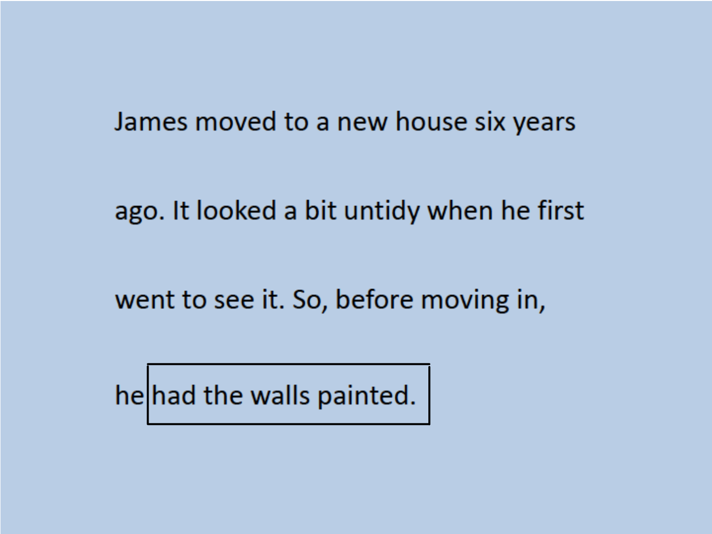


Figure 4.2. An example of the unenhanced input slides

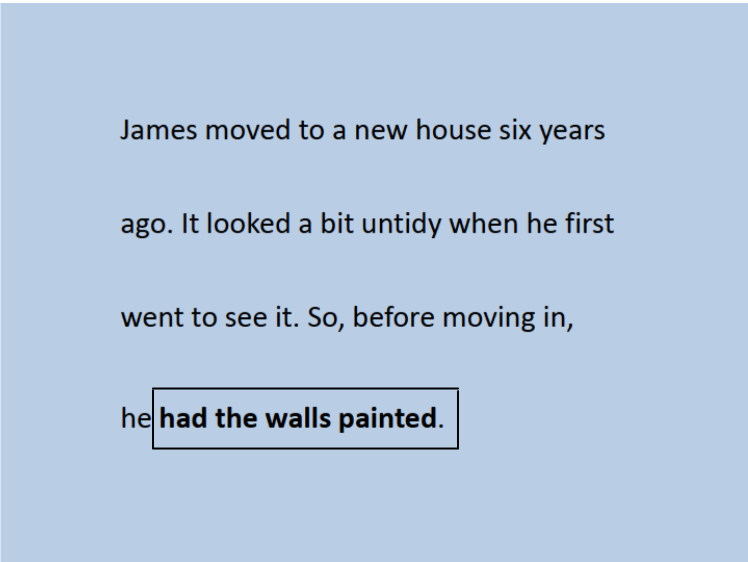


Figure 4.3. An example of the enhanced input slides

Each eye-tracking slide was comprised of either 4 or 5 lines of text. This depended on how many complete sentences could be included on a slide with double spacing. Double spacing was used in order to facilitate the retrieval of eye-tracking data on AOIs. Sentence halves (with a continuation on the next slide) were not included to avoid the necessity for the participants to move back to a previous slide. Then the PowerPoint slides were transferred to the Tobii eye-tracking software. They were saved as 12 different files. Table 4.4 highlights the number of eye-tracking files used in the experiment and the number of slides in each text.

Table 4.4

Eye-tracking files for each group

Group	Text 1	No. of slides	Text 2	No. of slides	Text 3	No. of slides
A	√ - enhanced	7	√ - enhanced	7	√ - enhanced	7
B	√ - enhanced	7	√ - enhanced	7	√ - enhanced	7
C	√ - enhanced	7	√ - enhanced	7	√ - enhanced	7
D	√- unenhanced	7	√- unenhanced	7	√ - unenhanced	7

The participants were informed that they were going to read a story on the computer immediately before they started reading. They were also told that they could move the slides when they had finished reading; however, they were not allowed to move back to a previous slide while reading the texts. A specific time was not allocated for reading, they could spend any amount of time on a slide reading and re-reading the text as they felt necessary.

At the beginning of each reading session, groups were told that they had to answer four CQs based on the text after reading. At the end of the reading task they read the four questions one by one on the slides (see Appendix R-W) and gave the answer orally for me to record. The participants were not allowed to go back to the text when answering the questions.

4.7.5 Pre and post-tests

The pre- and post-tests in the main study had different input modes: writing and listening. In both tests, the participants were given the SR task first. They read the instructions on the paper (see Appendix M/O) and wrote answers on the paper itself. The time allocated was 20 minutes then the papers were collected. Then, they received the GJ task answer sheet and I gave them oral instructions on how to complete the task (See Appendix M/O, Task 2 for the instructions that were read out). Thereafter, the recording was played for them to complete the trial questions, and it was paused to check the answers. Then, the recording was resumed and the participants completed the answer sheet. Finally, the third task was given to the participants; they read the instructions (Appendix M/O) and completed the task. The time allocated was 10 minutes. (All instructions were given in both L1 and L2).

4.7.6 WM tests

The data collection of the four WM tests was exactly the same as in the pilot study, discussed in 3.6.4. However, all instructions and explanations were given in the participants' L1.

4.8 Data analysis[22]

Each participant in the study was given a particular number which was used throughout the data collection/analysis process to identify the participants. After collecting all data,

[22] Some parts of this section were published in Indrarathne & Kormos (2016).

they were entered into a Microsoft Excel spreadsheet and later transferred to an SPSS 21 file. All the data were double checked for accuracy.

4.8.1 Eye-tracking data

The two main eye-tracking measurements used in the pilot study were also used in the main study. They were: TFD and the difference between observed TFD and expected TFD. The second measurement was taken based on the number of syllables in AOIs as a proportion of the number of syllables on the whole page (see 3.7.1).

The four experimental groups were exposed to the eye-tracking (n = *80*) and they read the three texts on three occasions. Eye-tracking data from all participants could not be included in the analysis due to their poor quality (Figure 4.4 shows a sample of a poor quality eye-tracking data). The eye-fixations that went beyond the screen for a considerable amount of time were also treated as poor quality.

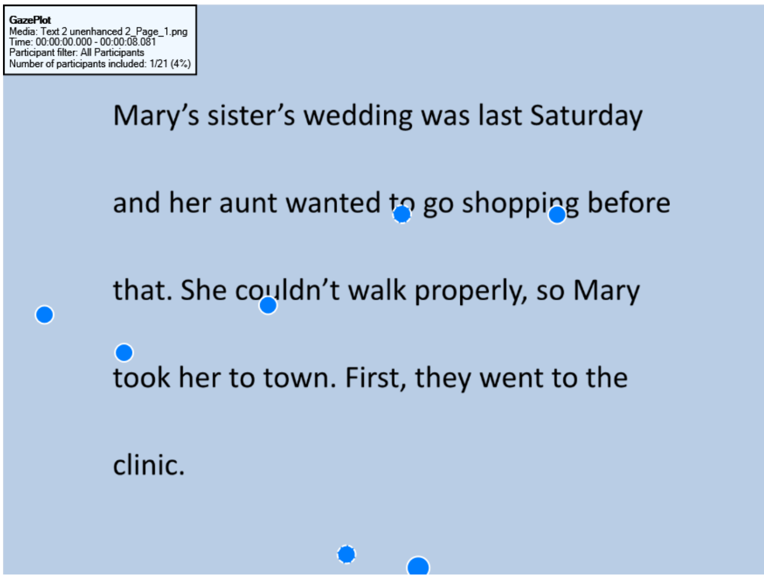


Figure 4.4. A sample of poor quality eye-tracking data

The data that were included in the analysis based on individual texts is as follows.

Table 4.5

Number of participants whose eye-tracking data was of sufficiently good quality in each stage

	T1	T2	T3
Group A	18	16	13
Group B	17	16	15
Group C	15	16	17
Group D	12	14	14
Percentage	77.5%	77.5%	73.75%

When I computed the composite variables for TFD and ΔOE based on all three texts, data of some more participants had to be excluded. For example, if Participant 1's data was of poor quality in Text 3, the participant was excluded from the composite variable even if their data in Texts 1 and 2 were of good quality. As a result, eye-tracking data of only $n = 45$ out of $n = 80$ were included in the composite variable. Most eye-tracking studies reviewed in Chapter 2 of this thesis have used fewer numbers of participants than those took part in this study. For example, Godfroid et al. (2013) $N = 28$, Godfroid and Schmidtke (2013) $N = 29$, Godfroid and Uggen (2013) $N = 40$, Issa et al. (2015) $N = 39$ and Smith (2012) $N = 18$. The only exception is Winke's (2013) study that used 80 participants initially and included eye-tracking data of $n = 55$. Therefore, it was assumed that $n = 45$ was a reasonable amount of data in this study. Table 4.6 highlights the amount of eye-tracking data included in the analysis for each group in composite measurements.

Table 4.6

Eye-tracking data of each group for analysis

Group	Number of participants whose data were included in the analysis
A	10
B	14
C	11
D	10
Total	45

When analysing the data, all AOIs in the three texts were highlighted on the eye-tracking slides and TFD was measured in milliseconds for all of them separately for each participant. The variables discussed in the pilot study data analysis (3.7.1) were used in the main study as well. They were, mean TFD of participants, mean TFD of AOIs, mean value of the difference between observed and expected TFD of participants and mean value of the difference between observed and expected TFD of AOIs.

$$\text{Mean TFD of participants (TFD-P)} = \frac{\text{TFD of a participant for all AOIs}}{\text{Number of AOIs}}$$

$$\text{Mean TFD of an AOI (TFD-AOI)} = \frac{\text{TFD of all participants of the AOI}}{\text{Number of participants}}$$

$$\text{Mean difference for a participant } (\Delta\text{OE-P}) = \frac{\text{TFD raw} - \text{EFD raw for a participant for all AOIs}}{\text{Number of AOIs}}$$

$$\text{Mean difference of an AOI } (\Delta\text{OE-AOI}) = \frac{\text{TFD raw} - \text{EFD raw of all participants of the AOI}}{\text{Number of participants}}$$

4.8.2 Pre-test and post-test

In both tasks 1 and 2, correct answers received 1 mark and incorrect answers were awarded a 0 score. Half marks were not allocated and as discussed in 3.7.2, spelling mistakes were ignored. First, the total score for both SR and GJ items was computed. Then, the gain score was computed separately for SR and GJ items as the difference between the total score of the post-test and the total score of the pre-test. All four variables: total SR score, total GJ score, SR gain score and GJ gain score were used in data analysis.

4.8.3 WM test data

WM test data analysis was carried out exactly the same way as in the pilot study discussed in 3.7.3

4.8.4 CQ data

There were 2 target CQs for each text. Since they were multiple choice items, a score of 1 (correct) or 0 (incorrect) was awarded. The score for each item for each participant was entered separately and then the total for texts 1, 2, and 3 were computed. A composite score for the three texts was also computed.

4.8.5 Post eye-tracking interview

The data collected in the interview was treated as quantitative data and used in the data analysis. For example, the answers given to Question 6 (Appendix P) – "What did you notice in the text?" were used to analyse whether the participants were able identify the target structure. If they had provided answers such a 'a grammar structure', an example of the target structure (e.g., She had the car repaired) or a metalinguistic explanation such as 'had … (art/n)... past participle' or 'had … (art/n).. ed', they were considered as evidence of awareness of the target construction. Analysis of the answers to Questions 7 (Did you notice any particular grammar structure in the texts?) and Question 8 (If your answer to question 7 is 'yes' can you please write down that structure?) were done together. If the answer to Question 7 was 'yes' and the answer to Question 8 was 'had …(art/n)... ed' or an example of the target structure, it was treated as awareness of the target construction. If the answer to Question 7 was 'yes', but the answer to Question 8 was not related to the target construction (e.g., if the answer was 'past tense'), such answers were not counted as evidence of awareness of the target construction.

4.9 Summary

This chapter outlined the methodology of the main study highlighting the context, participants, instruments, procedure and data analysis methods. Chapter 5 of this thesis includes the results obtained in the main study categorised into the three sets of research questions.

Chapter 5: Results

5.1 Introduction

Chapter 4 outlined the research questions to be answered in this study under three sections. The first set of research questions was aimed at analysing how attention to input in different input conditions influences the development of a TGC, the second set investigated the role of WM in input processing and the third set enquired into how awareness is related to input processing.

This chapter presents the results obtained in the study. An initial data screening was conducted in order to understand what type of statistical analyses should be performed on the data set and Section 5.2 discusses how data screening was carried out. The same section also gives the justification of statistical analyses performed. The chapter then discusses the results under the first set of research questions in 5.3 and the results under the second set of research questions in 5.4. Section 5.5 outlines the results relating to the third set of research questions. Section 5.6 summarises the findings of the free writing task that the participants took as a section of the pre-post-tests. The chapter ends with a summary of the results.

5.2 Data screening[23]

Before specific statistical analyses are conducted on a data set, it is necessary to screen data in order to understand what statistical analyses should be performed (Pallant, 2010). For example, certain assumptions are made, particularly by parametric tests such

[23] Some parts of this section were published in Indrarathne & Kormos (2016).

as t-test and ANOVA[24], when analysing data, and thus it is necessary to investigate whether those assumptions are fulfilled before applying such tests (ibid, p.51). In order to come to a conclusion on what type of data analyses can be performed on the data set, descriptive statistics and tests of normal distribution are used.

5.2.1 Pre and post-tests

5.2.1.1 Normality tests and descriptive statistics of the pre- and post-tests of the whole population

Pre and post-test results were scales and thus it was necessary to find out whether the data were reliable. Neither SR items nor GJ items in the pre-test showed sufficient levels of internal consistency when one inspects the Cronbach's Alpha values (-0.03 and -0.07 respectively). This was predicted since the participants were expected to have no or relatively little knowledge of the TGC. SR and GJ items in the post-test had a Cronbach's Alpha of 0.681 and 0.447 respectively, which showed that SR items had an adequate level of internal consistency but the GJ items were below the acceptable value (0.7). However, the results of the GJ items were included in further analysis based on several theoretical assumptions. First, Pallant (2010) notes that a test with a small number of items may not have a high Cronbach's Alpha value. Lance, Butts and Michels (2006) explain that Cronbach's alpha of 0.7 is not a universally acceptable cut-off point for all research. There can be fluctuations depending on the purpose/construct of the test. Therefore, the lower value of for the GJ task in this case might be due to the lower number of items or to the fact that participants' responses within the test might

[24] ANOVA: Analysis of variance. This involves one independent variable and different levels of the variable (Pallant, 2010).

not have been highly consistent (Lance, Butts and Michels, 2006). This was not unexpected as they were in the process of developing new knowledge. The mean for the post GJ task was also significantly below the 50% guessing range, $t(96) = -3.17$, $p < .002$.

The normality of the distribution of scores was calculated for the total of pre-test SR items, total of pre-test GJ items, total of post-test SR items and the total of post-test GJ items. The Kolmogorov-Smirnov values of the pre-test SR items and the pre-test GJ items indicated that scores were not normally distributed, $p < .001$, $N = 100$. Pallant (2010) notes that this is quite common in large samples and suggests considering other normality measurements such as skewness, kurtosis, histogram and normal Q-Q plots to see whether normal distribution is indicated by such measures. Table 5.1 presents the descriptive statistics of the pre-test SR items and the pre-test GJ items. The skewness (5.59) and the kurtosis values (29.89) of the SR items indicated asymmetrical and peaked distribution of scores. This is because all participants except three scored 0 in the pre-test SR items suggesting that they did not have prior knowledge of the target construction. However, the skewness value (-.16) of GJ items indicated nearly symmetrical distribution and the kurtosis (-.41) value showed a relatively flat distribution.

Table 5.1

Descriptive statistics of the pre-test items (N = 100)

	M	*SE*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Pre SR	.03	.08	.17	5.59	.24	29.89	.48
Pre GJ	4.55	.15	1.48	-.16	.24	-.41	.48

Neither the histogram (See Appendix X – Figure X1) nor the normal Q-Q plot (Appendix X – Figure X2) of pre-test SR items showed normal distribution. The boxplot (Appendix X – Figure X3) revealed the existence of three outliers, participants 22, 46 and 82 who scored one mark each out of six in the whole SR section of the test. Histogram and normal Q-Q plots of pre-test GJ items (Appendix X – Figures X4 and V5) showed normal distribution. The data obtained in the pre-test SR task verified the prediction that the participants did not have prior knowledge of the target construction.

Kolmogorov-Smirnov values for the post-test SR items, $p < .001$, $N = 100$, and the post-test GJ items, $p = .004$, $N = 100$, indicated that scores were not normally distributed even in the post-test. Table 5.2 highlights the descriptive statistics of the total post-test SR items and total post-test GJ items. The skewness (1.53) of the SR items suggested asymmetrical distribution of scores and the kurtosis (1.85) value indicated a peaked distribution. The kurtosis value was within the acceptable range of normal distribution (1.85) and the skewness was slightly higher than the acceptable value (1.53). However, the standard error of skewness indicated that the skewness of the SR items falls within the acceptable range. Skewness and kurtosis of post-test GJ items were also within the acceptable range of normality.

Table 5.2

Descriptive statistics of the post-test items (N = 100)

	M	*SE*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Post SR	1.02	.14	1.43	1.53	.24	1.85	.48
Post GJ	5.80	.20	2.02	.08	.24	-.33	.48

The histogram (Appendix X – Figure X6) of the post-test SR items showed an inverse curve which highlights that scores were not normally distributed, but the normal Q-Q plot (Appendix X – Figure X7) and the detrended normal Q-Q plot (Appendix X – Figure X8) indicated a reasonably normal distribution. The boxplot (Appendix X – Figure X9) shows that participants 22 and 46 were positive outliers. These were the participants who scored high in the whole test. This indicates that they were able to develop the knowledge of the target construction as a result of the experimental condition that they were in. Therefore, these outliers were not exempted from the analysis. Both the histogram and the normal Q-Q plot of the pre/post GJ items indicated normal distribution (Appendix X – Figures X4, X5, X10, X11).

Since some of the normality measures of the post-test SR items showed that scores were not distributed normally, an attempt was made to transform the data to make them normal. First, the post-test total SR score was transformed to an inverse score (Pallant, 2010), which is recommended for scores distributed in the shape of Weibull PDF (Γ = 0.5). The new variable is labelled as Inverse Post SR Total in Table 5.3. Since most participants also scored 0 in the post-test SR items, the new variable excluded more than half of the participants from the analysis. Therefore, the Visual Binning technique, which is recommended by Pallant (2010) for skewed data was used to collapse the continuous variable into groups. The new variable created was named as Binned Post SR Total. The histogram of the Binned SR Total still did not show normal distribution and thus the variable was again transformed to an inverse score (recommended by Pallant, 2010). This variable was named Binned Inverse Post SR Total. Table 5.3 includes the descriptive statistics of the three new variables.

As can be seen in Table 5.3, none of the new variables had an acceptable Kolmogorov-Smirnov value to consider the scores as normally distributed. The histograms (Appendix X – Figures X12-X14) of the new variables did not show normal distribution either. However, all three new variables had acceptable skewness and kurtosis values. Considering the descriptive statistics of the three variables created, Binned Post SR Total was considered a more reliable variable to be used in parametric tests. Thus, it was also used in the initial data analyses, in particular, to see whether it provided results that were different from the post-test SR items (raw) variable.

Table 5.3

Descriptive statistics of the transformed variables

	Kolmogorov-Smirnov	*N*	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Inverse Post SR Total	<.001	47	.68	.33	.00	.35	-1.83	.68
Binned Post SR Total	<.001	100	1.63	.75	.73	.24	-.85	.48
Binned Inverse Post SR Total	<.001	100	.79	.28	-.24	.24	-1.82	.48
Post SR Total - Raw	<.001	100	1.02	1.43	1.53	.24	1.85	.48

5.2.1.2 Normality and descriptive statistics of the pre and post-tests of the groups

Descriptive statistics for the five groups in the study i.e. control group and the four experimental groups were computed separately. Table 5.4 summarises the statistics.

Skewness and kurtosis values of the post-test GJ items in all groups and of the post-test SR items in groups A, B and D are in the acceptable range. However, asymmetrical and peaked distribution of the scores can be seen in the control group because most of the participants scored 0 in the post-test. The values for group C also indicate an asymmetrical and peaked distribution.

Table 5.4

Descriptive statistics for groups in pre/post-tests

	Test	*M*	*SE*	*SD*	Skewness	*SE*	Kurtosis	*SE*
control (E) (*n* = 20)	Pre SR	0	0	0	0	0	0	0
	Pre GJ	4.75	.42	1.86	-.36	.51	-.92	.99
	Post SR	.35	.21	.93	3.51	.51	13.43	.99
	Post GJ	4.85	.45	2.01	.09	.51	.05	.99
enhanced+ instructions (A) (*n* = 20)	Pre SR	.05	.05	.22	4.47	.51	20.00	.99
	Pre GJ	4.25	.36	1.62	.13	.51	-.89	.99
	Post SR	1.60	.41	1.82	.96	.51	-.00	.99
	Post GJ	6.05	.51	2.28	.34	.51	-.95	.99
enhanced+ instructions+ explanation (B) (*n* = 20)	Pre SR	.05	.05	.22	4.47	.51	20.00	.99
	Pre GJ	4.85	.32	1.42	.170	.51	.30	.99
	Post SR	2.20	.34	1.54	.39	.51	.64	.99
	Post GJ	6.85	.47	2.08	-.13	.51	-.69	.99
enhanced only (C) (*n* = 20)	Pre SR	0	0	0	0	0	0	0
	Pre GJ	4.15	.29	1.29	-.84	.51	.91	.99
	Post SR	.60	.22	.99	2.37	.51	6.73	.99
	Post GJ	5.75	.42	1.89	-.38	.51	-.61	.99
unenhanced (D) (*n* = 20)	Pre SR	.05	.05	.22	4.47	.51	20.00	.99
	Pre GJ	4.65	.26	1.18	-.29	.51	-1.41	.99
	Post SR	.35	.11	.49	.68	.51	-1.72	.99
	Post GJ	5.50	.31	1.39	-.77	.51	1.21	.99

There are two main reasons for some groups not having normally distributed data. The first is the lower number of participants in each separate group ($n = 20$) and the other is that most of the participants in some experimental conditions scored 0 in the SR task. These results (low scores in some groups) were vital initial predictors of the type of input provided to the participants in those groups and their ability to develop the knowledge of the target construction as a result of the input received.

As discussed earlier, the Binned Post SR Total variable was created because some of the normality measures of the post-test SR items did not show normal distribution. However, for further analyses, I applied both parametric and identical non-parametric tests for the raw total of the post-test SR items in the initial analyses. This was to identify whether the two types of tests would provide different results. Moreover, parametric tests were applied to the Binned Post SR Total variable as well. The results of all these different tests are reported in relevant sections of this chapter.

5.2.1.3 Normality and descriptive statistics of the gain scores of the whole sample and groups

The differences between the pre-test SR item total score and post-test SR item total score and the differences between the pre-test GJ item total score and post-test GJ item total score (gain scores) were used in further analyses of data. Table 5.5 summarises the descriptive statistics of the gain scores for the whole sample. The Kolmogorov-Smirnov value for the SR item gain score, $p < .001$, $N = 100$, and the GJ item gain score, $p < .001$, $N = 100$, indicated that scores were not normally distributed. Both skewness and kurtosis values of the latter and the kurtosis value of the former were within the acceptable range of normal distribution; however, skewness of the SR gain score was slightly higher than the acceptable value.

Table 5.5

Descriptive statistics of the gain scores (N = 100)

Test	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
SR	.99	1.37	1.33	.24	.72	.48
GJ	1.25	1.61	.28	.24	.42	.48

The histogram (Appendix X – Figure X15) of the SR item gain score did not indicate a normal distribution; however, both the normal Q-Q plot and the detrended normal Q-Q plot (Appendix X - X16, X17) of the same variable indicated normal distribution. Both the histogram and the normal Q-Q plot of the GJ item gain score showed a normal distribution (Appendix X – X18, X19).

Similar to the normal distribution of post-test SR item total score discussed in 5.2.1.1, the gain score of SR items demonstrated problems with normal distribution. Therefore, the Visual Binning technique (as discussed in 5.2.1.1) was used to make the data distribution more normal. Table 5.6 summarises the descriptive statistics of the raw SR gain score and Binned SR gain score. The new variable did not have an acceptable Kolmogorov-Smirnov value either; however, both skewness and kurtosis values fell within the acceptable range. The histogram (Appendix X – Figure X20) of the transformed variable did not show a normal distribution either and the normal Q-Q plot (Appendix X – Figure X21) or detrended normal Q-Q plot (Appendix X – Figure X22) of the new variable were not very different from those of the raw SR gain score. Thus, the standardized residuals and Cook's distances[25] were also calculated for the SR task. The Normal P-P plot of the standardized residuals indicated no deviation from normality and none of the residuals fell outside the +2/-2 range. All Cook's distance values were well below the critical threshold of 1 (Tabachnik & Fidell, 2007). These analyses indicated the possibility of applying parametric methods for further analyses in the case of SR gain score.

Table 5.6

Descriptive statistics of transformed variable – gain score of SR items

Test	Kolmogorov-Smirnov	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Pre-Post difference SR (raw)	<.001	.99	1.37	1.33	.24	.72	.48
Binned pre-post difference SR	<.001	1.62	.75	.76	.24	-.82	.48

The descriptive statistics of the gain scores of the five groups were also computed. According to the Shapiro-Wilk and Kolmogorov-Smirnov values, scores were not normally distributed. However, normal Q-Q plots and detrended normal Q-Q plots showed reasonably normal distribution in all variables. Moreover, as seen in Table 5.7, skewness and kurtosis of both gain scores in all groups except for the control and *unenhanced* groups fell within the acceptable range. Considering the lower number of participants in each group, all variables were treated as normally distributed data.

[25] Cook's distance: In least square regression analysis, Cook's distance is used to estimate the influence of a data point (Atkinson & Riani, 2000).

Table 5.7

Descriptive statistics of gain scores for groups

Group	Test	*M*	*SE*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Control (E) (*n* = 20)	SR	.35	.20	.93	3.51	.51	13.43	.99
	GJ	.10	.32	1.44	-.65	.51	-.33	.99
enhanced+ instructions (A) (*n* = 20)	SR	1.55	.38	1.70	.72	.51	-.90	.99
	GJ	1.80	.41	1.85	.43	.51	-1.90	.99
enhanced+ instructions+ explanation (B) (*n* = 20)	SR	2.15	.31	1.42	-.05	.51	-.39	.99
	GJ	2.00	.24	1.07	.000	.51	.23	.99
enhanced only (C) (*n* = 20)	SR	.60	.22	.99	2.36	.51	6.73	.99
	GJ	1.50	.35	1.57	1.03	.51	2.25	.99
unenhanced (D) (*n* = 20)	SR	.30	.10	.47	.94	.51	-1.24	.99
	GJ	.85	.30	1.34	.87	.51	.22	.99

5.2.1.4 Correlations

Since two types of test items were used in the pre and post-tests, the correlations of the scores between them were also computed to see if there is a relationship between the two types of test items. According to Table 5.8, a significant correlation cannot be seen between the total score of the pre-test SR items and the total score of the pre-test GJ items. However, there was a significant large correlation between total scores of the two types of items in the post-test, $r = .735$, $N = 100$, $p < .001$. Similarly, the gain score of the two types of items was also strongly correlated, $r = .613$, $N = 100$, $p < .001$.

Table 5.8

Correlation between pre-test and post-test scores (N = 100)

		Pre SR Total	Pre GJ Total	Post SR Total	Post GJ Total	Gain score of SR	Gain score of GJ
Pre SR Total	Pearson Correlation	1	.093	.408*	.279*	.301*	.265*
	p		.357	<.001	.005	.002	.001
Pre GJ Total	Pearson Correlation		1	.331*	.617*	.334*	-.147
	p			.001	<.001	<.001	.145
Post SR Total	Pearson Correlation			1	.735*	.993*	.618*
	p				<.001	<.001	<.001
Post GJ Total	Pearson Correlation				1	.732*	.688*
	p					<.001	<.001
Pre-Post difference of SR Total	Pearson Correlation					1	.613*
	p						<.001
Pre-Post difference of GJ Total	Pearson Correlation						1
	p						

*Correlation significant at 0.01 level

5.2.2 WM tests

5.2.2.1 Normality and descriptive statistics of the WM test data of the whole population

Table 5.9 reports the Shapiro-Wilk and Kolmogorov-Smirnov values of the four WM tests: Digit Span (DS), Keep Track (KT), Plus Minus (PM) and Stroop (ST). These statistics indicated that scores were not normally distributed.

Table 5.9

Tests of normality – WM tests

Test	Kolmogorov-Smirnov Sig.	Shapiro-Wilk Sig.
DS	<.001	.001
KT	.005	.012
PM	<.001	<.001
ST	<.001	<.001

Table 5.10 contains the descriptive statistics of the four WM tests. According to the data, the Digit Span and the Keep Track tests have skewness and kurtosis values within the acceptable range, but the Plus Minus and the Stroop tests do not. The Stroop test has a particularly skewed and peaked distribution.

Table 5.10

Descriptive statistics of the WM tests for the experimental sample (n = 80)

Test	*M*	*SE*	*SD*	Skewness	*SE*	Kurtosis	*SE*
DS	5.34	.17	1.56	.16	.27	-.82	.53
KT	74.21	1.74	15.53	-.46	.27	.16	.53
PM	19.03	1.73	15.51	1.51	.27	3.11	.53
ST	2057.78	72.09	664.77	2.15	.27	8.61	.53

The histograms, Q-Q plots and detrended Q-Q plots (Appendix X – Figures X23-X28) showed reasonably normal distribution of scores in the Digit Span and the Keep Track tests. The histograms of the Plus Minus and the Stroop tests (Appendix X – Figures X29 and X33) highlighted skewed data. There were five positive outliers in the Plus Minus test and one in the Stroop test indicated in the respective boxplots (Appendix X – Figures X32 and X36).

5.2.2.2 Normality and descriptive statistics of the WM tests of the experimental groups separately

Table 5.11 summarises the descriptive statistics of the four WM tests for the four experimental groups. It highlights skewed distribution of the Plus Minus test in Groups B, C and D, the Keep Track and the Stroop tests in Group D. A peaked distribution can also be noticed in the Stroop test score in Group D.

Table 5.11

Descriptive statistics of WM tests for groups

Group	Test	*M*	*SE*	*SD*	Skewness	*SE*	Kurtosis	*SE*
enhanced+ instructions (A) (*n* = 20)	DS	5.40	.30	1.35	.16	.51	-.85	.99
	KT	77.40	3.03	13.54	-.05	.51	-1.03	.99
	PM	18.75	3.83	17.11	.89	.51	-.32	.99
	ST	2127.30	135.38	605.45	.83	.51	-.43	.99
	Composite	.06		.92				
enhanced+ instructions+ explanation (B) (*n* = 20)	DS	5.70	.38	1.72	.04	.51	-1.20	.99
	KT	75.73	3.78	16.93	-.00	.51	-1.55	.99
	PM	23.90	4.39	19.66	1.65	.51	3.61	.99
	ST	1822.15	100.40	448.99	.52	.51	-.43	.99
	Composite	.29		.1.02				
enhanced only (C) (*n* = 20)	DS	5.20	.34	1.51	.24	.51	-.54	.99
	KT	73.45	3.19	14.28	-.08	.51	-1.25	.99
	PM	15.70	2.27	10.14	1.41	.51	3.99	.99
	ST	2131.50	119.38	533.89	.67	.51	.27	.99
	Composite	-.09		.99				
unenhanced (D) (*n* = 20)	DS	5.05	.37	1.67	.21	.51	-.63	.99
	KT	70.21	3.86	17.27	-1.22	.51	1.85	.99
	PM	17.75	3.01	13.47	1.13	.51	1.04	.99
	ST	2150.15	199.64	892.80	2.86	.51	10.17	.99
	Composite	-.23		1.96				

Considering the lower number of participants in the study and also that some normality measurements showed normal distribution in the four WM tests, the raw data of the four tests were used in further parametric analyses.

5.2.2.3 Correlations

Table 5.12 includes the correlation between the four WM tests. The Digit Span test significantly correlated with the Keep Track task, r = .818, n = 80, p < .001, and the Stroop task, r = -.530, n = 80, p < .001. The Keep Track task also significantly correlated with the Stroop task, r = -.455, n = 80, p < .001. The high correlation between the Digit Span test and the Keep Track test could be an indication that they measure similar functions of the WM.

Table 5.12

Correlations between WM tests (n = 80)

		DS	KT	PM	ST
DS	Pearson Correlation	1	.818*	.112	-.530*
	p		<.001	.322	<.001
KT	Pearson Correlation		1	.119	-.455*
	p			.295	<.001
PM	Pearson Correlation			1	-.046
	p				.684
ST	Pearson Correlation				1
	p				

*Correlation significant at 0.01 level

5.2.3 CQs' composite scores

5.2.3.1 Normality and descriptive statistics of CQs' composite scores for the whole sample

The study used two target CQs in each eye-tracking reading session (input Texts 1, 2 and 3) and based on the data, a composite score was calculated for each session. Both

Shapiro-Wilk and Kolmogorov-Smirnov values of all three composite CQ scores were significant ($p < .001$), which indicated that scores were not normally distributed. Histograms of the three variables did not show normality either; however, normal Q-Q plots and detrended normal Q-Q plots demonstrated a reasonably normal distribution (Appendix X – Figures X38, X39, X41, X42, X44 and X45). As shown in Table 5.13, the kurtosis value of all three variables fell within the acceptable range so did the skewness of Text 2 and 3 scores. The skewness of the Text 1 score was slightly higher.

Table 5.13

Descriptive statistics of the composite CQ scores of the experimental sample (n = 80)

Test	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Text 1	.11	.22	1.76	.26	2.20	.53
Text 2	.59	.37	-.32	.26	-1.14	.53
Text 3	.68	.36	-.72	.26	-.80	.53

5.2.3.1 Normality and descriptive statistics of CQs composite scores for groups

None of the Shapiro-Wilk or Kolmogorov-Smirnov values of the CQ composite scores in any group suggest a normal distribution. However, normal Q-Q plots and detrended normal Q-Q plots showed reasonably normal distribution in all groups, and the skewness and kurtosis values fell within the acceptable range in all but Group D (see Table 5.14).

Table 5.14

Descriptive statistics of composite CQ scores for groups

Group	Text	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
enhanced+ instructions (A) (*n* = 20)	1	1.00	.20	1.62	.51	.70	.99
	2	.95	.64	-.25	.51	-.44	.99
	3	.82	.24	-.68	.51	-1.72	.99
enhanced+ instructions+ explanation (B) (*n* = 20)	1	.12	.22	1.25	.51	-.49	.99
	2	.85	.23	-.94	.51	-1.24	.99
	3	.90	.20	-1.62	.51	.70	.99
enhanced only (C) (*n* = 20)	1	1.00	.20	1.62	.51	.70	.99
	2	.55	.39	-.19	.51	-1.31	.99
	3	.62	.39	-.49	.51	-1.15	.99
unenhanced (D) (*n* = 20)	1	.12	.27	2.24	.51	4.65	.99
	2	.35	.37	.55	.51	-.83	.99
	3	.40	.38	.37	.51	-1.13	.99

The next three sections of this chapter discuss how the variables discussed above were used to answer the research questions of the study.

5.3 Results of Section 1 research questions[26]

This section discusses the results related to the first set of research questions. They are:

RQ 1: How does knowledge of the form and meaning of the target syntactic construction *causative had*, as measured by a production and a comprehension task,

[26] These results were published in Indrarathne & Kormos (2016).

change under explicit and implicit instructional conditions in the case of Sri Lankan pre-intermediate/ intermediate level English language learners?

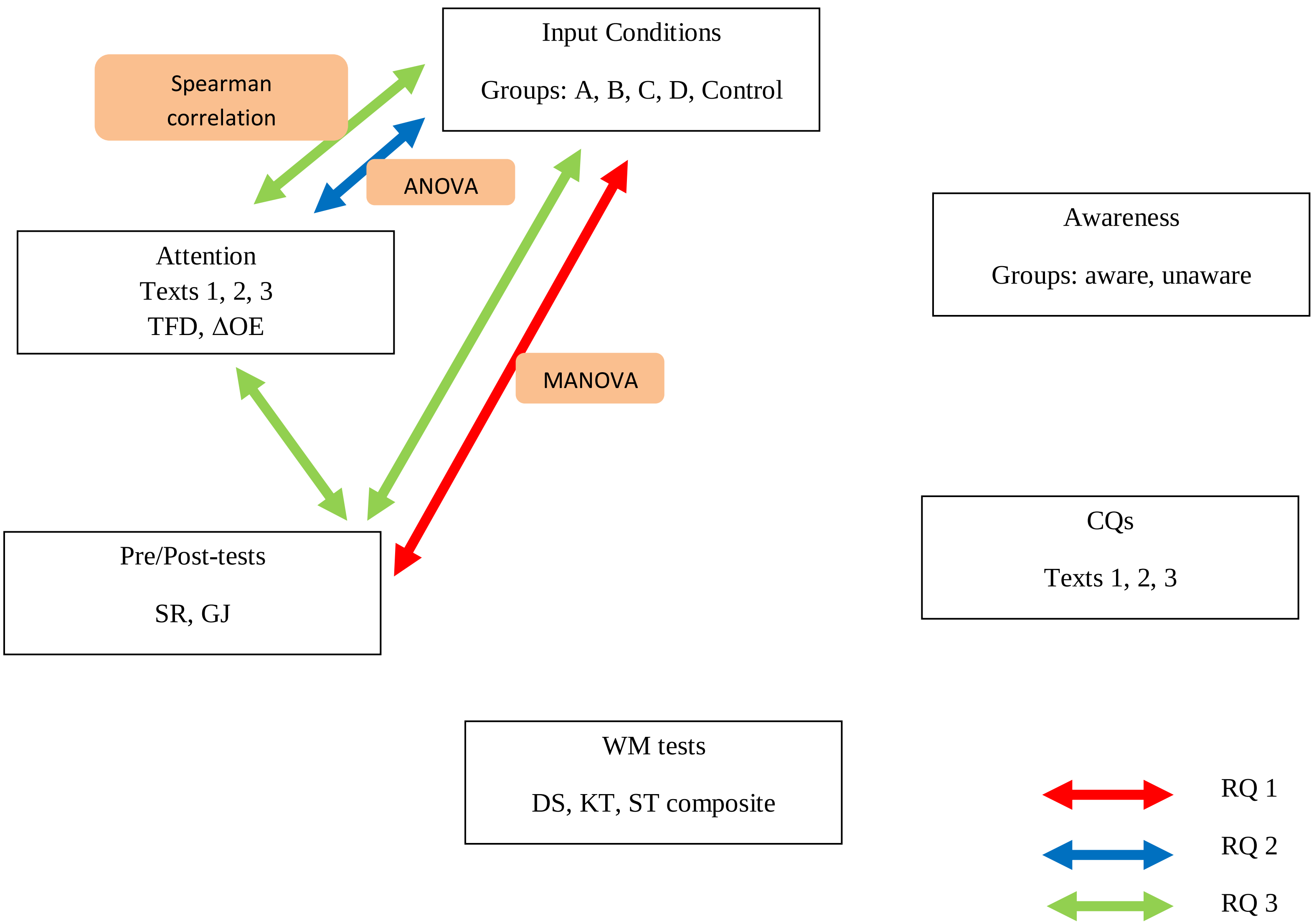
RQ 2: How does attention paid to examples of the *causative had* construction in the input texts differ under explicit and implicit instructional conditions?

RQ 3: How is the change in knowledge of the target construction under explicit and implicit instructional conditions related to the attention paid to examples of the *causative had* construction in the input texts?

In order to answer the first research question, I analysed the gains in knowledge as measured by the SR and GJ tasks across input conditions with repeated measures MANOVA[27]. In the analyses pertaining to the second research question, I compared the eye-tracking measures of the treatment groups by means of one-way analysis of variance. As for the third research question, the relationship between the gain scores and eye-tracking measures was examined with the help of Spearman rank-order correlations. The level of significance was set at $p < .05$ and in the case of multiple tests Bonferroni correction was applied to adjust the significance level and avoid Type I error. Effect sizes were calculated using partial eta square values. Based on Cohen (1988), partial eta square values above .06 were taken as indications of medium, and above .138 as large effect size (see Figure 5.1 for an illustration of the type of analysis done under the first set of RQs).

[27] MANOVA: Multivariate analysis of variance. This analysis contain more than one dependant variable (Pallant, 2010).

Figure 5.1. Data analysis methods of the first set of RQs



5.3.1 How does knowledge of the form and meaning of the target syntactic construction *causative had*, as measured by a production and a comprehension task, change under explicit and implicit instructional conditions in the case of Sri Lankan pre-intermediate/ intermediate level English language learners?

In order to answer this question, a repeated measure MANOVA analysis was conducted to find out whether there was an improvement in the SR and GJ task scores for the overall study sample, and whether an interaction between the effect of instructional conditions and the rate of improvement in the two tasks could be observed. This was followed by post-hoc Scheffe comparisons across the groups to detect where the differences between them lie.

A statistically significant increase of the scores from pre-test (M = .03, SD = .171) to post-test (M = 1.02, SD = 1.435) for the SR items, Wilks' Lambda = .57, $F(4, 95)$ = 69.81, $p < .001$, partial eta squared = .42, for the whole sample was found in the analysis. The mean increase was .99 with a 95% confidence interval ranging from .72 to 1.26. There was also a statistically significant difference between the total scores of the pre-test (M = 4.55, SD = 1.48) and the post-test (M = 5.80, SD = 2.02) of GJ items for the whole sample, Wilks' Lambda = .57, $F(4, 95)$ = 71.17, $p < .001$, partial eta squared = .42. The mean increase in this task was 1.25 with a 95% confidence interval ranging from .93 to 1.57. The mean values in the pre-test SR task indicate no existing knowledge of the target construction, whereas the mean for the GJ task is significantly below the 50% guessing range.

As illustrated in Table 5.15, the t-tests ran for each instructional condition (between pre-test and post-test) report significant improvement in both SR and GJ tasks in *enhanced+ instruction* (A), $t(19) = 1.55$, $p < 0.001$ and $t(19) = 1.80$, $p < 0.001$, *enhanced+ instructions+ explanation* (B) conditions, $t(19) = 2.15$, $p < 0.001$ and $t(19) = 2.00$, $p < 0.001$. However, only GJ task shows a significant improvement in *enhanced only* (C) group, $t(19) = 1.50$, $p < 0.001$, and the *unenhanced* (D) condition did not show a significant development in either of the tasks. The mean differences highlight the highest gains in both SR and GJ tasks in the *enhanced+ instructions+ explanation* (B) group followed by *enhanced+ instruction* (A) and *enhanced only* (C) groups.

Table 5.15

Paired t-test results of groups in the SR and GJ tasks

Group	SR – mean difference	*p*	GJ – mean difference	*p*
enhanced+ instructions (A) (*n* = 20)	1.55*	.001	1.80*	<.001
enhanced+ instructions+ explanation (B) (*n* = 20)	2.15*	<.001	2.00*	<.001
enhanced only (C) (*n* = 20)	.60	.014	1.50*	<.001
unenhanced (D) (*n* = 20)	.30	.010	.85	.011

In the case of the SR task, the repeated measures MANOVA analysis showed a significant between subject effect (in the post-test) with a large effect size, $F(4, 95) = 8.22$, $p < .001$ partial eta squared = .25, and a significant interaction between treatment conditions and improvement in this task also with a large effect size, Wilks Lambda = .71, $F(4, 95) = 9.62$, $p < .001$ partial eta squared = .288. As can be seen in Table 5.16 (also see Figure 5.1), a follow-up ANOVA with the gain score as the dependent variable and the post-hoc Bonferroni analysis showed that there is a significant mean difference in the gain scores for SR items between the control group and the *enhanced+ instructions* (A) group ($p = .018$), and between the control group and the *enhanced+*

instructions+ explanation (B) group ($p < .001$). A significant mean difference was also found between the *unenhanced* (D) group and the *enhanced+ instructions* (A) group ($p = .012$), and the *enhanced+ instructions+ explanation* (B) group ($p < .001$). Moreover, the *enhanced+ instructions+ explanation* (B) group showed a significant improvement ($p = .001$) compared to the *enhanced only* (C) group.

Table 5.16

Between-subject effect of the SR items (N = 100)

Group	Group	Mean difference	*p*
Control (E) (*n* = 20)	enhanced+ instructions (A)	-1.20*	.018
	enhanced+ instructions+ explanation (B)	-1.80*	<.001
	enhanced only (C)	-.25	1.000
	unenhanced (D)	.05	1.000
enhanced+ instructions (A) (*n* = 20)	Control (E)	1.20*	.018
	enhanced+ instructions+ explanation (B)	-.60	1.000
	enhanced only (C)	.95	.128
	unenhanced (D)	1.25*	.012
enhanced+ instructions+ explanation (B) (*n* = 20)	control (E)	1.80*	<.001
	enhanced+ instructions (A)	.60	1.000
	enhanced only (C)	1.55*	.001
	unenhanced (D)	1.85*	<.001
enhanced only (C) (*n* = 20)	control (E)	.25	1.000
	enhanced+ instructions (A)	-.95	.128
	enhanced+ instructions+ explanation (B)	-1.55*	.001
	unenhanced (D)	.30	1.000
unenhanced (D) (*n* = 20)	control (E)	-.05	1.000
	enhanced+ instructions (A)	-1.25*	.012
	enhanced+ instructions+ explanation (B)	-1.85*	<.001
	enhanced only (C)	-.30	1.000

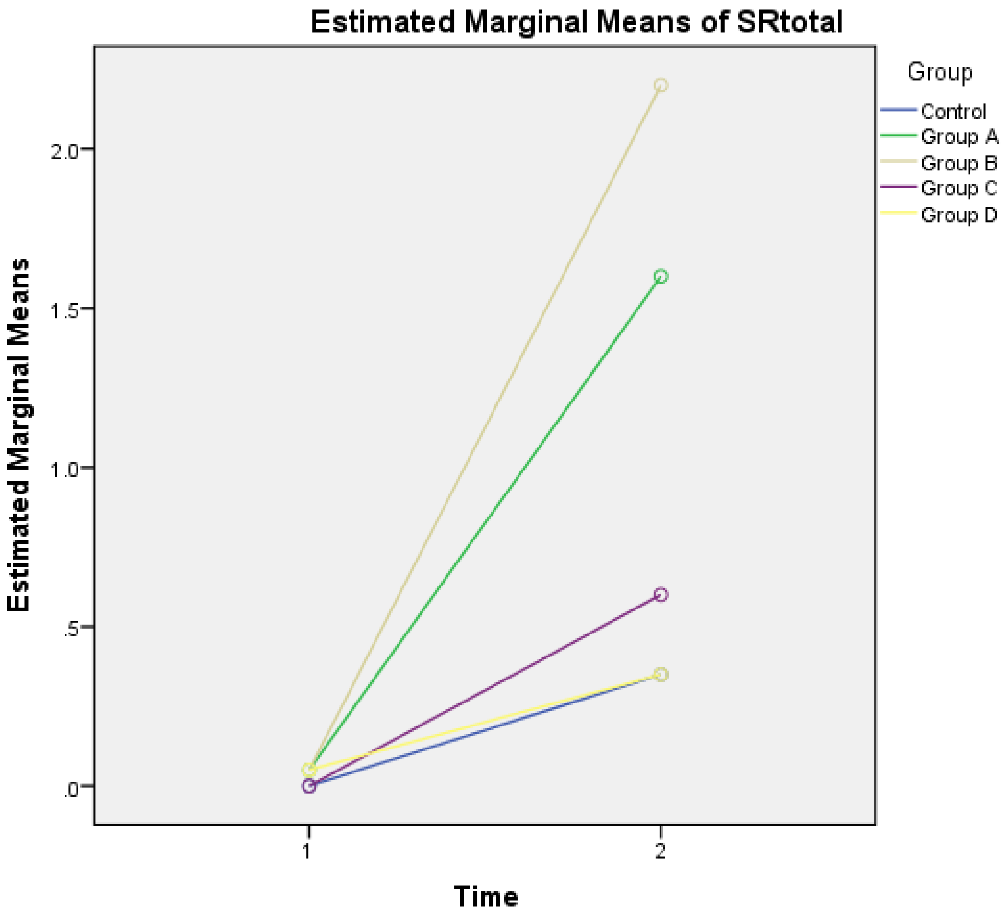


Figure 5.2: Between-subject effect of the total score of pre-test, post-test SR items

The repeated MANOVA analyses for the GJ task indicated no significant between group effect, Wilks lambda = 0.57, *F*(4, 95) = 1.28, *p* = .281, partial eta squared = .051. However, a significant interaction between treatment conditions and improvement over time, Wilks' Lambda = 0.81, *F*(4, 95) = 5.48, *p* = .001, partial eta squared = .188, was found. The comparison of the gain scores of the groups by means of ANOVA and post-hoc Bonferroni tests showed that *enhanced+ instructions* (A) group (*p* = .005), the *enhanced+ instructions+ explanation* group (B) (*p* = .001), and the *enhanced only* (C) group (*p* = .036), achieved significantly higher gains than the control group (see Table 5.17 and Figure 5.2). The pair-wise t-tests comparing the pre- and post-test GJ task scores revealed that the control group, *t*(19) = 0.30, *p* = 0.761, did not improve their scores over time. The *enhanced only* (C) group, $t(19) = 4.26$, $p < 0.001$, *enhanced+ instructions* (A) group, $t(19) = 4.34$, $p < 0.001$, and the *enhanced+ instructions+ explanation* group (B), $t(19) = 8.35$, $p < 0.001$, made significant gains. The improvement of the *unenhanced* (D) group did not reach the level of significance once Bonferroni correction was applied, $t(19) = 2.81$, $p = 0.011$.

Table 5.17

Between-subject effect of the GJ items (N = 100)

Group	Group	Mean difference	*p*
Control (E) (*n* = 20)	enhanced+ instructions (A)	-1.20*	.005
	enhanced+ instructions+ explanation (B)	-2.00*	.001
	enhanced only (C)	-.90*	.036
	unenhanced (D)	-.28	1.000
enhanced+ instructions (A) (*n* = 20)	control	1.20*	.005
	enhanced+ instructions+ explanation (B)	-.70	1.000
	enhanced only (C)	.15	1.000
	unenhanced (D)	.07	.551
enhanced+ instructions+ explanation (B) (*n* = 20)	control (E)	2.00*	.001
	enhanced+ instructions (A)	.70	1.000
	enhanced only (C)	.85	.910
	unenhanced (D)	.77	.314
enhanced only (C) (*n* = 20)	control (E)	.90*	.036
	enhanced+ instructions (A)	-.15	1.000
	enhanced+ instructions+ explanation (B)	-.85	.910
	unenhanced (D)	.07	1.000
unenhanced (D) (*n* = 20)	control (E)	.28	1.000
	enhanced+ instructions (A)	-.07	.551
	enhanced+ instructions+ explanation (B)	-.77	.314
	enhanced only (C)	-.07	1.000

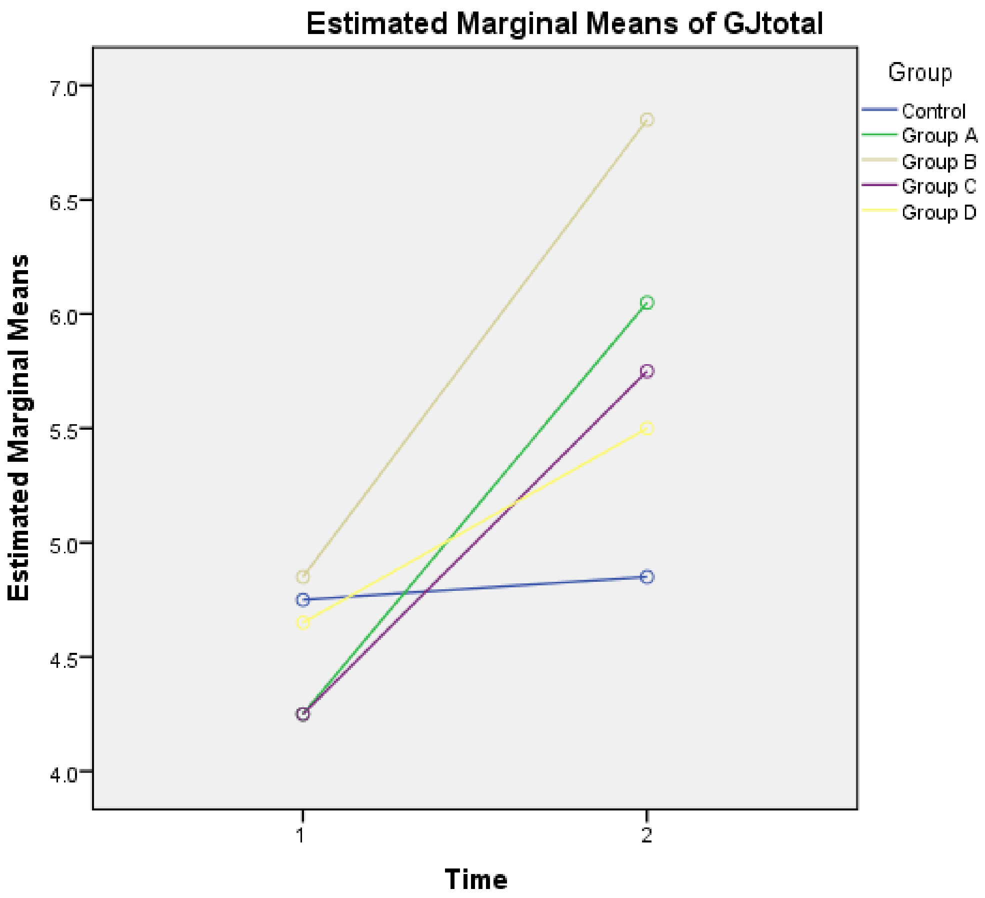


Figure 5.3: Between-subject effect of the total score of pre-test, post-test GJ items

Considering the fact that SR task scores are not normally distributed, a paired t-test was performed to compare the scores between the total score of the pre-test SR items and the transformed variable, Binned Post SR total (discussed in 5.2.1.1). A statistically significant increase from the total score of pre-test SR items (M = .03, SD = .171) to the post-test Binned Post SR total (M = 1.63, SD = .747) was observed, $t(99) = 22.07$, $p <$ .001. The mean increase was 1.60 with a 95% confidence interval ranging from 1.45 to 1.74. The eta squared statistic (.83) indicates a large effect size. This highlights that both the raw total of the SR task and the transformed total of the SR task provide similar results. A Wilcoxon Signed Rank test was also performed for the total score of the pre-test SR items and the total score of the post-test SR items to see whether a non-parametric test would provide different results. It also revealed a statistically significant increase in the total score of the post-test compared to the total score of the pre-test, z = 5.97, $p < .0001$, with a large effect size ($r = .59$). Since these additional analyses did not give different results, the further analyses were conducted based on the SR raw scores.

5.3.2 How does attention paid to examples of the *causative had* construction in the input texts differ under explicit and implicit instructional conditions?

In this study I also aimed to find the answer to the question whether the attention participants pay to the TGC differs under different treatment conditions. First, mean scores for both variables i.e. TFD for all AOIs (TFD-AOI) and the difference between the observed and expected TFD for all AOIs (∆OE-AOI) was computed. As can be seen in Table 5.18 (and Figure 5.4), the *enhanced+ instructions+ explanation* (B) group has the highest TFD-AOI and ∆OE-AOI mean values. This group is followed by the *enhanced+ instructions* (A) group, while the lowest value is recorded for the participants in the *unenhanced* (D) group. ∆OE-AOI values indicate that *enhanced+ instructions* (A), *enhanced+ instructions+ explanation* (B) and *enhanced only* (C) groups spent more time than expected gazing at the examples of target construction. The ∆OE-AOI is almost 0 in the *unenhanced* (D) group, suggesting that the participants in this group did not spend additional time than the normal reading time on the examples of the target construction.

Table 5.18

Descriptive statistics of the composite mean TFD-AOI and ΔOE-AOI for groups (n = 45)

Group	*n*	*M* TFD-AOI (in milliseconds) for all texts	*SD*	*M* ΔOE-AOI (in milliseconds) for all texts	*SD*
enhanced+ instructions (A)	10	2353.41	536.55	1485.28	593.45
enhanced+ instructions+ explanation (B)	14	2805.90	742.23	1824.28	717.71
enhanced only (C)	11	1502.33	216.88	506.20	306.14
unenhanced (D)	10	992.24	337.66	2.95	100.21

**Note*. Due to the poor quality of eye-tracking data, some participants had to be excluded from the analysis (please see Section 4.8.1). The number of participants included in the analysis is given under column *n*.

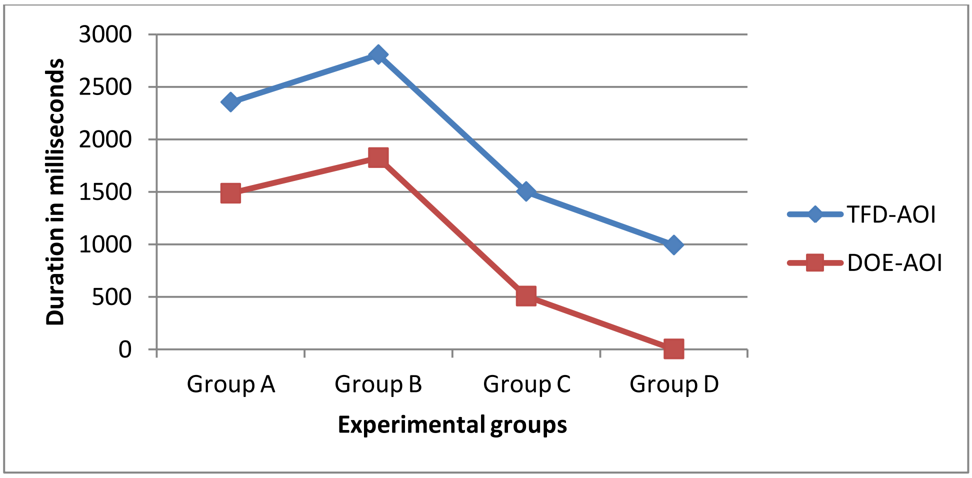


Figure 5.4: Composite mean value of the TFD-AOI and ΔOE-AOI for experimental groups

Figures 5.5 - 5.8 illustrate the mean fixation durations and the difference between the observed and expected fixation duration values for the three texts separately in each of the groups. In the *enhanced+ instructions* (A) group both TFD-AOI and ΔOE-AOI values are initially high for the first target item (see Figure 5.4 and Table 5.19). This is followed by a decrease for the second item, and then the values stabilize from the third item onwards. This trend is similar for all the three texts, with the exception of some fluctuations of the ∆OE-AOI values in Text 3 for items 5 and 6. As can be seen in Figure 5.6 (see Table 5.20 for the mean values), the fixation measures of the *enhanced+ instructions+ explanation* (B) group across the three texts show a similar pattern to those of the *enhanced+ instructions* (A) group. The only difference is that in the *enhanced+ instructions+ explanation* (B) group, the fluctuations in the ∆OE-AOI values for items 5 and 6 in Text 3 are considerably smaller.

Table 5.19

TFD-AOI and ∆OE-AOI for each target item in the three texts of enhanced+ instructions (A) group (n = 10)

	Text 1 TFD-AOI	Text 1 ∆OE-AOI	Text 2 TFD-AOI	Text 2 ∆OE-AOI	Text 3 TFD-AOI	Text 3 ∆OE-AOI
AOI1	3878.56	2923.56	3536.38	2635.33	3298.23	2837.67
AOI2	3042.56	2225.76	3078.99	2172.67	2550.57	2040.56
AOI3	2268.56	1280.67	2384.57	1667.74	1612.67	996.34
AOI4	2127.67	1369.10	2063.76	1317.34	1496.32	797.45
AOI5	2172.45	1318.88	1888.34	910.67	1380.45	150.23
AOI6	2062.67	1317.78	1690.78	1047.23	1572.56	1045.67
AOI7	1718.90	1044.77	1524.64	947.52	905.77	326.56

Note. Values are in milliseconds

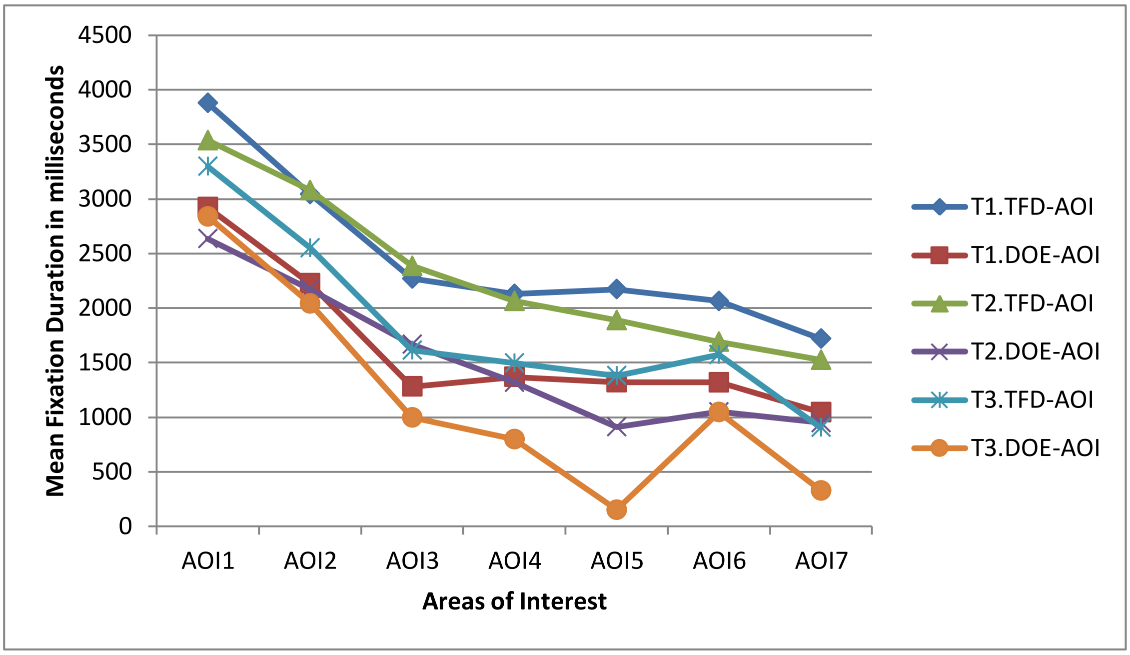


Figure 5.5. TFD-AOI and ΔOE-AOI values for each target item in the three texts in the *enhanced+ instructions* (A) group (*n* = 10)

Table 5.20

TFD-AOI and ΔOE-AOI values for each target item in the three texts of enhanced+ instructions+ explanation (B) group (n = 14)

	Text 1 TFD-AOI	Text 1 ΔOE-AOI	Text 2 TFD-AOI	Text 2 ΔOE-AOI	Text 3 TFD-AOI	Text 3 ΔOE-AOI
AOI1	3919.31	2964.98	5237.56	4115.78	4152.67	3255.23
AOI2	2820.52	1982.67	3892.99	2771.56	2795.85	1894.38
AOI3	2276.43	1271.45	2924.77	2038.45	2573.67	1711.44
AOI4	2232.78	1319.84	3023.84	1897.48	2696.56	1658.40
AOI5	2227.48	1181.96	2542.53	1400.35	1972.22	766.77
AOI6	2304.84	1512.66	2673.97	1905.75	1750.11	545.55
AOI7	1732.94	1123.67	2542.87	1903.86	1653.78	782.51

Note. *Values are in milliseconds*

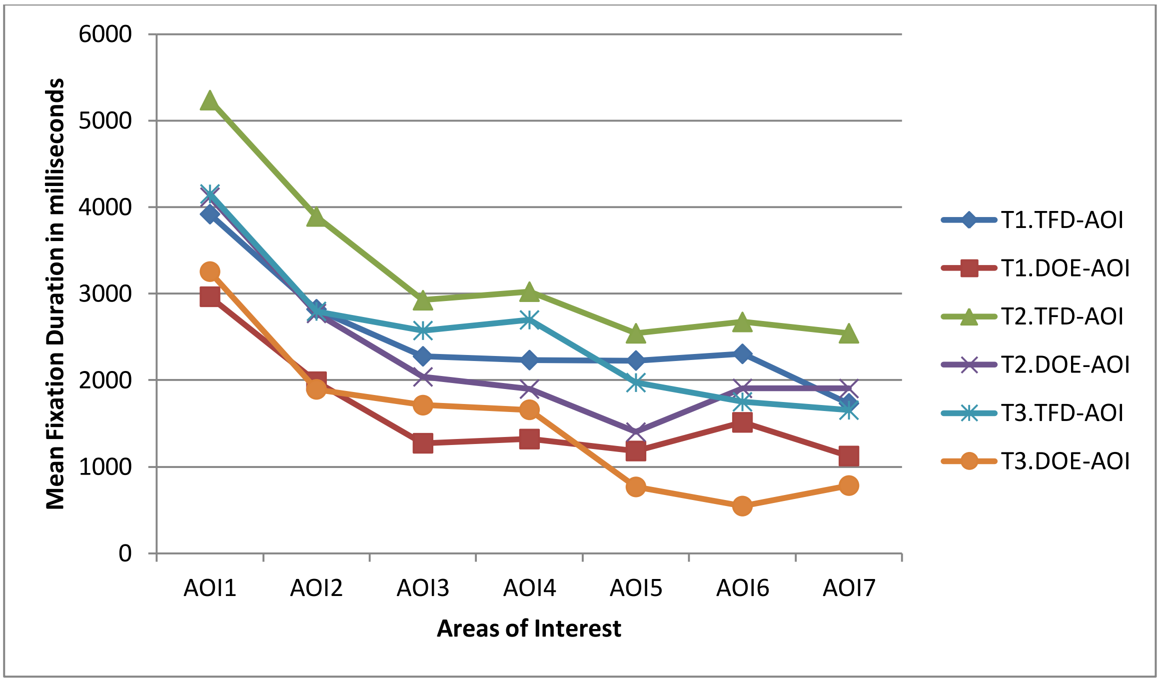


Figure 5.6. TFD-AOI and ΔOE-AOI values for each target item in the three texts in the *enhanced+ instructions+ explanation* (B) group (*n* = 14)

The TFD-AOI and the ΔOE-AOI values for the participants in the *enhanced only* (C) and *unenhanced* (D) conditions show a remarkably different pattern from the previous two treatment conditions (see Figures 5.7 and 5.8, Tables 5.21 and 5.22). In the *enhanced only* (C) group in Text 1, there is an initial fluctuation in fixations in the first 4 items, which is followed by a decrease for the remaining of the items in the text. In Text 2 there is a stable fixation pattern throughout, whereas for Text 3 there is an initially high fixation value for item 1, which then decreases. ΔOE-AOI is around 0 for item 5 in Text 3, and then sinks below 0 for item 7 indicating that the learners spent less time than expected on this item. The eye-tracking measures in the *unenhanced* (D) group show random fluctuation in all three exposures. Except for Text 2, the ΔOE-AOI of more than half of the items is a minus value, which is suggestive of lack of additional processing on these items. (Please see Appendix Y for total reading times for each participant separately and for groups in all texts).

Table 5.21

TFD-AOI and ΔOE-AOI values for each target item in the three texts of enhanced only (C) group (n = 11)

	Text 1 TFD-AOI	Text 1 ΔOE-AOI	Text 2 TFD-AOI	Text 2 ΔOE-AOI	Text 3 TFD-AOI	Text 3 ΔOE-AOI
AOI1	1618.22	569.01	1501.24	794.74	1819.70	1205.96
AOI2	1618.21	481.2	1491.10	784.61	1082.50	468.76
AOI3	1330.45	382.94	1479.35	914.15	1320.23	477.61
AOI4	1528.2	651.6	1561.96	867.34	1290.12	345.19
AOI5	1156.8	59.21	1503.61	270.33	1447.78	31.35
AOI6	1222.65	280.39	1459.02	850.09	1502.40	782.04
AOI7	1030.9	-189.2	1518.16	1074.86	1595.48	875.12

Note. Values are in milliseconds

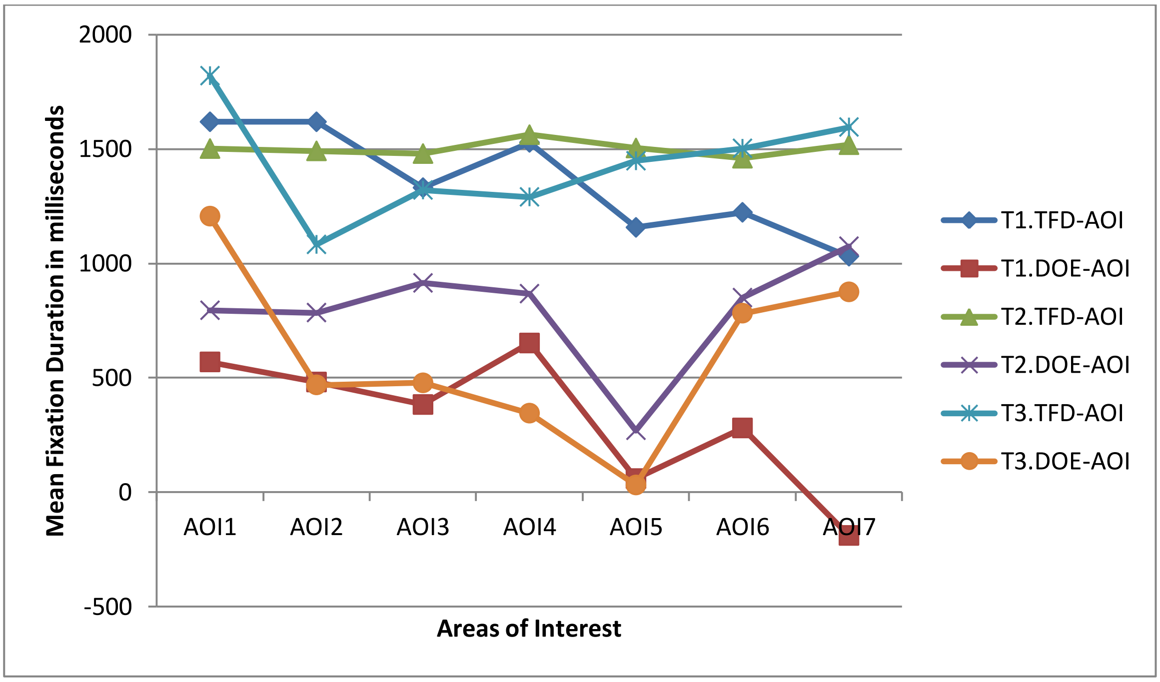


Figure 5.7. TFD-AOI and ΔOE-AOI values for each target item in the three texts in the *enhanced only* (C) group (*n* = 11)

Table 5.22

TFD-AOI and ΔOE-AOI values for each target item in the three texts of unenhanced (D) group (n = 10)

	Text 1 TFD-AOI	Text 1 ΔOE-AOI	Text 2 TFD-AOI	Text 2 ΔOE-AOI	Text 3 TFD-AOI	Text 3 ΔOE-AOI
AOI1	931.90	-82.28	1263.30	369.54	1434.91	553.29
AOI2	1551.70	565.07	997.41	103.65	749.23	-160.93
AOI3	984.84	-199.11	773.60	58.59	943.66	-142.32
AOI4	1158.95	150.68	628.30	-358.15	1023.62	-236.74
AOI5	751.33	-479.89	928.63	26.26	1117.92	-731.33
AOI6	1091.85	171.26	876.33	187.84	1195.87	197.10
AOI7	731.25	7.91	1019.17	449.57	811.28	-152.89

Note. Values are in milliseconds

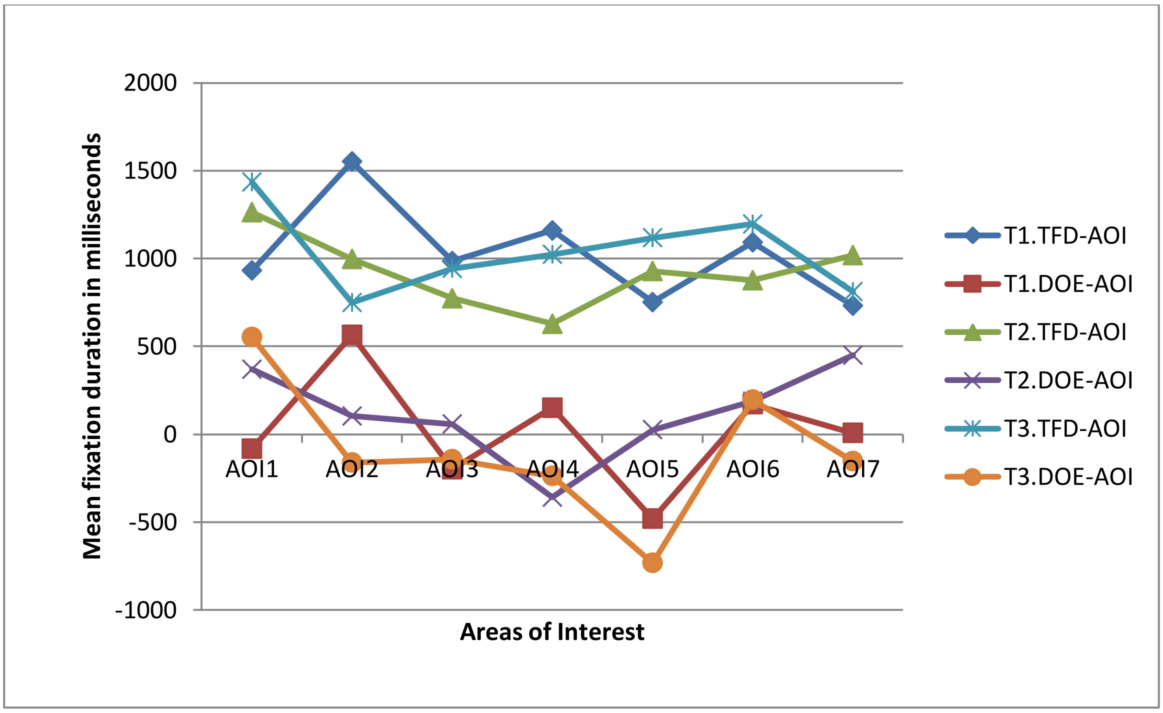


Figure 5.8. TFD-AOI and ΔOE-AOI values for each target item in the three texts in the *unenhanced* (D) group (*n* = 10)

An ANOVA analysis was also conducted to assess the impact of the treatment conditions on the TFD-P and ΔOE-P values for the three texts. The type of input was found to have an overall significant effect on TFD-P, $F(3, 41) = 27.06$, $p < .001$, partial eta squared = .67, and ∆OE-P, $F(3, 41) = 28.42$, $p < .001$, partial eta squared = .68. The effect size can be considered large for both variables. As can be seen in Table 5.23, the post-hoc Scheffe-procedure revealed that there is a significant mean difference in both the TFD-P and ∆OE-P values between the *enhanced+ instructions* (A) group and the *enhanced only* (C) and *unenhanced* (D) groups. Similarly, the *enhanced+ instructions+ explanation* (B) group demonstrated significantly higher TFD-P and ∆OE-P values than the *enhanced only* (C) and *unenhanced* (D) groups. No difference could, however, be detected between the *enhanced only* (C) and *unenhanced* (D) groups and between the *enhanced+ instructions* (A) and *enhanced+ instructions + explanation* (B) groups respectively.

Table 5.23

Differences in eye-tracking measures across groups

Group	Comparisons	Mean difference TFD-P	*p*	Mean difference ΔOE-P	*p*
enhanced+ instructions (A) (*n* = 10)	enhanced+ instructions+ explanation (B)	-452.48	.256	-342.99	.694
	enhanced only (C)	851.08*	.003	979.07*	.001
	unenhanced (D)	1361.17*	.<001	1482.32*	.<001
enhanced+ instructions+ explanation (B) (*n* = 14)	enhanced+ instructions (A)	452.48	.256	342.99	.694
	enhanced only (C)	1303.57*	.<001	1322.06*	.<001
	unenhanced (D)	1813.65*	.<001	1825.31*	.<001
enhanced only (C) (*n* = 11)	enhanced+ instructions (A)	-851.08*	.003	-979.07*	.001
	enhanced+ instructions+ explanation (B)	-1303.57*	.<001	-1322.06*	.<001
	unenhanced (D)	510.08	.186	503.25	.186
unenhanced (D) (*n* = 10)	enhanced+ instructions (A)	-1361.17*	.<001	-1482.32*	.<001
	enhanced+ instructions+ explanation (B)	-1813.65*	.<001	-1825.31*	.<001
	enhanced only (C)	-510.08	.186	-503.25	.186

Note. All values are in milliseconds

In a follow-up analysis I also examined if there is any interaction between the sequence of the reading text and the experimental condition. The results indicated that for both the variable of TFD-P, Wilks lambda = .51, $F(2, 40) = 5.11$, $p < .001$, partial eta squared =.280, and of ΔOE-P, Wilks lambda = .59, $F(2, 40) = 3.98$, $p = .002$, partial eta squared = .23, there is a significant interaction between the three reading texts and experimental condition (see Figures 5.8 and 5.9). The follow-up ANOVA analyses revealed that for

Text 1 the pattern of differences across conditions was the same as for the total TFD-P and ΔOE-P values. For Text 2, however, in addition to the differences found for the total TFD-P and ΔOE-P values, the *enhanced+ instructions+ explanations* (B) group had significantly longer TFD-P ($p < 0.001$), and higher ΔOE-P ($p = 0.17$).

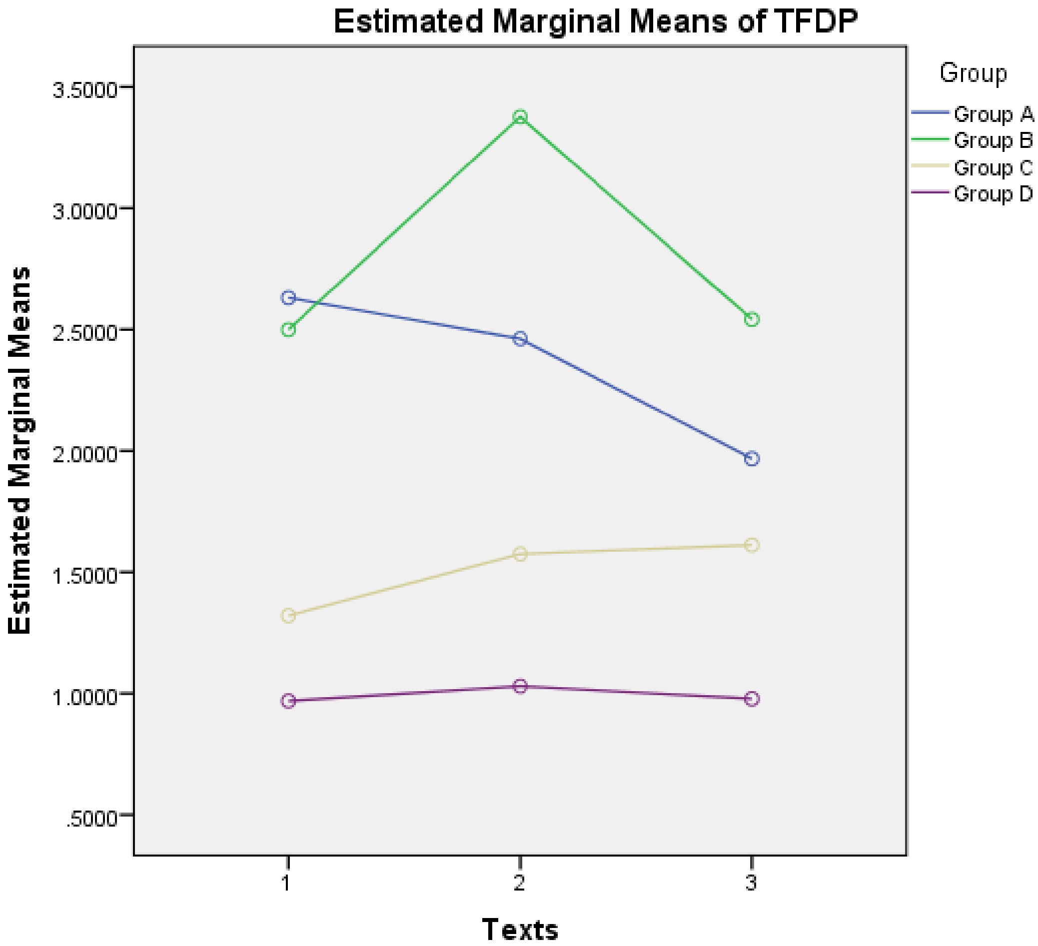


Figure 5.9. Between-subject effect of TFD-P

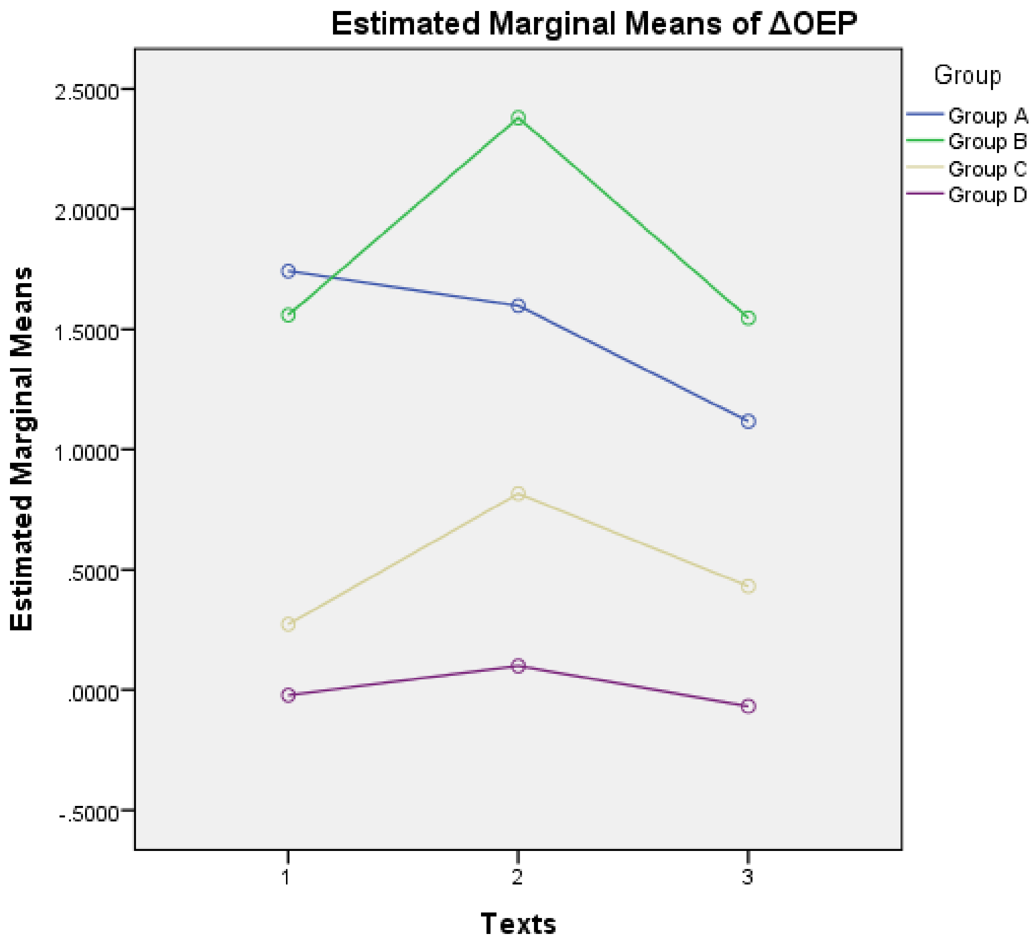


Figure 5.10. Between-subject effect of ΔOE-P

5.3.3 How is the change in knowledge of the target construction under explicit and implicit instructional conditions related to the attention paid to examples of the *causative had* construction in the input texts?

In this study I was also interested in exploring the relationship between the attention paid to the targeted grammatical construction and changes in knowledge in the different instructional conditions. For this purpose, I first computed Spearman rank-order correlations between the eye-tracking measures and the gain scores of the SR items and GJ items for the whole eye-tracking sample (n = 45). As can be seen in Table 5.24, there are strong positive correlations between the gain score of SR items and all variables in the eye-tracking data except between the TFD-P and SR gain score in Text 1. In Text 1 also, a moderate correlation was observed. The correlation between the gain score of the GJ items and ∆OE-P in Text 2 and Text 3 is also strong; but the correlation values between the GJ gain score and other eye-tracking variables only fall in the moderate range. The composite values of both TFD-P and ∆OE-P (for all three texts) show strong positive correlation with both the SR and the GJ gain scores.

Table 5.24

Correlations between eye-tracking measures and gain score of the experimental sample (n = 45)

		SR Gain score	GJ Gain score
TFD-P Text 1	Spearman rho	.454*	.429*
	p	<.001	<.001
TFD-P Text 2	Spearman rho	.699*	.442*
	p	<.001	<.001
TFD-P Text 3	Spearman rho	.508*	.320*
	p	<.001	<.001
TFD-P Composite	Spearman rho	.636*	.524*
	p	<.001	.001
∆OE-P Text 1	Spearman rho	.542*	.454*
	p	<.001	<.001
∆OE-P Text 2	Spearman rho	.686*	.507*
	p	<.001	<.001
∆OE-P Text 3	Spearman rho	.655*	.556*
	p	<.001	<.001
∆OE-P	Spearman rho	.688*	.584*
Composite	*p*	<.001	<.001

*Correlation significant at 0.05 level

Although the sample size in the experimental groups was relatively small, I computed Spearman rank-order correlations to investigate the relationship between the gain scores and the composite values of the eye-tracking measures in different input conditions (see Table 5.25). In the *unenhanced* (D) group no significant correlations were found, whereas for the other three groups the correlations between gain scores in both tasks and the ∆OE-P value were all statistically significant and strong. The TFD-P was also strongly associated with the improvement learners made in both of the tasks in the *enhanced+ instructions+ explanation* (B) and the *enhanced only* (C) conditions, but not in the *enhanced+ instructions* (A) group.

Table 5.25

Correlations between eye-tracking data and gain score in the experimental groups

Group			SR Gain score	GJ Gain score
TFD-P Composite	enhanced+ instructions (A) (*n* = 10)	Spearman rho	.583	.281
		p	.077	.431
	enhanced+ instructions+ explanation (B) (*n* = 14)	Spearman rho	.793*	.761*
		p	.001	.002
	enhanced only (C) (*n* = 11)	Spearman rho	.612*	.654*
		p	.046	.029
	unenhanced (D) (*n* = 10)	Spearman rho	-.242	-.272
		p	.334	.446
ΔOE-P Composite	enhanced+ instructions (A) (*n* = 10)	Spearman rho	.673*	.584*
		p	.008	.028
	enhanced+ instructions+ explanation (B) (*n* = 14)	Spearman rho	.882*	.644*
		p	<.001	.013
	enhanced only (C) (*n* = 11)	Spearman rho	.837*	.692*
		p	.001	.018
	unenhanced (D) (*n* = 10)	Spearman rho	-.266	-.044
		p	.458	.903

*Correlation significant at 0.05 level

The results related to the first set of research questions highlighted that the whole sample of the study together (N = 100) showed a significant improvement from the pre-test to post-test in both SR and GJ tasks. However, only the *enhanced+ instructions* (A) group and the *enhanced+ instructions+ explanation* (B) group improved in both tasks while the learners in the *enhanced only* (C) condition showed significant gains in only the GJ task. The *unenhanced* (D) group did not demonstrate significant learning gains in either of the tasks.

Attentional measurements revealed that the attention paid to target examples (TFD-P and ∆OE-P) significantly correlated with gain scores in both SR and GJ tasks for the whole experimental sample (n = 45). In the case of different conditions, the *unenhanced* (D) group did not show significant correlations between any of the eye-tracking measurements and the gain scores. The other three groups showed significant correlations between ∆OE-P and both SR and GJ gain scores. In addition, the *enhanced+ instructions* (A) condition and the *enhanced+ instructions+ explanation* (B) condition were able to draw the attention of the participants to the target examples significantly better than the *enhanced only* (C) condition and the *unenhanced* (D) condition.

5.4 Results of Section 2 research questions

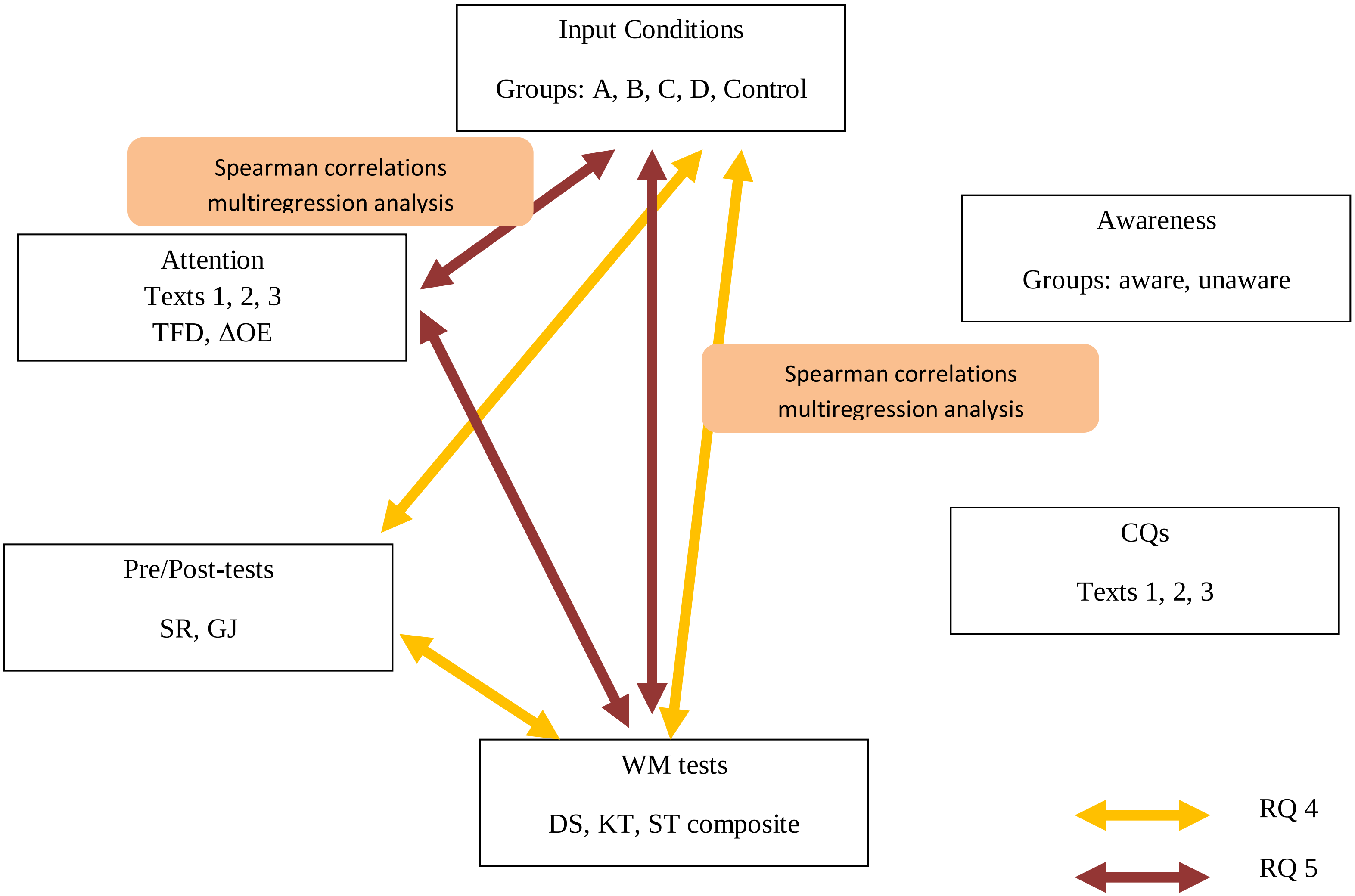
This set of research questions investigated the relationship between WM capacity and the change in knowledge of form and meaning of the target syntactic construction along with the relationship between WM capacity and the amount of attention paid.

RQ 4: How is the functioning of WM including both PSTM capacity and CE functions related to the change of knowledge of the TGC *causative had* in different input conditions?

RQ 5: How is the functioning of the WM related to attention paid to the target items in different input conditions?

Correlational and multiregression analyses were applied to answer both questions mainly based on a composite WM score (see Figure 5.11 for an illustration of the type of analysis done under the second set of RQs).

Figure 5.11. Data analysis methods of the second set of RQs



5.4.1 How is the functioning of WM including both PSTM capacity and CE functions related to the change of knowledge of the TGC *causative had* in different input conditions?

First of all, before analysing the role of PSTM capacity and CE functions in processing L2 input, I checked if the treatment groups were comparable in terms of their cognitive abilities. The one-way analysis of variance showed that the four groups were not statistically different from each other with regard to any of the cognitive tests (see Table 5.11 also for descriptive statistics). As discussed in Section 5.2.2.3 (Table 5.12), the Digit Span and the Keep Track tasks were very highly correlated, whereas performance on the Plus Minus task was not linked to the scores of any of the other tests. A factor analysis on all four tests also confirmed that the Plus-Minus task functioned differently from the other three tests (Its factor loading was .21). As this test did not correlate with any of the performance gains and eye-tracking measures either, it was excluded from further analysis.

There was a significantly high correlation between the Digit Span and the Keep Track tasks and a moderate correlation between the Stroop and the Digit Span and Keep Track tasks. When two or more variables are highly correlated with each other i.e. when there is *multicollinearity*, Type I error may occur when performing several correlational analyses (Tu, Kellett, Clerehugh & Gilthorpe, 2005). Thus, it was necessary to compute a composite WM variable. Creating a composite variable also minimized the danger of increasing the chance of Type I error. Therefore, I conducted a principal component analysis to ascertain whether the remaining three cognitive tests could be combined into one factor. The Kaiser-Meyer-Olkin measure of sampling adequacy was .63, which is higher than the recommended minimum value of .50 (Pett, Lackey & Sullivan, 2003). The Barlett's test of sphericity reached statistical significance ($p < .001$) supporting the factorability of the correlation matrix. The common variance shared by the three subtests was 73.96%, which indicates that the newly created factor accounts for a substantial proportion of the variance in the dataset. The three variables loaded on a single factor with an Eigenvalue of 2.21 and their factor loadings show that they contribute to the factor to a similar extent (Digit Span = .927, Keep Track = .900, Stroop = -.740). Therefore it was deemed to be appropriate to create a composite score using regression factor scores (Tabachnick & Fidell, 2001).

In order to answer the question how the functioning of WM is related to the change of knowledge of the TGC *causative had* in different input conditions, I first checked whether the treatment groups differed from each other in terms of their knowledge of the target construction at the beginning of the study. The one-way analysis of variance revealed that there was no initial significant difference among the groups (SR task pre-test, $F(3,76) = .333$, $p = .801$; GJ task $F(3,76) = .933$, $p = .429$) (see Section 5.2.1 for descriptive statistics on the test scores). Next I conducted a correlational analysis between the individual WM tests and the gains in SR and GJ tasks using the Spearman rank-order correlation. Then the same analysis was performed between the composite WM score and the gains scores of the SR and the GJ tasks. The Digit Span task significantly correlated with both SR and GJ gain scores in all experimental conditions. The Keep Track task also demonstrated a significant and strong correlation with both SR and GJ gain scores in all conditions except in the *enhanced+ instructions* (A) and the *unenhanced* (D) groups in the case of the GJ task. The Stroop task showed a significant correlation with the gains in all conditions except in the SR task in the *unenhanced* (D) group (see Table 5.26).

Table 5.26

Correlation between WM tests and gain scores in the whole experimental sample (n = 80) and in the experimental conditions

Test		*Experimental sample (n=80)*		*enhanced+ instructions (n = 20)*		*enhanced+ instructions+ explanation (n = 20)*		*enhanced only (n = 20)*		*unenhanced only (n = 20)*	
		SR gain score	GJ gain score	SR gain score	GJ gain score	SR gain score	GJ gain score	SR gain score	GJ gain score	SR gain score	GJ gain score
Digit Span	Spearman rho	.570*	.648*	.621*	.467*	.830*	.758*	.597*	.715*	.454*	.685*
	p	<.001	<.001	.003	.038	<.001	<.001	.005	<.001	.045	.001
Keep Track	Spearman rho	.519*	.576*	.631*	.386	.710*	.584*	.447*	.701*	.391	.693*
	p	<.001	<.001	.003	.093	<.001	.007	.048	.001	.088	.001
Plus Minus	Spearman rho	.088	.006	-.533	-.303	.311	.137	.264	-.176	.104	.358
	p	.436	.956	.116	.195	.182	.565	.260	.459	.661	.121
Stroop	Spearman rho	-.568*	-.547*	-.611*	-.444*	-.856*	-.765*	-.580*	-.544*	-.322*	-.520*
	p	<.001	<.001	.004	.050	<.001	<.001	.007	.013	.167	.019

*Correlation significant at 0.05 level

In the case of composite WM score, as Table 5.27 indicates, the correlations between the composite WM score and the gains in both tasks were statistically significant and were particularly strong in the *enhanced+ instructions+ explanation* (B) group. The weakest relationships emerged in the *unenhanced* (D) group. To assess whether there was a differential effect of the WM capacity across the four experimental conditions, I also ran a multiregression model with the experimental condition entered as the dummy variable and the WM composite score as the independent variable. There was one interaction term between the experimental condition and the WM composite score. The *unenhanced* (D) group was used as the reference category. In the case of the gains in the SR task, the multiregression model showed a significant interaction effect between the treatment condition and the composite WM score, Wald $\chi^2 = 23.089$, $p < .001$. The analysis also revealed that the relationship between the SR gain score and the composite WM score in the *unenhanced* (D) group was statistically different from the relationship observed in the case of the *enhanced+ instructions* (A), $\beta = 1.105$, $p < .001$, and the *enhanced+ instructions+ explanation* (B) group, $\beta = .973$, $p < .001$ (see Figure 5.12). This indicates that WM abilities played a stronger role in performance gains in the SR task in the *enhanced+ instructions* (A) and *enhanced+ instructions+ explanation* (B) groups than in the *unenhanced* (D) group. In the GJ task, however, I found no significant interaction between the WM composite score and differential instructional conditions score, Wald $\chi^2 = 1.114$, $p =.766$ (see Figure 5.13).

Table 5.27

Correlations between gains in the SR and GJ task and the composite WM score

Group	*n*	Correlation with WM composite score	SR gain score	GJ gain score
whole sample	80	Spearman rho	.614**	.658**
		p	.000	.000
enhanced+ instructions (A)	20	Spearman rho	.701**	.524*
		p	.001	.018
enhanced+ instructions+ explanations (B)	20	Spearman rho	.887**	.770**
		p	.001	.001
enhanced only (C)	20	Spearman rho	.533*	.714**
		p	.015	.001
unenhanced (D)	20	Spearman rho	.454*	.619**
		p	.04	.004

*Correlation significant at 0.05 level

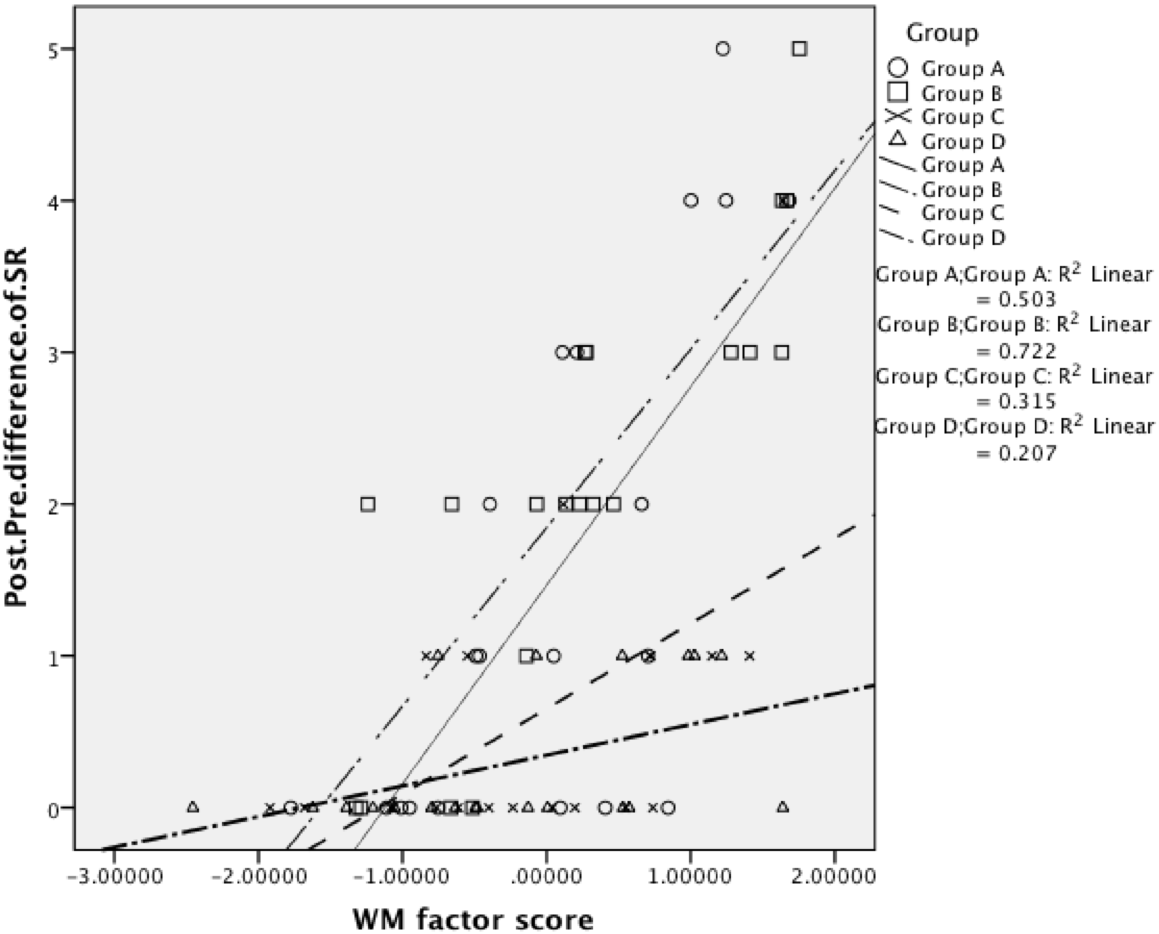


Figure 5.12. Interaction between composite WM value and instructional condition in the SR task

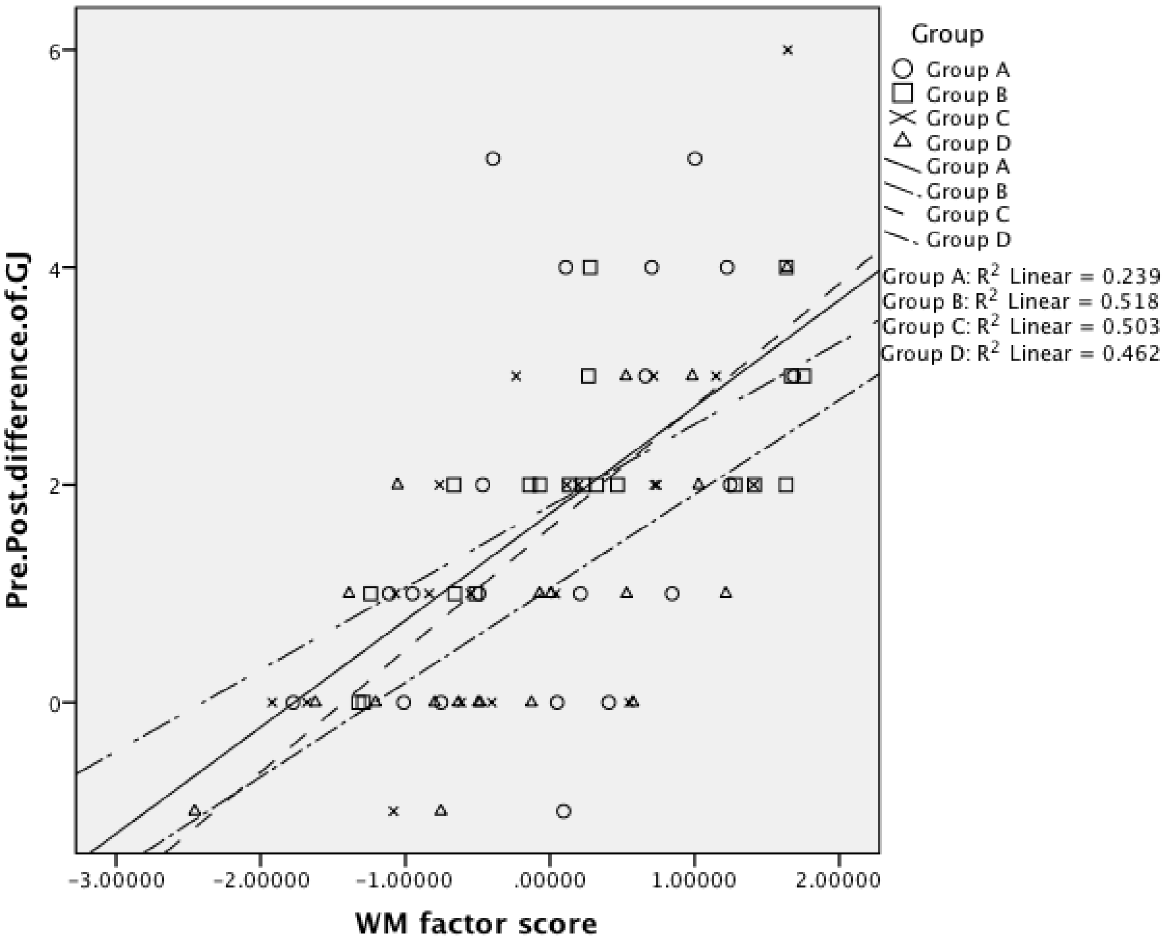


Figure 5.13. Interaction between composite WM value and instructional condition in the GJ task

5.4.2 How is the functioning of the WM related to attention paid to the target items in different input conditions?

In the second research question in this section, I aimed to investigate the nature of relationship between WM capacity and attention paid to the TL items in different input conditions. For this purpose I first conducted correlational analyses using Spearman rank-order correlations between individual WM tests and the mean TFD-P and ΔOE-P values for the whole experimental sample. As can be seen in Table 5.28, the Stroop task correlated significantly with both TFD-P and ΔOE-P values, but the Digit Span and the Keep Track tasks showed significant correlations only with the ΔOE-P value.

Table 5.28

Correlation between eye-tracking measurements and the WM test sub components for the whole experimental sample (n = 45)

		Digit Span	Keep track	Plus minus	Stroop
Mean TFD-P	Spearman rho	.250	.279	.026	-.307*
	p	.097	.064	.838	.040
Mean ΔOE-P	Spearman rho	.327*	.394*	.010	-.310*
	p	.028	.007	.950	.038

*Correlation significant at 0.05 level

When the composite WM score in the whole experimental sample was considered, the correlations between the two eye-tracking measures and the composite WM score were significant, but weak (Table 5.29). In the *enhanced+ instructions+ explanation* (B) group, however, the strength of relationship between the fixation durations and WM capacity was remarkably high. The ΔOE-P value, which has been argued to reflect added attentional processing, also showed a strong correlation with the composite WM score in the *enhanced+ instructions* (A) and the *enhanced only* (C) groups.

Table 5.29

Correlations between eye-tracking measures and the composite WM score

Group	*n*	Correlation with WM composite score	Composite TFD-P	Composite ΔOE-P
whole sample	45	Spearman rho	.324*	.394**
		p	.03	.007
enhanced+ instructions (A)	10	Spearman rho	.382	.697*
		p	.276	.025
enhanced+ instructions+ explanations (B)	14	Spearman rho	.859**	.824**
		p	.000	.000
enhanced only (C)	11	Spearman rho	.155	.673*
		p	.650	.023
unenhanced (D)	10	Spearman rho	-.455	-.030
		p	.187	.934

*Correlation significant at 0.05 level

For the examination of the interaction between WM capacity and experimental conditions, I ran a multiregression model with the experimental condition entered as the dummy variable and the WM composite score as the independent variable. The *unenhanced* (D) group was used as the reference category. In the case of the composite TFD-P, the multiregression model showed a significant interaction effect between the treatment condition and the composite WM score, Wald χ^2 = 34.49, p < .001. The analysis also revealed that the relationship between the composite TFD-P and the composite WM score in the *unenhanced* (D) group was statistically different from the relationship observed in the case of *enhanced+ instructions* (A), β = .274, p = .042, and the *enhanced+ instructions+ explanation* (B) group, β = .723, p < .001 (see Figure 5.14). As can be seen in Figure 5.15, similar findings were obtained for the composite ΔOE-P score, Wald χ^2 = 29.178, p < .001, where the relationship between the composite ΔOE-P and the composite WM score in the *unenhanced* (D) group also differed from the relationship observed in the case of *enhanced+ instructions* (A), β =.

395, p = .001, and the *enhanced+ instructions+ explanation* (B) group, β = .608, p < .001.

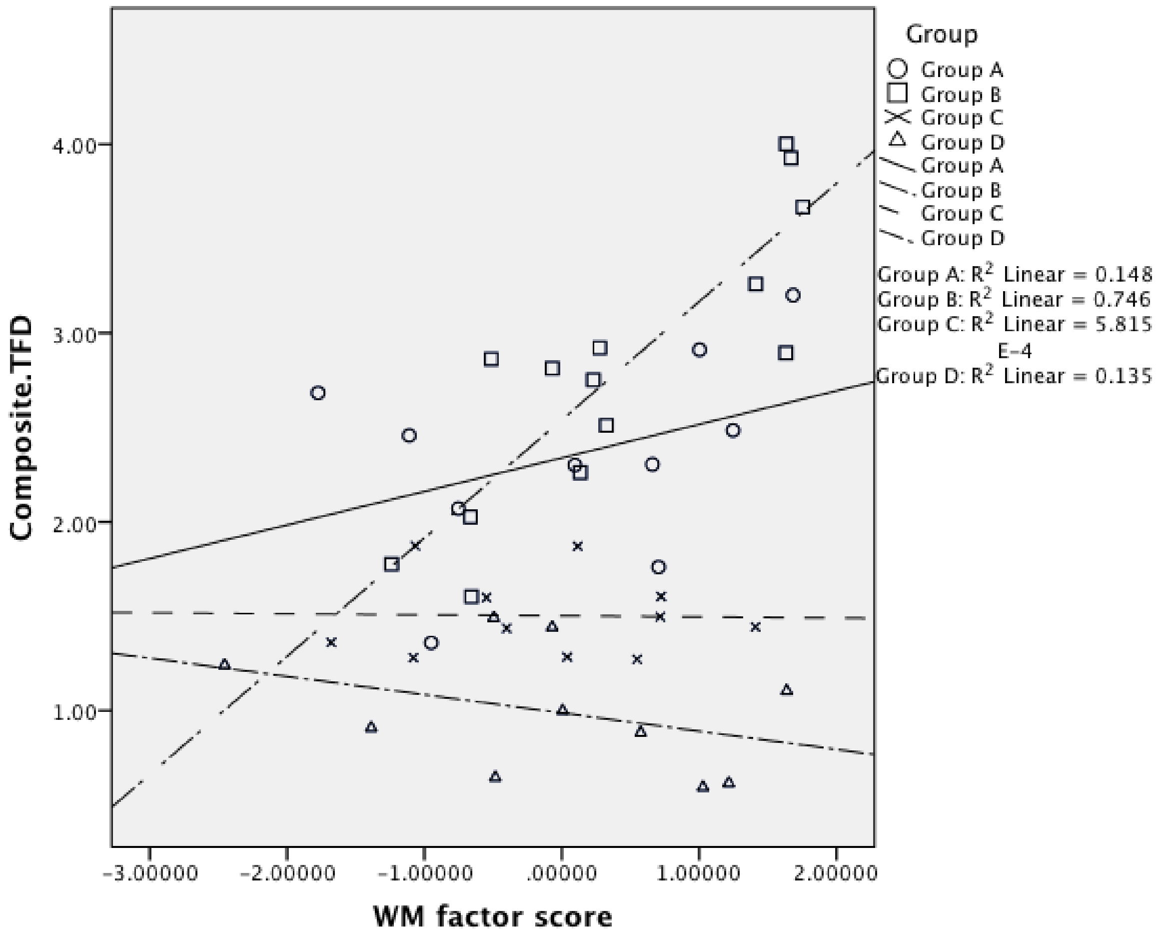


Figure 5.14. Interaction between composite WM value and instructional condition in the case of the TFD-P measure

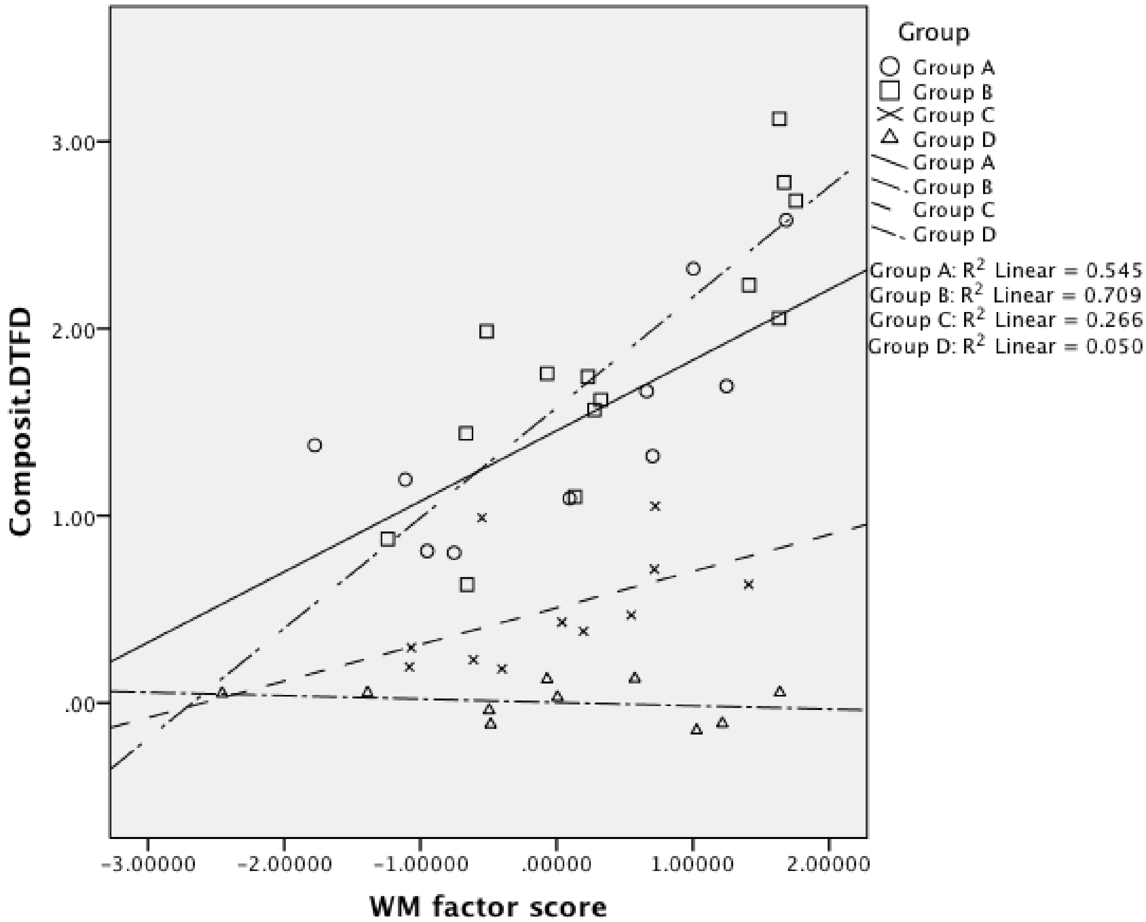


Figure 5.15. Interaction between composite WM value and instructional condition in the case of the ΔOE measure

In summary, the gain scores of SR and GJ tasks showed a significant correlation with the WM composite score with stronger correlations in the *enhanced+ instructions+ explanation* (B) group and weaker correlations in the *unenhanced* (D) group. The multiregression model analysis revealed that the SR gain score highly correlated with the WM composite score in the *enhanced+ instructions+* (A) and *enhanced+ instructions+ explanation* (B) groups compared to the *unenhanced* (D) group, but the interactional effect between the GJ gain score and WM composite score did not show significant difference between instructional conditions.

The TFD-P and ΔOE-P composite values showed only a moderate correlation with the

WM composite score in the case of the whole experimental sample (n = 45), however, when the data of different instructional conditions were analysed separately, the ∆OE-P value in particular showed very strong correlation in the *enhanced+ instructions+ explanation* (B) group and also strong correlations in the *enhanced+ instructions* (A) group and *enhanced only* (C) group. The between group analysis highlighted that there was a significant group difference between the *unenhanced* (D) group and the *enhanced+ instructions* (A) and *enhanced+ instructions+ explanation* (B) groups in the relationship between WM composite score and both TFD-P and ∆OE-P values.

5.5 Results of Section 3 research questions

The third set of research questions analysed whether the attention paid to input by the participants was sufficient for them to understand the meaning of the TGC and whether this has any relationship with the type of input received. The results reported in this section are based on the scores of the CQs that the participants answered during the exposure sessions, the gain scores reported in different input conditions and the answers that they provided in the post-exposure interview.

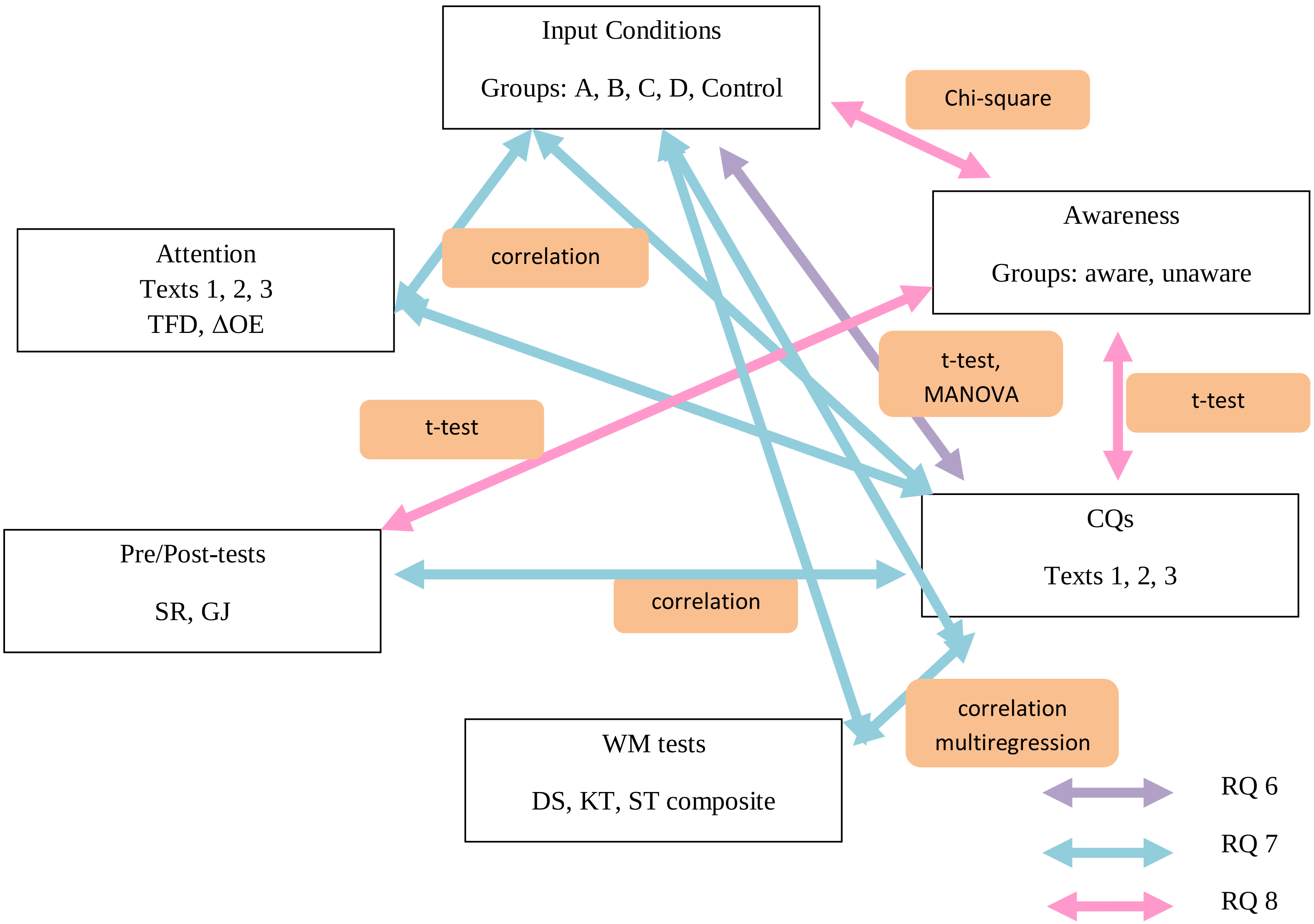
RQ 6: Is there an improvement in the CQ scores in different input conditions between the three exposures?

RQ 7: Is the improvement in the CQ task related to the eye fixation measures, the knowledge of the TGC *causative had* (SR and GJ tasks) and WM?

RQ 8: Is there a difference between the change in the knowledge of the TGC *causative had* (SR, GJ and comprehension tasks) between the participants who reported awareness and those who reported no awareness? Is the awareness reported in different input conditions significantly different to each other?

(see Figure 5.16 for an illustration of the type of analysis done under the third set of RQs).

Figure 5.16. Data analysis methods of the third set of RQs



5.5.1 Is there an improvement in the CQ scores in different input conditions between the three exposures?

In order to answer this question, first I computed a composite score for each exposure session (there were 2 target questions in each session). Table 5.30 summarises the paired t-test results between different exposures for the whole experimental sample (*n* = 80). Accordingly, there is a significant difference between the scores of exposure 1 and 2 and also exposure 3 and 1 in the experimental sample. Table 5.31 shows the highest mean value in exposure 3 and the lowest in exposure 1.

Table 5.30

t-test results of the CQs of three exposure sessions (n = 80)

		Mean difference	*p*	Effect size
Pair 1	Exposure 2 composite score – Exposure 1 composite score	.48*	<.001	.56
Pair 2	Exposure 3 composite score – Exposure 1 composite score	.57*	<.001	.63
Pair 3	Exposure 3 composite score – Exposure 2 composite score	.09	.058	.04

Table 5.31

Mean of CQ results of the three exposure sessions (n = 80)

		M
Pair 1	Exposure 2 composite score	.59
	Exposure 1 composite score	.11
Pair 2	Exposure 3 composite score	.68
	Exposure 1 composite score	.11
Pair 3	Exposure 3 composite score	.68
	Exposure 2 composite score	.59

A mixed between-within subject analysis of variance was conducted to analyse the relationship between the composite comprehension test scores of the three exposures and the different input conditions. There is a large significant interaction between composite score and type of input, Wilk's Lambda = .73, $F(6, 150) = 4.21$, $p = .001$, partial eta squared = .14. There is also a significant main effect for the composite score, Wilk's Lambda = .26, $F(2, 75) = 104.47$, $p < .001$, partial eta squared = .74. The between subject analysis also shows a large significant effect between the type of input provided and the comprehension test scores of the three exposures ($p < .001$) partial eta score = 88. According to Table 5.32 and Figures 5.17, there is a significant mean difference between the *enhanced+ instructions* (A) group and *unenhanced* (D) group ($p = .001$), on the one hand, and between the *enhanced+ instructions+ explanation* (B) group and the *enhanced only* (C) ($p = .004$) and *unenhanced* (D) ($p < .001$) groups on the other hand.

Table 5.32

Between-subject effect of the mean difference of CQ results - three exposure sessions

Group	Comparison	Mean difference	*p*
enhanced+ instructions (A) (*n* = 20)	enhanced+ instructions+ explanation (B)	-.11	.361
	enhanced only (C)	.09	.664
	unenhanced (D)	.22*	.001
enhanced+ instructions+ explanation (B) (*n* = 20)	enhanced+ instructions (A)	.11	.361
	enhanced only (C)	.20*	.004
	unenhanced (D)	.33*	<.001
enhanced only (C) (*n* = 20)	enhanced+ instructions (A)	-.09	.664
	enhanced+ instructions+ explanation (B)	-.20*	.004
	unenhanced (D)	.13	.129
unenhanced (D) (*n* = 20)	enhanced+ instructions (A)	-.22*	.001
	enhanced+ instructions+ explanation (B)	-.33*	<.001
	unenhanced only (C)	-.13	.129

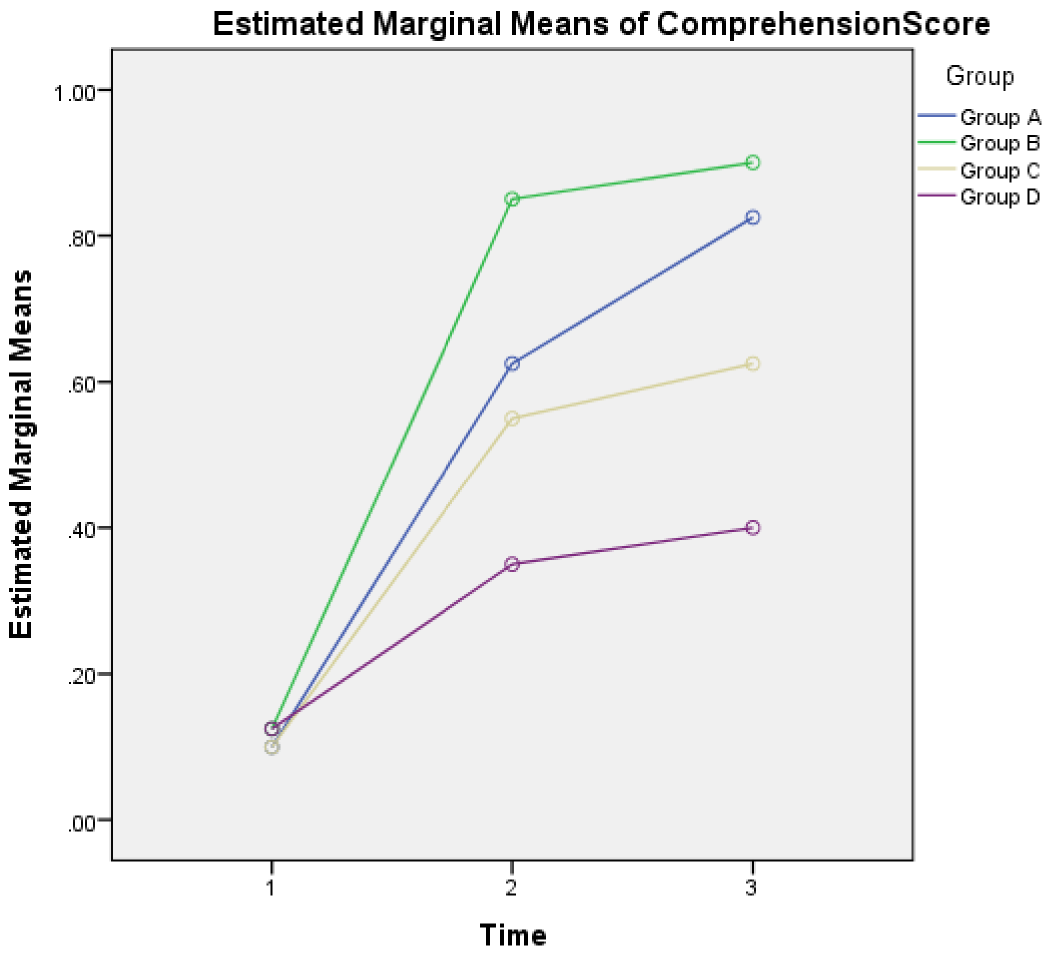


Figure 5.17. Between-subject effect of the mean difference of CQ results - three exposure sessions

The pair-wise t-tests were computed between the Exposure 1 composite CQ test score and the Exposure 3 composite CQ test under each input condition. The *enhanced+ instructions+ explanation* group (B), $t(19) = 10.10$, $p < .001$ and *enhanced+ instructions* (A) group, $t(19) = 10.72$, $p < .001$ made significant gains, but not the other two input conditions. The t-test analysis between Exposure 1 composite CQ test score and the Exposure 2 composite CQ test revealed that only *enhanced+ instructions+ explanation* group (B) made significant gains, $t(19) = 9.10$, $p = .001$. Between Exposures 2 and 3, none of the individual groups showed significant gains.

5.5.2 Is the improvement in the CQ task related to the eye fixation measures, the knowledge of the TGC *causative had* (SR and GJ tasks) and WM?

In order to gain an answer to this question, the correlation between the composite mean TFD-P, composite mean ΔOE-P and composite comprehension test score [28] (of Exposures 1, 2 and 3) was analysed. Table 5.33 reports the correlation of these variables for the whole experimental sample (n = 45). Accordingly, both TFD-P and ΔOE-P significantly correlate with the composite comprehension test score, TFD-P: $r(43) = .71, n = 45, p < .001$, ΔOE-P: $r(43) = .71, n = 45, p < .001$. In the case of experimental conditions, as highlighted in Table 5.34, both TFD-P and ΔOE-P show a significant and large correlation with the composite comprehension score in the *enhanced+ instructions+ explanation* (B) group. The ΔOE-P value and the composite comprehension test scores are also strongly correlated in the *enhanced+ instructions* (A) group.

Table 5.33

Correlations between composite comprehension test score and eye-tracking data (n = 45)

		TFD-P	ΔOE -P
Composite comprehension test score	Spearman rho	.712*	.708*
	p	<.001	<.001

*Correlation significant at 0.05 level

[28] Composite comprehension test score = This refers to the composite score of Exposures 1, 2 and 3 in the following discussions unless otherwise mentioned.

Table 5.34

Correlations between eye-tracking data and composite comprehension test score for each group

	n		TFD –P	∆OE – P
enhanced+ instructions (A)	10	Spearman rho	.486	.748*
		p	.154	.013
enhanced+ instructions+	14	Spearman rho	.622*	.629*
explanation (B)		*p*	.018	.016
enhanced only (C)	11	Spearman rho	.459	.186
		p	.155	.585
unenhanced (D)	10	Spearman rho	.491	-.557
		p	.150	.095

*Correlation significant at 0.05 level

The correlation between SR and GJ gain scores and the composite comprehension score was also analysed. As illustrated in Table 5.35, both SR and GJ gain scores show a significant and strong correlation with the composite comprehension score in the whole experimental sample, SR: $r(78) = .71$, $n = 80$, $p < .001$, GJ: $r(78) = .59$, $n = 80$, $p < .001$. The correlation analyses across input conditions indicate a significant and strong correlation between both the SR and GJ gain scores and the composite comprehension score in the *enhanced+ instructions* (A) and *enhanced+ instructions+ explanation* (B) groups. A strong correlation was also observed between the SR gain score and composite comprehension test score in the *enhanced only* (C) condition (see Table 5.36).

Table 5.35

Correlations between composite comprehension test score and SR and GJ gain score for the experimental sample (n = 80)

		SR Gain	GJ Gain
Composite comprehension test score	Spearman rho	.708*	.593*
	p	<.001	<.001

*Correlation significant at 0.05 level

Table 5.36

Correlations between composite comprehension test score and SR and GJ gain score in different input conditions

		SR Gain	GJ Gain
enhanced+ instructions (A) (*n* = 20)	Spearman rho	.699*	.644*
	p	.001	.002
enhanced+ instructions+ explanation (B) (*n* = 20)	Spearman rho	.648*	.689*
	p	.002	.001
enhanced only (C) (*n* = 20)	Spearman rho	.585*	.330
	p	.007	.155
unenhanced (D) (*n* = 20)	Spearman rho	.398	.314
	p	.082	.178

*Correlation significant at 0.05 level

Finally, the relationship between the performance in the comprehension test and the [29]WM capacity was investigated by performing correlational analysis between the composite comprehension test score and the composite WM factor score for the whole sample. A significant medium correlation was observed in the analysis. In the case of different groups, the composite comprehension test score and the composite WM score revealed a significant correlation in the *enhanced+ instructions* (A) and *enhanced+*

[29] The analyses related to WM in this section are based on the composite WM score.

instructions+ explanation (B) groups. The *enhanced only* (C) and *unenhanced* (D) groups did not show significant correlations (see Table 5.37).

Table 5.37

Correlations between composite comprehension test score and WM composite score

Group		WM composite score & CQ composite score
whole sample	Spearman rho	.453*
	p	.001
enhanced+ instructions (A)	Spearman rho	.513*
	p	<.001
enhanced+ instructions+	Spearman rho	.545*
explanation (B)	*p*	<.001
enhanced only (C)	Spearman rho	.402
	p	.120
unenhanced (D)	Spearman rho	.322
	p	.541

*Correlation significant at 0.05 level

A multiregression model analysis was also performed in order to assess if there was a differential effect of the WM capacity across input conditions in the case of composite comprehension test score. In this analysis, the experimental condition was used as the dummy variable and the WM composite score was the independent variable. The reference category was the *unenhanced* (D) group. The analysis highlighted that the relationship between the composite comprehension test score and the composite WM score in the *unenhanced* (D) group was statistically different from this relationship in the *enhanced+ instructions* (A), $\beta = .197$, $p < .001$, and the *enhanced+ instructions+ explanation* (B) group, $\beta = .294$, $p < .001$. The results highlighted that WM capacity has played a stronger role in the gains on CQs in the *enhanced+ instructions* (A) and the *enhanced+ instructions+ explanation* (B) groups than in the *unenhanced* (D) group.

5.5.3 Is there a difference between the change in the knowledge of the TGC *causative had* (SR, GJ and comprehension tasks) between the participants who reported awareness and those who reported no awareness? Is the awareness reported in different input conditions significantly different to each other?

For the purpose of answering this question, the participants who reported awareness [‘aware’] were separated from those who did not report awareness [‘unaware’]. Table 5.38 highlights the raw data of the reported awareness in each group. Thereafter, a t-test was performed for the ‘aware’ and ‘unaware’ group separately for the SR, GJ and comprehension items. In the case of comprehension items, the performance in test three (Text 3) was compared with the performance in test one (Text 1). The results show that both ‘aware’ and ‘unaware’ groups demonstrated a statistically significant development in the knowledge of the TGC in the SR, GJ and comprehension tasks (see Table 5.39).

Table 5.38

Raw number of reported awareness in different input conditions

Group	*n*	Reported
enhanced+ instructions (A)	20	4
enhanced+ instructions+ explanation (B)	20	7
enhanced only (C)	20	1
unenhanced (D)	20	0

Table 5.39

t-test results of 'aware' and 'unaware' groups

Group	Test	Mean difference	*t*	*p*
Aware (*n* = 12)	SR	3.42	10.17	<.001
	GJ	2.83	8.80	<.001
	Comprehension	1.75	13.40	<.001
Unaware (*n* = 68)	SR	.66	6.02	<.001
	GJ	1.03	6.26	<.001
	Comprehension	1.04	9.69	<.001

It was also deemed necessary to analyse if the reported awareness was related to the type of input that the participants received. Thus, a chi-square analysis was performed between the 'aware' and 'unaware' groups and the four experimental conditions. The results show that there is a significant relationship between the experimental conditions and reported awareness, $\chi^2(1, N = 100) = 11.76, p < .05$. According to Table 5.40, which provides the standard residuals in different input conditions, a significantly larger number of participants than expected in the *enhanced+ instructions+ explanation* (B) group reported awareness.

Table 5.40

Standard residuals reported in different input conditions

Group	Aware	Unaware
enhanced+ instructions (A)	.6	-.2
enhanced+ instructions+ explanation (B)	2.3	-1.0
enhanced only (C)	-1.2	.5
unenhanced (D)	-1.7	.7

In summary, the data analysis related to the research questions in Section 3 revealed that the whole sample showed a significant gain in exposure 2 and exposure 3 CQ score compared to exposure 1 CQ score. The between group comparison revealed that the *enhanced+ instructions* (A) group and the *enhanced+ instructions+ explanation* (B) group showed a significant gain compared to the *unenhanced* (D) group. The group difference is also significant in the *enhanced+ instructions+ explanation* (B) group compared to the *enhanced only* (C) group. CQ score of the whole experimental condition (n = 45) also significantly correlated with both eye-tracking measurements. The correlation was significant between the CQ score and ∆OE – P in the *enhanced+ instructions* (A) group and the *enhanced+ instructions+ explanation* (B) group and also in the case of TFD-P in the *enhanced+ instructions+ explanation* (B) group. Moreover, a significant correlation between the CQ scores and the WM capacity in the whole experimental condition (n = 80) was observed in the analysis. The relationship is stronger in the case of the *enhanced+ instructions* (A) and the *enhanced+ instructions+ explanation* (B) groups compared to the *unenhanced* (D) group. Both 'aware' and 'unaware' groups show a statistically significant improvement in SR, GJ and comprehension tasks although the improvement in the 'aware' group is significantly larger than in the 'unaware' group in all tasks. The reported awareness was significantly higher in the *enhanced+ instructions+ explanation* (B) group compared to the other three experimental conditions.

5.6 Results of the free writing task

A free production (writing) task was also given to the participants as the third task of both pre and post-tests. This was aimed to measure whether the participants were able to produce the TGC as a result of the exposure that they received. None of the participants in the pre-test used the target construction in this task. Furthermore, none of the participants in the control group, *enhanced only* (C) group or *unenhanced group* (D) were able to use the TGC in the free writing task. However, two participants in the *enhanced+ instructions* (A) group and three participants in the *enhanced+ instructions+ explanation* (B) group used the TGC. Table 5.41 summarises the number of occasions each participant used the target construction and the sentences that they had produced. Figure 5.18 illustrates the pre-test answer of Participant No.44 (Group B) of the free writing task and Figure 5.19 is the post-test answer of the same participant of the same task.

Table 5.41

Use of the TGC in the free writing task

Group	Participant No.	No. of occasions	Sample
enhanced+ instructions (A)	24	1	They had their passports made.
	38	1	I got the information collected.
enhanced+ instructions+ explanation (B)	44	3	I had the local money exchanged to foreign currency, I had the vehicle hired by an agent, I had the flight details checked by our agent
	46	2	I had them planned out, I had the tickets and visa arranged by the school union
	51	1	I had the tickets booked by someone else.

As things that I did for myself includes, getting all required material (things) needed for the trip, double check and pack them accordingly. Make sure my passport and air ticket was all done. Make sure the passports and ~~of~~ air tickets of the 5 scounts were ready as well. ~~A~~ Remind the 5 scouts about the required material (things) needed for the trip.
As things others & had done for me, arranging a transportation method of reaching the air port, getting required information about the scouts by requesting one of them to get it done, Arranging proper accomandation facilities.

Figure 5.18. Free writing task answer of Participant No. 44 (pre-test)

I got my passport ready and my other ~~need~~ required stuff. I had others to ~~do~~ pack my luggage. ~~and~~ I had the laundry man to clean my uniforms. I polished my shoes and prepared food needed for our journey. I had the local money exchanged to foreign currency by one of the junior scouts. I had the vehicle hired by an agent which we needed to go to the airport. ~~Before we the day o~~ I ~~go~~ had the flight details checked by our agent. Finally I made sure we be on time.

Figure 5.19. Free writing task answer of Participant No. 44 (post-test)

5.7 Summary

The aim of this chapter was to present the results obtained in the whole study. The initial data screening performed before the statistical analyses on the data set revealed that some of the data were not normally distributed. Therefore, several transformed variables were computed to seek the possibility of obtaining normal distribution. The initial statistical analyses performed on the non-normally distributed variables and their corresponding transformed variables did not provide different results. The parametric and non-parametric analyses conducted on the non-normally distributed variables did not indicate different findings either. However, the calculations of standardized residuals and Cook's distances highlighted that the data fell within normal distribution. As a result, the subsequent analyses of the data were based only on raw scores and mostly parametric analyses.

The key findings of the data analyses reveal that both SR and GJ gain scores increased significantly over the time for the whole sample with different extent of development between instructional conditions. The amount of attention paid to input measured by TFD-AOI, TFD-P, ∆OE-AOI and ∆OE-P also showed that different instructional conditions were influential in the amount of attention paid to the target examples. The WM composite value which was a combination of the capacity of the PSTM (Digit Span) and updating (Keep Track) and inhibition (Stroop) functions of the CE showed that it can be a strong predictor of the amount of attention paid and the learning gains made by the participants. The analyses of the comprehension test scores revealed that there was an influence of the input condition on the performance in the comprehension task. The results also indicated that awareness of the target syntactic construction of the participants had a minimal involvement with the development of knowledge in both SR and GJ tasks on this occasion (See Figures 5.20, 5.21 and 5.22 for illustrations of the results). Chapter 6 will discuss these results in detail with reference to the literature reviewed in Chapter 2.

Figure 5.20. Relationships between pre and post-tests (SR & GJ), attention (TFD & ΔOE) and WM in different input conditions (width of arrows represents the amount of change)

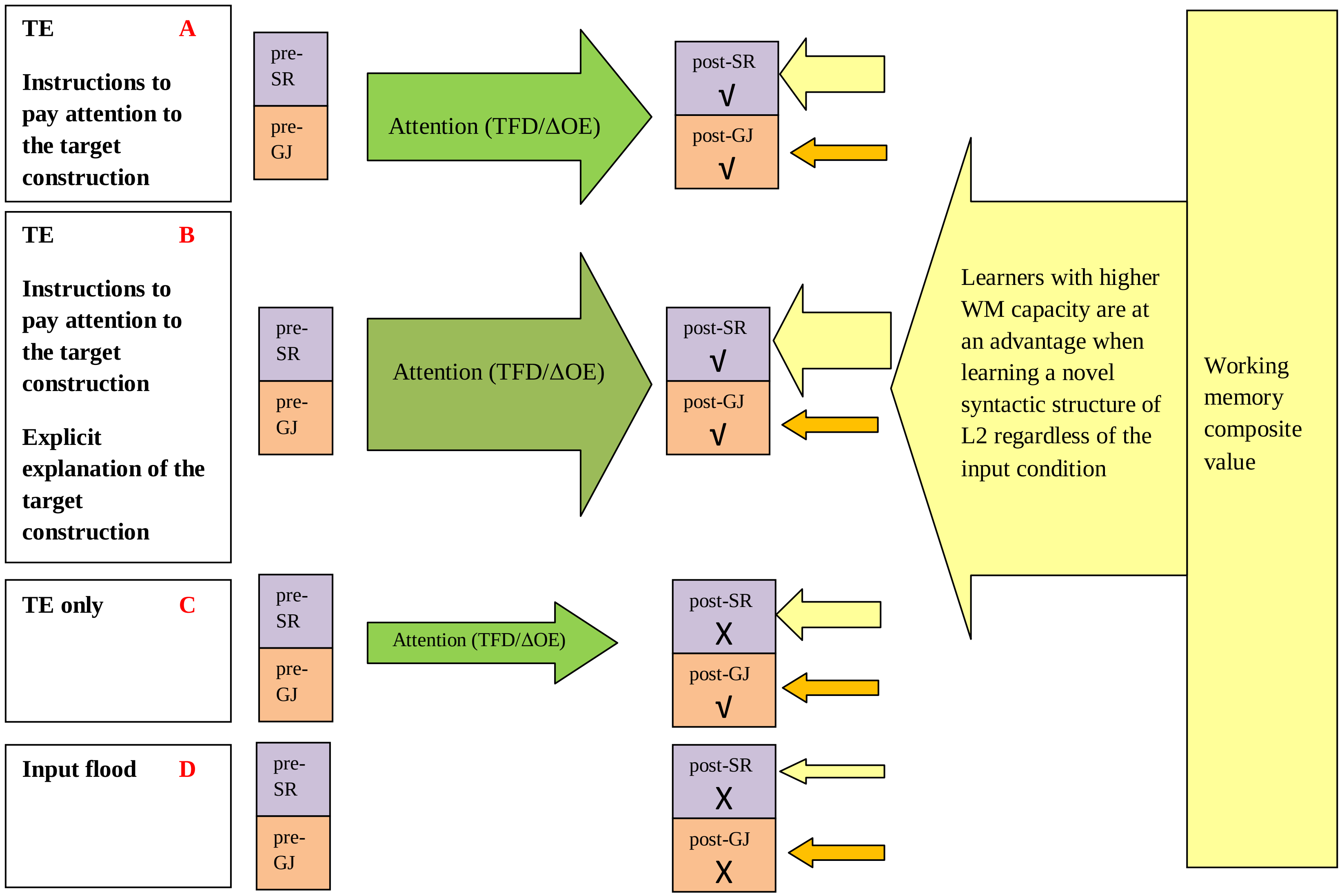


Figure 5.21. Relationship between attention (TFD & ΔOE), pre and post-tests (SR & GJ), WM and the comprehension task (width of arrows represents the amount of change)

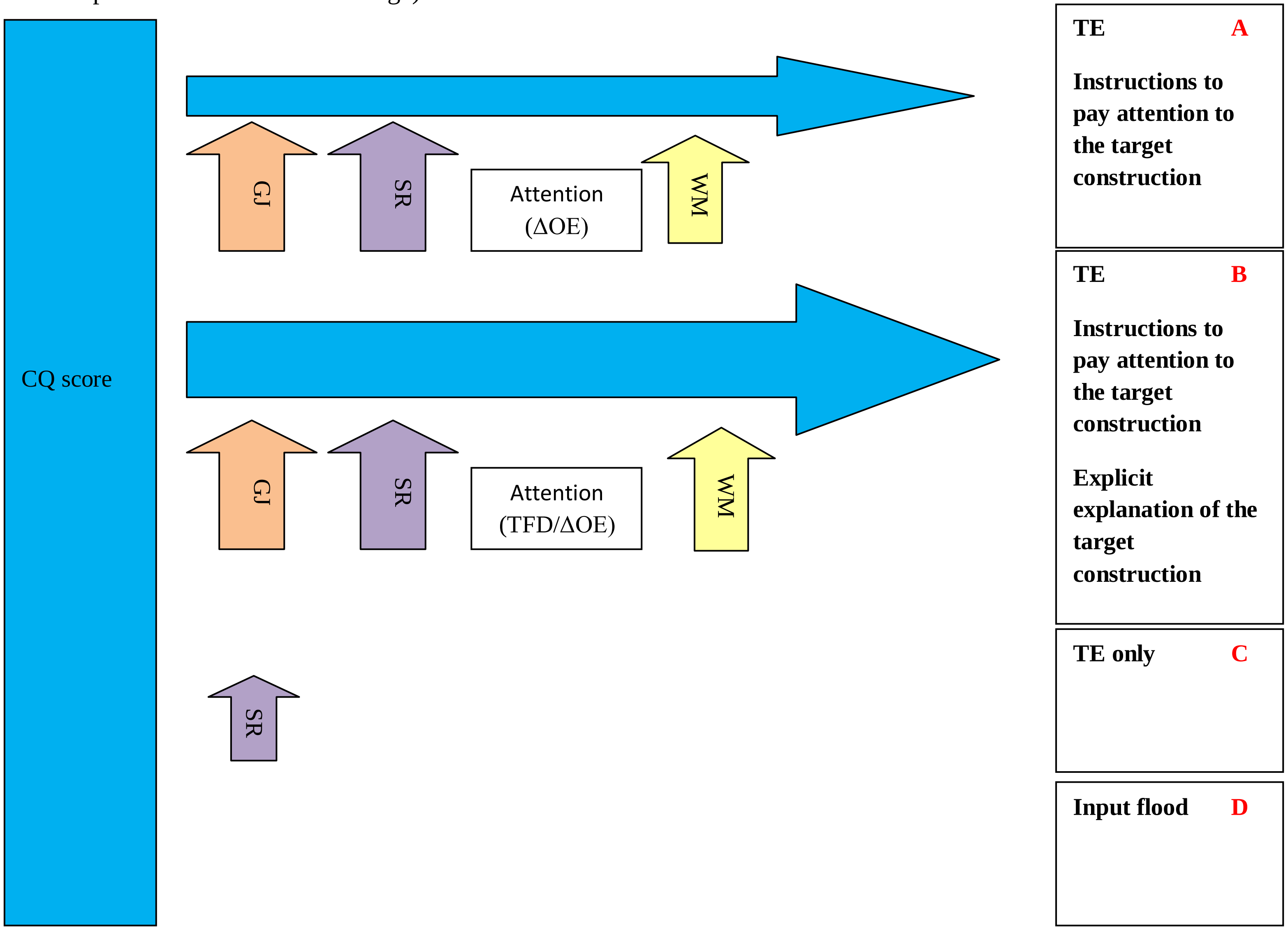
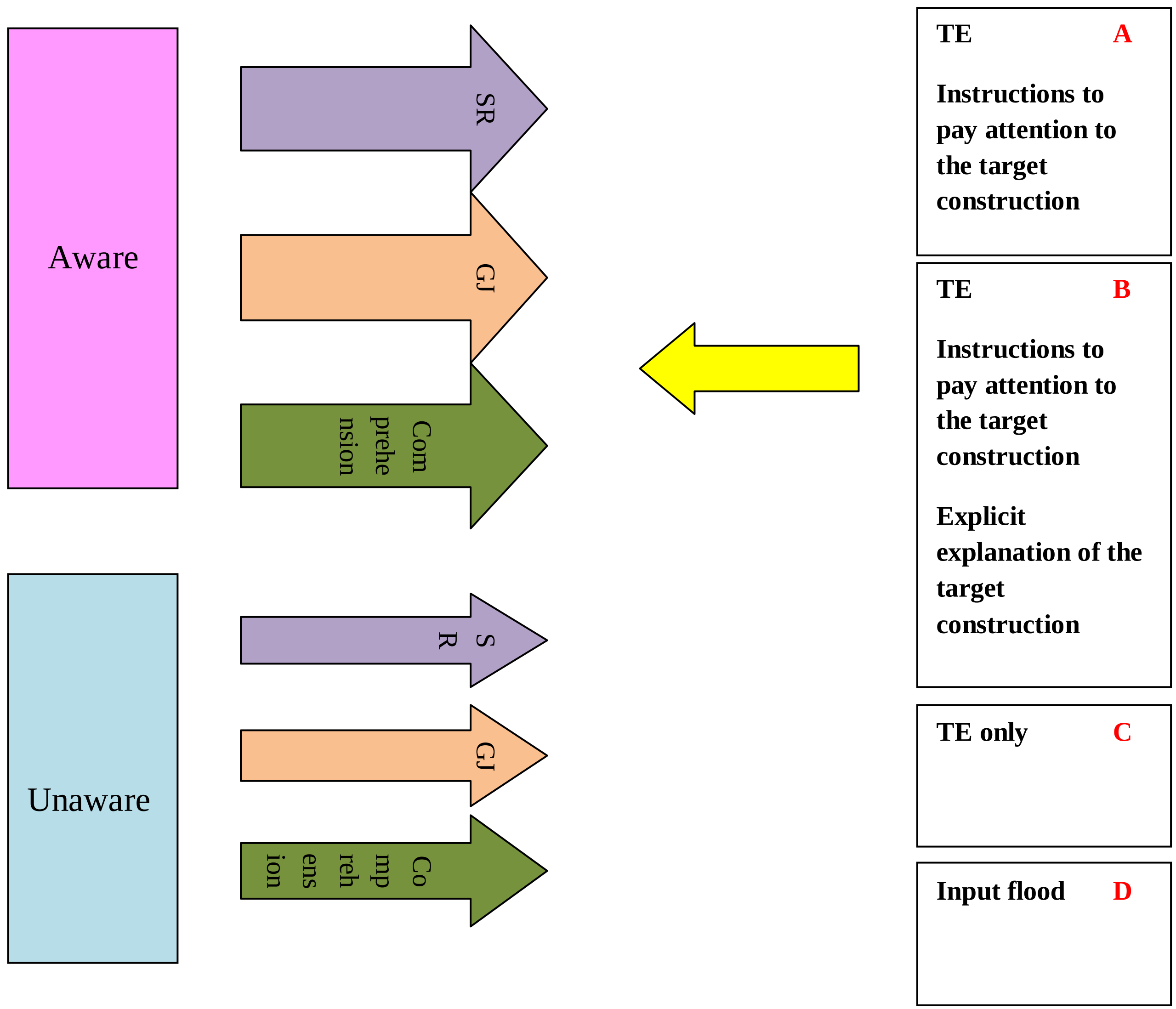


Figure 5.22. Relationship between reported awareness and SR, GJ, composite values in different groups (width of arrows represents the amount of change)



Chapter 6: Discussion

6.1 Introduction

The previous chapter presented the detailed results of the eight research questions that this study attempted to answer. The questions were categorized into three main sections. The first section aimed at analyzing the relationship between attentional processing in different input conditions and learning gains. The second section focused on attentional processing in different input conditions, learning gains and WM. The last section focused on the inter-relationship of attentional processing, learning gains, WM and awareness. The previous chapter also presented the results concerning the free writing task that constituted part of the pre- and post-tests. The results of this, however, were not discussed under a specific research question due to the limited data that it produced. This chapter of the thesis interprets the findings presented in Chapter 5 in relation to the literature reviewed in Chapter 2. The first three sections of this chapter (Sections 6.2, 6.3 and 6.4) present the discussion under the three sections of research questions and Section 6.5 discusses the findings of the free writing task. The chapter also includes a summary of the main findings in Section 6.6.

6.2 Relationship between attentional processing in different input conditions and learning gains[30]

The results of this relationship were presented with regard to the three research questions in the previous chapter. They are:

[30] Some parts of this section were published in Indrarathne & Kormos (2016).

RQ 1: How does knowledge of the form and meaning of the target syntactic construction *causative had* as measured by a production and a comprehension task change under explicit and implicit instructional conditions in the case of Sri Lankan pre-intermediate/ intermediate level English language learners?

RQ 2: How does attention paid to examples of the *causative had* construction in the input texts differ under explicit and implicit instructional conditions?

RQ 3: How is the change in knowledge of the target construction under explicit and implicit instructional conditions related to the attention paid to examples of the *causative had* construction in the input texts?

I will discuss each of these research questions in separate sections.

6.2.1 How does knowledge of the form and meaning of the target syntactic construction *causative had* as measured by a production and a comprehension task change under explicit and implicit instructional conditions in the case of Sri Lankan pre-intermediate/ intermediate level English language learners?

The data analysis with regard to this research question revealed that the knowledge of the form and meaning of the target syntactic construction *causative had* (measured by the SR and GJ tasks) changed from the pre-test to post-test in the different input conditions. The MANOVA analysis showed a significant pre-test to post-test change in the SR items for the whole sample. However, the follow up t-tests revealed that there was a significant pre-test to post-test change in the SR items in the *enhanced+ instructions* (A) and the *enhanced+ instructions+ explanation* (B) groups, but not in the *enhanced only* (C), the *unenhanced* (D) groups or the control group.

As Nassaji and Fotos (2011) argue, two of these input conditions, input flood (*unenhanced* - D) and TE only (*enhanced only* - C) can be classified as implicit FonF techniques, while the other two, instructions to pay specific attention to a grammatical construction (in addition to TE input in the case of the *enhanced+ instructions* – A group) and explicit explanation (with instructions and TE input in the case of the *enhanced+ instructions+ explanation* – B group) can be identified as explicit FonF techniques (Norris & Ortega, 2000). N. Ellis (2006), Sharwood Smith (1993) and Schmidt (1990) point out that salient features in the input might receive more attention and providing TL input frequently can increase salience of target features. Based on this argument, the salience of input was increased in this experiment by providing several examples of the target construction in the input flood condition (Nassaji & Fotos, 2011). The results show that, although the salience of the TGC was increased, the two implicit FonF techniques used in this study (input flood or TE only) did not result in a change of the knowledge of the TGCs of this sample of L2 learners in the SR task. It is worth noting here that input flood condition on this occasion can be regarded as an entirely incidental learning condition since there is no other manipulation of the target examples to draw the attention of the learners.

As discussed in Chapter 4, learners were expected to depend on the explicit knowledge of the target construction when constructing accurate sentences in the SR task. This assumption was made because the SR task required the participants to use their knowledge of the target construction productively. Moreover, the SR task was not performed under time-pressure, which allowed learners more time to work on their answers. Thus, as N. Ellis (2005) points out, the lack of time pressure is believed to have allowed learners to apply their explicit knowledge of the target construction in the

SR task. Although there is no consensus still on what type of knowledge explicit instruction can develop (MacWhinney, 1997), this finding indicates that the explicit input provided in this study has assisted the learners in developing their explicit knowledge. The learners in the implicit input conditions, however, were not able to develop their explicit knowledge of the target construction as the results of the SR task demonstrate.

The results also revealed a significant change between groups in the SR task, in particular, between the control group and the *enhanced+ instructions* (A) group and the *enhanced+ instructions+ explanation* (B) group. Similarly, the *enhanced+ instructions* (A) group and the *enhanced+ instructions+ explanation* (B) group showed a significant change compared to the *unenhanced* (D) group. In addition, the *enhanced+ instructions+ explanation* (B) group demonstrated a significant improvement compared to the *enhanced only* (C) group. The mean differences revealed the highest improvement in the *enhanced+ instructions+ explanation* (B) group followed by the *enhanced+ instructions* (A) group, the *enhanced only* (C) group and the u*nenhanced* (D) group. This finding further strengthens the previous argument that implicit input conditions were less beneficial in changing the explicit knowledge of the TGC among this sample of L2 learners.

In the case of the GJ task, the MANOVA analysis showed a significant change from pre-test to post-test for the whole sample. The follow-up t-test results of individual groups in the GJ task were slightly different from the results of the SR task. The *enhanced+ instructions* (A) group along with the *enhanced+ instructions+ explanation* (B) group and the e*nhanced only* (C) group showed a statistically significant improvement over time. The GJ task, which was presented to the learners in the auditory mode, was carried out under time-pressure. It is still a matter of debate whether a GJ task performed under time-pressure assesses L2 learners' explicit or implicit knowledge (e.g., Bowles, 2011; R. Ellis, 2005, R. Ellis & Loewen, 2007; Gutiérrez, 2013; Zhang, 2014). Nevertheless, in a recent discussion, R. Ellis (2015) suggests using timed GJ tasks to measure implicit knowledge. Evidence from an eye-tracking study by Godfroid et al. (2015) also seems to suggest that under timed-conditions L2 learners tend to apply their implicit knowledge. Taking this together with Bialystok's (1979, 1982) findings that under auditory conditions GJ task encourages learners to rely on their implicit knowledge, it can be hypothesized that in this study the participants primarily used implicit knowledge when making their GJs. Based on this argument, it can be assumed that the participants' implicit knowledge of the TGC changed not only in the explicit input conditions but also in the TE only condition in this study. MacWhinney (1997) points out that implicit instruction does not necessarily lead to implicit knowledge; nevertheless this finding highlights that in one of the implicit input conditions the participants were able to develop their implicit knowledge of the target construction. Although the *enhanced+ instructions* (A) group, the *enhanced+ instructions+ explanation* (B) group and the e*nhanced only* (C) group and the whole sample showed significant gains in the GJ task from the pre-test to the post-test, a between group effect was not observed. This can be explained in relation to the task type. The GJ task allowed the participants to guess the accuracy of the sentences so the overall performance of learners may have been influenced by participants guessing the answers in each input condition.

The above findings are in line with the type of knowledge that N. Ellis (2011) and Williams (2009) predict learners may develop through explicit and implicit learning. They note that explicit knowledge is gained through explicit instruction and implicit knowledge is developed by implicit instruction. This is evident in the results of both the SR and GJ tasks. The findings in the GJ task further revealed that the learners in the explicit input conditions were also able to develop the implicit knowledge required, supporting the claim that explicit learning may develop both explicit and implicit knowledge (MacWhinney, 1997).

The results of this study with regard to the differential effect of explicit and implicit learning conditions are in line with the outcomes of Goo et al.’s (2015), Norris and Ortega’s (2000) and Spada and Tomita’s (2010) meta-analyses, which showed that explicit FonF conditions have a stronger impact on L2 development than implicit ones. Both instructions to pay attention to a specific grammatical construction, which is similar to the rule-search condition in Robinson’s (1997) study, and instructions with explicit explanation resulted in significantly larger improvement in the SR task than input flood and TE. N. Ellis (2005) argues that explicit instruction can speed up the process of language development. This may have played a role on this occasion since the participants were exposed to the target construction within a short period of time and there was only a six day gap between the pre-test and the post-test. The findings with regard to the differential effects of the explicit and implicit FonF techniques seem to lend support to the argument that minimally guided instruction such as input flood and TE might not be beneficial if learners do not have previous familiarity with the targeted construction (Jahan & Kormos, 2015). It is noteworthy that the participants were exposed to the target construction three times within five days; therefore, as N.

Ellis (2005) argues, explicit processing may have been necessary to consolidate the novel constructions rapidly.

Although the *enhanced only* (C) group did not show any gains with time in the SR task, the GJ task revealed that they were able to develop their implicit knowledge. When this result is compared with the results of the input flood group (D), there is a notable effect of the TE technique in developing the implicit knowledge of the learners compared to the input flood technique. The finding is in contrast to the findings of Jourdenais et al. (1995) and Reinders and Ellis (2009), but in line with the findings of Robinson (1997) and Radwan (2005). Input flood provided in Reinders and Ellis' study included 36 examples of the target construction in both aural and written mode. The different modes of input and the amount of input provided may have influenced the learning gains in their study. In Jourdenais et al.'s study also 28 target verbs were included in the input flood condition. The higher number of target examples may have increased the salience of the target examples more successfully.

Overall, the findings reveal that the learners were able to develop both explicit and implicit knowledge of the target construction in explicit input conditions. However, in implicit input conditions, the learners were able to develop only implicit knowledge of the target structure. Moreover, the explicit input techniques resulted in larger gains in the post-test compared to the implicit input techniques. Among the implicit input techniques, TE seems to be more beneficial than input flood in developing learners' implicit knowledge. This differential effect of input conditions can be explained by analyzing the amount of attention paid to the examples of the target construction in each condition.

6.2.2 How does attention paid to examples of the *causative had* construction in the input texts differ under explicit and implicit instructional conditions?

The novelty of this study was that it did not only investigate the learning gains under the four input conditions but also analysed the eye-fixation durations to uncover how attentional resources are allocated under these conditions. The results revealed that the *enhanced+ instructions+ explanation* (B) group had the highest mean TFD-AOI and ΔOE-AOI followed by the *enhanced+ instructions* (A) group. The *enhanced only* (C) group showed the third highest mean values in both TFD-AOI and ΔOE-AOI. The most interesting finding is that the *unenhanced* (D) group demonstrated the lowest mean value in TFD-AOI and 0 mean value in ΔOE-AOI. If ΔOE-AOI values for individual AOIs are taken into consideration, both the *enhanced+ instructions+ explanation* (B) group and the *enhanced+ instructions* (A) group maintained the values above 0 in all AOIs. However, the ΔOE-AOI value went below zero in one AOI in the *enhanced only* (C) group while it fluctuated around and below zero in all AOIs in the *unenhanced* (D) group.

The findings highlight that under the input flood condition the participants' attention did not seem to be drawn to the examples of the target construction, and this was apparent in the fluctuation of the ΔOE-AOI around and below the value of 0. If the methodological argument that ΔOE represents the attention paid to input is taken into consideration, this finding suggests that the participants in the input flood condition did not perform additional attentional processing of the target examples in the input besides decoding them as part of the text. Therefore, in line with VanPatten's (1994) Primacy of Meaning principle, it seems that when learners are presented with a previously unfamiliar grammatical construction in a reading text, they might not direct their attentional resources to it. In other words, they might not exercise intentional control of the allocation of their attention to this feature of the input (Shiffrin, 1988; Styles, 2006). It is also possible that the limited capacity of their attentional processing (Desimone & Duncan, 1995; Shiffrin, 1988) while reading the text prevents them from selectively attending to this target grammatical item. The lack of selective attention to the target item is not unsurprising given previous studies that found limited or no learning gains under input flood conditions (e.g., Izumi, 2002, 2003; Jahan & Kormos, 2015; Winke, 2013).

Interestingly, the attentional processing of the learners in the TE condition does not seem to differ considerably from that of the participants in the input flood condition. Although the ΔOE-AOI value goes below 0 only once in the case of the TE condition and the overall mean values of TFD-AOI and ΔOE-AOI are slightly higher than those of the input flood condition, the overall TFD-AOI and ΔOE-AOI of the TE condition are not significantly different from the values found in the input flood condition. These findings align with those of Issa et al. (2015), but are different from the results obtained by Winke (2013) and Simard and Foucambert (2013). The reason for this difference might lie in the nature of the enhancement applied in this study, which was boldfacing the target examples in the TE condition, whereas Winke (2013) and Simard and Foucambert (2013) used underlining (Winke used underlining in addition to different colour font), which might have been more effective in creating an isolation effect (von Restorff, 1933), and consequently drawing learners' attention to the target items. It could also be possible that enhancement on this occasion had an opposite effect i.e. that the participants ignored the enhanced sections. This indicates that boldfacing as a form of TE in this experimental condition was not successful in inducing additional attentional processing. In addition to the reasons listed above concerning the type of enhancement, the primacy of meaning when processing input texts and the limited nature of attentional capacity, it is possible that the externally induced salience in this study might not have corresponded with learners' internally generated salience (Sharwood Smith, 1991, 1993). Furthermore, Lavie, Hirst, De Fockert and Viding's (2004) research in the field of cognitive psychology suggests that increasing perceptual load, such as the visual enhancement of a section of a text, results in an early exclusion of the stimulus from further cognitive processing unless the stimulus is deemed to be relevant for the task. Moreover, according to the capacity theory of attention (Kahneman, 1973), the selection of stimuli to be attended is made by the task demand. Chun, Golomb and Turk-Browne (2011) argue that attention involves selection of information that is most vital for the behaviour. It can thus be hypothesized that the participants in the TE condition did not consider the visually enhanced constructions essential for the comprehension of the text, and consequently did not engage in subsequent active attentional processing.

Based on the results of this study, the most effective experimental manipulation to draw attention of the learners to the examples of the target construction seems to be the instruction to pay attention to the grammatical construction embedded in the text. This is attested by the fact that both the *enhanced+ instructions* group (A) and the *enhanced+ instructions+ explanation* (B) group demonstrated significantly higher TFD-AOI and ∆OE-AOI values for the three texts jointly than the *enhanced only* (C) and the *unenhanced* (D) groups. Wickens' (2007) SEEV (salience effort expectancy value) model of attention can provide a possible explanation for these results. Wickens argues that four factors are important in determining what aspects of the incoming stimuli one attends to: salience, effort, expectancy, and value. The model predicts that when different pieces of information compete for attentional resources, information will be heeded that is salient, requires less effort to process, is expected in the given situation, and has high value in terms of the task to be solved. It can be hypothesized that by informing learners that a novel grammatical construction is embedded in the text, the expectancy value of the targeted structure was increased, and concomitantly it also became salient. Therefore, in these two experimental conditions learners can be assumed to have exercised both top-down attention control to search for the relevant feature in the input text and bottom-up control to select the necessary information among other stimuli (Koch & Tsuchiya, 2006). It is also very likely that in these two conditions attention also involved consciousness and the relevant stimuli were registered in the WM (Koch & Tsuchiya, 2006; Lamme, 2003).

The results of the study also suggest that the explicit explanation provided in the case of the *enhanced+ instructions+ explanation* (B) group after the exposure to Text 1 successfully induced additional attentional processing of the target construction in Text 2. This is evident by the highest TFD-AOI and ∆OE-AOI values shown in Text 2 in this group. The effect of the explicit explanation however did not seem to be lasting, because, as for Text 3, no significant differences between the *enhanced+ instructions* (A) and the *enhanced+ instructions+ explanation* (B) groups were observed.

Overall, both explicit input conditions used in this study were seen to be effective means of directing L2 learners' attention to the target feature. The most influential explicit technique seems to be informing learners that there is a target construction embedded in the input and they need to pay attention to the target examples. The results in the first research question showed that explicit and implicit learning gains were related to the type of input that the learners received and the results of the second research question revealed that attentional processing was also strongly associated with the learning gains. In order to have a better understanding of these results, the next section explores the relationship between learning gains, type of input and attentional processing.

6.2.3 How is the change in the knowledge of the target construction under different input conditions related to attention paid to the examples of the *causative had* construction in the input texts?

Spearman rho correlation analysis between the eye-tracking variables and the SR and GJ gain scores revealed strong positive correlations between the composite ∆OE-P and TFD-P values and the SR and GJ gain scores. The results of the correlational analyses point to a strong relationship between the eye-tracking variables and the learning gains made by the participants, in particular, in the case of the ∆OE-P measures and the improvement in the SR task. In line with the hypothesis that the difference between the observed and expected eye-fixation duration would be reflective of additional attentional processing, this value accounts for a somewhat larger proportion of variance in gain scores than the TFD-P value.

As already argued above, increased TFD and ∆OE can be regarded as indications of top-down and bottom-up attention control processes (Koch & Tsuchiya, 2006). These processes are likely to involve consciousness, which is a pre-condition for relevant information to be processed in the WM. Learning gains especially in the SR task, which assessed the establishment of explicit knowledge representations, are very strongly related to attentional processing of the target construction suggesting that conscious attentional processing is influential in explicit knowledge development. The somewhat weaker correlations between the eye-tracking measures and improvement in the GJ task can tentatively suggest that in the development of implicit L2 knowledge, conscious attentional processing plays a smaller role.

In the case of different experimental conditions, Spearman rho correlations revealed that the SR and GJ gain scores of the *enhanced+ instructions* (A) group, the *enhanced+ instructions+ explanation* (B) group and the *enhanced* only (C) group significantly correlated with the ∆OE-P composite value. The composite TFD-P value also significantly correlated with the SR and GJ gain scores of the *enhanced+ instructions+ explanation* (B) group and the *enhanced* only (C) group. The associations between improvement in the SR and GJ tasks and the eye-tracking variables under different conditions are similar to those observed for the whole learner sample with the notable exception in the *unenhanced* (D) group. As already pointed out, the eye-fixation durations (TFD-AOI and ∆OE-AOI) fluctuated randomly in this experimental group in all three texts, and therefore it is not surprising that no relationship between the eye-tracking measures and learning gains for the learners in this experimental condition emerged. This further strengthens the argument that unenhanced input flood or the entirely incidental learning condition in this experiment was not able to draw learners' attention to the target grammatical feature and consequently the participants in this group did not demonstrate any learning gains.

Apart from Winke's (2013) and Issa et al.'s (2015) studies, to my knowledge, there is no other eye-tracking study in the field of SLA that has investigated the effects of explicit and implicit input techniques on the amount of attention paid and the subsequent learning gains. Winke (2013) does not report an analysis between the gains and the attention data and Issa et al. (2015) report no relationship between the learning gains and the attention paid in the TE condition. However, in this study, I found that learning gains demonstrated by the group (both explicit and implicit) in the TE condition significantly correlated with the amount of attention paid. In contrast, in the input flood condition no correlation between the learning gains and the amount of attention paid emerged.

In summary, the findings with regard to the first set of research questions demonstrated that explicit input techniques seem to be more beneficial than the implicit input techniques in developing the explicit knowledge of novel grammatical constructions. However, both explicit and implicit input techniques, except input flood, seem to be able to develop the implicit knowledge of a novel grammatical construction. This suggests that providing several examples of a TGC without any external manipulation is less beneficial in developing either explicit or implicit knowledge of novel grammatical constructions. In line with this, it can also be suggested that external manipulation of input is necessary in drawing learners' attention to a target construction embedded in the input. Increased salience through input flood seems to have a minimal effect on drawing learners' attention to input. The results also suggest that paying attention to target input plays an important role in the development of knowledge of the target construction; more attention paid can result in better gains in both explicit and implicit knowledge.

The results discussed so far also demonstrated that conscious attention may have been associated with the amount of attention paid and subsequent learning gains particularly in the case of explicit input conditions. This suggests that the attended stimuli may have been registered in the WM. The next section of this chapter will examine the influence of WM capacity in attentional processing and learning gains in the four different input conditions.

6.3 Relationship between WM and attentional processing of input in different input conditions and learning gains

Chapter 4 presented the results of this relationship under two research questions. They are:

RQ 4: How is the functioning of the WM including both PSTM capacity and CE functions related to the change of knowledge of the TGC *causative had* in different input conditions?

RQ 5: How is the functioning of the WM related to attention paid to the target items in different input conditions?

The following paragraphs examine the results revealed under these two research questions in relation to the literature discussed in Chapter 2.

6.3.1 How is the functioning of WM including both PSTM capacity and CE functions related to the change of knowledge of the TGC *causative had* in different input conditions?

The Spearman rank-order correlational analyses conducted between the individual WM tests and the SR and GJ gain scores indicated that the Digit Span test correlated with both gain scores in all experimental conditions as well as in the whole experimental sample. The Keep Track task also showed similar results except in the *enhanced+ instructions* (A) and the *unenhanced* (D) groups in the GJ task. The correlation was not significant between the Stroop task and the SR gain score in the *unenhanced* (D) group. Spearman rank-order correlations performed between the composite WM score (composite score of the Digit Span, the Keep Track and the Stroop tasks) and the SR and GJ gain scores revealed significant correlations in the case of the whole experimental sample with a particularly strong correlation in the *enhanced+ instructions+ explanation* (B) group and a weak correlation in the *unenhanced* (D) group. The multiregression model revealed that the relationship between the SR gain score and the composite WM score in the *unenhanced* (D) condition is statistically different from the relationship in the *enhanced+ instructions* (A) and the *enhanced+ instructions+ explanation* (B) conditions. This indicates that WM played a stronger role in the *enhanced+ instructions* (A) and the *enhanced+ instructions+ explanation* (B) conditions when performing the SR task. However, the multiregression model did not show a differential relationship between the composite WM score and the experimental groups in the case of GJ gain score.

The use of tests of CE for assessing L2 learners' aptitude complexes was first recommended by Linck Osthus, Koeth and Bunting (2013), but in their study measures of CE had little predictive power for ultimate attainment in L2 reading and listening skills. To my knowledge, this study is the first one to show that the addition of tests of CE to traditionally used measures of the capacity of the PSTM can explain variations in the development of syntactic knowledge through exposure to exemplars in meaningful written texts. The strong relationship between the WM abilities and gains in both the SR and GJ tasks in the whole sample indicates that when the assessment scores of abilities relating to the PSTM capacity, updating and inhibition are combined, they account for a substantial variance in gain scores in these tests. This points to the significance of the ability to regulate attentional resources efficiently in learning from written L2 input. Higher WM capacity has been associated with better L2 reading skills in previous research (e.g., Alptekin & Ercetin, 2009; Ercetin & Alptekin, 2013; Harrington & Sawyer, 1992; Miyake & Freedman, 1998). Therefore, it is possible to hypothesize that learners with a high ability to hold and update verbal information in STM and to inhibit irrelevant stimuli are more efficient in text decoding, and consequently they have more attentional resources available for processing grammatical information in a reading text. In addition, the increased storage and processing functions of WM allow learners to manipulate longer chunks of language more efficiently. This assists pattern recognition as well as encoding grammatical knowledge in long-term memory (Martin & Ellis, 2012; Williams & Lovatt, 2003).

The results of the correlational and multiple regression analyses conducted with the whole sample also suggest that the PSTM and the CE functions account for a slightly higher percentage of variance in the GJ task than in the SR task. However, when the correlation co-efficients (SR gain r = .614 and GJ gain r = .658) are compared, no statistically significant differences emerge between them (*z* =-.46, *p* = .64). As discussed in 6.2.1, whereas the GJ task is presumed to have assessed implicit knowledge, the SR task mainly assessed the explicit knowledge. Based on this argumentation, the results seem to indicate that individual differences in PSTM capacity and attention regulation play an equally important role in the development of implicit and explicit knowledge. This finding is in line with that of Erctekin and Alptekin (2013) in whose study WM capacity as measured by a reading span test was also associated with performance both in timed GJ and untimed GJ tasks. The difference between their study and this, however, was that they assessed L2 learners' existing explicit and implicit knowledge of a variety of syntactic constructions, whereas in this research I aimed to measure learning gains through exposure to a specific novel syntactic structure. The results are also in line with Ellis and Sinclair's (1996) findings that PSTM is influential in both explicit and implicit knowledge development in grammar learning as well as with the findings of Saggar and Abbuhl's (2013) study on recasts which also reports significant relationships between PSTM and explicit/implicit knowledge gains. In summary, this study demonstrates that high WM abilities do not only facilitate the development of explicit but also implicit knowledge of a previously unknown syntactic structure.

The findings indicate that WM abilities contribute significantly to learning gains in all four learning conditions. The multiple regression analysis revealed that the effect of WM was very similar across groups in the case of the GJ task. Therefore learners with high PSTM capacity and good abilities to regulate their attentional resources seem to be at an advantage in acquiring implicit knowledge regardless of the type of instruction they receive. In the SR task, however, PSTM capacity and the CE functions seemed to play a different role in the different experimental groups. In the *unenhanced* (D) condition, which can be considered the entirely incidental learning context, the contribution of WM abilities to improvement in explicit knowledge gains was significantly smaller than in the *enhanced+ instruction* (A) and the *enhanced+ instruction+ explanation* (B) conditions, which were the two explicit learning contexts. The results of this study are similar to those of Robinson's (2005) study, in that, in explicit conditions participants with higher WM capacity achieved higher scores in a test of explicit knowledge. The findings of this study with regard to the facilitative role of WM abilities in explicit learning conditions when developing implicit knowledge, however, are new in the field of SLA research although a similar finding is reported in cognitive psychological research (Unsworth & Engle, 2005). It is also different to the assumption by Reber et al. (1991) that WM is less influential in implicit learning. R. Ellis (2009) notes that explicit learning imposes a heavy demand on the WM unlike implicit learning. However, the findings of this study indicate that even implicit learning can impose a demand on the WM.

Overall, the findings suggest that both PSTM capacity and CE functions (updating and inhibition) in combination seem to influence the development of explicit and implicit knowledge of a novel grammatical construction. This relationship is also likely to be independent of the type of input, i.e. WM may play a significant role in processing both explicit and implicit input, in developing both explicit and implicit knowledge. However, the development of explicit knowledge may be more strongly influenced by WM in the explicit input conditions than in the implicit conditions. The influence of WM on learning gains can be further discussed by analysing how learners pay attention to input in different input conditions.

6.3.2 How is the functioning of the WM related to attention paid to the target items in different input conditions?

The examination of the relationship between PSTM and the CE functions and the eye-tracking measures can be helpful to elucidate the potential mechanisms of how PSTM and the CE functions can affect learning in incidental and intentional learning conditions. When analysing the correlations between eye-fixation measures and the composite WM factor score and the individual PSTM and CE function tests, a significant but moderately strong correlation was observed in the case of ∆OE-P. The relatively weak relationship across the sample can be explained by the fact that, as discussed in Section 5.3.2, there is a large effect of the treatment condition on the eye-tracking measures, TFD-P, $F(3, 41) = 27.06$, $p < .001$, partial eta squared = .67, and ∆OE-P, $F(3, 41) = 28.42$, $p < .001$, partial eta squared = .68, with the *enhanced only* (C) and the *unenhanced* (D) groups demonstrating very low levels of attentional processing of the target input. The results reveal that in the most explicit learning condition, i.e. the *enhanced+ instruction+ explanation* (B) condition, individual differences in attention regulation and the capacity of the PSTM are highly predictive of how much attention L2 learners pay to the target input. The ∆OE measure, which has been argued to be an indicator of additional attentional processing, in addition to what would be expected if learners were just simply reading the text for meaning, was also strongly associated with the composite WM score in the *enhanced only* (C) and in the *enhanced+ instruction* (A) groups. This finding reveals that those learners who have high PSTM capacity and efficient attention regulation ability also engage in more attentional processing of the input if they are informed of the existence of a target syntactic construction in the input. What is interesting in this study is that despite apparent lack of increased attentional processing of the TL input in the *unenhanced* (D) condition, L2 learners with high PSTM capacity and higher updating and inhibition ability still made some learning gains in both explicit and implicit knowledge. This suggests that WM functions can potentially be involved in L2 learning processes that do not involve focal attention to input, in other words in statistical associative learning.

Recent theorizations of WM by Hassin et al. (2009) and Soto and Silvanto (2014) question the previously held assumption that WM can only operate on consciously attended input (e.g., Baddeley, 2003; Koch & Tsuchiya, 2006). Emerging evidence in Hassin et al.'s and Soto and Silvanto's work suggests that conscious awareness might not be necessary for information to enter WM and hence it can also be active in implicit learning. As Soto and Silvanto highlight, "the key aspect is that WM processes can be engaged without awareness, that WM can operate on nonconscious representations, and that the process through which WM contents reach awareness is subject to varying factors (i.e. attention, intention, motivation, heuristics)" (p. 524). The present study found a relatively strong association between the composite WM score and gains in a test of implicit knowledge in an incidental learning condition (input flood) in the absence of a link between attentional processing and WM abilities. This finding seems to lend additional support to the relevance of the PSTM and the CE functions in implicit L2 learning.

If the theorization of explicit and implicit learning processes is revisited, explicit learning is assumed to be a conscious process and implicit learning an unconscious process (R. Ellis, 2009; Williams, 2009). According to Lamme's (2003) attention models discussed in Chapter 2, only consciously attended stimuli are able to enter the

WM. If the learning gains reported in the implicit input conditions in this study occurred without consciousness, the input processing in these implicit conditions may not have made a significant demand on the WM of the learners. However, both explicit and implicit learning demonstrated by the participants in the implicit input conditions showed that their WM capacity was influential in the learning gains. Thus, it is possible to predict that implicit learning also has some involvement of consciousness. In other words, implicit input techniques on this occasion have drawn the conscious attention of the learners to the target construction.

In summary, the findings of the study indicate that L2 learners with higher PSTM capacity and better abilities to regulate their attentional resources in the CE are advantaged in learning a novel grammatical construction from written input both in explicit and implicit input conditions. WM abilities, however, were found to play a larger role in the acquisition of explicit knowledge in explicit input conditions than in implicit input conditions. Since both explicit input conditions received instructions to pay attention to the target examples, this finding also suggests that learners with high WM abilities might respond better to the instructions that they receive and subsequently that would enhance their learning. With regard to the development of implicit knowledge, there did not seem to be a difference in the strength of association between the composite WM score and learning gains in the different treatment groups, suggesting that regardless of the type of instruction, learners with high PSTM capacity and high updating and inhibition ability improve their implicit knowledge more successfully than those with lower levels of PSTM and attention regulation abilities. The eye-tracking data also suggests that learners with high PSTM capacity and attention regulation abilities might learn from L2 input that is not within the central focus of their attentional processing but through associative probabilistic learning.

The results of the relationship between attentional processing, learning gains and WM in different input conditions suggest that conscious processing of input may be involved particularly in the case of developing explicit knowledge. The results of implicit learning gains also imply the involvement of conscious attention to input. The discussion of the involvement of awareness, which is said to involve consciousness (Lamme, 2003), may provide more insights into how conscious processing of input influences learning gains. The next section of this chapter thus discusses the findings in relation to awareness in attentional processing, learning gains and WM.

6.4 Relationship between awareness and attentional processing, learning gains and WM

While providing input through three short stories, a set of CQs was also used in this experiment with the aim of providing the learners with a purpose for reading. Two of the questions given with each text were also designed to measure the knowledge of the form-meaning mapping of the target construction. As discussed under Section 1 RQs, SR task functioned as a written production task that assesses the understanding of form and meaning of the TGC as well. The CQ task also measured the understanding of form and meaning, but the learners had to choose an answer from a multiple choice i.e. it served as a form-meaning recognition comprehension task. The GJ task was also a comprehension task; however, due to the fact that it was performed under time pressure, it was assumed to have measured the implicit knowledge. The CQ task hence is a comprehension task that can be assumed to have measured the explicit knowledge.

In this section, the findings under CQ task are analysed in order to understand how learners performed in the comprehension task that assessed the explicit knowledge and how the results are related to attentional processing and WM. In addition, the results obtained in the post-exposure interview on awareness are also included in this section. The discussion is organized according to the following research questions.

RQ 6: Is there an improvement in the CQ scores in different input conditions between the three exposures?

RQ 7: Is the improvement in the CQ task related to the eye fixation measures, the knowledge of the TGC *causative had* (SR and GJ tasks) and WM?

RQ 8: Is there a difference between the change in the knowledge of the TGC *causative had* (SR, GJ and comprehension tasks) between the participants who reported awareness and those who reported no awareness? Is the awareness reported in different input conditions significantly different to each other?

6.4.1 Is there an improvement in the CQ scores in different input conditions between the three exposures?

The results of the analyses that compared the composite comprehension scores between exposures highlighted a significant improvement from exposure 1 to exposure 2 and exposure 1 to exposure 3 in the whole sample with the highest mean value in exposure 3. Since each question contained a multiple choice with three possible answers, there is a 33% possibility of participants guessing the answer. However, the t-test results revealed that with exposure, the participants were able to answer more accurately in the CQ task within the whole experimental sample (Between exposures 1 and 2 and 1 and 3). The development was non-significant between exposure 2 and exposure 3 which would mean that with time, the effectiveness of providing more exposure has declined. This was also evident in eye-tracking measures as well. In particular, in the *enhanced+ instructions* (A) group and the *enhanced+ instructions+ explanation* (B) group, the TFD-AOI and TFD-ΔOE values gradually declined over time with exposure (see Section 5.3.1 for the visual figures). The t-test results revealed that both groups did not show significant gains in the CQ task between exposures 2 and 3 either.

Between group analysis of data of the CQ composite scores revealed that the *enhanced+ instructions* (A) group demonstrated a significant improvement compared to the *unenhanced* (D) group while the *enhanced+ instructions+ explanation* (B) group showed a significant development compared to the *enhanced only* (C) and the *unenhanced* (D) groups. This is also in line with the development shown in SR and GJ task gain scores where the *enhanced+ instructions* (A) and the *enhanced+ instructions+ explanation* (B) groups showed higher overall gains than the *enhanced only* (C) and the *unenhanced* (D) groups. The individual t-test results revealed that only the two explicit input conditions demonstrated significant learning gains in the CQ task. In particular, the participants in the *enhanced+ instructions+ explanation* (B) group demonstrated significant gains between both Exposure 1 and 2 and Exposure 1 and 3. The target CQs required the learners to understand both the form and the meaning of the TGC in order to provide accurate answers. As DeKeyser (2005) notes, the form-meaning relationship depends on redundancy, optionality and/or opacity. The TGC in this study was not affected by any of those conditions. In other words, the form-meaning relationship was clear and thus knowing the form was necessary to understand the meaning of the construction.

According to the results, it is apparent that those participants who were exposed to explicit input were able to show improvement in the CQ task. On the one hand, this highlights that explicit knowledge was necessary to understand the form-meaning mapping in the case of CQs related to the TGC in this study. On the other hand, this finding indicates that explicit input leads to explicit knowledge if it is assumed that answering CQs requires explicit knowledge. These findings are similar to the results of the production task that required learners to apply their form-meaning knowledge (SR task). Although DeKeyser (2005) notes that less frequent input is sufficient for acquisition if the form-meaning relationship is transparent, frequency on its own seems to be insufficient in this experiment since the learners only in the externally manipulated input conditions were able to identify the form-meaning mapping in the CQs as well as in the SR task. The findings are also in contrast with Thornbury's (2005) argument that techniques such as input flood could raise the awareness of the form-meaning relationship. However, it is also possible that the amount of input provided on this occasion may have been insufficient and as a result the learners in the implicit conditions may have demonstrated little or no understanding of the form-meaning mapping of the target construction.

In summary, the discussion highlighted that the participants in the explicit input conditions were able to develop their ability to understand the form-meaning relationship of the target construction in this experiment. The results of the comprehension test score in different input conditions can be further examined by analysing the relationship that it has with the SR and GJ gain scores, attention data and WM data.

6.4.2 Is the improvement in the CQ task related to the eye fixation measures, the knowledge of the TGC *causative had* (SR and GJ tasks) and WM?

The correlation analysis between the composite comprehension test score and the SR and GJ gain scores for the whole experimental sample highlighted a significant and strong relationship between those variables. However, only the *enhanced+ instructions+ explanation* (B) and *enhanced+ instructions* (A) groups demonstrated a significant correlation between the composite comprehension test score and the SR and GJ gain scores in the group analysis. As pointed out previously, only these two explicit input conditions were effective in developing the explicit knowledge of the target construction. Therefore, this finding indicates that the knowledge gained by the learners in the explicit conditions was useful in performing the CQ task. Although I found a correlation between the SR gain score and the composite comprehension test score in the *enhanced only* group (C), the participants in this group did not show any development in the CQ task.

Spearman rank-order correlation analyses between the composite comprehension test score and composite values of the two eye-tracking variables revealed that both TFD-P and ∆OE-P strongly correlate with the composite comprehension test score. In the case of different input conditions however, the correlations did not show similar results. The composite comprehension test score strongly correlated with both TFD-P and ∆OE-P composite values in the *enhanced+ instructions+ explanation* (B) group. A similar relationship could be observed only between the ∆OE-P composite value and composite comprehension test score in the *enhanced+ instructions* (A) group. In the *enhanced only* (C) and the *unenhanced* (D) groups, no significant correlations emerged. These findings are in line with the gains in CQ task in different input conditions i.e. the gains were statistically significant in the explicit input conditions, but not in the implicit input conditions. This suggests that more attention to the TGC was drawn by the explicit input techniques and as a result the learners in those input conditions were able to significantly develop their knowledge of the form-meaning relationship of the TGC. As a result, they were able to apply that knowledge in performing the comprehension form-meaning recognition task more successfully than the participants in the implicit conditions. This finding is also in line with the findings under the production form-meaning task (SR).

Furthermore, WM capacity played a more significant role in the *enhanced+ instructions* (A) and the *enhanced+ instructions+ explanation* (B) groups in relation to the gains in the CQ task. This further denotes that WM may be more influential in developing explicit knowledge which subsequently leads to learners understanding the form-meaning mapping of a TGC.

Overall, the discussion under the CQ task further strengthen the argument that the understanding of the form-meaning relationship of a novel grammatical construction can be developed only through increased attention to the target structure in the input by providing explicit instructions. Furthermore, understanding of form-meaning relationship in the case of performing an explicit comprehension task seems to depend on the development of explicit knowledge of the target construction.

The next section of the discussion analyses the involvement of awareness when processing input.

6.4.3 Is there a difference between the change in the knowledge of the TGC *causative had* (SR, GJ and comprehension tasks) between the participants who reported awareness and those who reported no awareness? Is the awareness reported in different input conditions significantly different to each other?

As discussed in Section 2.2, there is a common agreement in literature that if a stimulus attended in the input gets registered in the WM or if a report on the stimulus attended can be offered, the attention on this occasion involves consciousness/awareness (Koch & Tsuchiya, 2006; Lamme, 2003). Based on this argument, two target questions in the post-exposure questionnaire were used to analyse the involvement of possible awareness when paying attention to input in the experimental conditions of this study. The t-test results discussed in Section 5.5.3 revealed that the group who reported awareness and the group who did not report awareness showed significant improvement in the SR, GJ and comprehension test tasks. This indicates that awareness was not a necessary condition for developing either explicit or implicit knowledge of the TGC. This finding is in line with Tomlin and Villa's (1994), Robinson's (1995) and Godfroid et al.'s (2013) argument that attention to input is sufficient for L2 learning and awareness is not a necessary condition. If both SR and comprehension tasks are considered to be assessments of explicit knowledge, this finding is in contrast to

Williams' (2009) argument, which claims that learners are aware of what they have learned if they develop explicit knowledge.

The analysis of the chi-square standard residuals in Section 5.5.3 further revealed that only the participants in the *enhanced+ instructions+ explanation* (B) group reported more awareness than expected. This implies that the participants in the input condition who received the maximum external manipulation of the examples of the TGC could increase their awareness as a result of paying attention to input. According to the t-test analysis, the mean difference in all three tasks (SR, GJ and comprehension) was higher in the 'aware' group compared to the 'unaware' group. These two findings together imply that the development of both explicit and implicit knowledge of the target syntactic construction was considerably larger among the participants who developed awareness as a result of paying attention to the input and that awareness can be increased through explicit instruction.

The findings of the study that indicate that awareness was not involved in the development of knowledge of the target construction could be due to the methodological issues of measuring awareness. If the definition of awareness is revisited, both in cognitive psychology and SLA, there is a common agreement that if awareness is involved when attending a stimulus, one should be able to provide a report of that stimulus. However, Koch and Tsuchiya (2006) highlight that there are other involvements of conscious awareness. For example, conscious awareness is necessary for a stimulus to get registered in the WM or to distinguish between stimuli in addition to having the possibility of offering a report. This indicates that even if learners are not able to provide a report, awareness may still have been involved in their attentional processing if the attended stimuli get registered in their WM or when they distinguish between stimuli. Allport's (1988) definition of awareness also notes that there are three indicators of awareness: "a show of some behavioural or cognitive change due to the experience, a report of being aware of the experience and a description of this subjective experience" (p.51). In this study I used the post-exposure interview in order to collect data on reported awareness, which is only one of the possible indicators of awareness. Therefore, although the results revealed that both the 'aware' and 'unaware' groups showed a significant development in both explicit and implicit knowledge, this does not necessarily rule out the possibility of the involvement of awareness. It can be hypothesized that the participants who did not report awareness could not verbalise awareness (VanPatten, 2011) and/or they went through a cognitive change that could not be reported.

Another methodological issue in measuring awareness lies in the type of questions used in the post-exposure interview of this study. The two questions used were:

Q7: Did you notice any particular grammar structure in the texts?

Q8: If your answer to question 7 is 'yes' can you please write down that structure?

A criticism that may arise in relation to Q7 is the use of the term 'grammar structure'. When the word grammar is used, one could argue that the question tested the metalinguistic awareness of the target construction rather than the knowledge of form-meaning mapping. Although the example sentences provided by the participants in answering Q8 (e.g., He had the tools delivered) were taken as accurate reports of awareness, some participants may have been confused by the term 'grammar'. In particular, in the input flood and TE conditions, the participants were not told that a particular language structure was embedded/ highlighted and as a result, when they were asked to report a ‘grammar’ structure they may have found it confusing. The reported awareness in these two groups was very low (1/20 in the TE condition and 0/20 in the unenhanced condition). This could suggest that the participants in these groups did not have metalinguistic awareness of the target construction. Nonetheless, this is not an accurate indication of the absence of awareness. R. Ellis (2009) points out that explicit instruction encourages the learners to develop their metalinguistic awareness. The results of my study reveal that reports of awareness by the participants in the most explicit instructional condition were significantly different from that of the other three input conditions. Therefore, it is possible to assume that the question on awareness may have measured the metalinguistic awareness of the target construction. The inability to report awareness may also be related to Hulstijn’s (2015) argument that proficiency is “independent of declarative knowledge of the rules of pedagogical grammar” (p. 38).

In summary, the results pertaining to the third set of research questions revealed that the learners in the explicit input conditions were able to develop their ability to make the form-meaning mapping of the target construction by paying more attention to the examples of the target construction in the input. According to the data collected in this study, awareness was not a necessary condition for the learning of a novel L2 syntactic construction; however, awareness in all its forms could not be measured. The methodological shortcomings of this study in measuring awareness have still kept the gap of knowledge open on the necessity of awareness for L2 language development.

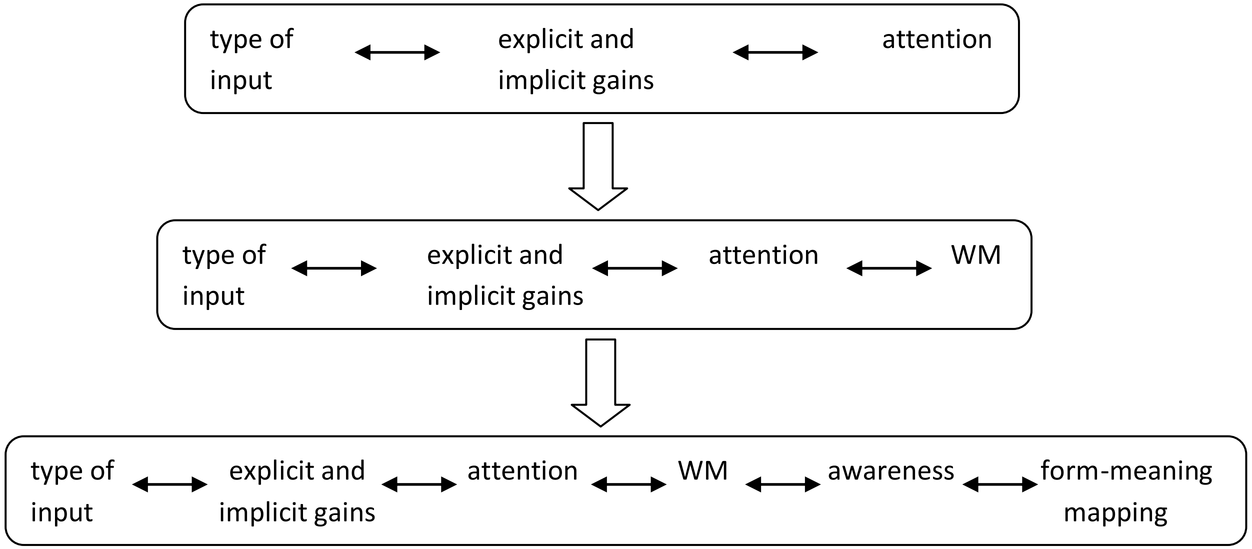
6.5 Free writing task findings

Doughty (2001) highlights the importance of developing learners' ability to use the form-meaning mapping that they have learned or acquired if L2 learning is to be successful. This study investigated the ability of the participants to use the target syntactic construction in a meaning based context in the form of a free writing task. As highlighted in Section 5.6, only $n = 5$ out of $N = 100$ were able to use the target construction in the post-test free writing task. Two out of the 5 who used the target construction in the production task belonged to the *enhanced+ instruction* (A) group and the rest to the *enhanced+ instructions+ explanation* (B) group. Despite poor performance in the whole sample, as Doughty (2001) highlights, the use of the target construction for meaningful communication at least by some participants indicates that the FonF approach has been able to make learners focus on form, meaning and use simultaneously. However, within the present study design, the exposure may have been insufficient for the participants to be able to use the target construction successfully. The findings further revealed that only the learners in the explicit input conditions were able to use the newly learned target construction at least to a certain extent. This indicates that explicit input techniques are more beneficial not only in drawing attention of the learners to the TGC, developing explicit and implicit knowledge of the target construction and developing their ability to understand the form-meaning mapping of the target construction, but also increasing the ability of using the target construction in meaningful communication.

6.6 Summary of the main findings

The study attempted to answer eight research questions under three main themes: (1) the relationship between attentional processing in different input conditions and explicit and implicit learning gains (2) the relationship between attentional processing in different input conditions, explicit and implicit learning gains and WM (3) the relationship between attentional processing, learning gains, WM, form-meaning mapping and awareness. The study also investigated if the learners could use the newly learned target syntactic construction in a meaningful communicative situation. Figure 6.1 summarises these relationships.

Figure 6.1. The relationships analysed in the study



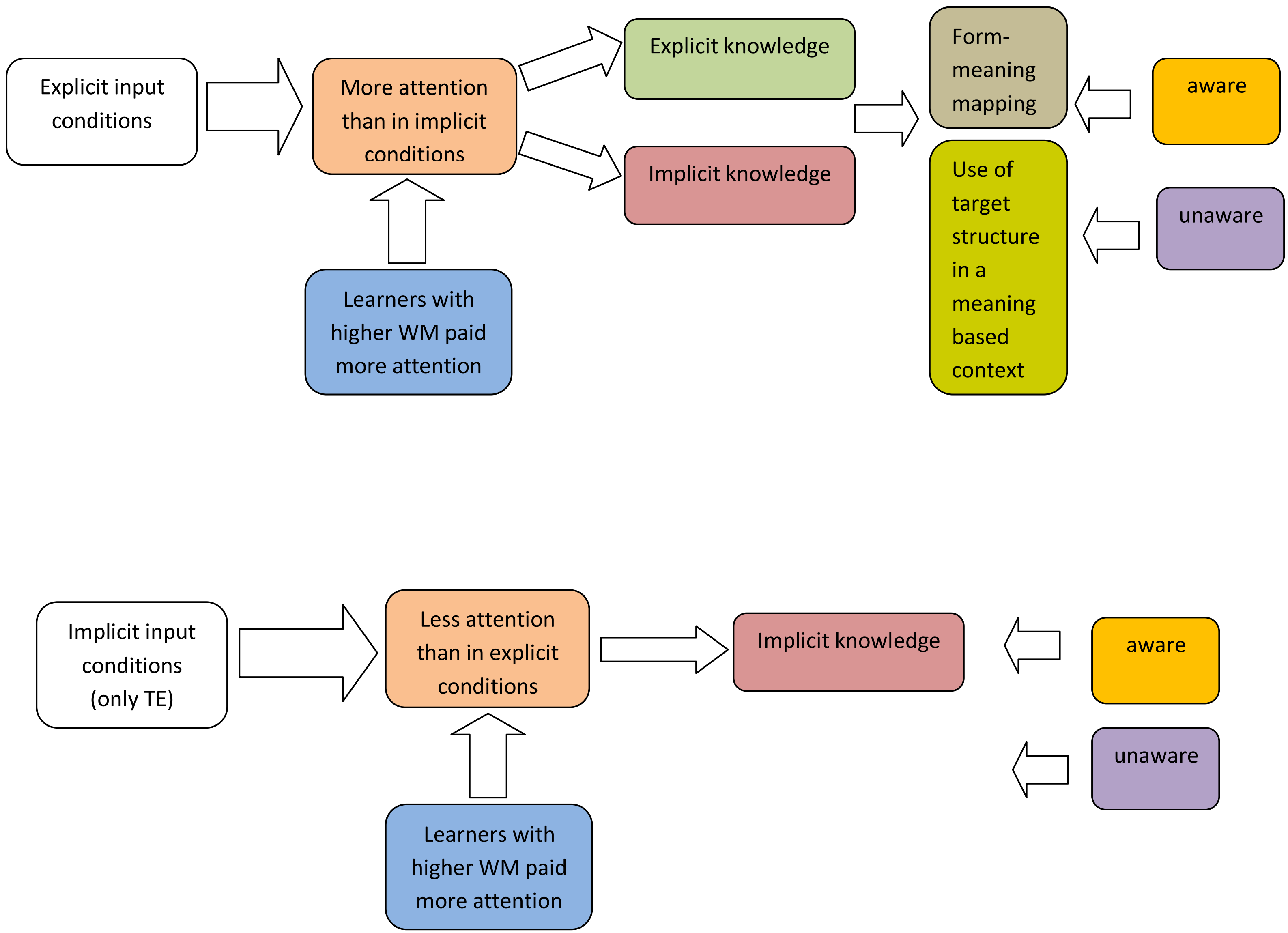
The main findings show that the learners in the explicit input conditions, i.e. in which they received specific instructions to pay attention to the target examples and explicit explanations of the form and meaning of the target syntactic construction, were able to develop their explicit and implicit knowledge of the target construction significantly. The participants in only one of the implicit input conditions i.e. TE could develop their implicit knowledge of the target construction. Moreover, none of the implicit input techniques were influential in developing the explicit knowledge of the target construction. The two explicit input conditions were able to draw the attention of the participants to the target examples more successfully than the implicit input condition. The participants in the TE condition also paid attention to the examples of the target to a certain degree. The learners who paid a significant amount of attention to the target examples could develop their explicit and implicit knowledge of the of the target construction, which indicates a strong relationship between the amount of attention paid and the learning gains.

The study also revealed that there is a strong relationship between the functions of the PSTM capacity and the updating and inhibition functions of the CE, the amount of attention paid to the target examples and the learning gains. There is a clear indication that the learners with high PSTM capacity and attention regulation abilities could develop both explicit and implicit knowledge of the target construction. Nevertheless, the impact of WM is more evident in developing explicit knowledge. The eye-tracking data also suggests that those with a high PSTM capacity and attention regulation abilities are able to direct their attention to the target constructions more successfully which subsequently leads to learning gains in both explicit and implicit input conditions.

The study indicated that gains of either explicit or implicit knowledge are independent of learners being aware of the target syntactic construction. However, methodological limitations in measuring awareness may have been influential in this regard. Moreover, the possible involvement of both top-down and bottom-up input processing especially in the explicit instruction conditions leave us with the question whether the involvement of awareness could be completely ignored.

The findings also highlighted that sufficient exposure to the target syntactic construction and/or an explanation of the meaning of the construction may be necessary for the learners to understand the form-meaning mapping of the target construction. Finally, only a very few learners in the sample (6.25%) were able to use the target construction in a meaningful communicative context highlighting the necessity of either more input, more explanation or more guided practice of the target construction. Figure 6.2 illustrates the main findings of the study.

Figure 6.2. Summary chart of the main findings



Chapter 7: Conclusions

This thesis so far discussed the theoretical rationale behind the present study, the methodology, the results obtained and the discussion of results. The results revealed that this study has made both theoretical and methodological contributions to the field of SLA and the findings in the research have also yielded implications for language teaching pedagogy. This concluding chapter summarises the main findings for the research questions, discusses the main theoretical and methodological contributions and highlights the pedagogical implications that each of the research questions have. It also provides an account of the limitations associated with the study and future research directions.

7.1 Main findings[31]

The study was based on eight research questions and this section summarises the main findings under each research question.

7.1.1 How does knowledge of the form and meaning of the target syntactic construction *causative had*, as measured by a production and a comprehension task, change under explicit and implicit instructional conditions in the case of Sri Lankan pre-intermediate/ intermediate level English language learners?

This question was investigated through a pre- and a post-test design that included SR items and GJ items. In the SR items, the participants had to rewrite the given sentences

[31] Some parts of this section were published in Indrarathne & Kormos (2016).

starting with the word/s provided so that the new sentences meant exactly the same as the first sentence. It was necessary for the participants to use the TGC *causative had* when reconstructing the target sentences. In the GJ task, the participants had to distinguish between well-formed and ill-formed sentences (accurate and inaccurate). SR items were hypothesized to measure the explicit knowledge of the target construction while the timed GJ task aimed to measure the implicit knowledge of the target construction. A MANOVA analysis revealed a statistically significant change in both explicit (SR) and implicit (GJ) knowledge of the target construction over time for the whole sample. Between group effects showed a statistically significant improvement in the case of the *enhanced+ instruction* (A) and the *enhanced+ instructions+ explanation* (B) conditions over time in the SR task while the *enhanced+ instruction* (A), the *enhanced+ instructions+ explanation* (B) and the *enhanced only* (C) groups demonstrated a significant improvement over time in the GJ task. The performance in the post-test revealed that the explicit input groups (instructions to pay attention to the target examples and the explicit explanation) were able to develop both explicit and implicit knowledge of the target construction while the TE condition, which is one of the incidental input mechanisms, only resulted in a change in implicit knowledge. The participants in the other incidental input condition, i.e. input flood, did not change either their explicit or implicit knowledge of the target construction. This finding further strengthens the results of Goo et al.’s (2015), Norris and Ortega’s (2000) and Spada and Tomita’s (2010) meta-analysis that explicit FonF techniques are more beneficial than implicit FonF techniques for L2 development.

7.1.2 How does attention paid to examples of the *causative had* construction in the input texts differ under explicit and implicit instructional conditions?

The attention paid to the examples of the target construction was measured through two eye-tracking measurements: TFD, and the difference between observed and expected TFD (∆OE). The second measurement was introduced in this study as a novel eye-tracking measurement to assess attention. The study also used TFD, one of the traditional measurements of attention in eye-tracking studies. According to the results, the *enhanced+ instructions+ explanation* (B) group demonstrated the highest mean TFD-AOI and ∆OE-AOI in the AOIs in all the input texts. The *enhanced+ instructions* (A) group and the *enhanced only* (C) group showed the second and third highest mean values respectively in terms of both TFD-AOI and ∆OE-AOI. The lowest mean value in TFD-AOI was found for the *unenhanced* (D) group with a 0 mean value in ∆OE-AOI. Paying attention to input means selecting certain features in the input and spending more time on them reading and rereading (Smith & Kosslyn, 2006). ∆OE was hypothesized to measure this possible extra time spent paying attention to the AOIs. Thus, the *unenhanced* (D) group showing 0 mean value in ∆OE-AOI highlights that the learners in this group did not spend extra time on the target items. In other words they did not pay attention to the target items in the input. This reveals that input flood was unable to draw the attention of the learners to the target examples. Although the mean values of TFD-AOI and ∆OE-AOI in the TE condition were higher than the values in the input flood condition, the difference was minimal suggesting that not even TE was able to draw the attention of the learners to the target examples successfully.

7.1.3 How is the change in knowledge of the target construction under explicit and implicit instructional conditions related to the attention paid to examples of the *causative had* construction in the input texts?

This research question was investigated through a correlational analysis of the eye-tracking variables and the gain scores of the SR and GJ tasks. A strong positive correlation was observed between both the SR and GJ gain scores and composite ∆OE-P and TFD-P values in the whole sample. The gain scores of both SR and GJ tasks significantly correlated with the ∆OE-P composite value in the *enhanced+ instructions* (A) group, the *enhanced+ instructions+ explanation* (B) group and the *enhanced* only (C) group and with the composite TFD-P value in the *enhanced+ instructions+ explanation* (B) group and the *enhanced* only (C) group. According to the findings, the learners in the two explicit input conditions (Groups A and B) were able to pay attention to the target constructions and as a result they developed both explicit and implicit knowledge of the target construction. The participants in one implicit condition (Group C – TE condition) were also able to pay attention to the input and simultaneously developed their implicit knowledge of the target construction. This finding is new to the SLA field since the previous studies that investigated the relationship between the attention paid and the learning gains through eye-tracking measurements have reported no relationship between attention and learning gains in TE conditions (Issa et al., 2015). The finding may have been linked to the novel eye-tracking measurement that was used in this study, which provides more accurate information on the amount of attention paid.

7.1.4 How is the functioning of WM including both PSTM capacity and CE functions related to the change of knowledge of the TGC *causative had* in different input conditions?

I used four WM tests for data collection in order to measure the capacity of the PSTM (Digit Span test) and three functions of the CE i.e. updating (Keep Track test), switching (Plus Minus test) and inhibition (Stroop test). The Plus Minus task did not show any correlation with either the SR or the GJ gain scores or the eye-tracking measurements in the initial analysis. Thus, this test was excluded from further analyses. A composite WM score of the Digit Span, the Keep Track and the Stroop tasks was then computed. This composite WM score as well as the three tests separately showed a significant correlation with both the SR and the GJ gain scores of the whole sample. The correlation was stronger in the *enhanced+ instructions+ explanation* (B) group and weaker in the *unenhanced* (D) condition. The multiregression model revealed that the relationship between the SR gain score and the composite WM score in the *enhanced+ instructions+ explanation* (B) group and the *enhanced+ instructions* (A) group was statistically different from the same relationship in the *unenhanced* (D) group. Such a difference was not observed in the case of the GJ task.

The significant correlations between the gain scores of both the SR and the GJ tasks and the composite as well as individual WM scores of the whole sample indicate that WM including the capacity of the PSTM and the CE functions play an important role in the development of both explicit and implicit knowledge. The results related to the SR task highlighted that WM plays a larger role in developing the explicit knowledge in the explicit input conditions than in the implicit input conditions. The multiregression analysis further revealed that WM played an equally important role in all input conditions in the case of the GJ task suggesting that learners with higher WM are at an advantage in developing the implicit knowledge of a novel grammatical construction in L2. The finding related to the facilitative role of WM in explicit input conditions in developing implicit knowledge is new in the field of SLA research.

7.1.5 How is the functioning of the WM related to attention paid to the target items in different input conditions?

A significant but moderately strong relationship between the eye-tracking variables and WM factor score was observed in the results. In particular, the ∆OE value showed a strong correlation with the WM factor score in the *enhanced+ instruction+ explanation* (B), the *enhanced+ instruction* (A) and the *enhanced only* (C) groups, indicating that the external manipulation of the target examples can draw the attention of learners to the target construction and such attention can be influenced by the WM capacity. In other words, L2 learners with higher capacity of the PSTM and CE functions can better regulate their attentional resources to the target input.

7.1.6 Is there an improvement in the CQ scores in different input conditions between the three exposures?

Four CQs were used with each input text in this study with the main purpose of giving the participants a purpose for reading. Two CQs out of the four (6 in total in the composite measurement) were used to assess the participants' understanding of the form-meaning mapping of the target construction when performing a form-meaning recognition comprehension task that assesses learners' explicit knowledge. A significant difference between the composite comprehension test score of exposure session 1 and 2 and 1 and 3 was observed. A significant mean difference was also noticed between the *unenhanced* (D) group and the *enhanced+ instructions* (A) group and the *enhanced+ instructions+ explanation* (B) group. There was also a significant mean difference between the e*nhanced only* (C) group and the *enhanced+ instructions+ explanation* (B) group. The performance in the comprehension task highlighted that the participants in the explicit input conditions were able to understand the form-meaning relationship of the target syntactic construction and apply it in performing a form-meaning recognition comprehension task that assessed their explicit knowledge of the TGC, but not the participants in the implicit input conditions. This is similar to the findings in the SR task in which the participants had to apply their explicit understanding of form-meaning relationship when producing the TGC sentences. This also suggests that learners need to develop explicit knowledge in order to understand the form-meaning relationship of a target construction.

7.1.7 Is the improvement in the CQ task related to the eye fixation measures, the knowledge of the TGC *causative had* (SR and GJ tasks) and WM?

A correlation analysis between the two eye-tracking measurements and the composite comprehension test score was performed in order to answer this research question. The three variables significantly correlated in the whole sample. In explicit input conditions, both TFD-P and ∆OE-P composite values strongly correlated with the composite comprehension test score in the *enhanced+ instructions+ explanation* (B) group. In the *enhanced+ instructions* (A) group, the correlation was significant between the CQ score and the ∆OE-P composite value. The t-tests revealed that only explicit input conditions demonstrated significant learning gains. This suggests that the participants in the explicit input conditions were able to develop the ability to recognise the form-meaning relationship of the target construction in the case of the CQ task since they paid more attention to the target construction. The participants in the implicit input conditions did not develop the ability to understand the form-meaning relationship on this occasion.

The correlation analysis performed between the composite comprehension score and the SR and GJ gain scores revealed that the correlation was significant in the explicit input conditions. Taking this finding and the gains in the CQ task together, it is possible to state that those who developed explicit knowledge were able to develop their understanding of the form-meaning relationship of the target construction. WM capacity measured by the composite WM score also demonstrated a significant influence only in the explicit input conditions when performing the comprehension task.

7.1.8 Is there a difference between the change in the knowledge of the TGC *causative had* (SR, GJ and comprehension tasks) between the participants who reported awareness and those who reported no awareness? Is the awareness reported in different input conditions significantly different to each other?

Awareness was measured in this study in terms of two questions asked in the post-exposure oral interview. According to the answers provided in the interview question, the participants were separated into two groups: 'aware' and 'unaware', those who reported awareness and those who did not. The t-test performed to analyse the improvement in the SR, the GJ and the comprehension tasks indicated that both groups significantly improved their explicit and implicit knowledge and the understanding of the form-meaning relationship of the target construction. This highlights that awareness was not a necessary condition in developing the explicit and implicit knowledge and understanding of the form-meaning relationship of the target construction in this experiment.

The findings also revealed that the mean differences reported in the t-test in the aware group were higher than the reported values in the unaware group. Moreover, the participants in the *enhanced+ instructions+ explanation* (B) group reported more awareness than expected in the standard residual analysis. These findings suggest that explicit input is more beneficial in increasing the awareness of the learners of a target construction than the implicit input. Furthermore, if learners become aware of the target construction while processing input, they are more likely to develop significantly better explicit and implicit knowledge than those who do not demonstrate awareness.

7.1.9 Free writing task

The purpose of including a free writing task in the pre- and post-tests was to analyse if the participants would be able to use the TGC that they learned during the experiment in a meaningful context. They showed poor performance in this task with only 5 participants in the two explicit input conditions using the target construction in the writing task. This highlights that the amount of input and the assistance provided to them might have been insufficient for the learners to be able to use the target syntactic construction in a meaningful communicative context.

7.2 Theoretical and methodological contribution of the study[32]

The main aim of the study was to investigate attentional processing of L2 input by learners and it was vital to understand the cognitive processes associated with attentional processing. The review of literature on attention and the related cognitive mechanisms pointed out several shortcomings in defining the cognitive processes associated with attention in SLA research. Thus, this study turned to the literature in the field of cognitive psychology in order to understand the theoretical concepts underlying attentional processing. As a result, the study avoided using the term noticing and treated attention and awareness as two different phenomena. It also considered that awareness involves conscious attention. In addition to identifying the relationship between these cognitive processes, the study also attempted to use the most accurate methods to measure attention and awareness available to date. Attention was assessed by means of eye-tracking while awareness was investigated with the help of verbal reports of attended stimuli.

Although eye-tracking has been used by several previous studies in the field of SLA research to measure attention, this study argued that the eye-tracking variables in such studies have ignored the selective nature of attention (Fuster, 2005). Thus, a new eye-tracking measurement named as ΔOE (difference between observed and expected TFD) was developed. The measurement was taken based on the number of syllables in an

[32] Some parts of this section were published in Indrarathne & Kormos (2016).

AOI as a ratio of the number of syllables on the whole text on the page. The rationale behind this was that if learners pay attention to features in the input, they should select such features to attend to and spend more time looking at them than the usual reading time they may have spent on them. Instead of taking ratio measurements based on the number of letters or words, the study opted for the number of syllables for two reasons. First, the number of letters or the length of AOIs did not provide any fruitful data in the piloting stage. Second, from a theoretical perspective, phonological awareness including syllables is treated as a strong predictor of reading ability (Bryant, Maclean, Bradley & Crossland, 1990; Chaney, 1998). Moreover, syllables are one of the basic units of word-level decoding in English language because there is no one-to-one correspondence between letters and sounds in English (Ziegler & Goswami, 2005). Although the difference between observed and expected reading time has been used in some cognitive psychological research, to my knowledge, this is the first study in SLA that has applied such a measurement to investigate attention to L2 input.

Unlike most eye-tracking studies that have investigated attentional processing with artificial grammar or individual word/sentence input, this study focused on FonF input in four different input conditions. It used three short stories as input texts in which the examples of the target syntactic construction were included. Unlike in artificial grammar studies, an existing English syntactic construction, *causative had* was embedded in the stories as the TL. There were two explicit input conditions. One of the groups received textually enhanced input with an explanation of the form and meaning of the target construction. In addition, the learners in this group were asked to pay attention to the highlighted target examples. The other group received the textually enhanced input with the instructions asking them to pay attention to the highlighted target examples. The study also had two implicit input conditions. One was an input flood condition i.e. the participants were given reading texts in which several examples of the target construction were embedded. The other implicit input condition, which is the TE group, received the reading text with the target examples highlighted by boldfacing. Thus, this is the first study in SLA that has combined two explicit input conditions and two implicit conditions and used eye-tracking to investigate attentional processing in these different input conditions.

The other important contribution that this study has made is the use of a battery of WM tests. The relationship between WM capacity and input processing has been measured before in SLA research. However, to my knowledge, none of those empirical studies has used a battery of tests to measure the capacity of both PSTM and the updating, switching and inhibition functions of the CE to analyse the relationship between WM and attentional processing. The findings in the study highlighted that three out of the four tests (Digit Span, Keep Track and Stroop) as a composite measurement can be a strong predictor of the relationship between WM capacity, input processing and learning gains.

With regard to the findings of the research, my study revealed that explicit input conditions are more beneficial in developing the explicit and implicit knowledge of a target syntactic construction than implicit input conditions. Moreover, explicit input conditions are able to draw the attention of learners to novel grammatical constructions significantly better than implicit input conditions; this subsequently leads to learning gains. Although implicit input conditions have demonstrated that they were less beneficial in drawing attention of the learners to the target examples, a new finding in this study was that the learning gains demonstrated by the participants in the TE condition are significantly correlated with the amount of attention paid. Moreover, this study demonstrated that high PSTM capacity and attention regulation abilities are influential in learning a novel syntactic construction when developing both explicit and implicit knowledge in both implicit and explicit learning conditions. Another novel finding in this study was the establishment of a relationship between WM capacity and implicit knowledge development in the explicit input conditions. Previous studies to date have identified that WM functions, mainly PSTM capacity, are related to explicit knowledge in explicit conditions and/or explicit/implicit knowledge in implicit conditions. The study also revealed that explicit knowledge gained through explicit input is necessary for the learners to recognise the form-meaning relationship of a novel grammatical construction.

7.3 Pedagogical implications[33]

The study also has several pedagogical implications for the second language classroom. One obvious implication of the findings concerns the effectiveness of different input conditions. The results of this study clearly demonstrated that explicit instruction is more productive in developing the knowledge of a novel syntactic construction of the L2 learners than implicit instruction. Although asking learners to pay attention to the target constructions in the input could develop the explicit knowledge of the target syntactic construction, the findings revealed that additional explicit explanation could bring even better results. For example, the metalinguistic explanation was able to draw the attention of the learners to the target construction immediately after the explanation

[33] Some parts of this section were published in Indrarathne & Kormos (2016).

was provided in the case of Text 2 in the experiment. Furthermore, the group who received the metalinguistic explanation indicated better performance in all the tasks: SR, GJ, comprehension and free writing. Thus, it seems that more support in the form of instructions asking learners to pay attention to the enhanced examples and providing an explanation of the target construction may be necessary for the acquisition of novel L2 constructions.

Another crucial pedagogical implication of the study is the importance of understanding individual differences among learners in terms of their WM capacity. The study clearly showed that WM capacity is related to learning gains. In other words, learners with higher PSTM capacity and attention regulation abilities showed the ability to process input more successfully than the learners with lower PSTM capacity and attention regulation abilities. This indicates that those who have a higher WM capacity could allocate more resources in their WM to process the input. Thus, it is important for language teachers to understand that learners with lower WM capacity may need additional time to process input and potentially they might also require additional support. This could be in the form of explicit explanation and/or more extensive exposure.

The further pedagogical implication of the study concerns the question whether learners are able to use the knowledge of the target construction in a meaning based communicative activity as a result of attentional processing of input. This study indicated that the learners' ability to use the newly learned target syntactic construction in a meaningful situation was minimal in both explicit and implicit input conditions. This suggests that language teachers should not test the ability of learners to use the TL forms in meaningful communication conditions without providing sufficient opportunities to practice using them. Moreover, the tasks that measure the accurate use of form and meaning of the target structure in isolation may not be indicative of learners' ability to use the target construction in a meaningful context. Providing an explicit explanation on one occasion or providing several examples of the target construction may not be sufficient for the learners to develop their ability to use the target construction in a meaningful context. More explicit explanations and more guided practice may thus be useful.

As I mentioned in Chapters 3 and 4, I created the input texts for the study due to the fact that no texts could be found that would have fulfilled all the requirements of this particular research (e.g., number of examples, vocabulary within the participants' proficiency level and similar length). Texts books, particularly at lower levels, often contain textually enhanced input texts, but such texts are less frequent in text books aimed at higher proficiency learners. Therefore, at higher proficiency levels, teachers may have to create enhanced texts if textual enhancement is used as an input technique.

One crucial finding relating to the analysis of the target items within the texts is that regardless of the input condition, the participants' attention to the target examples declined over time with exposure. This implies that providing a few examples of the target structure with explicit instructions may be sufficient for learning to take place. It also may be the case that textual enhancement of all examples in the text is unnecessary if explicit instructions are given. However, as Hulstijn (2015) also notes, it could still be necessary to flood learners with examples in implicit learning contexts.

Finally, the study indicated that explicit instruction can develop explicit and implicit knowledge and implicit learning conditions lead to implicit knowledge development. Therefore, it is important for language teachers to understand that only explicit methods or only implicit methods of assessing learners' knowledge may not provide a true picture of their knowledge. Thus assessment should be prepared to test both explicit and implicit knowledge of the learners.

7.4 Limitations[34]

This study is not without limitations, one of the most important is the fact that a large number of participants had to be excluded from the analysis of the eye-tracking data. This was due to the fact that a portable eye-tracker was used to access a sufficiently large number of participants from a homogenous language learning background. In future studies more sensitive equipment or even a larger number of participants would be needed to gain a more nuanced insight into the differences of attentional processing under different experimental conditions and to examine the relationship between eye-tracking measures and learning gains.

Due to data collection constraints, no delayed post-tests could be administered, and hence there is little knowledge about how long-lasting the learning gains were. Although I administered a production test, which would have required the participants to apply the TGC in a meaningful context in a written text, only a few learners actually used the construction in this test making it impossible to conduct statistical analyses. Therefore, in future studies other means for eliciting the productive use of the targeted

[34] Some parts of this section were published in Indrarathne & Kormos (2016).

construction would be needed. Another methodological limitation lies in the words used in the input texts. Although I controlled for unknown vocabulary in the texts, it does not guarantee that all the participants knew all the words used.

In addition, reading on a computer screen is different to real-life reading that usually takes place in classroom. For example, in the experimental condition, the participants did not see the whole story on one page, and hence they were not able to go back and forth while reading. This can make a difference to the amount of attention paid to the examples as well as the retention of information in memory.

In the discussion of the findings in this study, specific hypotheses have been set up about the kind of knowledge the test-tasks assessed. For example, I assumed that the SR task would measure explicit knowledge and the GJ task would measure implicit knowledge. Using confidence ratings or tracking the eye-movements of the participants while performing the tasks might yield more reliable information whether learners rely on explicit or implicit knowledge in these tasks.

A larger battery of WM tests could also have been applied so that conclusions about participants' WM abilities based on more sources of information could have been made. It would also be important to examine to what extent the conclusions concerning the role of WM abilities in grammar learning can be upheld with other syntactic constructions and when input is provided through an aural rather than written mode. Additional sources of information on whether particular instruments that gauge learning gains tap into explicit or implicit knowledge representations could also lend more support to the hypotheses with regard to the role of WM abilities in developing explicit and implicit L2 knowledge.

Due to practical reasons, this study used two WM tests (Keep Track and Stroop) in the L2. It is believed that the language of the WM test can be influential in accurately measuring the WM capacity. For example, Sanchez et al. (2010) administered a reading span and an operation span task only in English to a group of non-native speakers to analyse if a relationship between WM capacity and Raven's Advanced Progressive Matrices (RAPM) could be observed when WM tests are given in an L2. They found that the operation span task could reliably predict the participants RAPM, but not the reading span task. Thus, they note that the language of the WM test plays a role in reading span test. Therefore, it is important to administer WM tests in L1 in order to investigate if similar results can be obtained.

The calculation between the expected and observed TFD was based on word length measured in syllables, but this operationalization can be refined further to take various lexical characteristics into account such as word-frequency, concreteness and imageability that potentially influence how long a word is fixated on (Reichle et al., 2006).

The results obtained regarding the research question related to awareness also have limitations. The study used oral reporting as a measurement of awareness; however, as discussed in Chapter 6, there are other mechanisms that inform awareness other than oral reports. Moreover, the question used to collect data on awareness also has a methodological issue. It is likely that learners, in particular in the implicit input conditions, were confused by the use of the term ‘grammar structure’ in the question, which shows that the question also collected data on metalinguistic awareness of the learners.

Finally, memory may have influenced the performance in the CQ task. The participants had to read the whole text at once and answer the CQs without going back to the text. Thus, one could argue that more than understanding the form-meaning relationship, the participants had to rely on their memory of the storyline to answer the questions.

7.5 Future research directions

One of the future research possibilities is the investigation of the attentional processing of aural input. Aural input is the main input mode in L2 learning and thus it would be interesting to learn how different aural input types draw attention of the learners to the TL features and which of those types can be more successful in doing so.

The present study was based on learners in the pre-intermediate level of language proficiency and a future research direction is to replicate this study design at other proficiency levels in order to investigate if attentional processing and WM influence the learning of a novel grammatical construction in the same way at different proficiency levels. In particular, experiments at lower proficiency levels could provide more opportunities to explore the acquisition of a larger variety of syntactic constructions that are widely used in language teaching materials.

Not only WM capacity, but also other individual differences such as age, motivation and aptitude may impact the attentional processing of input. As such, measuring the relationship between other individual differences and attentional processing might provide additional insights into other cognitive mechanisms involved in attentional processing. Another suggestion is to use up to date WM tests that can be administered in more sophisticated modes (e.g., e-prime). Such tests would also provide more accurate data.

Involvement of awareness in attentional processing could also be investigated using more accurate data collection methods. This study indicated that awareness was not a necessary condition for the development of knowledge of a TGC; however, awareness could not be measured accurately since only one measurement of awareness was used. It would be worth investigating how to measure awareness more accurately. Such investigations would provide better insights into the involvement of awareness in language learning.

7.6 Concluding remarks

As I mentioned in the introduction of this thesis, the motivation behind this study was analysing why some adult learners process L2 input related to TGCs more successfully than others despite being exposed to the same input. I attempted to analyse if the answer lies in the type of input provided by including four different types of input in the study design. Moreover, my observation that adult learners prefer receiving explicit explanations on the TL features was influential in including an input condition in which learners received an explicit explanation of the target construction. Based on

DeKeyser's (2005) argument that individual differences play an important role in how grammar is learned, I also investigated the relationship between WM capacity and learning gains.

The study provided me with answers to the initial question. For example, the reasons behind some learners processing L2 input differently to others seem to lie in both the type of input provided as well as individual differences (WM in this case). Further, adult learners seem to develop their knowledge of target language more successfully if they receive an explicit explanation. This finding may explain why some even request explicit explanations from the teacher. Thus, it can be concluded that the type of input as well as individual differences should be taken into consideration when providing TL input.

References

Ahmadian, M. J. (2015). Working memory, online planning and L2 self-repair behaviour. In Z. Wen, M. Borges & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 160-174). Bristol, England: Multilingual Matters.

Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics, 29*, 175-193.

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in second language acquisition* (pp. 259–299). Honolulu, HI: University of Hawai'i Press.

Allport, A. (1988). What concept of consciousness? In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 159-182). London, England: Clarendon Press.

Allport, A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty five years. In D.E. Meyer & S. Kornblum (Eds.), *Attention and performance: synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp.183-218). Cambridge, MA: MIT Press.

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.

Andringa, S., De Glopper, K. & Hacquebord, H. (2011). Effect of explicit and implicit instruction on free written response task performance. *Language Learning, 61*, 868-903.

Atkinson, A., & Riani, M. (2000). Robust diagnostics and regression analysis. New York, NY: Springer.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation (Vol. 2)* (pp. 89–195). New York, NY: Academic Press.

Baars, B. (1988). *A cognitive theory of consciousness*. Cambridge, England: Cambridge University Press.

Bachman, T. (2006). A single metatheoretical framework for a number of conscious-vision phenomena. In Q. Jing, M.R. Rosenzweig, G. d'Ydewalle, H. Zhang, H. Chen & K. Zhang (Eds.), *Psychological science around the world* (pp. 229–242), Sussex, England: Psychology Press.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49*(A), 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 and language: An overview. *Journal of Communication Disorders, 36*(3), 189-208.

Baddeley, A.D. (2015). Working memory in second language learning. In Z. Wen, M. Borges & A. McNeill (Eds.). *Working memory in second language acquisition and processing* (pp. 17-28). Bristol, England: Multilingual Matters.

Baddeley, A. D., & Hitch, G. (1974). *Working memory.* In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory (Vol. 8)* (pp. 47–89). New York, NY: Academic Press.

Bao, M., Egi, T., & Han, Y. (2011). Classroom study on noticing and recast features: Capturing learner noticing with uptake and stimulated recall. *System, 39*, 215-228.

Baralt, M. (2015). Working memory capacity, cognitive complexity and L2 recasts in online language teaching. In Z. Wen, M. Borges & A. McNeill (Eds.). *Working memory in second language acquisition and processing* (pp. 248-269). Bristol, England: Multilingual Matters.

Barcroft, J. (2009). Effects of synonym generation on incidental and intentional L2 vocabulary learning during reading. *TESOL Quarterly*, *43*(1), 79-103.

Basturkmen, H., Loewen, S., & Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics*, *25*(2), 243–272.

Batterink, L. & Neville, H. (2013). The human brain processes syntax in the absence of conscious awareness. *Journal of Neuroscience, 33,* 8528-8533.

Becker, S., Kleinbohl, D., & Holzl, R. (2012). Awareness is awareness is awareness? Decomposing different aspects of awareness and their role in operant learning of pain sensitivity. *Consciousness and Cognition, 21*, 1073-1084.

Berquist, B. (1997). Individual differences in working memory span and L2 proficiency: capacity or processing capacity? In A. Sorace, C. Heccock & R. Shillcock (Eds.), *Proceedings of the GALA '97 conference on language acquisition* (pp. 468-473). Edinburgh, Scotland: The University of Edinburgh.

Bialystok, E. (1979). Explicit and implicit judgements of L2 grammaticality. *Language Learning, 29,* 81-103.

Bialystok, E. (1982). On the relationship between knowing and using forms. *Applied Linguistics, 3*, 181-206.

Biedroń, A., & Szczepaniak, A. (2012). Working memory and short-term memory abilities in accomplished multilinguals. *The Modern Language Journal, 96,* 290–306.

Bowles, M. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute? *Studies in Second Language Acquisition, 33*, 247-271.

Briscoe, J., & Rankin, P. M. (2009). Exploration of a 'double-jeopardy' hypothesis within working memory profiles for children with specific language impairment. *International Journal of Language & Communication Disorders, 44,* 236-250.

Broadbent, D. (1958). *Perception and communication*. London, England: Pergamon Press.

Bryant, P. E., Maclean, M., & Bradley, L. (1990). Rhyme, language, and children's reading. *Applied Psycholinguistics, 11*, 237-252.

Burgess, J. & Etherington, S. (2002). Focus on grammatical form: explicit or implicit? *System*, *30*, 433-458.

Carr, T. H., & Curran, T. (1994). Cognitive factors in learning about structured sequences: Applications to syntax. *Studies in Second Language Acquisition, 16,* 205-230.

Carrol, S. (2006). Salience, awareness and SLA. In M. O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th generative approaches to second language acquisition conference (GASLA 2006)* (pp.17-24). Somerville, MA: Cascadilla Proceedings Project.

Chaney, C. (1998). Preschool language and metalinguistic skills are links to reading success. *Applied Psycholinguistics, 19*(3), 433-46.

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

Chelazzi, L., Perlato, A., Santandrea, E., & Della Libera, C. (2013). Rewards teach visual selective attention. *Vision Research, 85*, 58-72.

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

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, *62*, 73-101.

Cognitivefun. (2008-2012). Cognitive tests: Color Reading Interference (Stroop) [Online test]. Retrieved from http://cognitivefun.net/test/2 [Accessed 01/04/2014].

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Lawrence Erlbaum.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*, 589-608.

Conway, C.M., & Christiansen, M.H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus specific representations. *Psychological Science, 17*, 905-912.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics in Language Teaching*, *5*, 161–170.

Council of Europe (2011). *Common European Framework of Reference for Languages*. Strasbourg: Language Policy Division, Council of Europe. Retrieved from http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf. [Accessed 12/12/2015]

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

Cowan, N. (2015). Second language use, theories of working memory and the vennian mind. In Z. Wen, M. Borges Mata & A. McNeill. *Working memory in second language acquisition and processing* (pp. 29-41). Bristol, England: Multilingual matters.

Cunningham, S. & Moor, P. (2013). *New cutting edge upper intermediate* (3rd ed.). Oxford, England: Pearson.

Curran, T., & Keele, S.W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology*: *Learning, Memory, and Cognition, 19*, 189–202.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour, 19,* 450-166.

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

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3,* 422–433.

de Abreu, E., Pascale, M. J., & Gathercole, S. E. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology, 104*(4), 974-986.

Dede, G., Caplan, D., Kemtes, K. A., & Waters, G. (2004). The relationship between age, verbal working memory, and language comprehension. *Psychology and Aging, 19,* 601-616.

de Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition, 19,* 249–276.

DeKeyser, R. M. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17*, 379-410.

DeKeyser, R.M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning, 55*(1), 1-25.

Desimone, R., & Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annual Reviews of Neuroscience 18,* 193–222.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–57). Cambridge, England: Cambridge University Press.

Doughty, C., & Williams, J. (1998). Issues and terminology. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 1-12). Cambridge, England: Cambridge University Press.

Egi, T. (2004). Verbal reports, noticing, and SLA research. *Language Awareness, 13*(4), 243-264.

Ellis, N.C. (1996). Sequencing in SLA: phonological memory, chunking, and points of order. *SSLA, 18*, 91-126.

Ellis, N.C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24,* 143-188.

Ellis, N.C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27,* 305–352.

Ellis, N.C. (2006). Phenomena in L2 acquisition: contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 164–194.

Ellis, N. C. (2011). Implicit and explicit SLA and their interface. In C. Sanz & R. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA & bilingualism* (pp. 35-47). Washington, DC: Georgetown University Press.

Ellis, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.) *Implicit and explicit learning of languages* (pp. 3-25). Amsterdam, The Netherlands: John Benjamins.

Ellis, N. C., & Sagarra, N. (2010). Learned attention effects in L2 temporal reference: The first hour and the next eight semesters. *Language Learning, 60*(2), 85-108.

Ellis, N. C., & Sagarra, N. (2011). Learned attention in adult language acquisition: A replication and generalization study and meta-analysis. *Studies in Second Language Acquisition*, *33*(4), 589-624.

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: putting language in good order. *The Quarterly Journal of Experimental Psychology, 49A*(1), 234-250.

Ellis, R. (1991). Grammatically judgments and second language acquisition. *Studies in Second Language Acquisition, 13*(2), 161-186.

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

Ellis, R. (1999). Input-based approaches to teaching grammar: A review of classroom-oriented research. *Annual Review of Applied Linguistics, 19*, 64-80.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A Psychometric Study. *Studies in Second Language Acquisition, 27*, 141-172.

Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis (Ed.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-25). Bristol, England: Multilingual Matters.

Ellis, R. (2015). Form-focused instruction and the measurement of implicit and explicit L2 knowledge. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 417-442). Amsterdam, The Netherlands: John Benjamins.

Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus-on-form. *System, 30*, 419–432.

Ellis, R., & Loewen, S. (2007). Confirming the operational definitions of explicit and implicit knowledge in Ellis (2005): Responding to Isemonger. *Studies in Second Language Acquisition, 29*, 119-126.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. London, England: Routledge.

Engle, R. W., Laughlin, J. E., Tuholski, S. W., & Conway, A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309-333.

Erçetin, G., & Alptekin, C. (2013). The explicit/implicit knowledge distinction and working memory: Implications for second-language reading comprehension. *Applied Psycholinguistics*, *34*, 727-753.

Erickson, A. P., (2013). The ABC's of font: effects of changing default fonts. *EWU Masters Thesis Collection.* Paper 108.

Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A.

Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 257-297). Cambridge, England: Cambridge University Press.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, *9*, 147–171.

Fuster, J. M. (2005). The cortical substrate of general intelligence. *Cortex, 41*, 228-229.

Gascoigne, C. (2006). Explicit input enhancement: Effects on target and non-target aspects of second language acquisition. *Foreign Language Annals, 39*, 551-564.

Gass, S. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics, 9*, 198-217.

Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.

Gass, S., & Lee, J. (2011). Working memory capacity, inhibitory control, and proficiency in a second language. In M. S. Schmid & W. Lowie (Eds), *Modeling bilingualism: From structure to chaos* (pp.59-84). Amsterdam, The Netherlands: John Benjamins.

Gass, S. M., & Mackey, A. (2015). Input, interaction and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: an introduction* (pp 180-206). New York, NY: Routledge.

Gass, S., & Selinker, L. (2008). Second language acquisition*:* an introductory course, 3rd ed. London, England: Routledge.

Gass, S., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning, 53,* 497–545.

Gathercole, S., & Baddeley, A. D. (1993). *Working memory and language* (Essays in cognitive psychology). Hove, England: Lawrence Erlbaum.

Gathercole, S. E., Willis, C. S., Baddeley, A. D. & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2*, 103-127.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for word: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition, 35*(3), 483-517.

Godfroid, A., Loewen, S., Jung, S., Park, J., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgment measure distinct types of knowledge: Evidence from eye-movement patterns. *Studies in Second Language Acquisition, 37*(2), 269-297.

Godfroid, A., & Schmidtke, J. (2013). What do eye movements tell us about awareness? A triangulation of eye-movement data, verbal reports, and vocabulary learning scores. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 183–205). Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center.

Godfroid, A., & Uggen, M. S. (2013). Attention to irregular verbs by beginning learners of German: An eye-movement study. *Studies in Second Language Acquisition*, *35*(2), 291-322.

Goldschneider, J., & DeKeyser, R. (2005). Explaining the natural order of L2 morpheme acquisition in English: A meta-analysis of multiple determinants, *Language Learning, 55*, 27-77.

Goo, J., Grañena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning: Norris & Ortega (2000) revisited and updated. In P.

Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443-482). Amsterdam, The Netherlands: John Benjamins.

Greene, J., & Hicks, C. (1984). *Basic cognitive processes*. Milton Keynes, England: Open University Press.

Grey, S., Williams, J. N., & Rebuschat, P. (2015). Individual differences in incidental language learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences, 38,* 44-53.

Guo, X., Zheng, L., Zhu, L., Yang, Z., Chen, C., & Zhang, L. (2011). Acquisition of conscious and unconscious knowledge of semantic prosody. *Consciousness and Cognition*, *20*, 417–425.

Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit L2 knowledge. *Studies in Second Language Acquisition, 35*(3), 423-449.

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

Hassin, R. R., Bargh, J. A., Engell, A. D., & McCulloch, K. C. (2009). Implicit working memory. *Consciousness and Cognition, 18,* 665–678.

Heatley, A., Nation, I.S.P. & Coxhead, A. (2002). Range and frequency programs. Retrieved from http://www.victoria.ac.nz/lals/staff/paul-nation.aspx.

Hernández, T. A. (2008). The effect of explicit instruction and input flood on students' use of discourse markers on a simulated oral proficiency interview. *Hispania, 91,* 665–75.

Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., & De Wiijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford, England: Oxford University Press.

Hull, R., Martin, R. C., Beier, M. E., Lane, D., & Hamilton, A. C. (2008). Executive function in older adults: A structural equation modeling approach. *Neuropsychology, 22*, 508–522.

Hulstijn, J. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds), *The handbook of second language acquisition* (pp. 349-381). Oxford, England: Blackwell.

Hulstijn, J. (2015). Explaning phenomena of first and second language acquisition with the constructs of implicit and explicit learning. The virtues and pitfalls of a two system view. In P. Rebuschat (Ed.) *Implicit and explicit learning of languages* (pp. 25-47). Amsterdam, The Netherlands: John Benjamins.

Hummel, K. M. (1998). *Working memory capacity and L2 proficiency.* In Complementary perspectives on second language research. Second language research forum conducted at University of Hawai'i, Honolulu, HI.

Hunt, A. R., & Kingstone, A. (2003). Covert and overt voluntary attention: linked or independent? *Cognitive Brain Research, 18*, 102 – 105.

Hutton, U., & Towse, J. (2001). Short-term memory and working memory as índices of children's cognitive skills. In S. Gathercole (Ed.), *Short-term and working memory* (pp. 383-394). Sussex, England: Psychology Press Ltd.

Indrarathne, B., & Kormos, J. (2016). Attentional processing of input in explicit and implicit learning conditions: an eye-tracking study. *Studies in Second Language Acquisition*. DOI: http://dx.doi.org/10.1017/S027226311600019X

Isaki, E. & Plante, E. (1997). Short-term and working memory differences in language/learning disabled and normal adults. *Journal of Communication Disorders, 30*(6), 427-437.

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, The Nertherlands: John Benjamin Publishing Co.

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. (2003). Visual input enhancement as focus on form. *Sophia Linguistica, 13*, 1–30.

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, *34*(2), 239-278.

Iwanaka, T., & Takatsuka, S. (2007). Influences of attention and noticing on second language acquisition. *Languages & Literatures, 2,* 21-30.

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*(1), 46-66.

James, W. (1890). *The principles of psychology*. New York, NY: Henry Holt.

Jeffries, S., & Everatt, J. (2004). Working memory: Its role in dyslexia and other specific learning difficulties. *Dyslexia, 10*(3), 196-214.

Jonides, J., & Smith, E.E. (1997). The architecture of working memory. In M. D. Rugg (Ed.), *Cognitive Neuroscience* (pp. 243-276), Sussex, England: Psychology Press.

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, HI: University of Hawai'i Second Language Teaching and Curriculum Centre.

Juan, C., Shorter-Jacobi, S., Schall, J., & Sperling, G. (2004). Dissociation of Spatial Attention and Saccade Preparation. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(43), 15541-15544.

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research, 21*(2), 121–151.

Juffs, A. (2006). Working memory, second language acquisition and low-educated second language and literacy learners. *Low-educated adult second language and literacy acquisition: Proceedings of the Inaugural Symposium, Occasional Series*, 89–104.

Juffs, A., & Harrington, M. W. (2011). Aspects of working memory in L2 Learning. *Language Teaching: Reviews and Studies*, *42,* 137-166.

Just , M. A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.

Kachinske, I., Osthus, P., Solovyeva, K., & Long, M. (2015). Implicit learning of a L2 morphosyntactic rule, and its relevance for language teaching. In P. Rebuschat (Ed.) *Implicit and explicit learning of languages* (pp. 387-417). Amsterdam, The Netherlands: John Benjamins.

Kahneman, D. (1973). *Attention and effort*. New Jersey, NJ: Prentice-Hall.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29-61). Orlando, FL: Academic Press Inc.

Kentridge R. W. (2011). Attention without awareness: A brief review. In C. Mole, D. Smithies & W. Wu (Eds.), *Attention: Philosophical and psychological essays* (pp. 228-246). Oxford, England: Oxford University Press.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1999). Insights into working memory from the perspective of the EPIC architecture for modeling skilled perceptual-motor and cognitive human performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 183-223). New York, NY: Cambridge University Press.

Kim, S., Christianson, K., & Packard, J. (2015). Working memory in L2 character processing: The case of learning to read Chinese. In Z. Wen, M. Borges Mata & A. McNeill. *Working memory in second language acquisition and processing* (pp. 85-104). Bristol, England: Multilingual matters.

Koch, C., & Tsuchiya, N. (2006). Attention and consciousness: two distinct brain processes. *TRENDS in Cognitive Sciences, 11*(1), 16-22.

Kormos J., & Safar A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 2*, 261-271.

Krashen, S. D. (1985). *The input hypothesis: issues and implications.* London, England: Longman.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods, 9*(2), 202–220.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *TRENDS in Cognitive Sciences, 7*(1), 12-18.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*(3), 339-354.

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.

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. (1997). Attention, awareness and foreign language behaviour. *Language Learning, 47* (3), 467-505.

Leow, R. (2000). A study of the role of awareness in foreign language behaviour. *Studies in Second Language Acquisition, 22*, 557-584.

Leow, R. (2001). Do learners notice enhanced forms while interacting with the L2?: An online and offline study of the role of written input enhancement in L2 reading. *Hispania, 84*(3), 496-509.

Leow, R. (2013). Attention, Noticing, and Awareness in Second Language Acquisition. In C. A. Chapelle (Ed.) *The Encyclopedia of Applied Linguistics* (pp. 1-7). Oxford, England: Blackwell Publishing Ltd.

Leow, R., Egi, T., Nuevo, A., & Tsai, Y. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning, 13* (2), 1–16.

Leow, R. & 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. (2015). Working memory, language analytical ability and L2 recasts. In Z. Wen, M. Borges & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 139-159). Bristol, England: Multilingual Matters.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2013). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin and Review*, 1-23.

Liversedge, S., Gilchrist, I. D., & Everling, S. (2011). *The Oxford handbook of eye movements [electronic resource]* ([Oxford library of psychology]). Oxford, England: Oxford University Press.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, D. Coste, C. Kramsch, & R. Ginsberg (Eds.), *Foreign language research in a cross cultural perspective* (pp. 39–52). Amsterdam, The Netherlands: John 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, England: Cambridge University Press.

Lovett, M. C., Reder, L. M., & Lebiere, C. (1999). Modeling working memory in a unified architecture: An ACT-R perspective. In: A. Miyake & P. Shah (Eds), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 135-182). New York, NY: Cambridge University Press.

Marcel, A. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*, 197-237.

Massaro, D. W. (1975). *Experimental psychology and information processing*. Chicago, IL: Rand- McNally.

McBride-Chang, C., & Kail, R. (2002). Cross-cultural similarities in the predictors of reading acquisition. *Child Development, 73,* 1392– 1407.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics, 27*(3), 405–430.

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. 181-210). Amsterdam, The Netherlands: John Benjamins Publishing Company.

Mackey, A., & Sachs, R. (2012), Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning, 62*, 704–740.

MacWhinney, B. (1997). Implicit and explicit processes. *Studies in Second Language Acquisition, 19,* 277-281.

Martin, K. I. & Ellis, N. C. (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.

Massaro, D. W., (1975). *Experimental psychology and information processing*. Chicago, IL: Rand-McNally.

Max Velmans, G. (2009). How to define consciousness—and how not to define consciousness. *Journal of Consciousness Studies, 16*(5), 139-156.

McFarland, D. (2006). *Dictionary of animal behavior*. Oxford, England: Oxford University Press.

McMurray, B., & Hollich, G. (2009). Core computational principles of language acquisition: can statistical learning do the job? Introduction to special section. *Developmental Science, 12,* 365–368.

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

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (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.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* New York, NY: Cambridge University Press.

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning, 62,* 302-331.

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G.A. (2013). Forward and backward span for verbal and visuo-spatial data: Standardization and normative data from an Italian adult population. *Neurological Sciences, 34*, 749–754.

Mojavezi, A., & Ahmadian, M. J. (2014). Working memory capacity and self-repair behavior in first and second language oral production. *Journal of Psycholinguistic Research, 43*(3), 289-297.

Morais, J. (2003). Levels of phonological representation in skilled reading and learning to read. Reading and writing: *An Interdisciplinary Journal, 16*, 123–151.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning, 60*, 154–193.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review, 83*, 435-450.

Nassaji, H. (1999). Towards integrating form-focused instruction and communicative interaction in the second language classroom: Some pedagogical possibilities. *The Canadian Modern Language Review, 55*, 385–402.

Nassaji, H., & Fotos, S. (2007). Issues in form-focused instruction and teacher education. In S. Fotos & H. Nassaji (Eds), *Form-focused instruction and teacher education* (pp. 7-16). Oxford, England: Oxford University Press.

Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms*. New York, NY: Routledge.

Nezami, S. H. (2012). *A study of errors, corrective feedback and noticing in synchronous computer mediated communication*. (Master's thesis). Retrieved from https://www.academia.edu/3313857/A_Study_of_Errors_Corrective_Feedback_and_Noticing_in_Synchronous_Computer_Mediated_Communication [Accessed on 15/03/2014]

Nissen , M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*, 1 – 32.

Niżegorodcew, A. (2007). *Input for instructed L2 learners: the relevance of relevance*. Clevedon, England: Multilingual Matters.

Norman, D., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. Davidson, R. G. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (pp. 1–18). New York, NY: Plenum Press.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–528.

O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics, 27,* 377–402.

O'Regan, J. K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*, 939–1031.

Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, *30*(4), 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.

Ostrosky-Solís, F., & Lozano, A. (2006). Digit span: Effect of education and culture. *International Journal of Psychology, 41*(5), 333–341.

Overstreet, M. (1998). Text enhancement and content familiarity: The focus of learner attention. *Spanish Applied Linguistics, 2*, 229–258.

Pallant, J. (2010). *SPSS survival manual: A step by step guide to data analysis using SPSS* (4th ed.). Maidenhead, England: McGraw-Hill.

Papagno, C., & Vallar, G. (1995). Short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology, 48*(A), 98-107.

Park, E. S. (2007). *Learner-generated noticing of L2 input: An exploratory study*. Columbia, NY: Teachers College, Columbia University.

Park, E. S. (2010). Learner-generated noticing of written L2 input: What do learners notice and why? *Language Learning, 61*(1), 146–186.

Park, E. S. (2013). Learner-generated noticing behavior by novice learners: tracing the effects of learners' L1 on their emerging L2. *Applied Linguistics, 34*(1), 74-98.

Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal*, *20*(1), 7-32.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233-238.

Pett, M., Lackey, N., & Sullivan, J. (2003). *Making sense of factor analysis*. Thousand Oaks, CA: Sage Publications Inc.

Pickering, S. J., & Gathercole, S. E. (2001). *Working memory test battery for children*. London, England: Psychological Corporation.

Posner, M. I., & Rothbart, M. K. (1992). Attentional mechanisms and conscious experience. In A. D. Milner & M. D. Rugg (Eds.), *Foundations of neuropsychology series* (pp. 91-112). New York, NY: Academic Press.

Puric, D., & Pavlovic, M. (2012). Executive function of shifting: Factorial structure and relations to personality and intelligence domains. *Suvremena Psihologija, 15*(2), 177-191.

Radwan, A. A. (2005). The effectiveness of explicit attention to form in language learning. *System, 33*, 69-87.

Rathvon, N. (2004). Early reading assessment. New York, NY: The Guilford Press.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219-235.

Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(5), 888-896.

Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning, 63*(3), 595-626.

Rebuschat, P., & Williams, J. N. (2012a). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics, 33* (4), 829-856.

Rebuschat, P., & Williams, J. N. (Eds.) (2012b). *Statistical learning and language acquisition*. Berlin: Mouton de Gruyter.

Redston, C. & Cunningham, G. (2013). *Face to face upper intermediate* (2nd ed.). Cambridge, England: Cambridge University Press.

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*(1), 4-22.

Reinders, H., & Ellis, R. (2009). The effects of two types of input on intake and the acquisition of implicit and explicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 281–302). Bristol, England: Multilingual Matters.

Revesz, A. (2012). Working memory and the observed effectiveness of recasts on different l2 outcome measures. *Language Learning, 62*(1), 93–132.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the saccade task. *Journal of Experimental Psychology: General*, *123*, 374–393.

Roberts, L., & Siyanova-Chanturia**,** A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 sentence and discourse processing. *Studies in Second Language Acquisition*, *35*(2), 213-235.

Robinson, P. (1995). Attention, memory and the noticing hypothesis. *Language Learning, 45*(2), 283–331.

Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition, 19*, 223–247.

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., Mackey, A., Gass, S. M., & Schmidt, R. (2012). Attention and awareness in second language acquisition. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition (pp.247-267).* London, England: Routledge.

Rogers, R. D., Monsell, S. (1995). The cost of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology*: *General*, *124*, 207–231.

Rogers, J., Revesz, A. & Rebuschat, P. (2015). Challenges in implicit learning research: Validating a novel artificial language. In P. Rebuschat (Ed.) *Implicit and explicit learning of languages* (pp. 275-301). Amsterdam, The Netherlands: John Benjamins.

Rosa, E., & Leow, R. (2004). Awareness, different learning conditions, and second language development. *Applied Psycholinguistics, 25*, 269–292.

Rosa, E., & O'Neill, M. (1999). Explicitness, intake and the issue of awareness: Another piece to the puzzle. *Studies in Second Language Acquisition, 22,* 511–566.

Russo, J., Johnson, E., & Stephens, D. (1989). The validity of verbal protocols. *Memory and Cognition, 17*, 759-769.

Saggara, 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* (pp. 229-248). Oxford, England: Oxford University Press.

Sagarra, N., & Abbuh, R. (2013). Optimizing the noticing of recasts via computer delivered feedback: evidence that oral input enhancement and working memory help second language learning. *The Modern Language Journal, 97*(1), 196-216.

Sanchez, C. A., Wiley, J., Miura, T., Colfesh, G. H., Ricks, T. R., Jensen, M. S., & Conway, A. R. A. (2010). Assessing working memory capacity in a non-native language. *Learning and Individual Differences, 20,* 488–493.

Santamaria, K., & Sunderman, G. (2015). Working memory in processing instruction: The acquisition of French clitics. In Z. Wen, M. Borges & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 205-223). Bristol, England: Multilingual Matters.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206-226.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp 1-63). Honalulu, HI: University of Hawaii, Second Language Teaching and Curriculum Centre.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge, England: Cambridge University Press.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, *Proceedings of CLaSIC 2010*, *Singapore, December 2-4* (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.

Schneider, W. (1999). Working memory in a multilevel hybrid connectionist control architecture (CAP2) In A. Miyake & P. Shah (Eds). *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 340-374). New York, NY: Cambridge University Press.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception and Psychophysics, 74*(1), 5-35.

Sch¨utze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago, IL: University of Chicago Press.

Segalowitz, N., & Frenkiel-Fishman, S. (2005). Attention control and ability level in a complex cognitive skill: Attention shifting and second-language proficiency. *Memory and Cognition*, *33* (4), 644-653.

Shahabi, S. R., Abad, F. J., & Colom, R. (2014). Short-term storage is a stable predictor of fluid intelligence whereas working memory capacity and executive function are not: A comprehensive study with Iranian schoolchildren. *Intelligence, 44*, 134–141.

Shanks, D. & John, M. (1994). Characteristics of dissociable human learning system. *Behavioural and Brain Sciences*, *17*, 367-447.

Sharwood-Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research, 7,* 118–132.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*, 165 – 179.

Shettleworth, S. J. (1998). *Cognition, evolution, and behavior*. New York, NY: Oxford University Press.

Shiffrin, R. M. (1988). Attention. In R. A. Atkinson, R. J. Herrnstein, G. Lindzey & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology Vol. 2, Learning and Cognition* (pp.739-811). New York, NY: Wiley.

Shin-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Science, 12*(5), 182–186.

Shiu, L. J. (2011). *EFL learners' perceptions of grammatical difficulty in relation to second language proficiency, performance, and knowledge*. (Doctoral dissertation) Retrieved from https://tspace.library.utoronto.ca/bitstream/1807/29869/1/Shiu_LiJu_201106_PhD_thesis.pdf [Accessed on 05/01/2014]

Shook, D. (1994). FL/L2 reading, grammatical information, and the input-to-intake phenomenon. *Applied Language Learning, 52*, 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.

Smith, B. (2010). Employing eye-tracking technology in researching the effectiveness of recasts in CMC. In F. M. Hult (Ed.), *Directions and prospects for educational linguistics* (pp. 79 – 97). New York, NY: Springer Verlag.

Smith, B. (2012). Eye-tracking as a measure of noticing: A study of explicit recasts in SCMC. *Language Learning and Technology, 16*(3), 53-81.

Smith, B., & Renaud, C. (2013). Eye-tracking as a measure of noticing corrective feedback in computer-mediated instructor-student foreign language conferences. In K. McDonough & A. Mackey (Eds.), *Interaction in diverse educational settings* (pp. 147-165). Philadelphia, PA: John Benajamins.

Smith, E. E., & Kosslyn, S. M. (2006). *Cognitive psychology: Mind and brain*. Upper Saddle River, NJ: Pearson/Prentice Hall.

Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist*, *27,* 5-32.

Soars, J. & Soars, L. (2014). *New headway upper intermediate* (4th ed.). Oxford, England: Oxford University Press.

Soto, D., & Silvanto, J. (2014). Reappraising the relationship between working memory and conscious awareness. *Trends in Cognitive Sciences, 18,* 520-525.

Spada, N., & Lightbown, P. M. (2008). Form-focused instruction: Isolated or integrated? *TESOL Quarterly, 42, 181-207.*

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60,* 263-308.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology, 59*(4), 745-59.

Styles, E. A. (2006). *The psychology of attention,* 2nd ed. Hove, England: Psychology Press.

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 House.

Swanson, H. L., Orosco, M. J. & Lussier, C. M. (2015). Growth in literacy, cognition, and working memory in English language learners. *Journal of Experimental Child Psychology, 132*, 155-188.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics*. Boston, MA: Allyn and Bacon.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston, MA: Pearson.

Tagarelli, K.M., Borges Mota, M., & Rebuschat, P. (2011). The role of working memory in implicit and explicit language learning. In L. Carlson, C. Holscher & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2016-2066). Austin, TX: Cognitive Science Society.

Tagarelli, K. M., Borges Mota, M., & Rebuschat, P. (2015). Working memory, learning conditions and the acquisition of L2 syntax. In Z. Wen, M. Borges & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 224-247). Bristol, England: Multilingual Matters.

Tamnes, C. K., Ostby, Y., Walhovd, K. B., Westlye, L.T., Due-Tonnessen, P., & Fjell, A. M. (2010). Intellectual abilities and white matter microstructure in development: A diffusion tensor imaging study. *Human Brain Mapping, 31*, 1609-1625.

Tamnes, C. K., Walhovd, K. B., Grydeland, H., Holland, D., Østby, Y., Dale, A.M., & Fjell, A. M., (2013). Longitudinal working memory development is related to structural maturation of frontal and parietal cortices. *Journal of Cognitive Neuroscience, 25* (10), 1611–1623.

Thornbury, S. (2005). *Uncovering grammar*. Oxford, England: Macmillan.

Tode, T. (2007). Durability problems with explicit instruction in an EFL context: the learning of the English copula be before and after the introduction of the auxiliary be. *Language Teaching Research, 11*(1), 11–30.

Tomlin, R., & Villa, H. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*, 183–203.

Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition, 15*, 181–203.

Treisman, A. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behaviour, 3* (6), 449–201.

Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory and analytical ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 171-196). Oxford, England: Oxford University Press.

Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research, 14* (2), 103–135.

Truscott, J. (2015). *Consciousness and second language learning*. Bristol, England: Multilingual Matters.

Truscott, J., & Sharwood Smith, M. (2011). Input, intake, and consciousness: the quest for a theoretical foundation. *Studies in Second Language Acquisition, 33*, 497–528.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285-318.

Tu, Y.K., Kellett, M., Clerehugh, V. & Gilthorpe, M.S. (2005). Problems of correlations between explanatory variables in multiple regression analyses in the dental literature. *British Dental Journal, 199*(7), 457.

Unsworth, N., & Engle, R. W. (2005). Working memory capacity and fluid abilities: Examining the correlation between operation span and raven. *Intelligence, 33,* 67– 81.

Uggen, M. S. (2012). Reinvestigating the noticing function of output. *Language Learning, 62*(2), 506-540.

Van der Linden, M., Brédart, S., & Beerten, A. (1994). Age-related differences in updating working memory. *British Journal of Psychology, 85*, 145–152.

VanPatten, B. (1994). Evaluating the role of consciousness in second language acquisition: Terms, linguistic features & research methodology. *AILA Review, 11*, 27-36.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.

VanPatten, B. (2002). Processing instruction: An update. *Language Learning, 52,* 755–803.

VanPatten, B. (2004). Input processing in SLA. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–32). Mahwah, NJ: Lawrence Erlbaum Associates.

VanPatten, B. (2011). Stubborn syntax: How it resists explicit learning and teaching. In C. Sanz & R. Leow (Eds.), *Proceedings of the 2009 Georgetown University Roundtable on Linguistics.* Washington, DC: Georgetown University Press.

VanPatten, B. & Rothman, J. (2015). What does current generative theory have to say about the explicit-implicit debate? In P. Rebuschat (Ed.) *Implicit and explicit learning of languages* (pp. 89-117). Amsterdam, The Netherlands: John Benjamins.

Vejnović, D., Milin, P. & Zdravković, S. (2010). Effects of proficiency and age of language acquisition on working memory performance in bilinguals. *Psihologija, 43*, 219-232.

von Restorff, H. (1933). Über die wirkung von bereichsbildungen im spurenfeld (The effects of field formation in the trace field). *Psychological Research, 18,* 299–342.

Vygotsky, L. S. (1962) *Thought and language.* Cambridge, MA: MIT Press.

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

Ward, G., & Allport, A. (1997). Planning and problem-solving using the 5-disc Tower of London task. *Quarterly Journal of Experimental Psychology, 50,* 49-78.

White, J. (1998). Getting the learner's attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 85–113). Cambridge, England: Cambridge University Press.

White, L., Spada, N., Lightbown, P., & Ranta, L. (1991). Input enhancement and L2 question formation. *Applied Linguistics, 12,* 416-432.

Wickens, C. D. (2007). Attention to attention and its applications: A concluding view. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Attention: From theory to practice.* Oxford, England: Oxford University Press.

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Personality Science and Individual Differences, 4,* 433-440.

Williams, J. (1999). Learner generated attention to form. *Language Learning, 49(4*), 583-625.

Williams, J. (2001). The effectiveness of spontaneous attention to form. *System, 29,* 325-340.

Williams, J. N. (2009). Implicit learning in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319–353). Bingley, England: Emerald Press.

Williams, J. (2013). Attention, Awareness, and Noticing in Language Processing and Learning. In J. M. Bergsleithner, S. Frota, & J. K.Yoshioka (Eds.). *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 39-61). Honolulu, Hawaii: University of Hawai‘i, National Foreign Language Resource Center.

Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139–155). Cambridge, England: Cambridge University Press.

Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning, 53*(1), 67-121.

Wilson, M. P., & Garnsey, S.M. (2009). Making simple sentences hard: Verb bias effects in simple direct object sentences. *Journal of Memory and Language, 60*(3), 368-392.

Winke, P.M. (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*(2), 323-352.

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*(1), 254–275.

Winke, P. & Lim, H. (2015). ESL essay raters' cognitive processes in applying the Jacobs et al. rubric: An eye-movement study. *Assessing Writing, 25*(2), 37-53.

Wolfe, J. M., & Horowitz, T.S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews, 5,* 1-7.

Wong, W. (2003). Textual enhancement and simplified input: Effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning, 13,* 109-132.

Yoshida, M. (2008). Think-Aloud protocols and type of reading task: the issue of reactivity in L2 reading research. In M. Bowles, R. Foote, S. Perpinan & R. Bhatt (Ed.), *Selected proceedings of the 2007 second language research forum* (pp. 199-209). Somerville, MA: Cascadilla Proceedings Project.

Yilmaz, Y., & Granena, G. (2015). The role of cognitive aptitudes for explicit language learning in the relative effects of explicit and implicit feedback. *Bilingualism, Language and Cognition, 19*(1), 147-161.

Zhang, R. H. (2014). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition, 37*(3). 457-486.

Zheng, D., Dong, X., Sun, H., Xu, Y., Ma, Y. & Wang, X. (2012). The overall impairment of core executive function components in patients with amnestic mild cognitive impairment: A cross-sectional study. *BMC Neurology*, *12*, 138-150.

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin, 31*, 3–29.

Appendices

Appendix A: Ethics documents – pilot study

Date: July 2013

INFORMATION SHEET

As part of my Doctoral studies in the Department of Linguistics and English Language, I am conducting a study that involves eye tracking data and short term memory test data.

I have approached you because I am interested in investigating how non-native speakers of English engage in reading English language texts. The data will be collected individually in two sessions. In the first session you will have to take a written test (about 20 minutes) and fill in a questionnaire. After that you will have to take four tests to measure your short term memory capacity (One test is paper based and the others are computer based). The short term memory tests will take approximately 40 minutes. After this first session, I will ask you to meet me again on a later date for the second session. In the second session you will have to do a reading activity on the computer that will track your eye movements while reading. After the reading activity, I will ask some questions from you and the conversation that we have will be recorded. Then you will have to take another paper based test on the same day. The second session will take about one hour.

You will be rewarded with a £10 Amazon voucher. Your participation will not affect your performance on your course and your relationship with the university.

I would be very grateful if you would agree to take part in my research.

You are free to withdraw from the study at any time but not later than one month after the second session. If you decide to withdraw within this time period, you data will be destroyed and not included in the study. If you decide to take part, at every stage, your name will remain confidential. The data will be kept securely and will be used for academic purposes only. The audio-recordings will be password protected so that no unauthorised person would be able to listen to them.

If you have any queries about the study, please feel free to contact myself at h.indrarathne@lancaster.ac.uk or my course supervisor, Dr. Judit Kormos, who can be contacted at j.kormos@lancaster.ac.uk or by phone on 01524 593039. If at any stage of the study you wish to speak to an independent person about this project, you are welcome to contact the Head of Department, Prof. Elena Semino, at e.semino@lancaster.ac.uk or by phone on 01524 594176.

Signed

Bimali Indrarathne

h.indrarathne@lancaster.ac.uk

UNIVERSITY OF LANCASTER

Department of Linguistics and English Language

Consent Form

Project title: **Learner generated noticing of written L2 input and its relationship with working memory capacity**

1. I have read and had explained to me by **Bimali Indrarathne** 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 described in the Information Sheet in so far as they relate to my participation.

3. I understand that my participation is entirely voluntary and that I have the right to withdraw from the project any time but not later than one month after the second session. If you decide to withdraw within this time period, you data will be destroyed and not included in the study.

4. I have received a copy of this Consent Form and of the accompanying Information Sheet.

Name:

Signed:

Date:

Appendix B: Background questionnaire – pilot study

Background Questionnaire

1. Number: ....
2. Age: ....
3. Gender: Male ☐ Female ☐
4. IELTS score: ....
5. Nationality: ....
6. First language: ....
7. When did you start learning English? (E.g. 2 years old)..
8. How many years have you studied English? ....
9. Are you a
   - ☐ monolingual?
   - ☐ bilingual?
   - ☐ trilingual?
10. Do you speak
    - ☐ Mandarin?
    - ☐ Cantonese?
    - ☐ both Mandarin and Cantonese?
11. What languages do you speak other than your first language? ....

Appendix C: Pilot study reading text

Pilot study reading text

I moved to a new house six years ago. It looked untidy when I went to see it. So I had painted the house by the time I moved into it. I had replaced the window panels as well. But, I had the roof repaired because I thought I could not do that myself. And also I had the furniture polished since I didn't want to ruin them.

Later, my girlfriend Sarah also moved into the house. By the time she arrived, I had replaced the curtains. Also I had the garden designed by a professional. She was quite happy with what I had done. She had the carpets replaced after her arrival.

One day I had my hair cut at a local salon. The hairdresser, Peter told me that he wanted someone to paint his house. He said he had bought paint as well. I said I could give it a try.

After I started painting his house I had to do many other things. Sarah also helped me in many things. We had the roof repaired because it was leaking. By the time the company that we hired finished doing the roof, I had removed the old wall papers. By the end of the first week Sarah had ordered new curtains. Peter had the materials delivered whenever we needed them.

By the end of the second week, I had painted the walls. We had the carpets changed because I didn't have the necessary tools. In two and a half weeks, I could finish everything and the house looked new and beautiful. Peter was so glad, he even paid me extra.

Then I realized that I could do this as a business. I started getting more offers, so Sarah and I formed a company of DYI Activities (Do It Yourself). When we moved to this new house, we had all the DIY activities done by our company.

Appendix D: Pilot study pre-test

Pilot study - Pre Test

Task 1

Rewrite the following sentences starting with the word/s given so that they mean exactly the same as the first sentence. Do not add apostrophe (‘) to the given word/words.

1. John speaks English much better than Amy.
   Amy ....
2. That beach was the most beautiful beach that I’d ever seen.
   I ....
3. I haven't received a call from Karen yet.
   Karen ....
4. Paul’s uncle is teaching him Spanish.
   Paul ....
5. People say that Mexico is a beautiful country.
   Mexico ....
6. My sister paid someone to paint her house.
   My sister had ....
7. It's too cold for a picnic today.
   Today ....
8. There isn’t any milk left.
   We ....
9. That’s the best meal I’ve ever eaten.
   I ....
10. She read more than 100 pages before going to bed.
    By the time she ....
11. The rain stopped the tennis match.
    The tennis match ....
12. What Joe really hates is driving at night.
    Joe ....
13. It may be hot, so take some sun-screen.
    Take ....
14. We got some people to landscape our garden for us last year.
    We had ....

15. "How about going to the cinema tonight?", said my dad.
My dad ....

16. Barry wasn't strong enough to lift the box.
Barry ....

17. I was made responsible for foreign sales.
They ....

18. Jenny started learning to drive two years ago.
Jenny ....

19. A mechanic repaired my car yesterday.
I had my ....

20. What a shame. There is no food left!
I ....

21. I regret not speaking to Kevin sooner.
I ....

22. We went to the party at 7 but Mary was there before us.
By the time ....

23. I think we will arrive at 7pm if the weather remains good.
We ....

24. This is the first time I have ever drunk rum.
I ....

25. A company framed our photographs for us.
We had ....

26. The bank closed before I arrived.
When ....

27. How long have you had your new car?
When ....

28. Leave the shop or I will call the police.
I ....

29. During the film on TV, the phone rang.
While ....

30. I saw him before he killed the girl.
When I ....

Task 2

Decide whether the following sentences are grammatically correct or incorrect. Put a ✓ in the appropriate column.

No.	Sentence	Correct	Incorrect
1	The farmer bought two pig at the market.		
2	That was the most amazing painting I had ever seen.		
3	A bat flied into our attic last night.		
4	I had my lunch delivered to my office.		
5	They forgot to turn in their papers.		
6	We have discussed the movie in class before watching it.		
7	The girl who answered the phone she was polite.		
8	She played cricket with the boys.		
9	The student who performs best in the competition will get the prize.		
10	By yesterday, I have received ten applications.		
11	We needed five mice for the experiment.		
12	She had her garden watered.		
13	When Kevin had moved to Japan Judi lived there for five years.		
14	She bought a set of knives at the store.		
15	We had our house painted last week.		
16	Anna and Pat have been married for more than 25 years.		
17	I asked for his book, but he did not lend me.		
18	He has got a new girlfriend who she works in a bar.		
19	Mary had finished her homework when we reached home.		
20	Did you had your shirt ironed?		
21	She left favorite earrings in the room.		
22	She bought five ink pens for school.		
23	She has been watching too much television lately.		
24	Are you going to be waiting for her when her plane arrives tonight?		
25	My dad has all the trees in our backyard sawed yesterday.		
26	She drank eight glasses of water.		
27	Colin has driven 100 miles when he reached home yesterday.		
28	He knew that she would be arriving late.		
29	The child flossed between all of her tooths before bedtime.		
30	I felt bad to sell my car because I had owned it for ten years.		
31	We had delivered pizza to our home last evening.		
32	I was so happy to receive a gift from my grandparents.		
33	Joe had cooked breakfast when we got up.		
34	You will be tired out after you have been working all night.		
35	The Jeffersons went the grocery shopping Sunday.		
36	He had all his money deposited in the bank while he was in hospital.		
37	If I had to be at the meeting, I would have done this earlier.		
38	I wash my clothes tomorrow.		
39	Tim has been worked as a teacher for over 25 years.		
40	I tried to interrupt into their conversation but was told on.		
41	The cashier told me that Sue had paid the bill.		
42	She had cut her hair by her sister.		
43	Who did you quit college because you hated?		
44	She sells her house in the auction.		
45	As Phil had a broken arm he had his secretary typed the		

	letter.		
46	Last year I worked for two days at a restaurant.		
47	I think that my friend had decided a name for the company.		
48	We knew that we'd be late for the meeting.		
49	Last year our office party we had organized by a professional.		
50	I'd be delighted to work with you.		

Appendix E: Pilot study post eye-tracking task

Pilot study post eye-tracking task

Now you are going to read the story again. Please read it and answer the questions given below the text.

I moved to a new house six years ago. It looked a bit untidy when I first went to see it. So, before I moved in, I had it painted. Before the holidays, I had replaced the window panels as well. I noticed some damp places in the house, so I had the roof repaired because I knew I couldn’t do that myself. I also had the chimney replaced.

My friend Sarah decided to move in with me and by the time she arrived, I had replaced the curtains. Also, I had redesigned the garden. Sarah was quite happy with what I had done. But she had the carpets replaced after her arrival.

One day I had my hair cut at a local salon. The hairdresser, Peter, told me that he wanted someone to paint his house. By then he had bought the paint as well. I said I would give it a try. But I soon realised there were other things that needed to be done.

I had removed the old wall paper, when the paint was delivered. We had the roof repaired because it was leaking. By the end of the first week, Sarah had ordered new curtains. Peter had the materials delivered whenever we needed them.

By the end of the second week, I had painted the walls. We had the carpets changed because I didn’t have the necessary tools. In two and a half weeks, I finished everything and the house looked new and beautiful. Peter was so pleased, he even paid me extra.

Then I realised that I could do this as a business. I started getting more job offers, so Sarah and I formed our own DYI (Do It Yourself) company. Two years later, when we bought a new house, we had all the DIY activities done by our company.

1. Who painted James’ house?

2. Who replaced the curtains in James’ house?

3. Who designed James’ garden?

4. What did Sarah do after moving to James' house?

5. What did the first thing that James do in Peter's house?

6. Who repaired the roof of Peter's house?

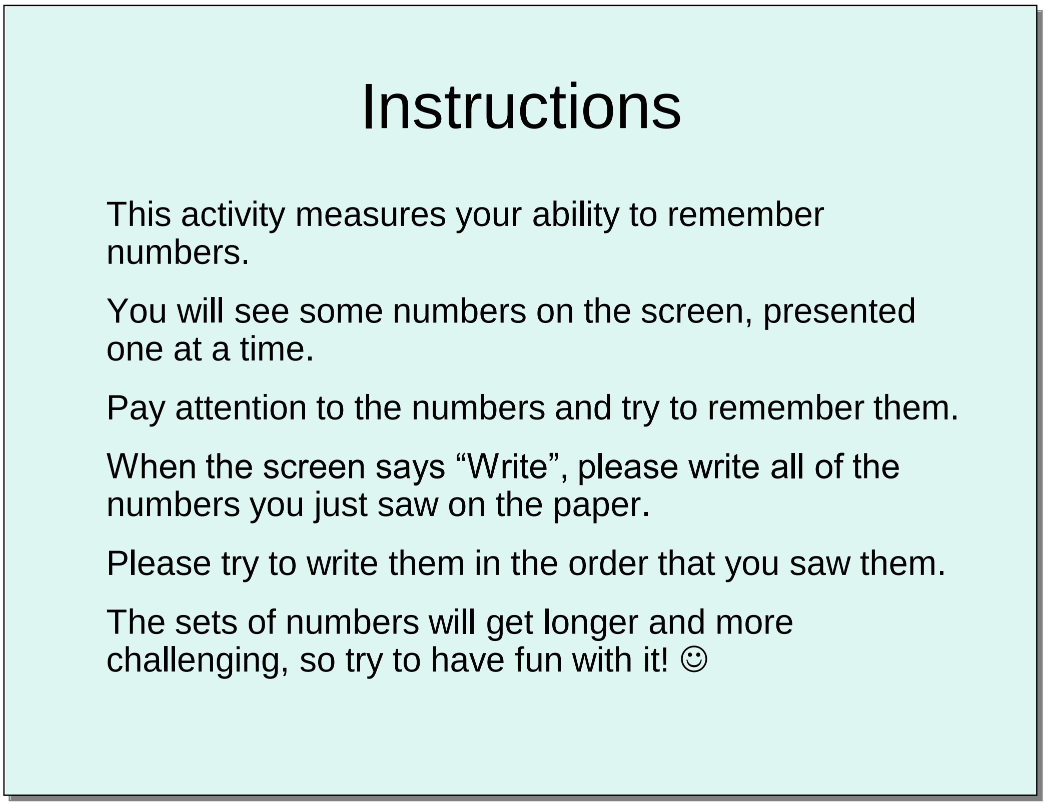
7. What did Sarah do to help James when he worked in Peter's house?

8. Who delivered materials to Peter's house?

9. How long did James take to finish repairing Peter's house?

10. Why did James start a DIY company?

Appendix F: Digit Span Test Slides



Slide 2

Digit Span Test

Instructions

This activity measures your ability to remember numbers.

You will see some numbers on the screen, presented one at a time.

Pay attention to the numbers and try to remember them.

When the screen says “Write”, please write all of the numbers you just saw on the paper.

Please try to write them in the order that you saw them.

The sets of numbers will get longer and more challenging, so try to have fun with it! ☺

Let’s practice!

4

7

5

Write

You should have written:
4 7 5

8

3

7

Write

You should have written:
8 3 7

Are you ready?

When you click, the test will begin.

3

8

6

Write

6

1

3

Write

3

5

1

7

Write

6

1

5

8

Write

beep!

8

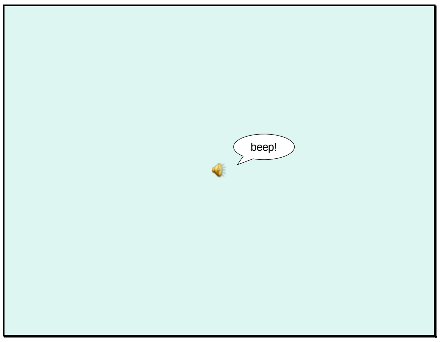
4

2

9

7

Write



5

2

1

8

6

Write

3

6

9

1

7

4

Write

7

9

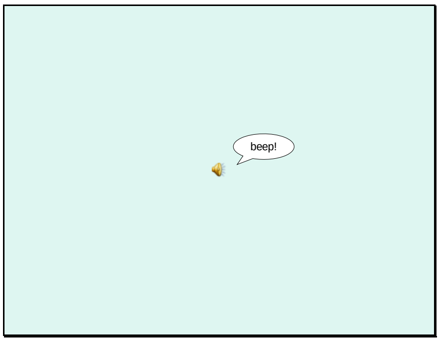
6

4

8

3

Write



5

1

7

4

2

6

8

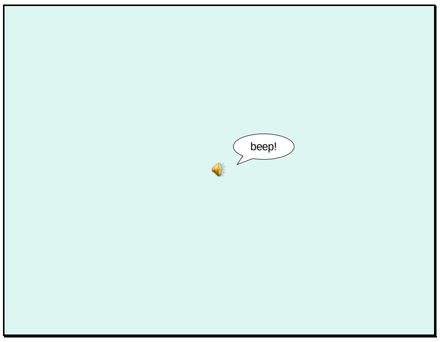
Write

beep!

9

8	5
2	1
6	3

Write



1

6

4

7

9

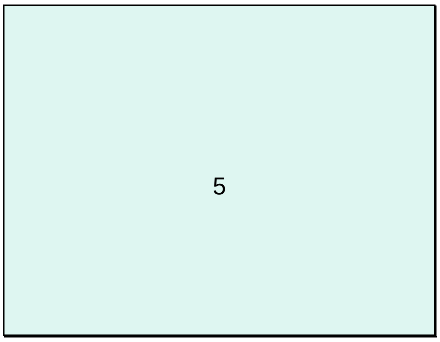
3

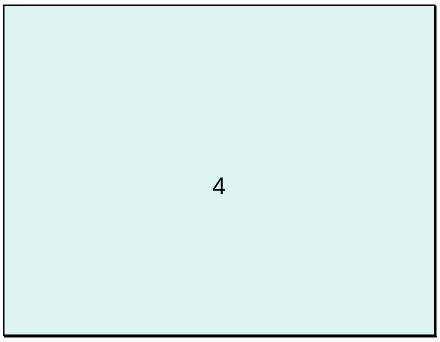
7

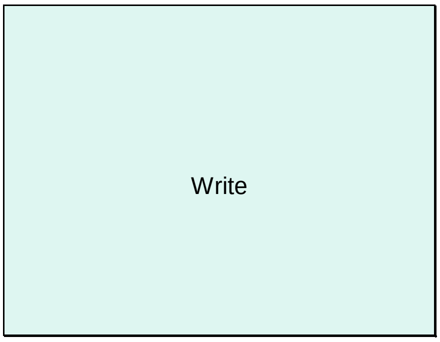
6

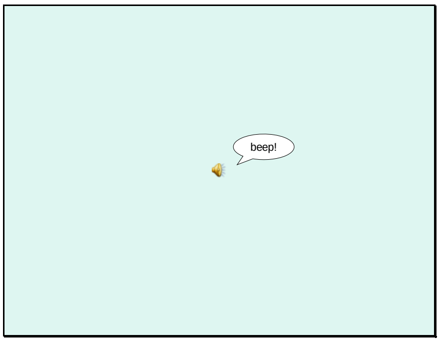
Write

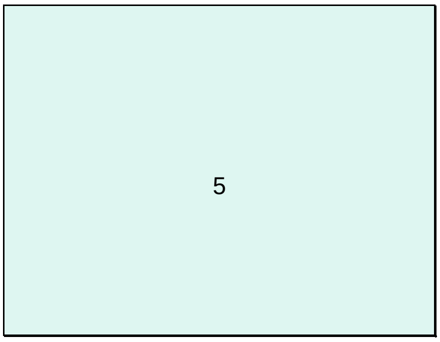
2	6
9	3
7	1

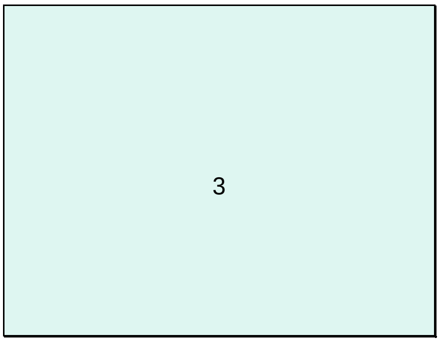

5


4


Write


beep!


5


3

8	7
7	4
1	6

9

Write

4

2

6

9	1
7	5
3	5

Write

9

5

3

9

7

2

3

7

6

5

Write

2

1

2

5

3	5
4	7
7	9

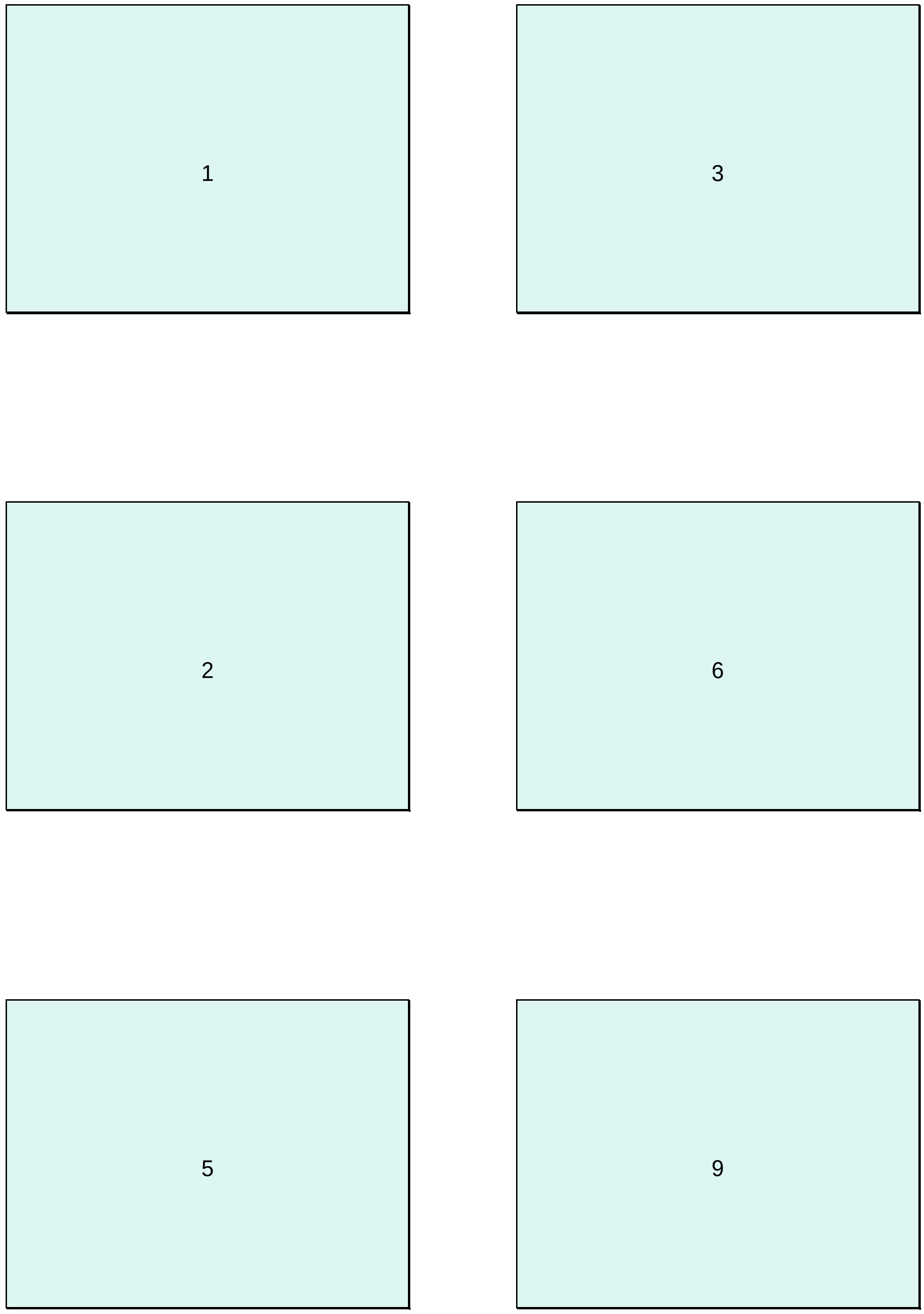
Write

6

4

7

6


1
3
2
6
5
9

8

Write

8

3

7

3	6
5	2
1	7

9

1

Write

☺ THE END ☺

Appendix G: Plus Minus Task

Plus-minus task

Instructions: Add 3 to each number in the first column and subtract 3 from each number in the second column. Add 3 to first number, subtract 3 from second number, add 3 to third number, subtract 3 from fourth number and complete the list in the third column following the same pattern.

Step 1

47 + 3 50

Column 1	
	Answer
11	
24	
74	
41	
61	
33	
57	
49	
70	
36	
85	
65	
16	
89	
52	
17	
82	
44	
93	
22	
63	
48	
72	
38	
97	
55	
78	
29	
80	
23	

Step 2

21 – 3 18

Column 2	
	Answer
30	
68	
12	
39	
60	
51	
25	
66	
42	
15	
77	
99	
73	
54	
87	
34	
92	
45	
59	
18	
94	
28	
79	
88	
21	
83	
47	
95	
32	
98	

Step 3

56 + 3	59
31 – 3	28
22 + 3	25
11 – 3	8

Column 3	
	Answer
37	
26	
50	
69	
13	
56	
75	
14	
84	
40	
81	
27	
62	
43	
31	
90	
67	
86	
35	
76	
96	
53	
91	
19	
58	
71	
46	
64	
20	
49	

Appendix H: Keep Track Task Slides

Slide 2

In this task, you will see some words that belong to the following categories

countries, metals, relatives,
colours, animals, distances

Slide 3

Study the following words that belong to the categories that you have just seen.

animals	colours	countries	distances	metals	relatives
horse	red	Canada	metre	gold	sister
cat	purple	Japan	centimetre	iron	daughter
cow	blue	Iran	mile	silver	uncle

- You will always see the categories at the bottom of the screen.
- You will see example words under some categories appearing on the screen. You have to remember the last word of each category. At the end, you have to write the last words of each category that appeared on the screen.

Slide 4

Keep-track task

In this task, you will see some words that belong to the following categories

countries, metals, relatives, colours, animals, distances

Study the following words that belong to the categories that you have just seen.

animals	colours	countries	distances	metals	relatives
horse	red	Canada	metre	gold	sister
cat	purple	Japan	centimetre	iron	daughter
cow	blue	Iran	mile	silver	uncle

- You will always see the categories at the bottom of the screen.
- You will see example words under some categories appearing on the screen. You have to remember the last word of each category. At the end, you have to write the last words of each category that appeared on the screen.

Let's practise.

Iran

countries

Canada countries	Japan countries
iron metals	gold metals
silver metals	uncle relatives

sister

relatives

daughter

relatives

Write.

You should have written:

Japan, silver, daughter

red

colours

blue

colours

purple

colours

horse

animals

cow

animals

cat

animals

iron

metals

silver

metals

gold

metals

Write.

You should have written:

purple, cat, gold

Are you ready?

When you click, the test will begin.

cat

animals

cow

animals

horse

animals

metre

distances

mile

distances

centimetre

distances

sister

relatives

uncle

relatives

daughter

relatives

blue

colours

purple

colours

red

colours

Write.

Japan

countries

Iran

countries

Canada

countries

horse

animals

cat

animals

cow

animals

iron

metals

silver

metals

gold

metals

mile

distances

centimetre

distances

metre

distances

Write.

purple

colours

blue

colours

red

colours

cat

animals

horse

animals

cow

animals

daughter

relatives

uncle

relatives

sister

relatives

silver

metals

iron

metals

gold

metals

Write.

Iran

countries

Canada

countries

Japan

countries

metre

distances

centimetre

distances

mile

distances

cat

animals

deer animals	horse animals
gold metals	silver metals
iron metals	blue colours

red

colours

purple

colours

Write.

cat

animals

cow

animals

horse

animals

purple

colours

blue

colours

red

colours

daughter

relatives

uncle

relatives

sister

relatives

Canada

countries

Japan

countries

Iran

countries

centimetre

distances

mile

distances

metre

distances

Write.

gold

metals

iron

metals

silver

metals

Japan countries	Iran countries
Canada countries	daughter relatives
sister relatives	uncle relatives

cow

animals

horse

animals

cat

animals

purple

colours

blue

colours

red

colours

Write.

☺ THE END ☺

Appendix I: Pilot study eye-tracking slides

You are going to read a story told by James. Please read it first. Then you will have to answer some questions about the story.

I moved to a new house six years ago. It looked a bit shoddy when I first went to see it. So, before I moved in, I had it painted. Before the holidays, I had replaced the window panels as well. A few shingles were also missing so I had the roof repaired because I knew I couldn’t do that myself. I also had the driveway paved and the chimney replaced.

My friend Sarah decided to move in with me and by the time she arrived, I had replaced the curtains. Also, I had the garden designed. Sarah was quite happy with what I had done. But she had the carpets replaced after her arrival.

One day I had my hair cut at a local salon. The hairdresser, Peter, told me that he wanted someone to paint his house. He said he had already bought the paint. I said I would give it a try. But I soon realised there were other things that needed to be done.

When the paint was delivered, I had already removed the old wall paper. We had the roof repaired because it was leaking. By the end of the first week, Sarah had ordered new curtains. Peter had the materials delivered whenever we needed them.

By the end of the second week, I had painted the walls. We had the carpets changed because I didn't have the necessary tools. In two and a half weeks, I finished everything and the house looked new and beautiful. Peter was so pleased, he even paid me extra.

Then I realised that I could do this as a business. I started getting more job offers, so Sarah and I formed our own DYI (Do It Yourself) company. Two years later, when we bought a new house, we had all the DIY activities done by our company.

Appendix J: Ethics documents – main study

Date: July 2014

INFORMATION SHEET

As part of my Doctoral studies in the Department of Linguistics and English Language at Lancaster University, I am conducting a study that involves eye tracking data and short term memory test data.

I have approached you because I am interested in investigating how non-native speakers of English engage in reading English language texts. The data will be collected individually in six sessions. In the first session you will have to take a test (about 40 minutes) and fill in two questionnaires. After this first session, I will ask you to meet me again on later dates for the second, third and fourth sessions. In these three sessions you will have to do a reading activity on the computer that will track your eye movements while reading and answer some questions based on the texts. Each of these sessions will take approximately 15 minutes. In the fifth session you will have to take another test similar to the one that you took in session 1 (40 minutes) and respond to some questions orally. Your responses will be audio recorded. In the final session, you will have to take four working memory tests that measure short term memory capacity. The short term memory tests will take approximately 40 minutes.

Your participation will not affect your performance on your course and your relationship with the university. However, if you participate in all sessions and complete all tasks, you will get additional 5 marks for the second semester end English language paper (How you perform in the experiment will not be counted – E.g. whether your answers for the questions in the test are correct or incorrect will not be considered).

I would be very grateful if you would agree to take part in my research.

You are free to withdraw from the study at any time but not later than one month after the experiment. If you decide to withdraw within this time period, your data will be destroyed and not included in the study. If you decide to take part, at every stage, your name will remain confidential. The data will be kept securely and will be used for academic purposes only. The audio-recordings will be password protected so that no unauthorised person would be able to listen to them.

If you have any queries about the study, please feel free to contact myself at h.indrarathne@lancaster.ac.uk or my thesis supervisor, Dr. Judit Kormos, who can be contacted at j.kormos@lancaster.ac.uk or by phone on 01524 593039. If at any stage of the study you wish to speak to an independent person about this project, you are welcome to contact the Head of Department, Prof. Elena Semino, at e.semino@lancaster.ac.uk or by phone on 01524 594176.

Signed

Bimali Indrarathne

h.indrarathne@lancaster.ac.uk

UNIVERSITY OF LANCASTER

Department of Linguistics and English Language

Consent Form

Project title: **Learner generated noticing of written L2 input and its relationship with working memory capacity**

1. I have read and had explained to me by **……………………………………** 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 described in the Information Sheet in so far as they relate to my participation.

3. I understand that my participation is entirely voluntary and that I have the right to withdraw from the project any time but not later than one month after the experiment.

4. I have received a copy of this Consent Form and of the accompanying Information Sheet.

Name:

Signed:

Date:

Appendix K: Background questionnaire – main study

Background Questionnaire

1. Number: ....

2. Age: ....

3. Gender: ☐ Male ☐ Female

4. Class: ....

5. Stream of study: ....

6. Nationality: ....

7. First language: ....

8. When did you start learning English? (E.g. 2 years old).....................................

9. How many years have you studied English? ....

10. Are you a

    ☐ monolingual?

    ☐ bilingual?

    ☐ trilingual?

    If bilingual or trilingual, what are the languages you speak? ....

11. What languages do you speak other than your first language? ....

12. Did you have difficulties learning how to read and write in your first language? ....

    If yes, please explain the kind of difficulties that you had.

Appendix L: Three reading texts

Three reading texts

Text 1

New house

James moved to a new house six years ago. It looked a bit untidy when he first went to see it. So, before moving in, he **had the walls painted**. He noticed water on the floor in some places, so he **had the roof repaired** because he knew he couldn't do that himself. He also **had the curtains replaced**. Later, his friend Sarah decided to move in. She didn't like the garden very much. So, they **had the garden cleaned**.

One day, Peter, their neighbour told James that he wanted someone to paint his house. James said he could give it a try. But he soon realized that there were other things to be done in the house. He made a list of tools that he needed and Peter **had the tools delivered** the same week. By the end of the second week, James finished painting the walls and repairing the windows. They **had the carpets changed** because James didn't have time to do that.

In two and half weeks, James finished everything he could do except repairing the back door. Later, Peter **had a door made** and he also changed some windows. The house then looked new and beautiful. Peter was so pleased, he even paid James extra. James started getting more job offers, so later Sarah and he formed their own company that took orders for house repairing.

Text 2

Mary's aunt's shopping

Mary's sister's wedding was last Saturday and her aunt wanted to go shopping before that. She couldn't walk properly, so, Mary took her to town. First, they went to the clinic. Mary's aunt **had her blood tested** there and **had the reports sent** to her doctor. Then they had to wait one hour to see the doctor.

After that her aunt **had her hair cut** for two hours. While having the hair cut, she told Mary that she wanted to invite her best friend to the wedding. So she **had a letter written** to her friend in the salon. Then she reminded Mary of her lost passport. So after the hair cut Mary's aunt **had her photograph taken** for a new passport. Next, they went to a coffee shop because both of them were hungry. In the coffee shop, Mary's aunt **had sugar added** to her coffee although the doctor warned her in the morning to be careful with sweets.

Mary’s aunt actually didn’t do any shopping, but wanted to walk in town. When she was ready to go back home it was dark and it was time for the last bus to leave. Mary’s aunt couldn’t walk fast, so Mary **had the bus stopped** at the bus station until they reached it. They went home quite late, but her aunt was happy about her day out.

<u>Text 3</u>

<u>Joe’s interview</u>

Joe was not ready to face his first job interview as a photographer. The day before the interview he told his brother James that he didn’t have a car to go there. James had a friend who got a car. So he **had that car left** at a car park near the train station for Joe to pick up in the morning. He told Joe not to worry because he **had the car checked** so that it would not break down on this important day.

Then Joe checked his camera because he had to take it to the interview. But, it didn’t work properly and he didn’t know how to repair it. So he called a friend and **had the camera fixed.** Then only he remembered that he had to print a letter from his former employer. But he had so many other things to do, so James walked to the town and **had the letter printed.**

Joe was also asked to bring some of the photographs that he had taken. Luckily, he **had his photographs developed** at a local store the day before. The next morning, Joe was late to wake up so he **had his breakfast put** into his bag. Suddenly he realized that his shoes were worn out so he **had his shoes mended** on the way. In the end he got the job. After this, he was never late for anything.

Appendix M: Main study pre-test

Pre-test

<u>Task 1: 20 minutes</u>

Reconstruct the following sentences starting with the word or words given so that the new sentence means exactly the same as the first sentence.

Example 1

My friends have painted my new house.
New sentence: My new house has been painted by my friends.

Example 2

Jane had finished cleaning the house before her sister came home.
New sentence: By the time Jane's sister came home, Jane had finished cleaning the house.

Example 3

Our teacher made Camilla the head of the team.
New sentence: Camilla was made the head of the team by our teacher.

1. John speaks English much better than Amy.
   Amy ....

2. Sera got someone to print invitation cards for her party.
   Sera had ....

3. I haven't received a call from Karen yet.
   Karen has ....

4. A new teacher is teaching Spanish to Paul.
   Paul is ....

5. Mary asked Simon to put the lights on at home.
   Mary had the ....

6. It's too cold for a picnic today.
   Today is a ....

7. It may be hot, so take some sun-screen.
Take ....

8. My parents paid someone to paint our house last year.
My parents had ……..

9. She read more than 100 pages before going to bed.
By the time she ....

10. Jane got Fred to clean the office after the meeting.
Jane had the ....

11. How long have you had your new car?
When did ....

12. He paid someone to repair my car before the trip.
He had ………..

13. What Joe really hates is driving at night.
Joe really ....

14. Leave the shop or I will call the police.
I ....

15. We hired Williams to make a new window last week.
We had a ....

16. Jenny started learning to drive two years ago.
Jenny has ....

17. During the film on TV, the phone rang.
While ....

18. I regret not speaking to Kevin sooner.
I ....

19. I saw him before he killed the girl.
When I ....

20. I was made responsible for foreign sales.
They made ....

Task 2

Decide whether the sentences you hear are grammatically correct or incorrect. Put a ✓ for correct sentences and a **X** for incorrect sentences in your answer sheet.

Listen to the example sentence.

She has gone to a party last night.
Now tick the appropriate boxes on the answer sheet.
The answer is: incorrect

Listen to another example sentence.

I read ten book last year.
Now tick the appropriate boxes on the answer sheet.
The answer is: incorrect

Listen to the last example sentence.

Did you have a good time at the party yesterday?
Now tick the appropriate boxes on the answer sheet.
The answer is: correct

Let's start the test.

No.	**Sentence**
1	The farmer bought two pig at the market.
2	That was the most amazing painting I had ever seen.
3	A bat flied into our attic last night.
4	My dad had his lunch delivered to his office yesterday.
5	They forgot to turn in their papers.
6	We have discussed the movie in class before watching it.
7	When I was eating yesterday, I had added more salt to the plate.
8	Phil was busy yesterday, so he had the documents typing.
9	The student who performs best in the competition will get the prize.
10	By yesterday, I have received ten applications.
11	We needed five mice for the experiment.
12	She had a taxi stopped for her at the gate.
13	When Kevin had moved to Japan Judi lived there for five years.
14	She bought a set of knives at the store.
15	My sister had her new house decorate before the weekend.
16	Anna and Pat have been married for more than 25 years.
17	I asked for his book, but he did not lend me.
18	Tom had repair his leather bag by a nearby shop.

19	Mary had finished her homework when we reached home.
20	Did you had your shirt ironed?
21	She left favourite earrings in the room.
22	Janet had her photo taken for a magazine last year.
23	She has been watching too much television lately.
24	Are you going to be waiting for her when her plane arrives tonight?
25	My dad had our family history written by his secretary.
26	She played cricket with the boys.
27	Colin has driven 100 miles when he reached home yesterday.
28	He knew that she would be arriving late.
29	The child flossed between all of her tooths before bedtime.
30	I felt bad to sell my car because I had owned it for ten years.
31	My mom had a pizza deliver to our home last evening.
32	I was so happy to receive a gift from my grandparents.
33	Joe had cooked breakfast when we got up.
34	You will be tired out after you have been working all night.
35	Sean had an x-ray taken in the clinic by doctors.
36	The Jeffersons went the grocery shopping Sunday.
37	If I had to be at the meeting, I would have done this earlier.
38	She had cut some papers for a project by her sister.
39	Tim has been worked as a teacher for over 25 years.
40	I'd be delighted to work with you.

Task 3: 10 minutes

Suppose you were a senior scout and you had to take 5 junior scouts to India for a camp. You had to plan everything for the trip. There were things that you could do yourself, but there were other things that you had to have done for you. Write a paragraph about how you planned the trip including:

- the things that you did yourself
- you had others done for you

Appendix N: Pre-test task 2 answer sheet

Pre-Test Task 2 = Answer sheet

	Question No.	✓ or **X**
	Example 1	
	Example 2	
	Example 3	
	1	
	2	
	3	
	4	
	5	
	6	
	7	
	8	
	9	
	10	
	11	
	12	
	13	
	14	
	15	
	16	
	17	
	18	
	19	
	20	
	21	
	22	
	23	
	24	
	25	
	26	
	27	
	28	
	29	
	30	
	31	
	32	
	33	
	34	
	35	
	36	
	37	
	38	
	39	
	40	

Appendix O: Main study post-test

Post-Test

Task 1: 20 minutes

Reconstruct the following sentences starting with the word or words given so that the new sentence means exactly the same as the first sentence.

Example 1

My friends have painted my new house.
New sentence: My new house has been painted by my friends.

Example 2

Jane had finished cleaning the house before her sister came home.
New sentence: By the time Jane's sister came home, Jane had finished cleaning the house.

Example 3

Our teacher made Camilla the head of the team.
New sentence: Camilla was made the head of the team by our teacher.

Let's start the test.

1. Andy speaks Chinese much better than Helen.
   Helen ....

2. Ann got someone to print her certificates for the interview.
   Sera had ....

3. They haven't received a message from their son yet.
   Their son ....

4. Peter's sister is teaching him tennis.
   Paul ....

5. Emily asked Jane to put the lights on at home.
   Mary had ....

6. It's too hot for us to play outside today.
   Today ....

7. It may be cold, take some warm clothes with you.
   Take ....

8. Our neighbours paid someone to paint their house last week.
   They had ....

9. Chris read more than three guide books before going China.
   By the time he ....

10. Harris got John to clean the house after the party.
    Harris had the ....

11. How long has she had your new umbrella?
    When ....

12. Someone repaired my sister's van last week after the accident.
    My sister had ....

13. What I really like is swimming in the sea.
    I ....

14. Leave the house I will call the police.
    I ....

15. My parents hired Edward to make a new book shelf.
    My parents had ....

16. I started learning to ride horses one year ago.
    I ....

17. The doorbell rang during the meeting.
    While ....

18. Jessica regrets not helping Charlotte when she was ill.
    Jessica ....

19. We saw Garry before he hit the boy.
    When we ....

20. Diane was made responsible for decorations.
    They ....

Task 2

Decide whether the sentences you hear are grammatically correct or incorrect. Put a ✓ for correct sentences and a **X** for incorrect sentences in your answer sheet.

Listen to the example sentence.

She has gone to a party last night.
Now tick the appropriate boxes on the answer sheet.
The answer is: incorrect

Listen to another example sentence.

I read ten book last year.
Now tick the appropriate boxes on the answer sheet.
The answer is: incorrect

Listen to the last example sentence.

Did you have a good time at the party yesterday?
Now tick the appropriate boxes on the answer sheet.
The answer is: correct

Let's start the test.

No.	**Sentence**
1	Janet bought two egg at the shop.
2	That was the most beautiful park I had ever seen.
3	A crow flied into our living room yesterday.
4	My sister had her new jeans delivered to her office.
5	Fiona's class forgot to turn in their assignments.
6	The students have discussed the drama in class before watching it.
7	When he was eating lunch, he has put more salt to the dish.
8	Fred was ill, so he had the letter typing yesterday.
9	The dancer who performs best in the last round will get the prize.
10	By last Monday, we have received ten requests.
11	They needed ten chicken for the experiment.
12	He had a car stopped for him at the entrance.
13	When I had moved to China, my brother lived there for five years.
14	I bought a piece of cheese at the store.
15	My uncle had his old car clean day before yesterday.
16	My parents have been married for more than 30 years.
17	Jen asked for Pat's car, but she did not lend her.
18	Phil had repair his shoes by a shoe shop yesterday.

19	Jessie had done her cooking when her husband reached home.
20	Did she had her letter sent?
21	He left favourite tie in the office.
22	Clara had her daughter's picture taken for a cover page.
23	My sister has been watching too much football lately.
24	Is John going to be waiting for Elena when her flight arrives tomorrow?
25	The Symonds had their company history written by a professional.
26	Sally played tennis with her friends last evening.
27	Pauline has travelled 400 miles when she reached home last week.
28	I knew that my friends would be arriving early.
29	Andy flossed between all of his tooths before going to bed.
30	My dad felt bad to sell his car because he owned it for ten years.
31	I had two cakes deliver to our home last Monday.
32	Sally was so happy to get a gift from me.
33	My mom had cooked before we got up.
34	She will be tired out after she has been working for so long.
35	My grandma had her blood taken for a test yesterday.
36	My neighbours went the grocery shopping Saturday.
37	If I had to go to the cafe, I would have left earlier.
38	James had cut some cardboard by his sister for school.
39	Nancy has been worked as a nurse for 30 years.
40	She'd be happy to work with her new colleagues.

Task 3: 10 minutes

Suppose you were a senior scout and you had to take 5 junior scouts to India for a camp. You had to plan everything for the trip. There were things that you could do yourself, but there were other things that you had to have done for you. Write a paragraph about how you planned the trip including:

- the things that you did yourself
- you had others done for you

Post-Test Task 2 = Answer sheet

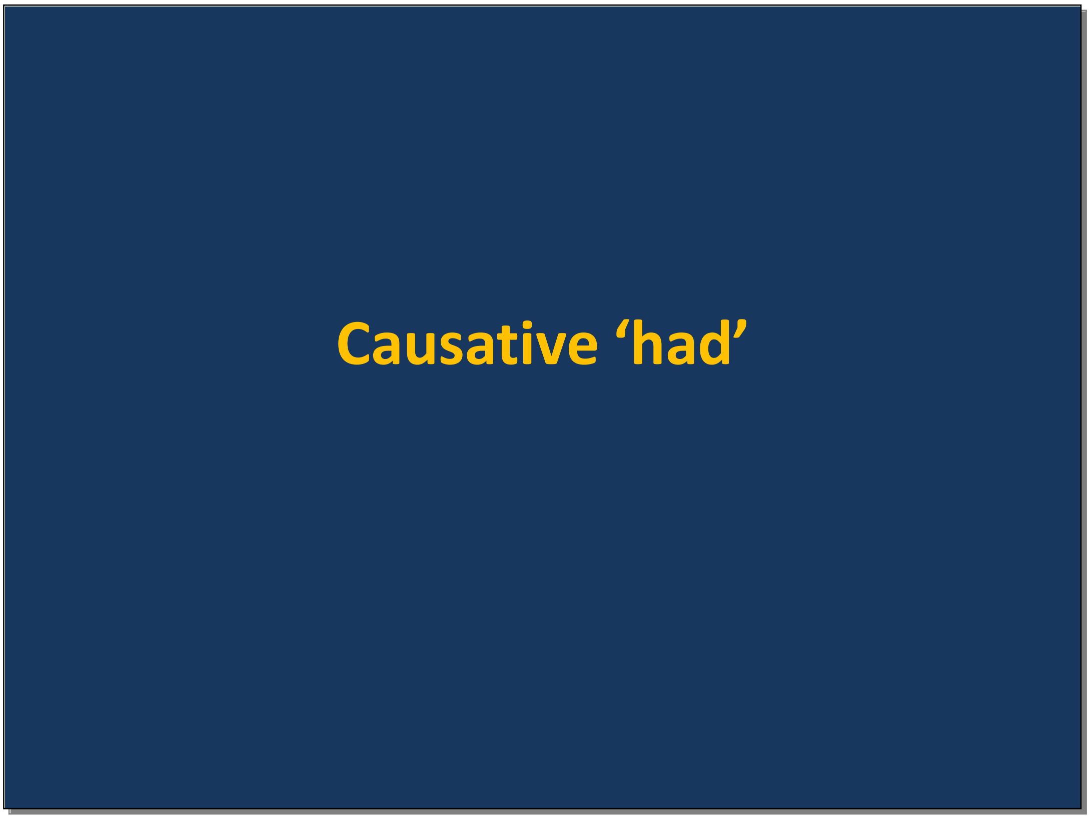
	Question No.	✓ or **X**
	Example 1	
	Example 2	
	Example 3	
	1	
	2	
	3	
	4	
	5	
	6	
	7	
	8	
	9	
	10	
	11	
	12	
	13	
	14	
	15	
	16	
	17	
	18	
	19	
	20	
	21	
	22	
	23	
	24	
	25	
	26	
	27	
	28	
	29	
	30	
	31	
	32	
	33	
	34	
	35	
	36	
	37	
	38	
	39	
	40	

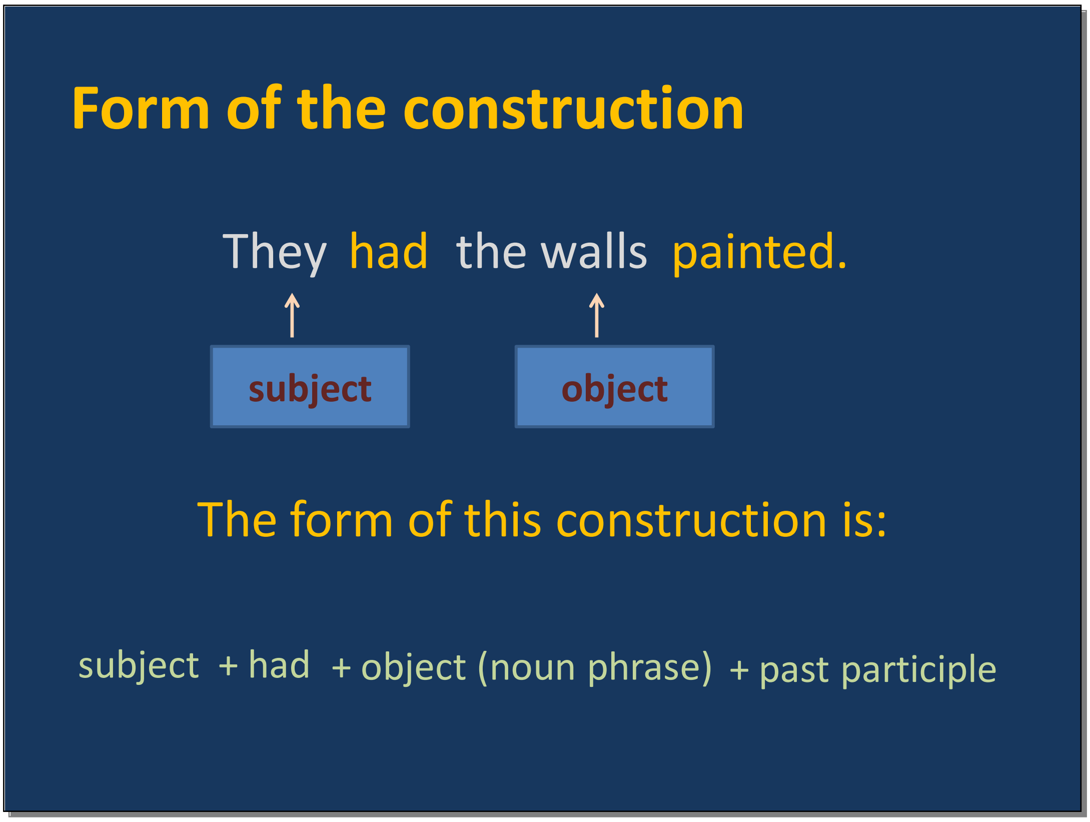
Appendix P: Post eye-tracking interview questions

Post eye-tracking interview

1. Were the reading tasks difficult for you?
2. If your answer is ‘yes’, why did you find them difficult?
3. Which text took longer to read? (1, 2 or 3)
4. Why did it take longer?
5. Did you find it difficult to read on computer than on paper? Please explain your answer.
6. What did you notice in the texts?
7. Did you notice any particular grammar structure in the texts?
8. If your answer is ‘yes’ can you please write down that structure?

Appendix Q: Slides of the PowerPoint presentation shown to Group B





Form of the construction: more examples

subject	had	object (noun phrase)	past participle
They	had	the walls	painted.
He		the roof	repaired.
Peter		the tools	delivered.
Peter		a door	made.

Meaning of the construction

They had the walls painted.

This has a passive meaning. *They* arranged the walls to be painted by someone. According to this sentence we don't know who painted the walls.

Lets' look at another example.

James had the roof repaired.

According to this sentence, James arranged someone to repair the roof. We don't know who repaired it.

James had the roof repaired by his friend.

It is possible to add the 'doer' to this sentence. According to this sentence, the doer or the person who painted the walls is 'his friend'. However, this addition is optional in the construction.

Appendix R: Reading text 1 – eye-tracking slides (enhanced)

<u>New house</u>

James moved to a new house six years ago. It looked a bit untidy when he first went to see it. So before moving in he **had the walls painted**.

He noticed water on the floor in some places, so he **had the roof repaired** because he knew he couldn't do that himself.

He also **had the windows replaced**.

Later his friend Sarah decided to move in.

She didn't like the garden very much.

So they **had the garden cleaned**. One day, Peter, their neighbour told James that he wanted someone to paint his house.

James said he could give it a try. But he soon realized that there were other things to be done in the house. He made a list of tools that he needed and Peter **had the tools delivered** the same week.

By the end of the second week, James finished painting the walls and repairing the windows. They **had the curtains changed** because James didn't have time to do that.

In two and half weeks, James finished everything he could do except repairing the back door. Later, Peter **had a door made** and he also changed some windows. The house then looked new and beautiful.

Peter was so pleased, he even paid James extra. James started getting more job offers, so later Sarah and he formed their own company that took orders for house repairing.

When did Sarah move in?

a. Before James

b. After James

c. With James

Who delivered tools to Peter's house?

a. Peter

b. James

c. Someone else

Who made a door for Peter?

a. James

b. Sarah

c. Someone else

Why did James and Sarah form a company?

a. because they didn't have jobs

b. because they liked repairing houses

c. because they got more job offers

Appendix S: Reading text 1 – eye-tracking slides (unenhanced)

New house

James moved to a new house six years

ago. It looked a bit untidy when he first

went to see it. So before moving in

he had the walls painted.

He noticed water on the floor in some places, so he had the roof repaired because he knew he couldn’t do that himself.

He also had the windows replaced.

Later, his friend Sarah decided to move in.

She didn’t like the garden very much.

So they had the garden cleaned. One day, Peter, their neighbour told James that he wanted someone to paint his house.

James said he could give it a try. But he soon realized that there were other things to be done in the house. He made a list of tools that he needed and Peter had the tools delivered the same week.

By the end of the second week, James finished painting the walls and repairing the windows. They had the curtains changed because James didn't have time to do that.

In two and half weeks, James finished everything he could do except repairing the back door. Later, Peter had a door made and he also changed some windows. The house then looked new and beautiful.

Peter was so pleased, he even paid James extra. James started getting more job offers, so later Sarah and he formed their own company that took orders for house repairing.

When did Sarah move in?

a. Before James

b. After James

c. With James

Who delivered tools to Peter's house?

a. Peter

b. James

c. Someone else

Who made a door for Peter?

a. James

b. Sarah

c. Someone else

Why did James and Sarah form a company?

a. because they didn't have jobs

b. because they liked repairing houses

c. because they got more job offers

Appendix T: Reading text 2 – eye-tracking slides (enhanced)

<u>Mary’s aunt’s shopping</u>

Mary’s sister’s wedding was last Saturday

and her aunt wanted to go shopping before

that. She couldn’t walk properly, so Mary

took her to town. First, they went to the

clinic.

Mary’s aunt **had her blood tested** there

and **had the reports sent** to her doctor.

Then they had to wait one hour to see the

doctor. After that her aunt **had her hair cut**

for two hours.

While having the hair cut, she told Mary

that she wanted to invite her best friend

to the wedding.

So she **had a letter written** to her friend in

the salon.

Then she reminded Mary of her lost passport. So after the hair cut Mary’s aunt **had her photograph taken** for a new passport. Next, they went to a coffee shop because both of them were hungry.

In the coffee shop,

Mary’s aunt **had sugar added** to her coffee although the doctor warned her in the morning to be careful with sweets.

Mary's aunt actually didn't do any shopping, but wanted to walk in town. When she was ready to go back home it was dark and it was time for the last bus to leave.

Mary's aunt couldn't walk fast, so Mary **had the bus stopped** at the bus station until they reached it. They went home quite late, but her aunt in the end was happy about her day out.

Why did Mary’s aunt want to go shopping?

a. because of Mary’s sister’s wedding
b. because she had to go to the clinic
c. because she had to buy some clothes

Who wrote a letter in the salon?

a. Mary’s aunt
b. Mary
c. Someone else

Who stopped the bus?

a. Mary

b. The driver

c. Someone else

Was Mary's aunt careful of her health?

a. Yes

b. no

Appendix U: Reading text 2 – eye-tracking slides (unenhanced)

<u>Mary’s aunt’s shopping</u>

Mary’s sister’s wedding was last Saturday

and her aunt wanted to go shopping before

that. She couldn’t walk properly, so Mary

took her to town. First, they went to the

clinic.

Mary’s aunt had her blood tested there and had the reports sent to her doctor. Then they had to wait one hour to see the doctor. After that her aunt had her hair cut for two hours.

While having the hair cut, she told Mary that she wanted to invite her best friend to the wedding.

So she had a letter written to her friend in the salon.

Then she reminded Mary of her lost passport. So after the hair cut

Mary’s aunt had her photograph taken for a new passport. Next, they went to a coffee shop because both of them were hungry.

In the coffee shop,

Mary’s aunt had sugar added to her coffee although the doctor warned her in the morning to be careful with sweets.

Mary’s aunt actually didn’t do any shopping, but wanted to walk in town. When she was ready to go back home it was dark and it was time for the last bus to leave.

Mary’s aunt couldn’t walk fast, so Mary had the bus stopped at the bus station until they reached it. They went home quite late, but her aunt in the end was happy about her day out.

Why did Mary's aunt want to go shopping?

a. because of Mary's sister's wedding

b. because she had to go to the clinic

c. because she had to buy some clothes

Who wrote a letter in the salon?

a. Mary's aunt

b. Mary

c. Someone else

Who stopped the bus?

a. Mary

b. The driver

c. Someone else

Was Mary's aunt careful of her health?

a. Yes

b. no

Appendix V: Reading text 3 – eye-tracking slides (enhanced)

<u>Joe's interview</u>

Joe was not ready to face his first job interview as a photographer. The day before the interview he told his brother James that he didn't have a car to go there. James had a friend who got a car.

So he **had that car left** at a car park near the train station for Joe to pick up in the morning. He told Joe not to worry because he **had the car checked** so that it would not break down on this important day.

Then Joe checked his camera because he had to take it to the interview. But it didn't work properly and he didn't know how to repair it. So he called a friend and **had the camera fixed.**

Then only he remembered that he had to print a letter from his former employer. But he had so many other things to do, so James walked to the town and **had the letter printed.**

Joe was also asked to bring some of the photographs that he had taken. Luckily, he **had his photographs developed** at a local store the day before.

The next morning, Joe was late to wake up so he **had his breakfast put** into his bag. Suddenly he realized that his shoes were worn out so he **had his shoes mended** on the way.

In the end he got the job. After this, he was never late for anything.

Whose car did Joe take to the interview?

a. His own car

b. His brother’s car

c. His brother’s friend’s car

What did Joe have to take to the interview?

a. a camera only

b. a camera, photographs and a letter

c. application form and a letter

Why did James and Sarah form a company?

a. because they didn't have jobs
b. because they liked repairing houses
c. because they got more job offers

Who mended Joe's shoes in the morning?

a. His brother
b. Joe himself
c. Someone else

Appendix W: Reading text 3 – eye-tracking slides (unenhanced)

<u>Joe’s interview</u>

Joe was not ready to face his first job interview as a photographer. The day before the interview he told his brother James that he didn’t have a car to go there. James had a friend who got a car.

So he had that car left at a car park near the train station for Joe to pick up in the morning. He told Joe not to worry because he had the car checked so that it would not break down on this important day.

Then Joe checked his camera because he had to take it to the interview. But it didn't work properly and he didn't know how to repair it. So he called a friend and had the camera fixed.

Then only he remembered that he had to

print a letter from his former employer. But

he had so many other things to do, so

James walked to the town and

had the letter printed.

Joe was also asked to bring some of the

photographs that he had taken. Luckily

he had his photographs developed at a

local store the day before.

The next morning, Joe was late to wake up so he had his breakfast put into his bag. Suddenly he realized that his shoes were worn out so he had his shoes mended on the way.

In the end he got the job. After this, he was never late for anything.

Whose car did Joe take to the interview?

a. His own car

b. His brother’s car

c. His brother’s friend’s car

Where was the car park?

a. near a train station

b. near a bus station

c. near Joe’s house

What did Joe have to take to the interview?

a. a camera only

b. a camera, photographs and a letter

c. application form and a letter

Who put the lunch in Joe's bag?

a. Joe

b. His brother

c. Someone else

Appendix X: Graphs of the descriptive statistics

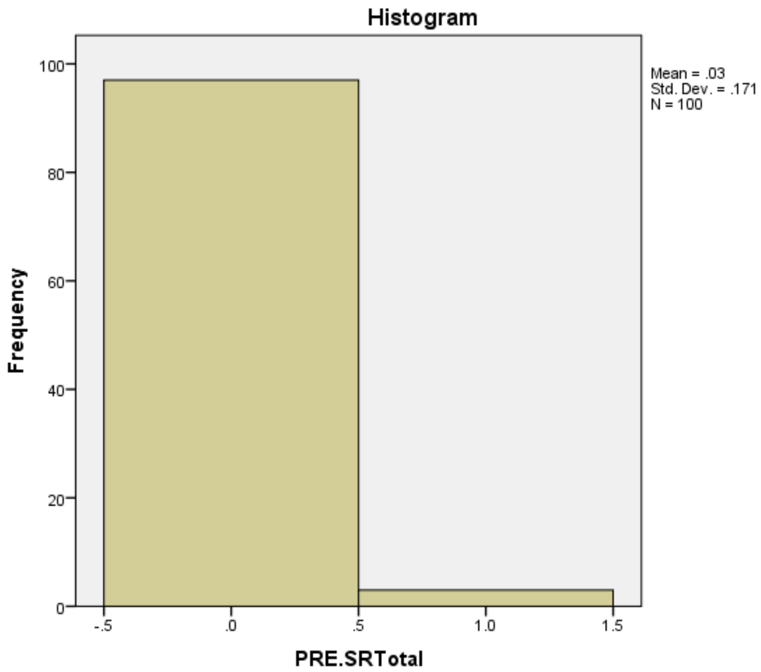


Figure X1: Histogram of pre-test SR items

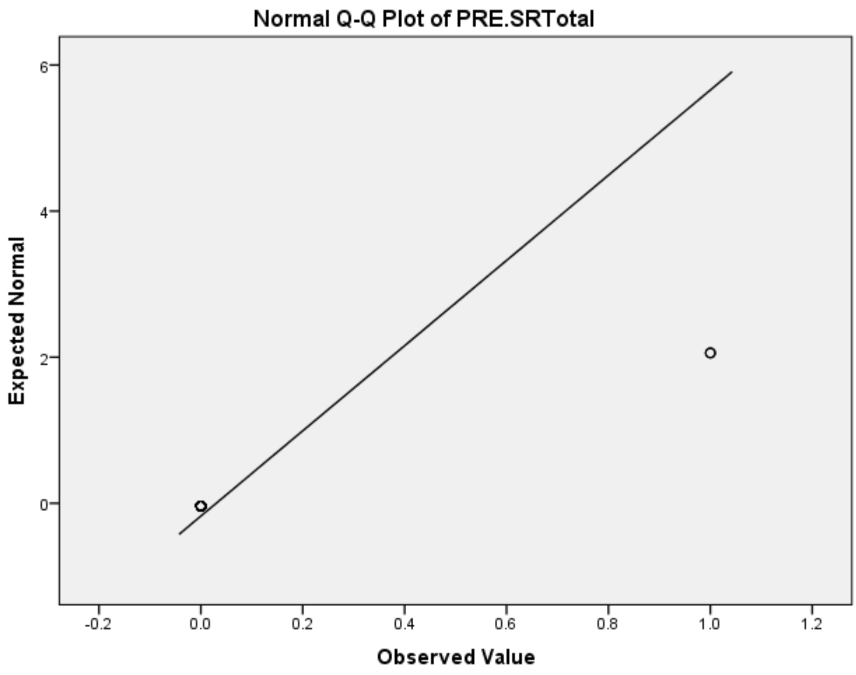


Figure X2: Normal Q-Q plot of pre-test SR items

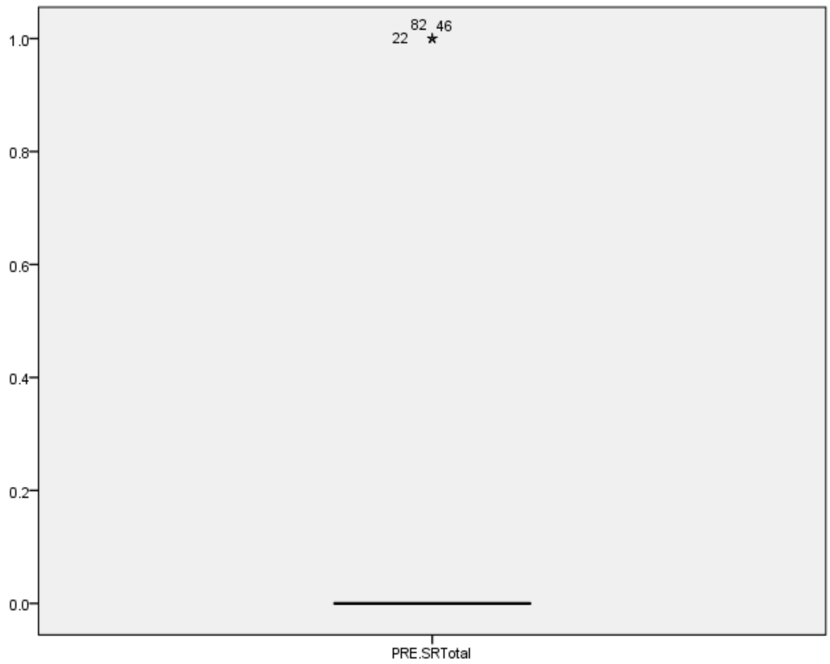


Figure X3: Boxplot of pre-test SR items

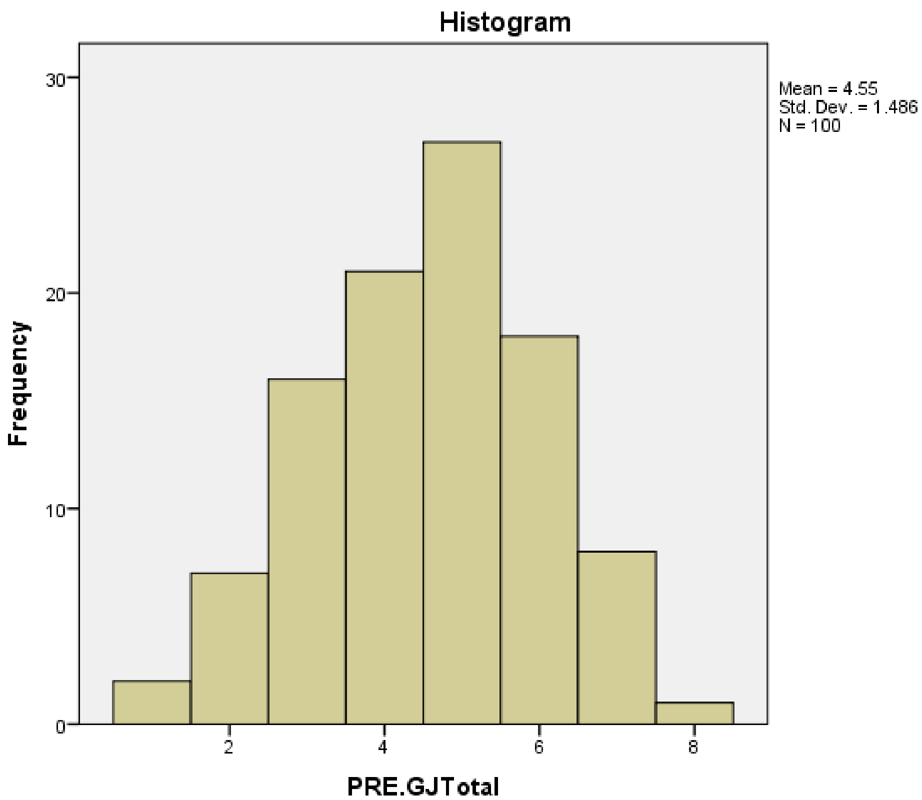


Figure X4: Histogram of pre-test GJ items

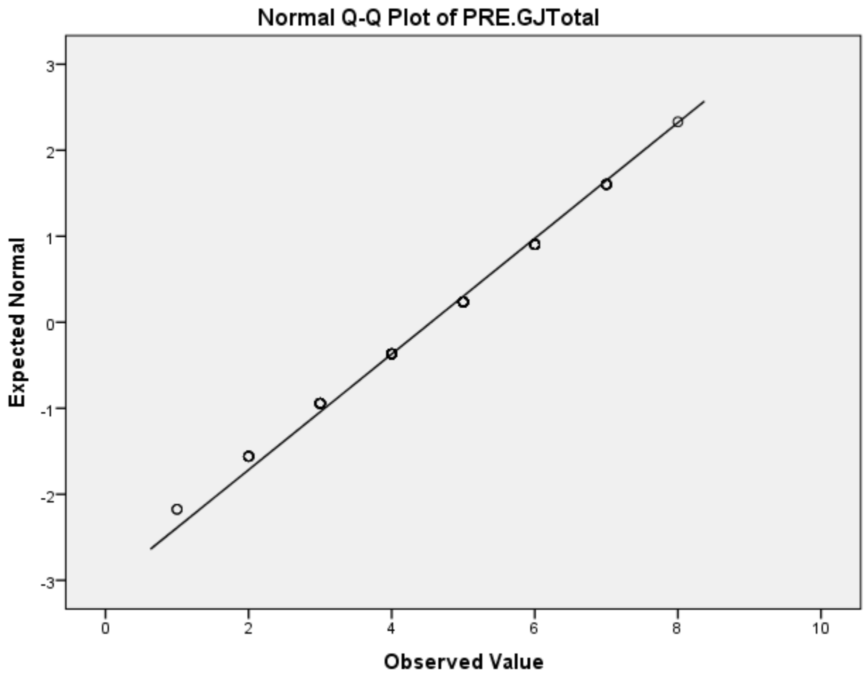


Figure X5: Normal Q-Q plot of pre-test GJ items

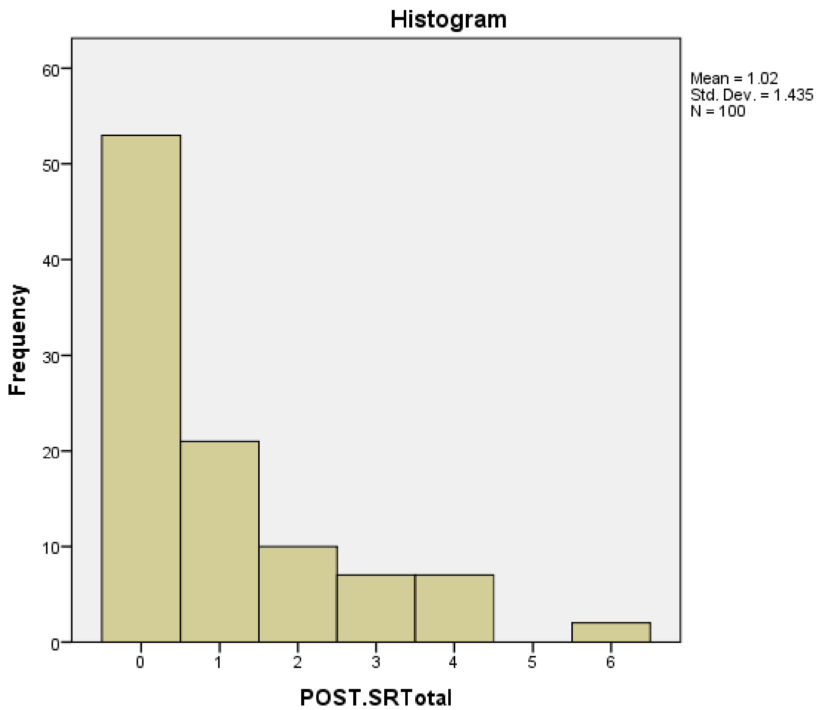


Figure X6: Histogram of the post-test SR items

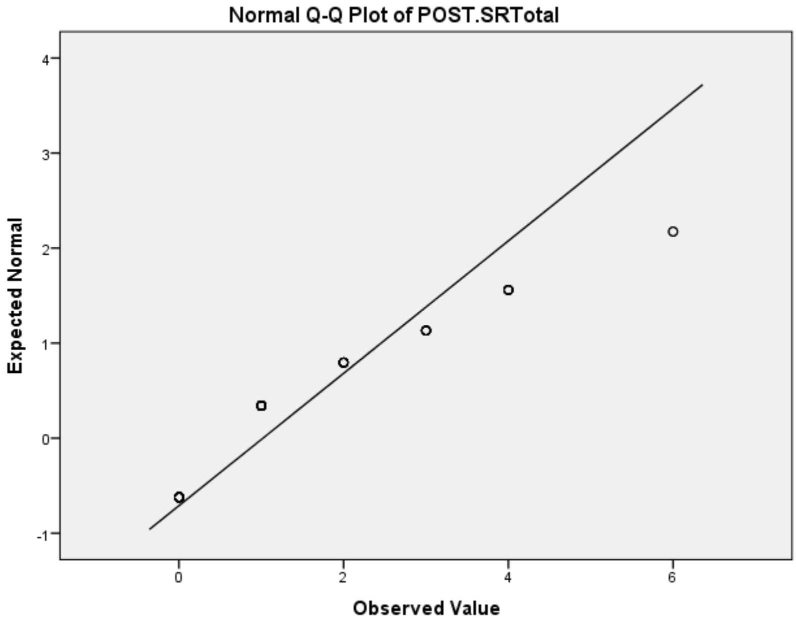


Figure X7: Normal Q-Q Plot of post-test SR items

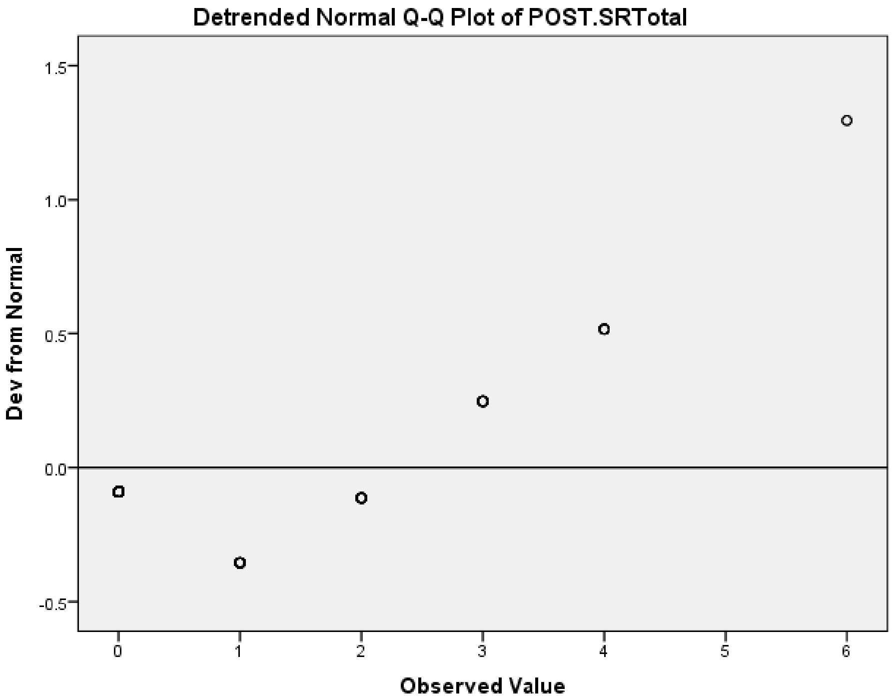


Figure X8: Detrended Normal Q-Q Plot of post-test SR items

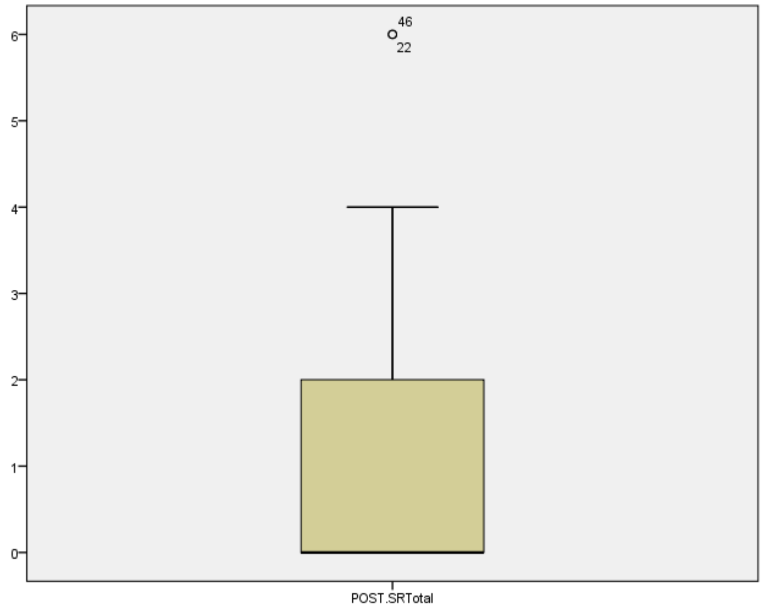


Figure X9: Boxplot of the post-test SR items

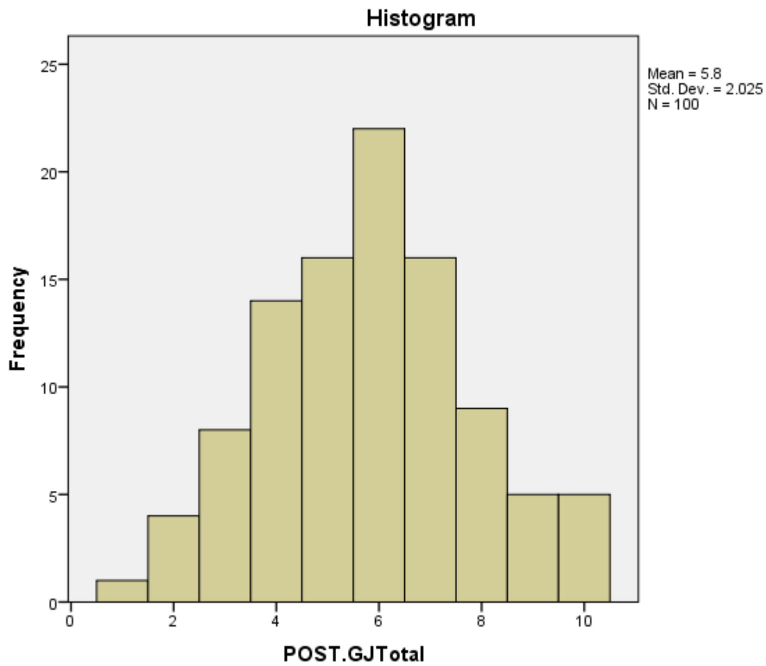


Figure X10: Histogram of the post-test GJ items

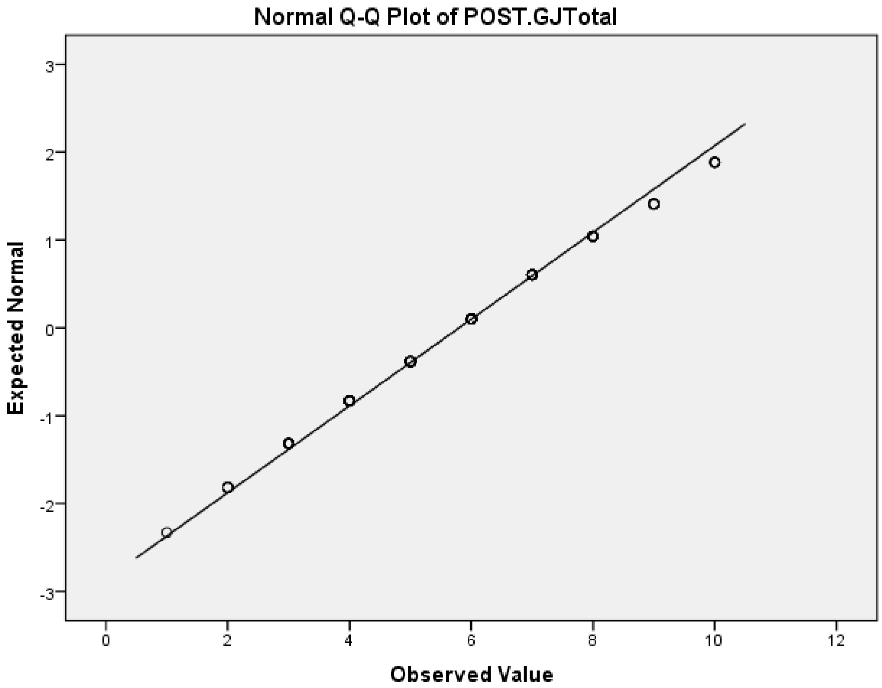


Figure X11: Normal Q-Q Plot of post-test GJ items

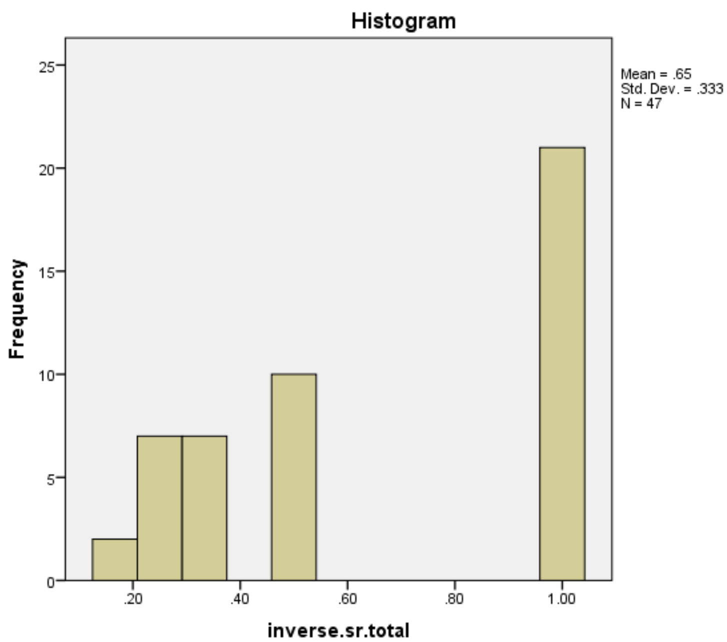


Figure X12: Histogram of Inverse Post SR Total

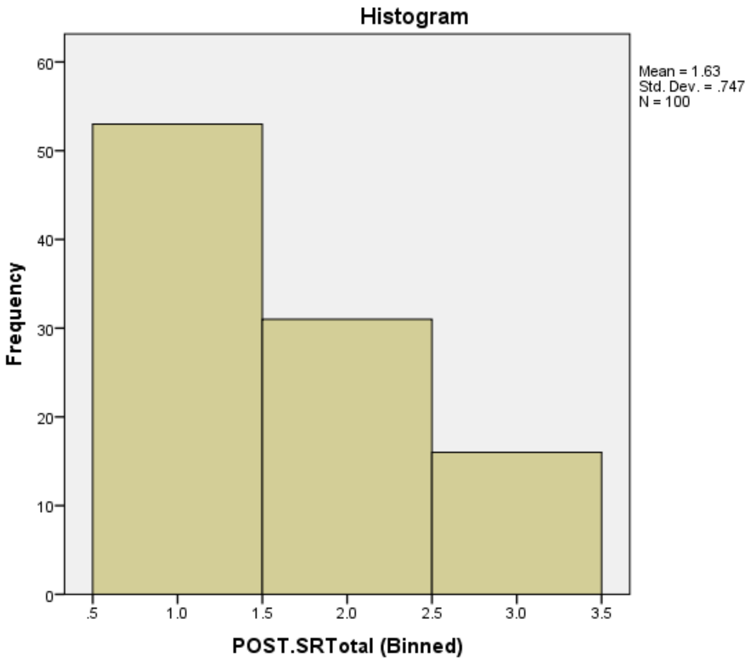


Figure X13: Binned Post SR Total

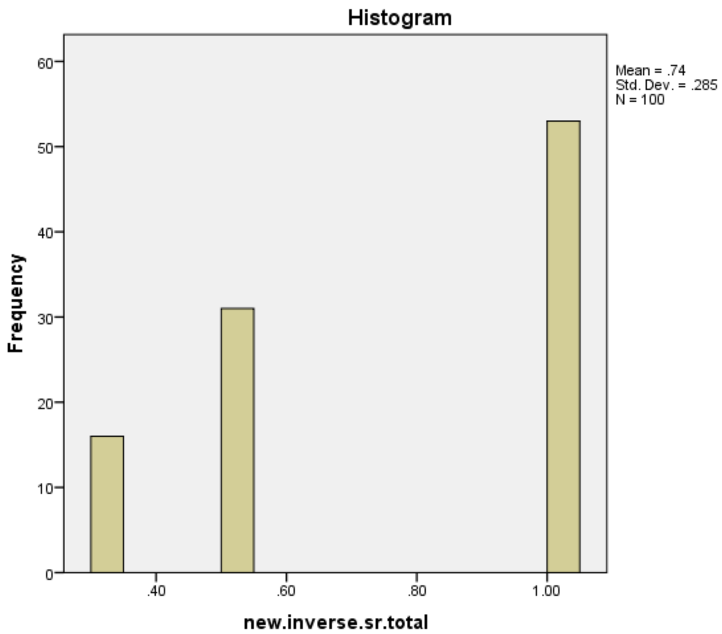


Figure X14: Binned Inverse Post SR Total

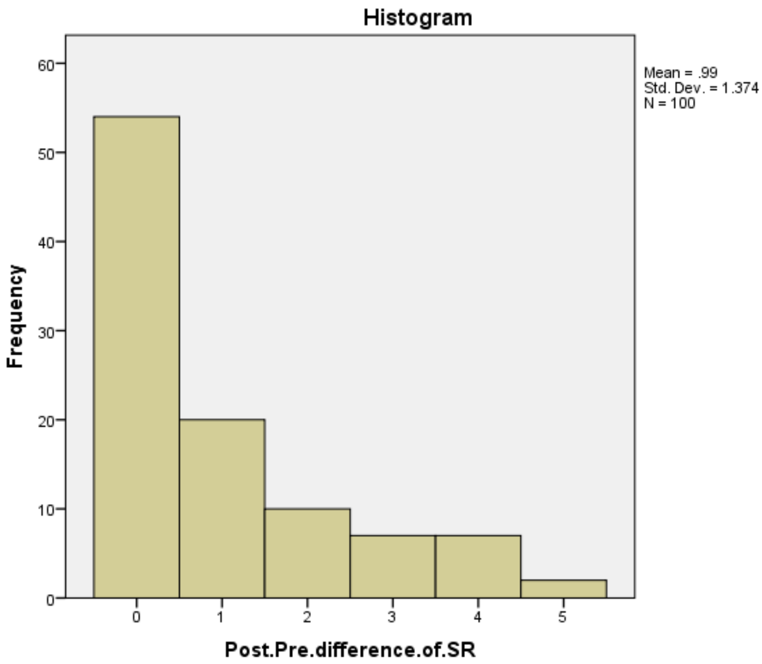


Figure X15: Histogram of SR gain score

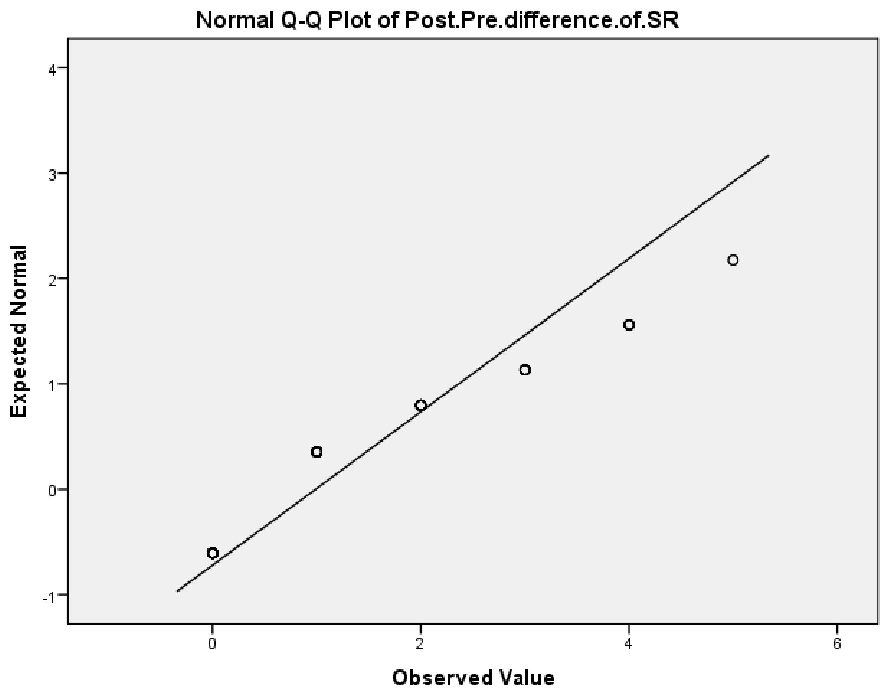


Figure X16: Normal Q-Q Plot of SR gain score

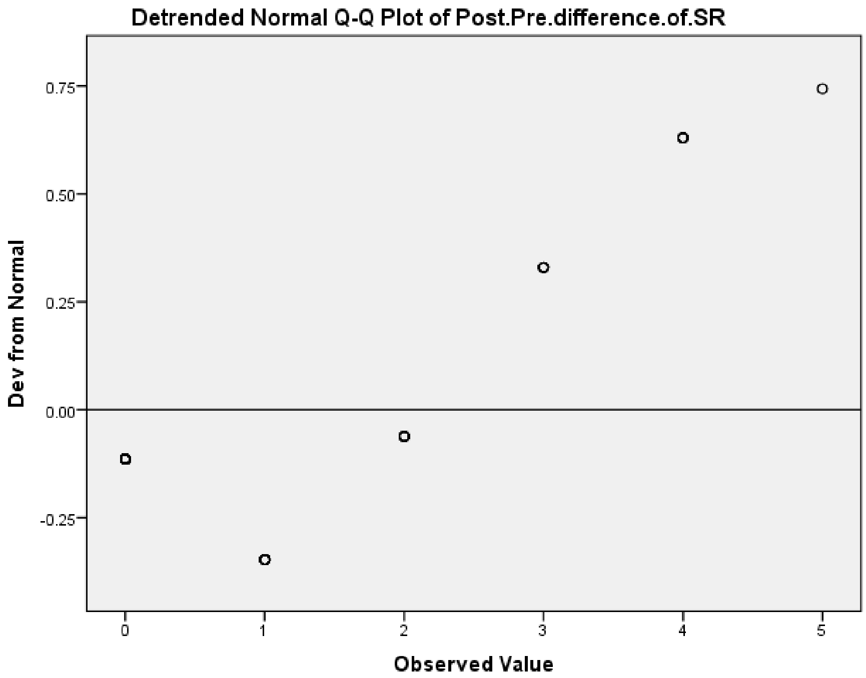


Figure X17: Deterended Normal Q-Q Plot of SR gain score

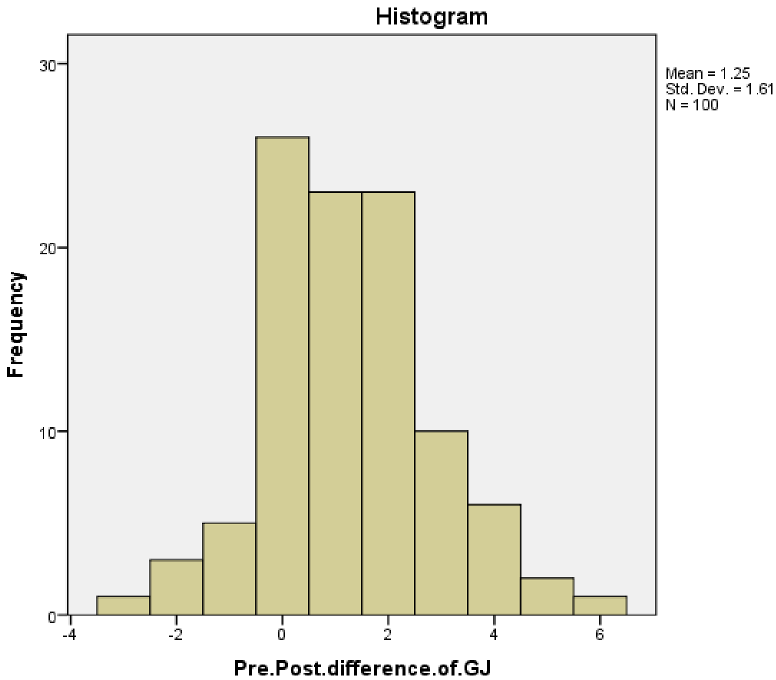


Figure X18: Histogram of the GJ gain score

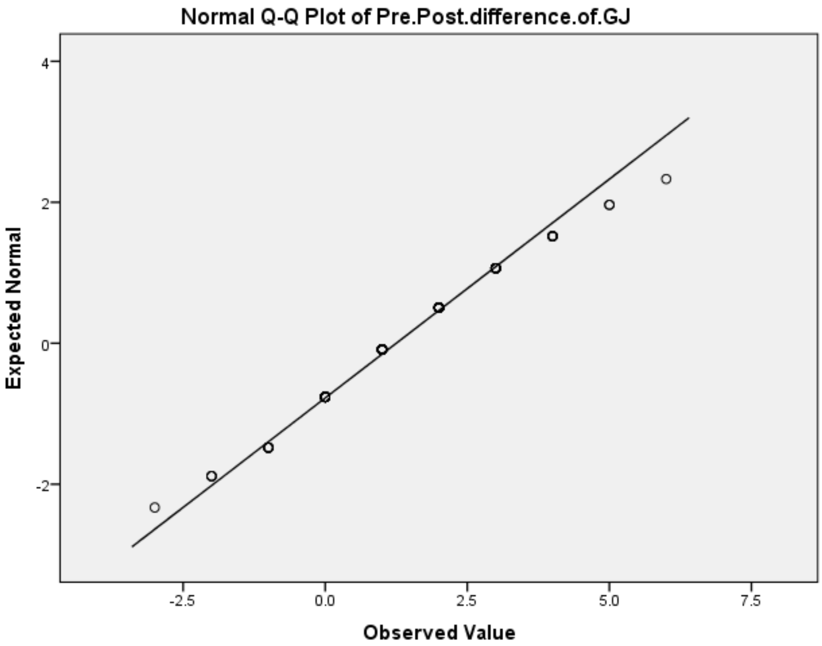


Figure X19: Normal Q-Q Plot of the GJ gain score

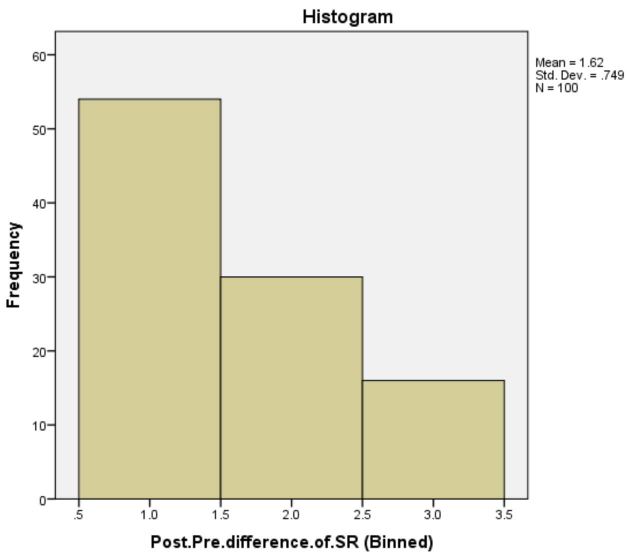


Figure X20: Histogram of SR Binned gain score

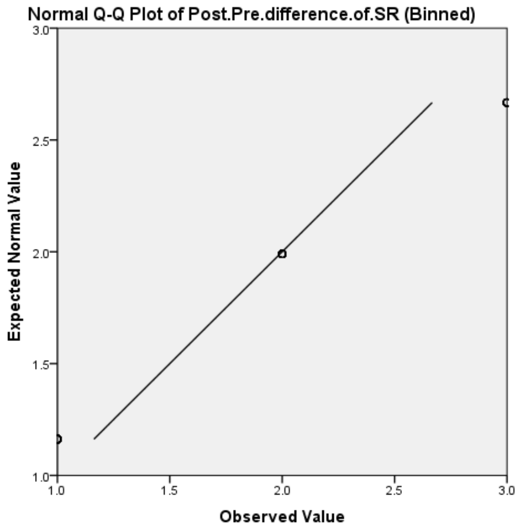


Figure X21: Normal Q-Q plot of SR Binned gain score

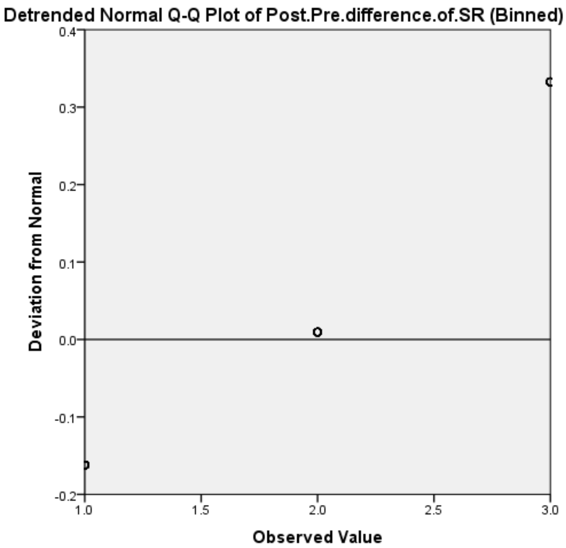


Figure X22: Detrended Normal Q-Q plot of SR Binned gain score

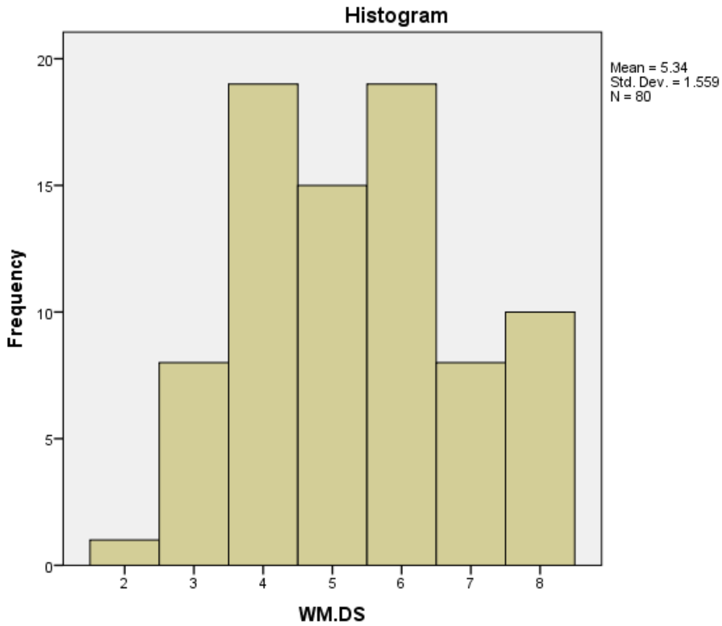


Figure X23: Histogram of DS test

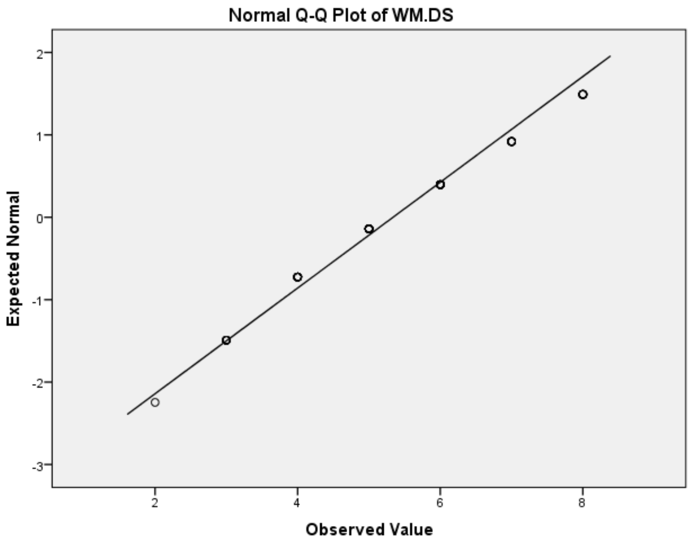


Figure X24: Normal Q-Q plot of DS test

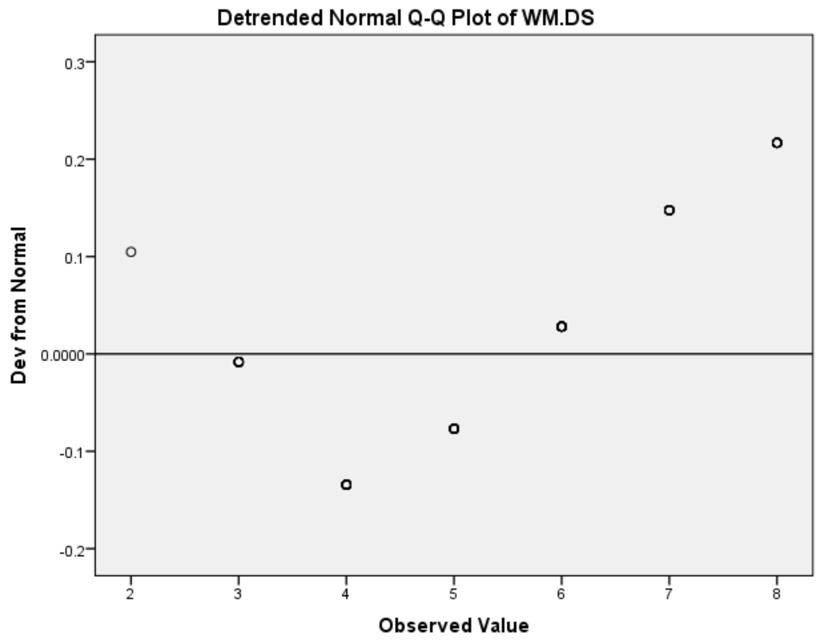


Figure X25: Detrended Normal Q-Q plot of DS test

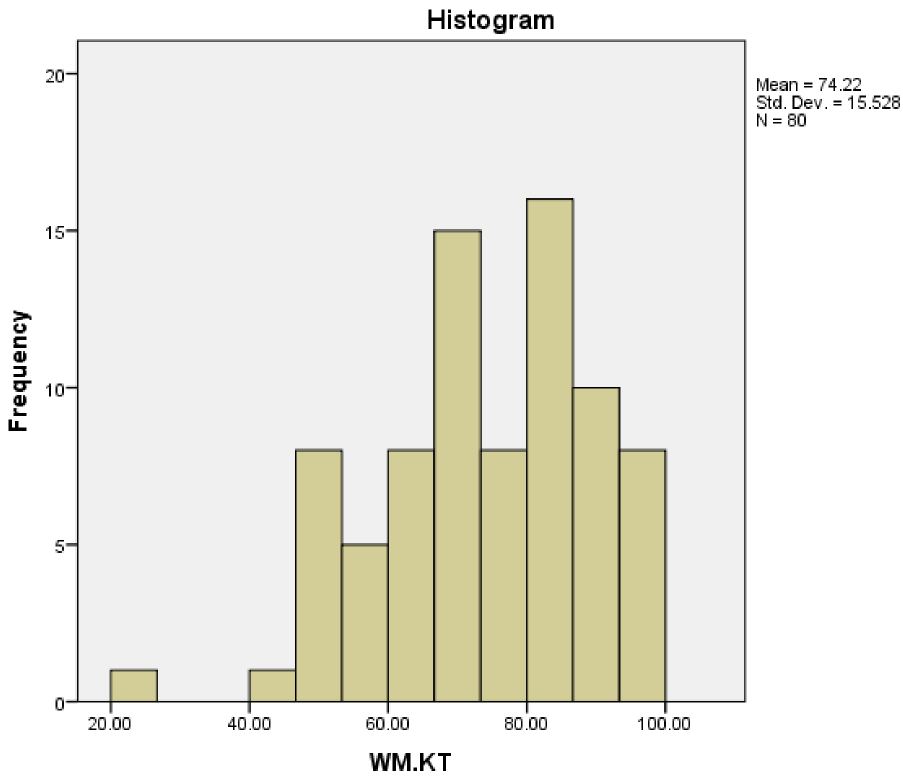


Figure X26: Histogram of KT test

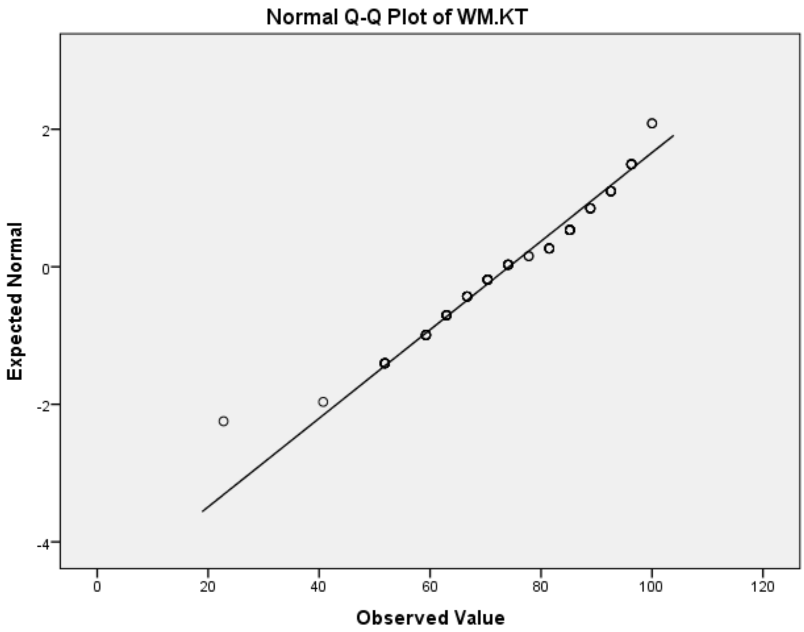


Figure X27: Normal Q-Q plot of KT test

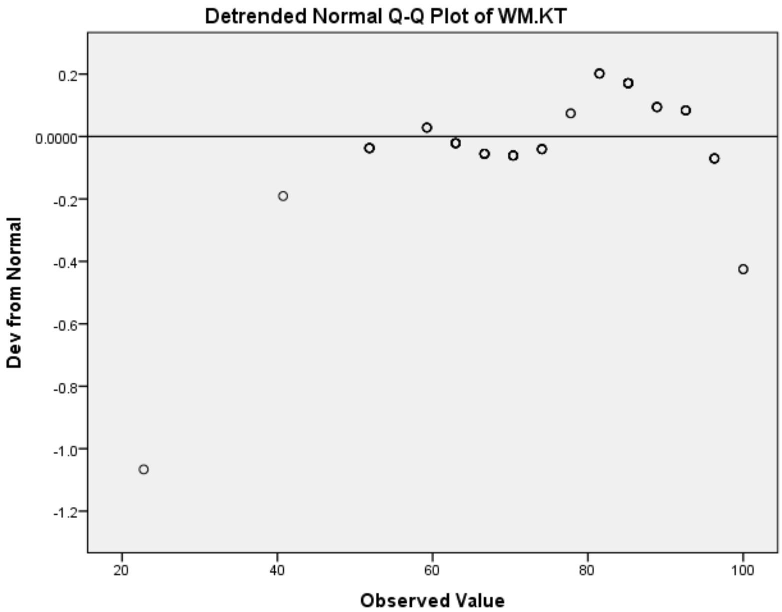


Figure X28: Detrended Normal Q-Q plot of KT test

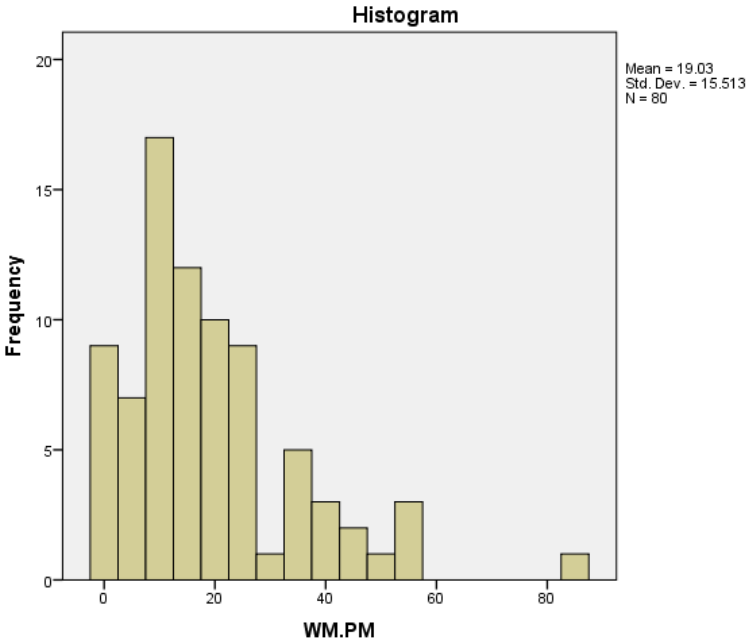


Figure X29: Histogram of PM test

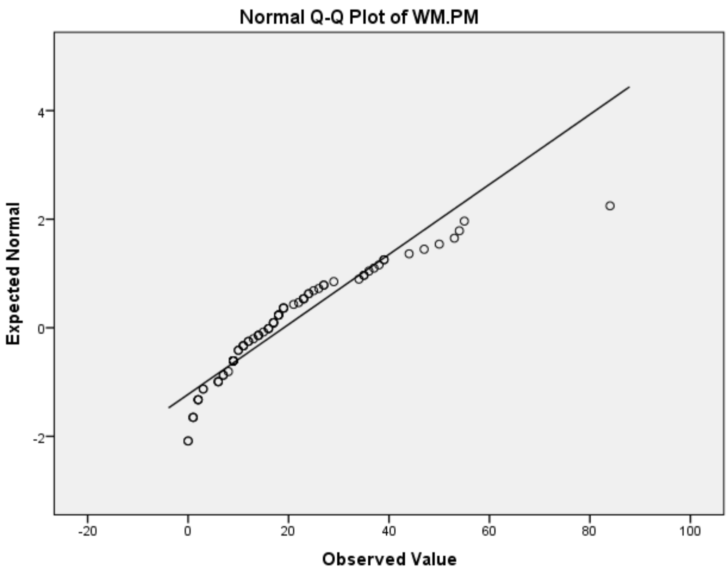


Figure X30: Normal Q-Q plot of PM test

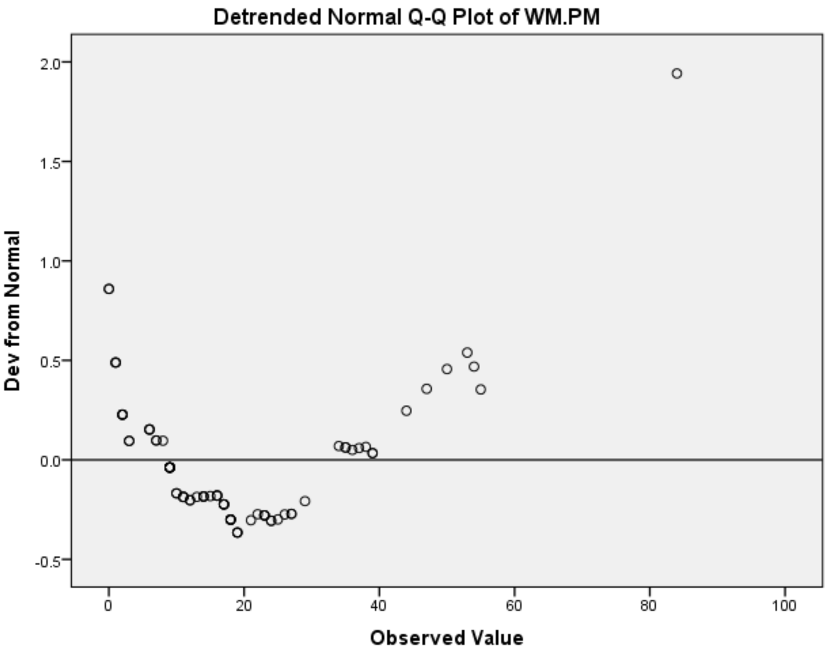


Figure X31: Detrended Normal Q-Q plot of PM test

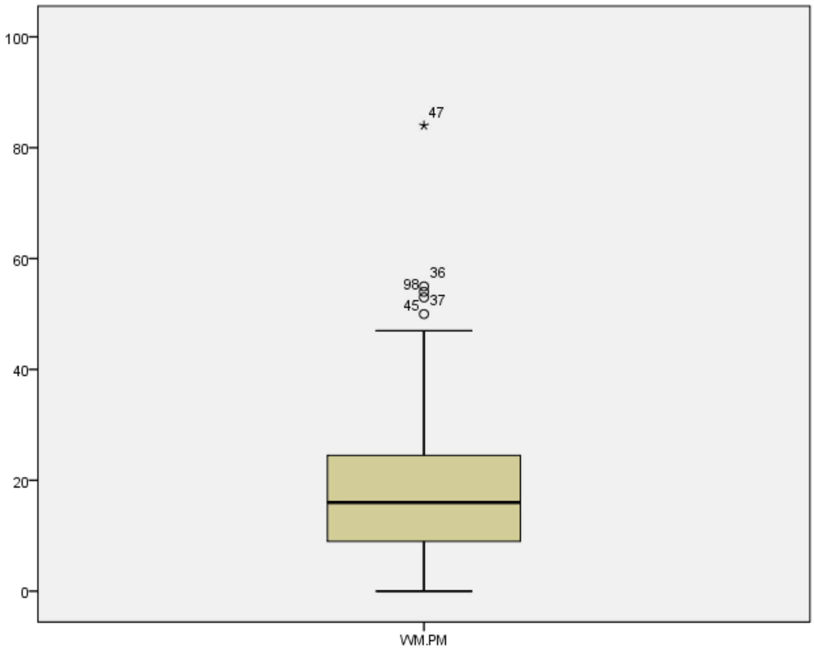


Figure X32: Boxplot of PM test

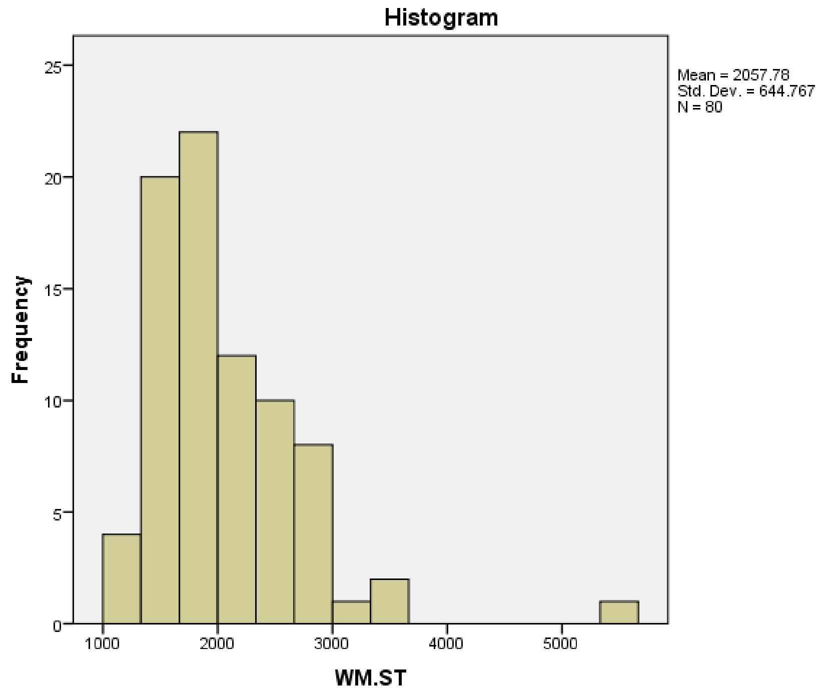


Figure X33: Histogram of ST test

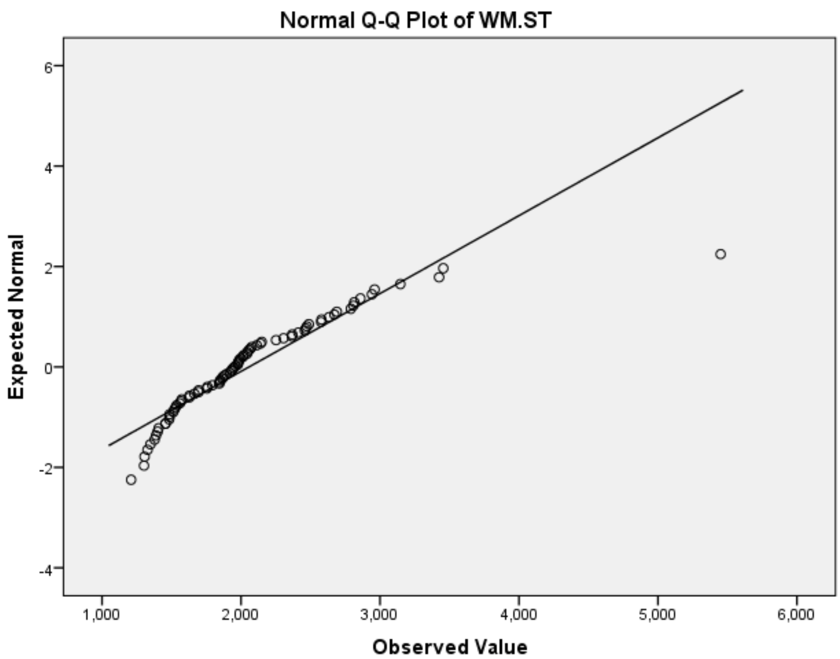


Figure X34: Normal Q-Q plot of ST test

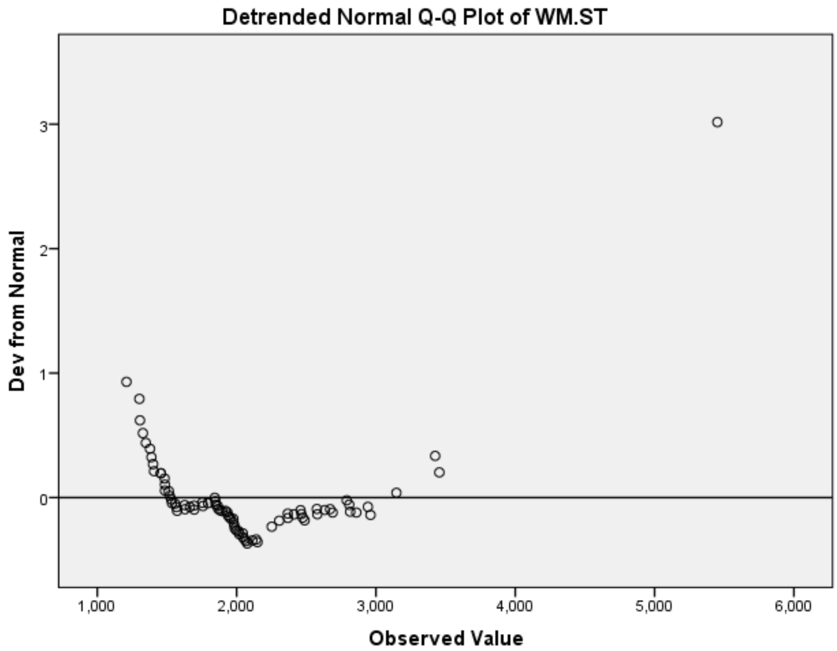


Figure X35: Detrended Normal Q-Q plot of ST test

Figure X36: Boxplot of ST test

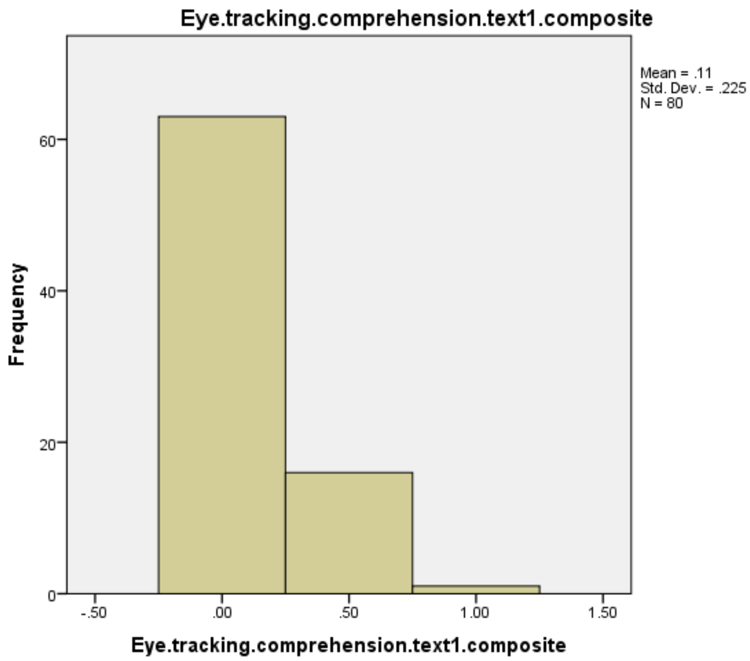


Figure X37: Histogram of the Composite comprehension test score of Text 1

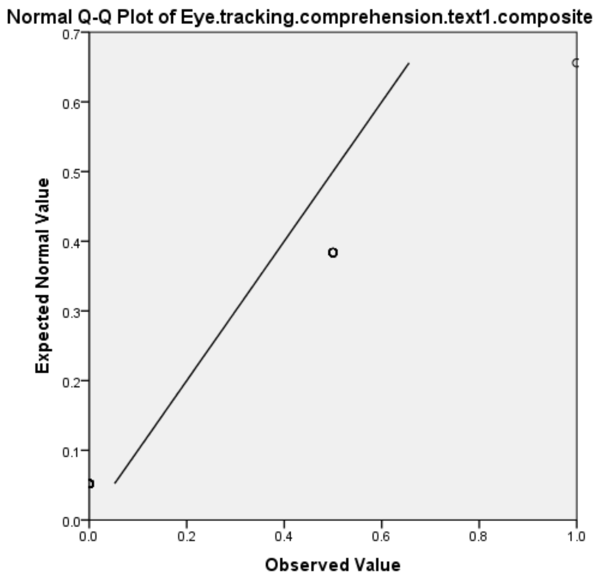


Figure X38: Normal Q-Q plot of the Composite comprehension test score of Text 1

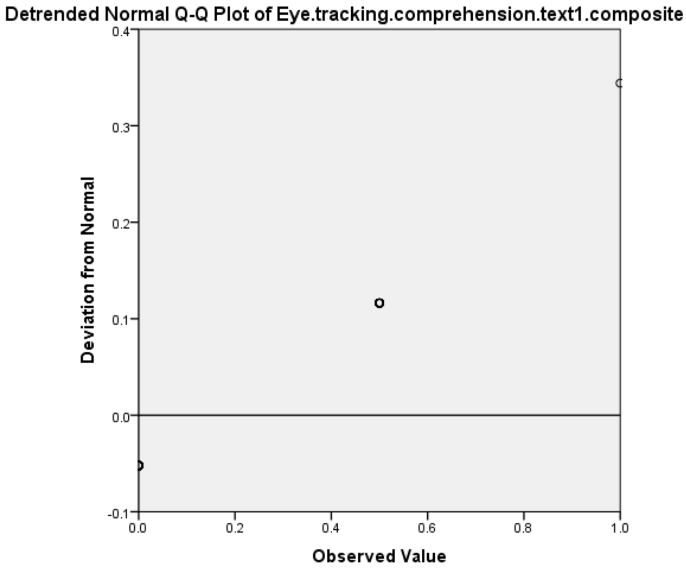


Figure X39: Detrended Normal Q-Q plot of the Composite comprehension test score of Text 1

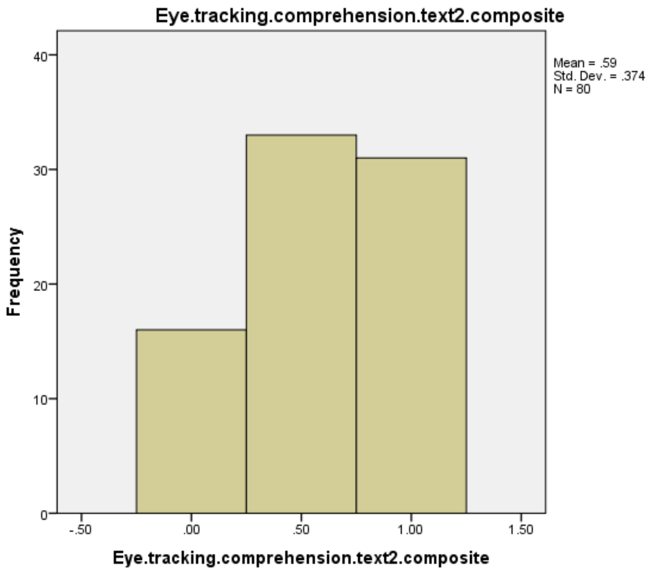


Figure X40: Histogram of the Composite comprehension test score of Text 2

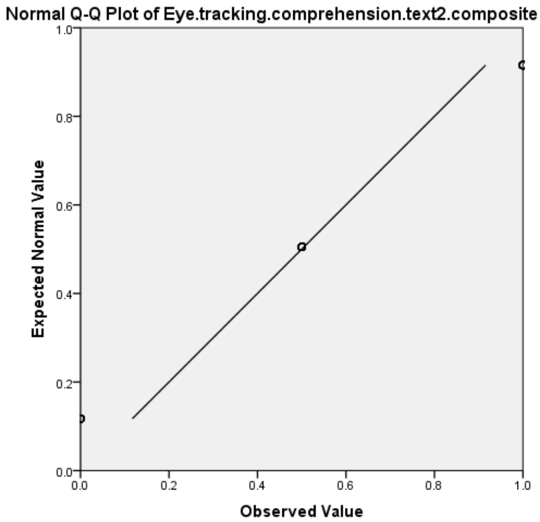


Figure X41: Normal Q-Q plot of the Composite comprehension test score of Text 1

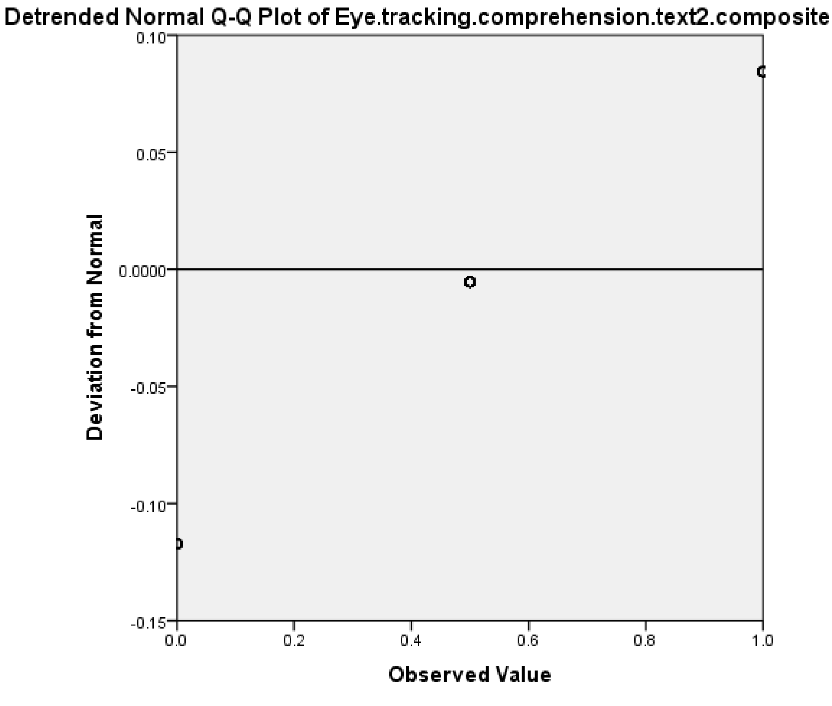


Figure X42: Detrended Normal Q-Q plot of the Composite comprehension test score of Text 1

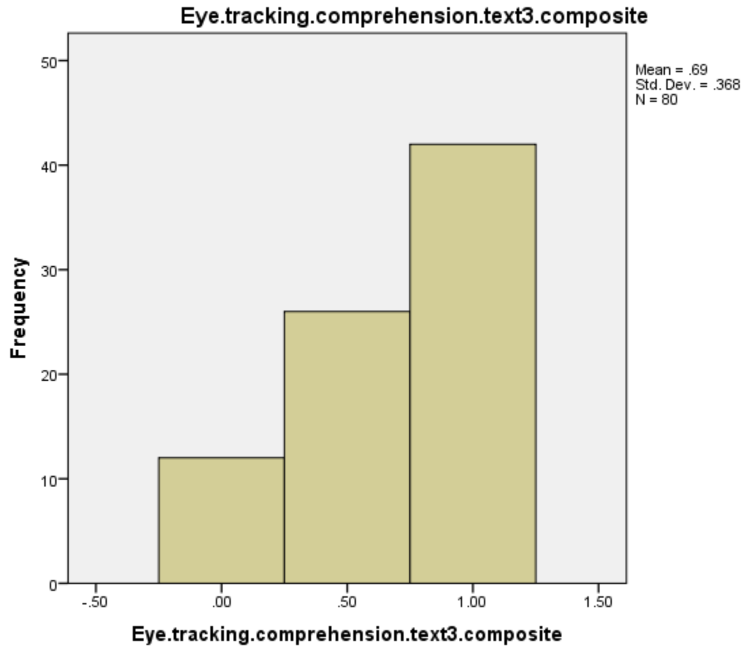


Figure X43: Histogram of the Composite comprehension test score of Text 3

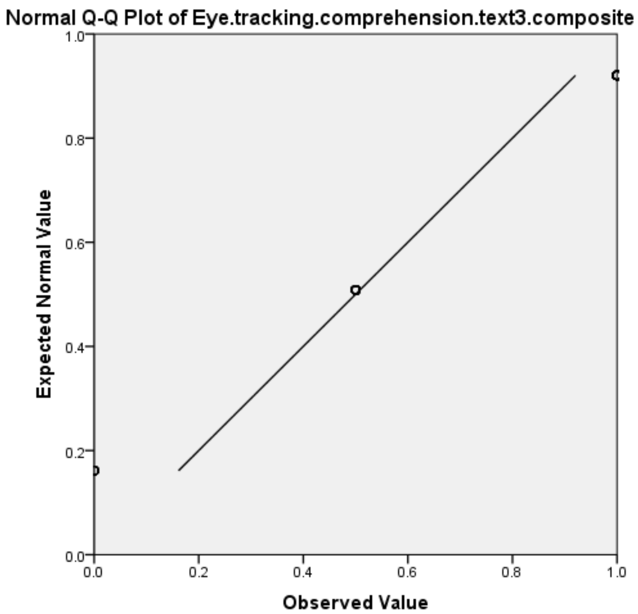


Figure X44: Normal Q-Q plot of the Composite comprehension test score of Text 1

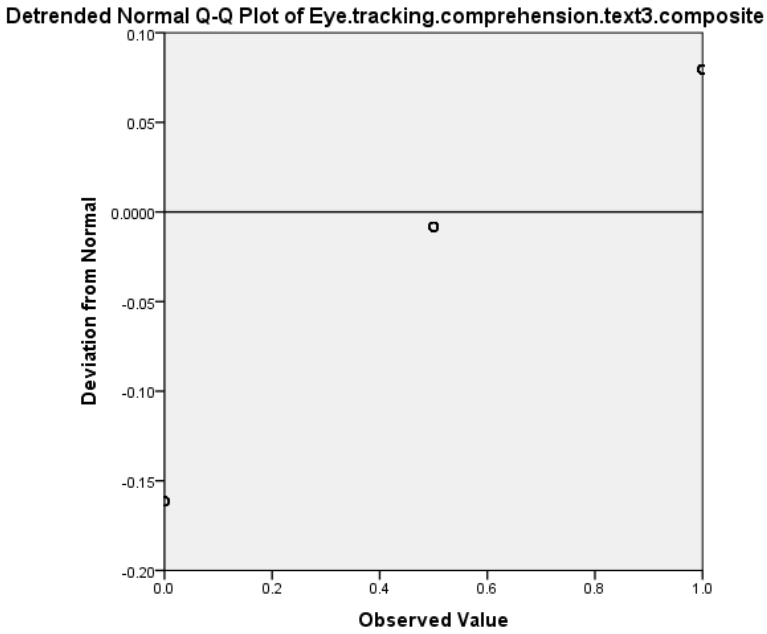


Figure X45: Normal Q-Q plot of the Composite comprehension test score of Text 1

Appendix Y: Total reading times of the three texts

Total reading times of the three texts

Table Y1.

Group A (enhanced+ instructions)

Participant	Text 1	Text 2	Text 3
	min.sec.ms	min.sec.ms	min.sec.ms
1	1.43.802	1.30.529	1.41.087
2	2.33.186	1.37.033	2.17.659
3	2.26.454	2.08.044	2.11.345
4	2.34.797	2.02.392	2.10.145
5	2.31.305	2.17.432	2.40.368
6	1.48.805	1.37.033	1.59.899
7	1.55.738	1.37.226	1.36.784
8	6.42.404	3.11.460	2.17.407
9	3.18.104	3.03.972	2.38.335
10	2.02.835	2.08.232	2.13.490
11	3.10.452	2.15.610	2.59.145
12	4.31.033	4.24.073	3.13.904
13	3.25.390	2.48.886	2.35.932
14	3.35.103	3.38.321	2.36.109
15	3.04.081	2.34.979	2.31.915
16	3.46.173	3.07.569	2.45.534
17	3.16.672	2.45.262	3.06.201
18	2.15.728	2.24.419	2.28.199
19	4.02.733	3.10.844	2.52.572
20	2.28.910	1.57.440	2.00.676

Table Y2.

Group B (enhanced+ instructions+ explanation)

Participant	Text 1	Text 2	Text 3
	min.sec.ms	min.sec.ms	min.sec.ms
1	2.48.300	2.45.707	2.33.286
2	4.22.632	2.43.429	3.13.678
3	4.10.423	3.26.543	3.13.027
4	2.29.902	2.06.696	2.05.780
5	1.57.035	2.44.426	2.11.608
6	2.31.729	2.31.138	2.34.267
7	2.10.307	2.18.280	2.10.610
8	3.15.890	3.04.200	2.54.921

9	2.31.968	2.27.886	2.40.398
10	2.44.550	2.46.500	2.32.630
11	3.28.160	3.03.626	2.54.766
12	3.24.014	3.58.201	3.45.376
13	3.56.686	3.21.874	3.41.267
14	3.20.725	2.21.027	2.31.344
15	3.40.322	2.34.296	3.05.168
16	2.32.251	2.43.853	2.46.228
17	3.42.390	3.34.663	3.28.235
18	2.46.396	2.53.936	2.35.325
19	2.49.580	2.59.670	2.40.987
20	3.15.173	3.03.425	2.54.687

Table Y3.

Group C (enhanced only)

Participant	Text 1	Text 2	Text 3
	min.sec.ms	min.sec.ms	min.sec.ms
1	2.52.263	2.42.331	4.20.234
2	2.45.172	2.05.147	2.31.953
3	2.21.626	1.46.224	1.57.111
4	3.30.809	2.44.429	2.55.425
5	2.19.274	2.10.079	2.16.685
6	3.08.499	2.53.274	2.14.694
7	2.20.285	2.06.673	2.15.239
8	4.33.149	3.50.310	3.40.429
9	3.18.739	2.51.210	2.14.862
10	1.50.456	1.45.493	1.48.476
11	2.10.903	2.22.450	2.26.962
12	2.29.856	3.07.525	3.13.043
13	3.47.727	2.25.937	4.01.749
14	4.15.999	2.52.001	3.07.151
15	4.55.223	2.38.508	3.08.340
16	3.52.509	2.44.435	2.52.145
17	5.03.227	4.25.081	3.57.437
18	2.15.315	2.24.265	2.24.489
19	2.18.985	2.16.465	2.18.917
20	2.42.208	2.51.381	2.39.326

Table Y4.

Group D (unenhanced)

Participant	Text 1	Text 2	Text 3
	min.sec.ms	min.sec.ms	min.sec.ms
1	2.40.464	2.27.507	2.16.391
2	2.10.294	3.06.129	2.55.713
3	3.32.412	2.35.139	2.33.565
4	2.35.428	2.58.730	4.46.762
5	2.47.006	2.53.583	2.28.759
6	2.08.912	2.09.241	2.50.863
7	2.31.425	2.14.594	2.25.522
8	2.35.522	2.20.163	2.09.610
9	2.36.430	2.27.283	2.30.485
10	2.24.154	2.24.698	2.18.383
11	2.32.315	2.29.563	2.12.275
12	2.56.082	2.49.880	2.36.590
13	3.09.966	3.26.247	3.21.372
14	2.40.099	2.23.920	2.20.594
15	3.18.036	2.18.724	2.12.083
16	2.44.720	2.23.945	2.54.642
17	2.41,779	2.32.241	2.42.663
18	2.18.930	2.24.882	2.55.557
19	3.28.880	3.20.510	3.56.778
20	3.14.636	2.54.776	2.33.417

Table Y5.

Total reading time of each group (in milliseconds)

Group	Text1		Text2		Text 3	
	Mean in milliseconds	*SD*	*Mean in milliseconds*	*SD*	*Mean in milliseconds*	*SD*
Enhanced+ instructions (A)	143123.67	4872.57	130953.46	4549.45	126867.39	2637.36
Enhanced+ instructions + explanation (B)	145453.78	4046.78	160708.67	2806.12	129651.46	2809.25
Enhanced only (C)	110535.17	3848.24	118902.55	3820.05	124900.56	4100.56
Unenhanced (D)	124801.67	2470.66	118357.34	2210.98	119159.47	3861.81