Focus on Form(s) and Meaning in a Technology- Enhanced Language Learning Environment (TELL)

Submitted by
Rola Naeb

Supervised by
Prof. Martha Young-Scholten
Mr. Scott Windeatt

Doctor of Philosophy
School of English Literature, Language and Linguistics
Newcastle University
April, 2015

Abstract

Despite decades of Instructed Second Language Acquisition research, there is still a dearth of research on the applicability of findings in different learning environments, particularly self-accessed technology-enhanced environments (TELL). In ISLA, types of input available in the classroom can be categorized as Focus on Meaning (FoM), Focus on Form (FoF) and Focus on FormS (FoS), (Doughty & Williams 1998). In traditional classrooms, research indicates superiority of FoF and FoS (Spada & Tomita 2008, 2010). The question still remains, though, of which type of input is most effective in TELL. One assumption about TELL is that it enhances input quantity and quality. That is, input is delivered in greater quantity and when the learner can make best use of it, better quality. Moreover, the type of interaction in TELL (human-software) is different to classroom interaction (human-human). Such differences are likely to affect both the learner's output (product) and the learner's behaviour during learning (process). A study of 71 ESL learners, divided into three groups, was conducted to investigate the effectiveness of FOM, FOF, and FoS in a TELL

Learner performance on a construction selected for its difficulty for L2 learners of English (indirect speech) was taken as a measure of intake. Data on patterns of behaviour were obtained through log files to gauge participants' awareness of form during task completion.

Results revealed that all learners improved their performance on the construction selected after the treatment. However, the FoF group outperformed the other two groups. In terms of the contributing factors, task type, modality of input, processing time and number of trials were identified as effective factors. Contrary to what studies of classroom learners have shown, learners in the FoF and FoS groups chose not to focus on form even when they were stuck. They mostly behaved instead like FoM learners. This behaviour vitiates the effectiveness of FoF or FoS in a TELL environment. The behaviour of one learner from each group was examined to arrive at a more nuanced picture of these differences. These three learners exhibited flip-flop behaviour where they kept switching between items. However, the FoS learner showed a more confident route which, however, resulted in lower attainment. The FoM learner displayed a more confused route. Finally the FoF learner showed a mixed pattern that ultimately led to better attainment on the target construction.

Dedication

To my late dad, for believing in me, you will always be my inspiration,

To my mum for your endless and unconditional love,

To my husband for always being there for me.

Acknowledgment

I would like to express my deepest gratitude to my my supervisors, Prof. Martha Young Scholten and Mr Scott Windeatt for their guidance and support.

Special thanks to all my family members for their never ending love and patience. Thank you Lisa, Wasel, Reem, Banan and Sinan.

This thesis would not have been possible without the support of two very special men; my husband, Mawoud and my brother, Rakan. Thank you for everything you did for me over the years.

Thanks to each and every one who has helped me in any way.

Focus on form(s) and meaning in a technology-enhanced language learning environment

Table of Contents

Lists of tables

List of figures

List of abbreviations

CAI	Computer-Assisted Instruction,
CAL	Computer-Assisted Learning
CALL	Computer-Assisted Language Learning
CBE	Computer-Based Education,
CBI	Computer-Based Instruction
CFI	Communication-Focused Instruction
CLT	Communicative Language Teaching
CMC	Computer-Mediated Communication
COLT	Communicative Orientation of Language Teaching
EFL	English as a Foreign Language
ELT	English Language Teaching
ESL	English as a Second Language
FFI	Form-Focused Instruction
FL	Foreign Language
FoF	Focus on Form
FoM	Focus on Meaning
FonF	focus on form
FoS	Focus on FormS
FTF	Face-To-Face
GJ	Grammaticality judgment tasks
GSLA	Generative Second Language Acquisition
ISLA	Instructed/Second Language Acquisition
ICALL	Intelligent Computer-Assisted Language Learning
IIO	Input, Interaction and Output
IP	Input processing
ISLA	Instructed Second Language Acquisition
L1	First language
L2	Second language
LAD	Language Acquisition Device
MOGUL	Modular Online Growth and Use of Language
NLP	Natural Language Processing
PDA	Personal Digital Assistants
PPP	presentation, practice, production
PT	Processability Theory
SLA	Second Language Acquisition
SPELL	Spoken Electronic Language Learning
TBLL	Task-Based Language Learning
TBLT	Task-Based Language Teaching
TELL	Technology-Enhanced Language Learning
TL	Target Language
UG	Universal Grammar

Chapter One

Introduction

This introduction sets out the theoretical framework for the current study and states the motivation for undertaking an investigation of input-based approaches to second language learning in a technology-enhanced language learning environment. Section 1.1 presents the reasons behind undertaking the study. Section 1.2 introduces the wider research field of Second Language Acquisition (SLA). Section 1.3 introduces the debate about theory of language in SLA. Section 1.4 discusses different approaches to SLA research. Sections 1.5 and 1.6 present some of the key arguments in SLA and section 1.7 introduces the field of Instructed Second Language Acquisition (ISLA). Section 1.8 then sets out the rest of the thesis and summarizes the argumentation.

1.1 Motivation for the study

Among the reasons Mitchell and Myles (2004) mention for why we study how second language (L2) acquisition takes place are two that guide this thesis. First, such studies are interesting in themselves because they help us understand the human mind better through understanding the nature of language and the mechanisms of human learning. Spada and Lightbown (2002), Cook (2008) and de Graff and Housen (2009) argue that SLA studies have practical and theoretical significance. From a theoretical point of view, they help in 'understanding how the brain processes linguistic input of various kinds to arrive at linguistics representations in the mind' (de Graff and Housen, 2009, p.727). Second, better knowledge about how L2 acquisition takes place and what triggers acquisition can benefit millions of teachers and learners all over the world. From a practical point of view, 'a better understanding of how instruction affects L2 learning may lead to more effective teaching (Cook, 2008; de Graff and Housen, 2009). By understanding what leads to effective acquisition and what does not, teachers and educationalists will be able to modify their teaching methods and practices, thus enhancing the learning environment. This is at the heart of the current study in addition to the added aim of looking at how second language acquisition theories and technology can work together in self-accessed learning environments. The present study is not intended to be a second language acquisition study in the sense that it does not focus on the question of whether the learners have acquired the target construction. Reasons for this include the fact that the amount of input provided to learners is influenced by many factors that might affect acquisition of the target construction (see section 7.2 for the limitations of the study and the experiment). In fact, the present study is more of an exploratory nature. One of its main goals is to explore how learners deal, by type, with the input they receive. In order to obtain a detailed record of learners' activities while processing input, it was necessary to limit experimentally the amount of input provided.

Studies that investigate how learners deal with the input they receive are typically conducted in traditional classrooms using pre-post-test design to investigate the effectiveness of the input and think-aloud protocols to investigate the learners' behaviour or learning processes. Even when technology is used in these studies, it is simply used as a medium to deliver the treatment or later to analyse the data. Hulstijn (2000) has highlighted the need for studies that utilise innovative methodologies to deliver the input or elicit data and he remarked that; 'what is needed, are studies investigating the bottom-up processing of oral L2 stimuli. The multimedia computer is the ideal tool to present linguistic stimuli, both in spoken and written form, and to register all reactions of learners in terms of both accuracy and time' (2000, p.39). Nearly 1 ½ decades on from Hulstijn' s (2000) remark, there are still very few studies in language acquisition that have attempted to use technology in an innovative way to elicit these sorts of data from learners. More recently, technology-based methods have been used in SLA to measure reaction/response time when investigating the role of working memory (e.g. Wright, 2010, 2012) or the real-time processing of grammaticality judgment tasks (e.g. Marinis, 2010). There is still a dearth of literature regarding studies that use methods to register all reactions of learners when they deal with input, particularly in self-accessed technology–enhanced learning environments. Doing so is one of the main aims of the present study. As Hulstijn (2000) puts it, computer-aided tools have the potential to provide researchers with a closer look at the processes of language acquisition and use. On the other hand, there is a dearth of research that looks at interaction in technology-enhanced learning environments from an applied linguistic perspective. The call for such research was made by Chapelle (1990) who used discourse analysis to look at how learners deal with computer-based language learning materials. The present study therefore investigates the processes that determine the effectiveness of input in a technology-enhanced learning environment and examines the factors that contribute to the expected variability in learners' performance. The study uses log files that register all actions of learners divided into three groups based on the type of input. The same technology is used to deliver the input, elicit and collect data and then to analyse the data.

The present study adopts task-based language learning (TBLL) and teaching methodology (Nunan, 1989, 2004, 2006). The dominance of task-based learning has increased in recent years and its potential advantages have been the focus of many studies including technology-based studies (Reinders, 2005). It seemed logical, in this sense, to use tasks as the medium of delivering the input. Furthermore, since tasks are increasingly used in learning environments nowadays, research that investigates the task features that work will be particularly significant for pedagogical applications.

The following sections will briefly introduce the key concepts in SLA and ISLA research that inform the current study. The focus is primarily on topics that will be of relevance to the literature in the subsequent chapters rather than a comprehensive review of the literature mentioned.[1]

1.2 SLA research

This section introduces the field of SLA. This introduction is essential in order to understand the nature of second language acquisition and the conditions under which input is effective. Section 1.2.1 defines what is meant by language *acquisition* and section 1.2.2 establishes the main arguments in the field of SLA.

1.2.1 What is language acquisition

Research on child language acquisition informs the early studies of SLA and I, therefore, start by considering it. Child language acquisition, or as it is commonly referred to L1 acquisition, is the study of the linguistic competence that human beings acquire and the verbal medium they use to communicate in their early childhood. L1, mother tongue, native language are all used to refer to the language first spoken by children although these terms carry different connotations (the L1 might not be the mother's language). Since the late 1950s, first language acquisition research has been

[1] In these sections, the reader is directed to relevant sources for further information.

dominated by the generative approach led by Noam Chomsky. Chomsky (1957, 1980, 1985) argues that regardless of the language to which they are exposed, children are born with a human language faculty and during their childhood, they have access to *Universal Grammar (UG),* a unified system that is responsible for linguistic knowledge. Tomasello (1995) and Abbot Smith and Tomasello (2006) criticise Chomsky's innateness theory and argue for a functional usage-based account of language acquisition. They suggest that language is composed of conventional symbols and these symbols are acquired by children in the context of 'culturally constituted event structure' (Tomasello, 2003, 2006). The importance of such discussion of the nature of acquisition is critical to the understanding of SLA and its implications for the classroom. Both generative and functional approaches have informed ISLA theories and consequently they inform the teaching methodologies used in the study (this will be discussed briefly in section 1.5 below). The next section addresses the main issues in SLA in light of L1 acquisition.

1.2.2 SLA

There is no doubt that L2 learners are different from L1 learners. They are already familiar with at least one linguistic system. They have already developed a cognitive system, and in the case of adults, a comprehensive one. Therefore, it is logical to assume that L2 acquisition, particularly by adults, is different from L1 acquisition. It is important then to define what is meant by second language acquisition before any discussion of the role of input or the processes involved in acquiring a second language takes place. *Target* language, *second* language, *foreign* language have all been used to refer to the language acquired after the establishment of at least one language system, the L1. These terms carry different associations and it could be easily argued that they are not the same since the target language could be the third or so on. According to Cook (2008), an L2 is simply a 'language acquired by a person in addition to his mother tongue' (Cook, 2008, p.2). This definition is the one adopted in this study. What is noteworthy here is that the hot debates in SLA research have immensely affected how researchers define 'acquisition' in L2 acquisition, and this has varied from native-like knowledge to native-like performance to ability to communicate and so on.

One of the earliest and most widely-known discussions of the nature of SLA is Krashen's distinction between *acquisition* and *learning* (Krashen, 1982; see also Schwartz, 1993). Krashen differentiates between unconscious *acquisition* and conscious *learning*. According to Krashen, acquisition means acquiring a second language in the same way children acquire L1. *Acquisition* is natural, unconscious and learners end up knowing the language but not necessarily knowing about the language. *Learning* is a conscious formal process and learners end up knowing about the language. In more precise terms, *acquisition*, in Krashen's argument, is the implicit knowledge of a language while learning is the explicit knowledge of a language (see section 2.3.1). Krashen's distinction has been used widely to highlight the differences between acquiring a language in a natural setting and learning it in the classroom. However, not all SLA researchers make this differentiation (e.g. see Ellis, 1997).

Such a view is held by Schmidt. In his study examining Wes, a Japanese learner of English, Schmidt (1983) seems to define acquisition in relation to what native speakers do. So Schmidt claims that Wes has acquired a form if he shows the same patterns of language use as a native speaker. One of the problems with this definition is that it does not account for the fact that *knowledge* is different to *use*. The learner might know a certain form but might not produce it for different reasons. Ellis (1997) defines second language acquisition as 'the ways in which people learn a language other than their mother tongue, inside or outside of a classroom' (p.3). Ellis, here, does not make a distinction between conscious and unconscious learning, or between acquiring a language in the classroom or naturally, contrary to Krashen (1982; 1985). Ellis' definition also does not link acquisition to outcome; it only refers to the process itself.

Gass and Selinker (2008) provide definitions of acquisition and list the criteria by which researchers determine if/when a form is acquired. Two of the most common ones are:

1. the first occurrence of a correct form
2. the percentage of correct/accurate forms produced

However, they state clearly that considering language forms only is insufficient as one needs to consider the stages through which learners pass to acquire a certain form. The need to do so is echoed in the increased interest in learners' processes in the field of SLA, a prime interest of the present study. Saville-Troike (2012) defines the scope of SLA as any phenomena involved in learning an L2 but she distinguishes between *second*, *foreign*, *library* and *auxiliary* language in terms of the function the L2 will serve.

In brief, SLA as a field of inquiry could be simply defined as investigation into the knowledge and use of any language in addition to the first language. So in this sense, the term 'second language acquisition' is widely used to refer to the systematic study of how people acquire languages other than their L1. These definitions are the ones that are applied in this study. Although, the study is mainly focused on instructed acquisition, unless indicated otherwise, *second language or L2* is used as a cover term that refers to both second (learned in an immersion setting) and foreign languages (learned in the classroom). Likewise, no distinction is made in this thesis between *conscious learning* and *unconscious acquisition* and the two terms are used interchangeably throughout. However, a distinction is made between SLA in general and *Instructed* SLA (see section 1.8 below) because the present study concerns the latter. The next three sections summarize some of the main debates in SLA that aid the understanding of the context of the current study.

1.3 Property and transition theories of language acquisition

It is important to consider theories of SLA, and Mitchell and Myles' (2004) definition is useful. As an 'abstract set of claims about the units that are significant within the phenomenon under study, the relationships that exist between them and the processes that bring about change', a theory of language should not only aim at describing the language system itself, it should also focus on the processes or stages that 'bring about change', i.e. lead to the acquisition of this system and the processes that affect this change. Cummins (1983) first drew a distinction between property and transition theories where the former is one that explains the properties of a system while the latter explains change in a state of a system. In the field of SLA, the same argument is also echoed by Ellis (1999) where he argues that in order to have a complete theory of language acquisition, we need to have a property theory and a transition theory (see also Gregg, 1993; Robinson 2001; Jordan, 2004).

Mitchell and Myles (2004) define a property theory as one that is 'primarily concerned with modelling the nature of the language system that is to be acquired' (p.7). Robinson (2001) argues that a property theory should describe the principles underlying the 'instantiation of linguistic knowledge' in learners' minds. In other words, it is a theory that identifies the elements of the system that is required for SLA to happen. A transition theory, on the other hand is one that is 'primarily concerned with modelling the change or developmental processes of language acquisition' (p.7), i.e. it explains the processes of SLA (Mitchell and Myles, 2004). In this sense, a study in the SLA field, ideally, should answer two questions;

1. What is the nature of the L2 knowledge/acquisition?
2. How does this knowledge come into being?

These two questions summarize the main debate in generative- and cognitive-linguistics driven SLA research, and finding answers to cover both the property and transition aspects of the questions is still the main focus in SLA research (see e.g. Piske and Young-Scholten, 2009). The present study attempts to answer the second question by examining the processes that contribute to effectiveness of input and the factors that contribute to attainment using user-behaviour tracking technology. The study is thus not concerned with the sort of knowledge learners come to acquire nor the question of whether the learners ultimately acquire the target construction; it is rather aimed at exploring how tracking user behaviour in a technology enhanced environment can contribute to the understanding of the processes involved in acquisition. A quick review of the approaches adopted by SLA researchers to answer the previous two questions is necessary here. These approaches are briefly summarized in the next section.

1.4 Approaches to SLA

Reviewing the literature on SLA, the following main approaches can be identified: generative, cognitive, functional/pragmatic, interactionist and sociocultural (see Fig. 1.1 below). Generative SLA focuses on property theories while cognitive SLA focuses on the transition theories. Both approaches, though, concentrate on the internal aspects of SLA. On the contrary, sociocultural SLA is more concerned with the external social and socio-cultural aspects of the SLA process and particularly with SLA use.

Generative SLA researchers argue that Universal Grammar (UG) is the only well-developed property theory available (Eubank and Gregg, 1995; Gregg, 2003, Slabakova, 2009). As noted above, UG-based language acquisition theories describe the elements and characteristics of the LAD (Language Acquisition Device) that is argued to control child language acquisition as well as second language acquisition (e.g. Gregg, 1993, 1996; White 1989, 2003). Here researchers have been mainly concerned with whether learners have access to UG and how they reset parameters when acquiring new languages. UG primarily addresses the 'logical problem'[2] of SLA rather than the 'developmental one'[3] (Pienemann, 2005, p.36). That is, generative SLA researchers do not claim to provide an account of the mechanisms involved in the SLA process so they do not commonly seek to include a transition theory. Carroll (1996, 2001), Gregg (1996) and White (2003) have all pointed out the generative SLA is far from being able to provide a transition theory and do not deny the importance of a developmental account of SLA. Gregg (1996) emphasizes the importance of such a developmental account and states that a transition theory should describe how linguistic input brings about grammar change. White (2003) claims that Carroll's attempt to construct a transition theory in her 1993 and 2001 publications is perhaps the most detailed one in generative SLA.

As noted above, cognitive SLA has been more concerned with how knowledge is represented and processed mentally, i.e. it falls into the transition theory category. Cognitive SLA researchers believe that in order to understand how learning takes place, we need to look at how the brain process and deals with information in real time. Central questions in cognitive SLA are what strategies learners use to deal with input in real time and why some learners are better than others. Mitchell and Myles (2004) divide cognitive SLA into two approaches, the *processing* approach, which focus on how learners process linguistic input and which includes researchers such as Pienemann, Towell and Hawkins, and the *emergenist* or *constructionist* approach which focuses on how language acquisition is driven by usage. This includes researchers such as N. Ellis, MacWhinney and Tomasello. Although cognitive SLA is mainly concerned with transition theory rather than property theory, it is widely argued that Herschensohn's (2000) *constructionism* provides both representational and developmental accounts of L2.

[2] Input alone cannot account for the infiniteness and variance of utterances. It is too inconsistent and incomplete to determine acquisition.
[3] Why are some features acquired earlier than others? How does acquisition take place?

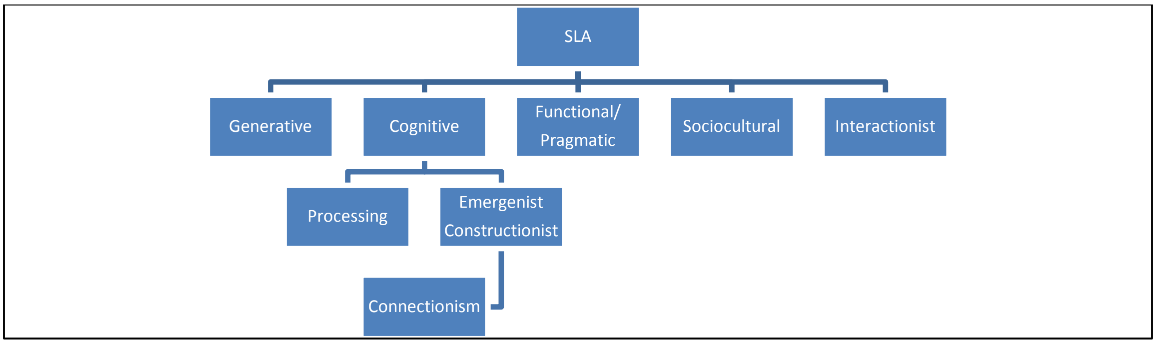


Figure 1.1 Approaches to SLA (based on Mitchell and Myles, 2004)

Functional or pragmatic approaches to SLA are more concerned with the learner's interlanguage rather than the linguistic system as is the case in generative and cognitive approaches. Functional researchers are interested in how learners make meanings within their immediate social, physical and discourse contexts (DeKeyser, 1995, 1997, 2007, N. Ellis, 2003, Tomasello, 1998, 2000, 2003). A central question in functional approaches is how form and function relate to one another (Gass and Selinker, 2008)

Sociocultural approaches focus on the learner as a social individual. The central assumption in socio-cultural theory is that of mediation. Socio-cultural researchers argue that all human activities are, in fact, mediated by higher level mental artifacts (see Vygotsky 1962, 1978; Lantolf, 2000, 2006; Lantolf and Thorne 2006; Zuengler and Miller, 2006). In this sense, language acquisition is believed to be socially mediated and dependent on face-to-face interaction and problem solving. (See Lantolf, 2000, Swain et al., 2003; Lantolf and Thorne, 2006; Zuengler and Miller, 2006 for further information about the socio-cultural approaches.)

Similar to sociocultural approaches, interactionist approaches to SLA view language in social terms. The main focus of the interactionist approach is how language use contributes to language development. Interactionist researchers place emphasis on the role of input and output in interlanguage development. The interactionist hypothesis (Long, 1981, 1983, 1996) and the output hypothesis (Swain 1985, 1995) reflect the main concepts of the interactionist approaches. Some of these hypotheses will be discussed later in Chapter Two when reviewing the literature on the role of attention and noticing in ISLA.

Having established the basics of each of these approaches, it can be argued that although different in their focus, both generative and cognitive SLA approaches focus on individual learners, unlike the sociocultural, functional and interactionist approaches which focus on the learner as a member of society. What is noteworthy here is that these approaches do not exist as separate entities; they overlap in most models of SLA. Furthermore, generative and cognitive SLA studies can be classified under psycholinguistic studies while sociocultural and functional SLA fall under sociolinguistic (see section 2.4.1 about Ellis's discussion of research in ISLA).

The importance of these approaches to the present study lies in the different roles that each approach assigns to input and input processing. This will be discussed in Chapter Two. Also, as the present study is concerned with examining the processes that learners exhibit when dealing with different types of input and also with which type of input is most effective in a TELL based environment, it draws more from cognitive approaches to SLA. It is expected that the type of data that is collected in this study using user behaviour tracking technologies will provide valuable insights into the learning processes which consequently will implicate SLA and in turn pedagogy. Sections 1.6 and 1.7 elaborate on the central concepts of competence and performance and variability in attainment.

1.5 Competence and performance

One of the major distinctions in generative linguistics and in generative SLA is that of *competence* and *performance*. Chomsky (1965) differentiates between underlying internal knowledge which he called *competence* and the manifestations of this knowledge through utterances and written production which he called *performance*. The distinction is widely accepted in generative SLA research, although with modification in some cases (see the variable competence models Fulcher, 1995; Tarone, 1987, 1985, 1983).

The competence-performance distinction is crucial when collecting data in SLA research. Some SLA researchers, mainly generative ones, argue that any research aiming at testing or falsifying a particular hypothesis about the nature of the language system should aim at testing the learners' competence rather than the performance. However, competence can only be accessed indirectly and only under controlled conditions through particular tasks such as grammaticality judgments. Performance data is an imperfect representation of internal knowledge (Mitchell and Myles, 2004). However, this is not accepted by all researchers; competence may differ from performance but because competence generates performance, we can use performance to study and describe competence. Cognitive SLA researchers argue that one way of getting closer to competence is by looking more closely at the processes that learners exhibit while dealing with input. These processes indicate if learners are drawing from their internal automatized knowledge (associated with competence) or temporary storage (associated with performance) (see section 2.3.1). On the other hand, functional and interactionist SLA researchers believe that language should be studied through language use and production. They argue that performance varies according to the conditions that are in operation at the time of production. This distinction will be revisited in Chapter Two when discussing the role of noticing. What is noteworthy here is that although accepting the distinction between competence and performance, the present study does not seek to investigate the type of knowledge the learners draw from when dealing with the input presented to them in the experiment, as interesting as this question is. The present study looks at what contributes to the effectiveness of input in instructed environment as determined by improvement in their performance on a specific construction. As R. Ellis stated that when researching educational issues, 'it is not the learners' competence that is important but their proficiency' (Ellis, 1994, p.156). Therefore, the focus of the present study is learners' performance, which is, however, taken to represent their language knowledge under the internal and external conditions operating at the time of data collection.

The next section highlights some of the factors that are held to lead to variability in performance.

1.6 Variability in L2 attainment

One perplexing issue in SLA research is that of variability. Two types of variability have been the focus of SLA research: variability within the learner's interlanguage and across learners' interlanguages. Questions have been raised about why the same learner's performance varies and also about why learners vary in how successfully they learn a second language even when they come from the same L1 background, are the same age and acquire the target language under the same conditions. The focus of the current research is on such differences across learners. Different accounts have emerged over the years to explain variability. Some linguists associate it with linguistic or cognitive differences. Others argue that it is misleading to claim that variability is the result of linguistic or cognitive differences only; many social, cultural, and other factors play a role.

Factors that have been identified to contribute to variability are age, gender, culture, length of exposure to the language, type of input, instruction, attention, working memory, processing, learning styles, learning strategies and affective filters, among others. (See Skehan, 1989; Ellis, 1994; Robinson, 2002; Dornyei 2005, 2006; Dewaele, 2009 for further discussion.)

One of the central accounts of variability across learners is the environment in which the language is acquired, i.e. the difference between uninstructed and instructed acquisition and in how input is presented; this is the focus of the present study. This should not be confused with Krashen's distinction between *acquisition* and *learning* (see section 1.2.2 above and 1.8 below). The uninstructed/instructed distinction is mainly concerned with whether the language is acquired under formal or naturalistic conditions.

As the current study was conducted in a technology-enhanced learning environment, it is based in the instructed second language acquisition field (ISLA). Therefore, the focus of the discussion in the following chapters is only on the factors that are relevant to instructed learning conditions. In the process of identifying the framework for the study, the next section will define what ISLA is.

1.7 Instructed Second Language Acquisition/ ISLA

ISLA is defined as 'any systematic attempt to enable or facilitate language learning by manipulating the mechanisms of learning and/or the conditions under which these occur' (Housen and Pierrard, 2005, p.2). Going back to Krashen's (1981, 1982, 1985) distinction, ISLA appears to be what Krashen refers to as *learning*. An underlying assumption is that one needs to be aware of the process of learning to manipulate it. However, and as stated above, , the distinction between conscious learning and unconscious acquisition is not adopted in the present study for the simple fact that unconscious acquisition could take place in the classroom where learners are consciously attending to the language system (see section 2.2.4 about the role of noticing). As already noted, a distinction is made, though, between SLA which occurs in natural settings and ISLA which takes place in formal settings. Since the term was first used in the1980s, ISLA research has focused on identifying the conditions that best facilitate language learning and result in ultimate acquisition of the target language in formal settings. ISLA is the link that connects theory to practice, SLA to language teaching. The main focus of ISLA is the role of instruction in SLA. However, researchers have still to agree on whether instruction is effective in promoting language learning (Ellis, 2008). Chapter Two will provide detailed view of the debate surrounding the role of input and instruction in the learning process.

1.8 Organization of the thesis

This chapter aimed at providing a brief introductory view of the main theoretical concepts underlying the current research. Most of these concepts will be revisited in chapter two and three in more details.

Briefly, the present study is cognitive in nature and aims to add to our understanding of the learning process in general and in a TELL setting in particular. In theoretical terms, the present study explores what triggers learning in a TELL environment and how our understanding of the SLA processes can be enhanced through technology. In more pedagogical terms, the study investigates what works in a self-accessed TELL environment. The present study cannot be classified as an SLA study since its focus is not on whether the learners acquire the construction or not but rather on how they deal with the input in their attempt to acquire the target construction. It is, however, worth noting here that in order to understand what features/factors/processes contribute to performance in a TELL environment, it is important to examine performance to allow comparison among different input types.

The thesis is organized into seven chapters. Chapter One provides preliminary insights into the theoretical field the research is conducted in. Chapter Two presents an in-depth review of input, the role of instruction and attention in learning, the teaching methodologies informing the design of the materials and the reasons for conducting the research in a TELL environment. Chapter Three is a review of the related literature on the technological aspects of the study. Chapter Four details the design, data collection and methodology adopted in the present study. Chapter Five presents the analysis of data and results of the study. Chapter Six summarizes the findings in relation to the literature and finally Chapter Seven draws a number of conclusions and implications, presents the limitation of the study and provides directions for further research.

Chapter Two

Instructed Second Language Acquisition and Task-Based Language Learning

2.1 Introduction

This chapter provides a detailed review of the studies and theoretical concepts underpinning the present study. As discussed in Chapter One, the present study examines the effectiveness of different types of input in a Technology-Enhanced Language Learning (TELL) environment and aims to identify the factors that account for why one type of input might be more effective in facilitating language learning in a TELL environment. The chapter is organised as follows: Section 2.2 details the issues related to input in ISLA and section 2.3 examines the role of instruction. The discussion turns, then, to the effectiveness of different types of input in section 2.4. The relationship between Instructed Second Language Acquisition (ISLA) research and teaching methodology is discussed in section 2.5; here Communicative Language Teaching/ CLT and Task-based Learning and Teaching/TBLL are the focus. The principles and main issues surrounding the CLT and TBLL are described. I turn then to discuss research on individual differences in relation to performance in section 2.6 and finally section 2.7 presents the case for conducting the present study, in a TELL environment.

2.2 Input in ISLA

In the context of on-going debate in the SLA field, there is a consensus on one point: to learn a language, you need exposure to that language (Gass, 1997; Ellis, 1997; Carroll, 2001; Gregg, 2001). The verdict is still out, though, on the characteristics of input that best facilitate language learning. Gass (2003) has identified three types of language data that learners have access to: positive, negative and indirect negative evidence. Positive evidence is any form of language the learner is exposed to that does not involve metalinguistic information such as correction. Negative evidence is the information provided to learners concerning their incorrectness. Indirect negative evidence[4] is the correctness of learners' forms without providing direct instruction, which is also referred to as implicit negative evidence. The next section elaborates on what is meant by input.

2.2.1 What is input?

First, being the focus of the present study, it is important to define what is meant by input within the field of SLA. Input has been defined as the 'samples of language to which a learner is exposed' (Ellis, 1997, p. 5). More recently, Barcroft and Wong (2013) define *input* as the meaningful samples of a target language to which a language learner is exposed in a meaningful context. This definition is limited in the sense that it excludes any naturalistic input and classroom input that learners do not understand. It is easy to assume then, based on these definitions, that any auditory or visual L2 data learners encounter in or outside the classroom is *input*. However, not all researchers adopt this wide-ranging definition. Sharwood Smith (1993) argues that *input* is 'the potentially processible language data which are made available by chance or by design, to the language learner' (Sharwood Smith, 1993, p.167). Sharwood Smith's definition does not differentiate between positive and negative input but it, interestingly, links input to processing. On similar grounds, Carroll (2001) rejects use of the term *input* to refer to the raw material from which learners learn. For her, there is a difference between stimuli which are the 'observable instantiations' of the second language; and input which is the processed stimuli that enter the brain. She argues for a more cognitive-based definition and uses the term input to refer to 'any mental representation which is analysed by a processor' (Carroll, 2001, p. 16) that is, after the raw acoustic stimuli have been received by the learner. Carroll's definition treats input not as physical information but as a mental system.

The terms *input*, *positive evidence*, *intake*, *stimuli* have all been used to refer to the visual and auditory L2 data that learners encounter in or outside the classroom. And in addition to the varying definitions of input alone referred to in the previous paragraph, these terms cannot be used interchangeably as they carry different connotations. For example, input is not only restricted to positive evidence, for the simple fact that

[4] The final category differs from Chomsky's (1981) indirect negative evidence which he argues is information about which sentences have not appeared in the input. Gass' definitions are referred to here as they match the types of input in the present study. For further discussion of other definitions, see Chomsky 1981, Schwartz 1993).

linguistic communication by nature includes instances of correction and repair whether it is in L1 or L2 acquisition or among native speakers. The importance of defining the term input is highlighted by the fact that most ISLA researchers distinguish between *input* and *intake* where *intake* refers to the processed input (Corder, 1967, 1978; Krashen, 1981; Gass et al. 1998; VanPatten 1993, 1996). This distinction was first introduced by Corder in (1967) where he called data that is available for learner *input* and what goes in the mind *intake*. This distinction will be revisited when discussing the role of noticing in section 2.2.4. For the time being, it is worth noting that input has also been defined and classified based on the orientation of the research (see Fig 2.1); for example, distinctions are made between, *authentic* and *modified* input, *simplified* and *elaborated* input, *explicit* and *implicit* input.

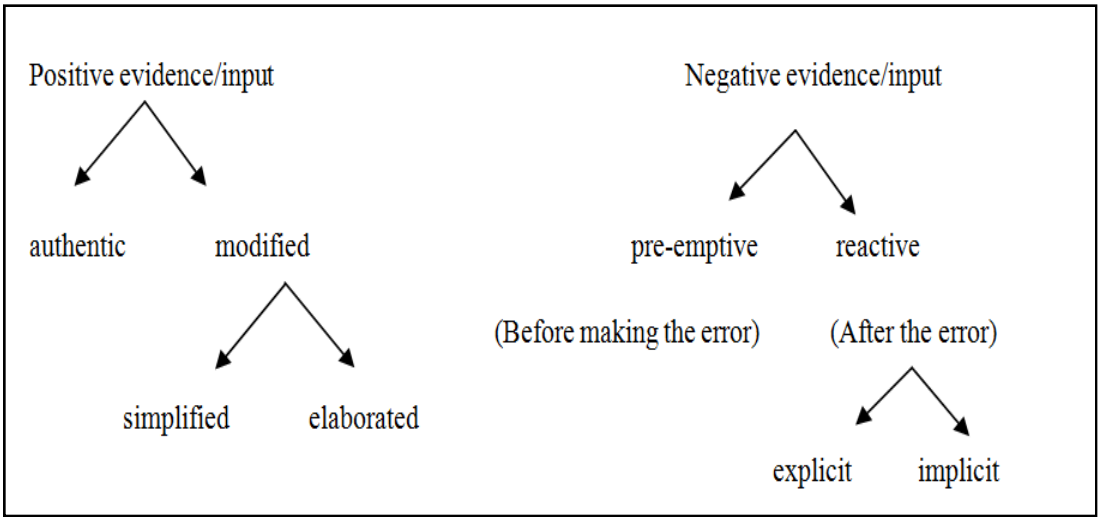


Figure 2.1: Types of input/evidence in SLA

While a variety of definitions of the term *input* has been suggested, the working definition adopted in the present study is Reinders' (2012) definition of input as *linguistic data*, mainly because this all-inclusive umbrella definition comprises all the types of input including the types used in the present study.

Having established definitions and connotation of the term *input*, the next section aims to review studies of the role of input in language learning to better ground the reader understands of the theoretical underpinnings of input-based second language acquisition theories which will be discussed later in section 2.2.3.

2.2.2 Role of input

Theories of second language acquisition have changed dramatically over the years and, as suggested above, different approaches to SLA have assigned different roles to input. However, it is not only the role of input that has been at the centre of research. There is an on-going debate on what type of input, how important each type of input is and the amount of input that should be provided to learners. In the middle of the twentieth century, second language learning approaches echoed the views of the prevailing psychological approach, Behaviourism. Fries (1957, p. vii) viewed language learning as habit formation based on associations that stemmed from the input. At this stage, input was of supreme importance as the main source of what was imitated by the learner. Later, when Behaviourism was rejected in SLA, the role of input oscillated between learners receiving sufficient comprehensible input (Krashen, 1985) to engaging in interaction to obtain input (Long, 1981, 1991; Pica, 1994), to receiving modified input designed to facilitate processing (VanPatten, 1990, 1996). Krashen's (1985, 1994) 'Comprehensible Input' model has had a predominant influence on both second language teaching practice and on theories and hypotheses of SLA. According to Gass (1997), Krashen's model assumes a central, dominant role for input, while later studies placed within a Universal Grammar framework assume less important but still central role for input. UG researchers claim that input by itself is insufficient for acquisition to take place, mainly because input does not include all information that learners need to acquire a language[5]. Therefore, other mechanisms, i.e. *Universal Grammar*, are operating (see e.g. White, 1989). On the other hand, Krashen (1982, 1985) argues that as long as the input is at the right level, i.e. i+1, it is sufficient (see section 2.2.3 below for discussion of Krashen's model). While Krashen (1994) focused merely on one-way comprehensible input, others, such as Long (1985) and Pica (1994), have taken an interactionist position in emphasising the role of two-way communication. Input in the interactionist approach also holds a central role as the main assumption is that learners will acquire L2 through interacting and noticing certain aspects of the input (see section 2.2.3 below for full discussion). On the same grounds, VanPatten's Input Processing assumes a fundamental role for input. For him, input should be modified by the teacher/materials developer according to certain principles to allow maximum processing of the target structures (see section 2.2.3 below for full discussion; also see

5 This assumption is referred to as the 'logical problem' of language acquisition (Bley-Vroman, 1989).

VanPatten and Cadierno, 1993; VanPatten 1996, 2002, 2003, 2004; VanPatten and Oikkenon, 1996).

In brief, over the years, SLA theory under any approach has assumed an important role for input but these approaches differ in their assumptions about the quality and quantity of input and what is, in addition, required. This section has provided a general picture of the role of input in SLA research. In the next section, prominent input-based approaches will be discussed in more detail.

2.2.3 Input based approaches

As mentioned in section 2.2.2 above, theories, models and hypotheses accounting for input include Comprehensible Input, the Interaction Hypothesis, and Input Processing among others.[6] Following on from Corder's (1967) distinction between *input* and *intake* discussed above, Krashen made a further distinction between the resulting knowledge from different types of input, that is, between what learners *learn* and what they *acquire* as part of their linguistic competence. For Krashen, learning is a conscious and planned process and leads to metalinguistic knowledge while acquisition is an unconscious and natural process and leads to implicit knowledge. On this basis, Krashen argued that comprehensible input triggers acquisition and that formal instruction does not trigger acquisition but instead involves learning. For Krashen, the Comprehensible Input which leads to acquisition is i+1, i.e. input that represents structures slightly above the learner's current competence level. Krashen did not assign input any role for production, for output, or processing. In the 1970s and 1980s, his model underwent criticism by UG SLA researchers for being vague since i+1 cannot easily be measured and since input which leads to acquisition might not be comprehensible (see White 1987) by instructed SLA researchers for downplaying the role of instruction. One of the main shortcomings pointed out of Krashen's Comprehensible Iinput hypothesis were his views of acquisition and learning as one-way processes where the learners are passive recipients of input.

In response, in 1981, Long introduced several new ideas. First, he argued that comprehensible input is essential but he criticised Krashen's hypothesis on the basis

6 Other theoretical approaches focus on the role of input in acquisition such as Newport's (1990) Less is More hypothesis or Pienemann's (1984, 1985,1988, 1989, 1998) Processability Theory. However, these are not discussed in the present study as they either focused on first language acquisition (Newport's) or on natural rather than instructed input (Pienemann's).

that it failed to account for the importance of the social aspect of the learning process, i.e. as a two-way process. Long (1981, 1996) highlighted the role of interaction in this process in his *Interaction Hypothesis*. Contrary to Krashen, Long views learning as one in which learners are active participants who negotiate meaning and form to arrive at better understanding of the input for the speaker and the output for the listener. Long points out that in their attempt to make form-meaning connections, learners are actually exposed to more comprehensible input.

Schmidt (1990, 1993, 1994) takes learners' attempts to make form-meaning connections further in arguing that in order for input to become intake, noticing must take place. In his *Noticing Hypothesis,* he proposes that learners must notice the forms in the input in order to acquire them. Schmidt (2010) argues that there are many factors that lead to learners noticing or not noticing the linguistic features in the input. These include motivation, aptitude and language learning history. It is important to note here that Schmidt does not claim that noticing will definitely lead to acquisition but that it is an essential part of the process. The role of noticing is detailed below when discussing input processing (see sections 2.2.4 and 2.2.5).

In 1996, VanPatten suggested that although interaction is important to make form-meaning connections, input should also be modified through the use of materials that maximise the opportunities of forming these connections and consequently transferring *input* into *intake*. In his publications at the time and subsequently (VanPatten and Cadierno, 1993, VanPatten 2002, 2003, 2004, VanPatten and Oikkenon, 1996), VanPatten presents the details of his evolving *Input Processing* theory and offers principles that should be implemented to prompt the transfer of input into intake.

Gass (1997) elaborated on VanPatten's *Input Processing* in her *Input, Interaction and Output* model (IIO) to include the interaction aspect of the acquisition process. For Gass, acquisition takes place through a number of stages starting from raw input and ending in output. Figure 2.2 below illustrates Gass' model of the acquisition process.

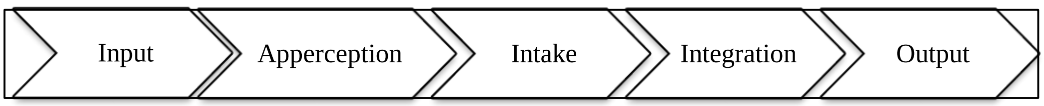


Figure 2.2: Input, Interaction, Output Model (based on Gass, 1997)

In the IIO model, comprehension (apperception) precedes intake and if the input is not comprehended, it does not become intake. Various filters are also in operation at the different stages; for example, frequency and saliency play major roles in apperceiving the input while the learner's personality and attitude affect output, which is an essential part of the acquisition process in her model.

Swain (1985, 2000) argues that output should come at an earlier stage in the process of acquisition; learners have to be pushed to produce output that forces them to utilize the structures they have noticed in the input if they want to produce more target-like utterances.

This section provided an overview of the major and multifaceted accounts of input-based approaches. There is still continued debate on the type of input and conditions which best facilitate acquisition, and in the next section, the role of noticing on the learning process is examined.

2.2.4 The role of noticing

A conclusion that much SLA points to is that even when learners are provided with input, regardless of the type and amount, it is not guaranteed that acquisition will take place. As stated in the previous section, many researchers agree that acquisition takes place when input becomes intake. There is an unvoiced consensus among instructed SLA researchers that greater attention results in more learning. Most of the models and theories mentioned in section 2.2.3 above include a role of noticing, though sometimes not stated explicitly. Yet, noticing is not a straightforward concept. The term *attention*, which is associated with noticing in I/SLA research, embodies a multitude of concepts including awareness, alertness, consciousness and noticing. It is necessary, then, to define these terms.

For Schmidt, noticing is attention and it 'necessarily entails conscious registration of the contents of focal attention' (Schmidt, 1994, p.17). He argues that in order for input to be processed and transformed into intake. Learners need to notice or attend to the linguistic forms in the input and process them (Robinson, 1996, Schmidt, 1990, 1994, 2001). According to Schmidt, noticing is conscious and is the driving force of learning; one can distinguish between intentionality, attention, awareness and control. Here he separated consciousness as *awareness*, which refers to different levels of awareness, and consciousness as *control*, which refers to automatic and controlled processing (Schmidt 1990, 1993, 1994, 2001)

Tomlin and Villa (1994) divide attention into alertness, orientation and detection. Detection, which is their equivalent of noticing, is a process 'by which particular exemplars are registered in memory' (Tomlin and Villa, 1994, p.192). Unlike Schmidt however, they argue that attention does not necessarily involve conscious noticing. They assert that there is enough evidence from studies to show that 'information can be cognitively detected even though the individual is not aware of it having occurred' (Tomlin and Villa, 1994, p. 193). Reviewing both arguments, Robinson (1995) combines both of these approaches to noticing and maintains that detection takes place first and then encoding by allocating the attentional resources. At that point noticing happens. Conversely, VanPatten (1996) rejects Tomlin and Villa's argument and undermines their conclusion on the basis that their learners were not involved in processes where they need to make form-meaning mappings, an essential element in the learning process according to VanPatten. Based on the definitions proposed by Tomlin and Villa (1994), Gass (1988) and Robinson (1995, 2003, 2005), Schmidt (2001) adopted a more restricted approach to noticing that does not involve awareness.

The importance of noticing in the learning process is evident in all prominent models of instructed SLA. Gass's (1977) *IIO model* involves an apperceived input stage (see above), similar to Tomlin and Villa's detection stage. Gass and Selinker (2001) point out that learners need to notice units (forms, words, sounds, etc.,) in order for input to become intake. Once the forms are part of the intake, this means that the forms are now internalized / automatized and can be used to form new structures later (in Gass's integration stage). In the *Interaction Hypothesis*, learning is driven by interaction, but the idea is that L2 learners are forced to notice the forms in the input when they are required to negotiate meaning with other learners in order to overcome a communication breakdown, complete a task or solve a problem. This is well documented in the literature that advocates the importance of the negotiation of meaning in noticing and consequently acquiring L2 forms (Long, 1981, 1990; Pica, 1994; Gass, 1997; Doughty, 1998; Blake, 2000, 2005). Finally, VanPatten's *Input Processing* is based on the principle that to learn a form, you must detect it in the input you hear and/or read, i.e. connect the form to a meaning or function (see section 2.2.5 below for further discussion).

Gass and Selinker (2008) claim that noticing has become one of the vital constructs of SLA and the relationship between noticing and learning has now been widely investigated in SLA (Shook, 1994; Leow, 1997, 1998, 2000, 2001; Wong, 2001; Gass, Svetics and Lemelin, 2003; Rosa and Leow, 2004; Williams, 2004). Researchers investigating the role of noticing in the learning process focus on input processing and the difference between input and intake. Some of the models that account for the role of noticing are reviewed in section 2.2.5 below when reviewing input processing.

In brief, the role of noticing, conscious or unconscious, in second language acquisition is claimed to be crucial and cannot therefore be overlooked, particularly when looking at input processing - the focus of the next section. Many factors contribute to whether learners notice the forms in the input or not. One of these factors is the type of instruction; this will be examined in detail in section 2.3. At this point, I can now state that noticing will be used in the present study to refer to the learner attending to the surface forms in the input regardless of whether attention is conscious or unconscious.

2.2.5 Input processing in ISLA

The relationship between input and production is not a straightforward one. One of the most debated questions in SLA is 'why would learners process some elements in the input and not others?' Input processing has been the focus of studies tackling this question from cognitive and psycholinguistic perspectives. In fact, many researchers (Krashen, 1982; Gass, 1988; Carroll, 2001) would argue that input processing is the most important component of the learning process though their views vary on whether processing is conscious or unconscious with respect to acquisition versus learning. Psycholinguistically-informed SLA research has greatly increased our understanding of the learning process, yet it is still the case that 'little is known about how language learners comprehend or produce language in real time' (Clahsen and Felser, 2006, p.3). Reviewing the literature on grammatical processing in SLA in a series of articles (Clahsen and Felser, 2006 and Fesler and Clahsen, 2009), they note that many researchers have proposed models about grammatical processing in language learners (Clahsen, 1984; Anderson, 1993; VanPatten, 1996, 2004; Pienemann, 1998; Carroll, 2001; Hulstijn, 2002; Gregg 2003), and they argue that these models provide little psycholinguistic evidence on the mechanisms learners employ to process the target language in real time, a main concern of the present study.

One approach researchers adopt to look at the processes involved in L2 learning is to compare them to L1 processes. Some L2 processing studies suggest that in their processing strategies L2 learners are influenced by their L1 and they do not use processing strategies used by speakers of the L2 (Frenck-Mestre and Pynte, 1997; Juffs, 1998, 2005; Frenck-Mestre, 2005). For example, Frenck-Mestre's results indicate that English and Spanish learners of French at the beginning level resort to L1 strategies when processing sentences with relative pronouns. However, other studies show that there is no influence of L1 processing strategies on the processing of L2 (Schachter, 1989; Johnson and Newport, 1989, 1991; Williams, Möbius, and Kim, 2001; Felser, Roberts, Gross, and Marinis, 2003; Papadopoulou and Clahsen, 2003; Roberts, Marinis, Felser, and Clahsen, 2004). Other studies reveal that L2 learners behave in a very similar way to native (L2) speakers when processing syntax while reading (Juffs and Harrington, 1996, Juffs, 1998). As opposed to the UG view which sees L1 and L2 acquisition involving linguistic mechanisms, there are also researchers who argue that language acquisition, whether L1 or L2, is driven only by general learning mechanisms (N. Ellis 1994, 2001; Robinson 1996, 2001, 2003; DeKeyser 2000, 2003; McDonald 2006;). VanPatten and Keating (2007) argue on the basis of their study of tense, for example, that L2 learners start with universal processing principles based on general cognitive mechanisms rather than linguistic mechanisms and that learners do not rely on L1 processing strategies. This and other differences in research findings regarding processing could be a result of the specific form being studied. For example, tense is assumed to be simpler to process than relative clauses, and it could be the case that learners start with universal processing principles when processing simple forms and then move on to use L1 processing strategies with more complex forms (Gass and Selinker, 2008). The difference could also be related to difference in the L2 proficiency levels of subjects, with the assumption that more advanced learners resort to L2 processing strategies or vice versa.

As important as the link between L2 processing and L1 processing is, the attention in this section is directed to L2 input processing and the models and theories that attempt to describe it, particularly those that focus on instructed input which is the focus of the present study.

Carroll (2001, 2006), Pienemann (2007), Schmidt (1990), Sharwood Smith (1993) and VanPatten, (1996, 2004) are among the researchers who have looked at L2 learning processes. Carroll's Autonomous Induction Theory (2006) aims at explaining how internal mental structures are responsible for learners' interlanguage development. It focuses on naturalistic rather than instructed learning contexts, however, and is therefore of limited relevance to the present study. Pienemann's Processability Theory traces the sequence of acquisition by examining the language processor (Pienemann, 2007). Similarly to Carroll's model, it also focuses on naturalistic rather than instructed contexts. Truscott and Sharwood Smith's (2004, 2011; Sharwood Smith and Truscott, 2014) processing model is of greater relevant. Their Modular Online Growth and Use of Language (MOGUL) model is an account of processing in real time, They argue that there is one unified processing system for L1 and L2 and, importantly, for both naturalistic and instructed learning contexts. Within this system, linguistic features in the input compete for attention, and only items that are activated in the system are acquired. Under this model, priming (input enhancement; see Sharwood Smith 1981, 1993) aids acquisition by raising the potential of activation of the target items, i.e., the chances of the learner noticing the form and consequently processing it. Sharwood Smith (1981, 1993) and Rutherford and Sharwood Smith (1985) argue that learners' attention should be directed to form by raising their awareness of the target features in the input. Examining Sharwood Smith's claims was Fotos (1993, 1994) who conducted several studies to examine the effect of conscious raising on input processing. She used different pedagogical techniques (e.g. learners working in small groups) to raise the learners' awareness. The results suggest that learners are more likely to notice target features in consciousness raising activities. In her studies, the consciousness raising group performed as effectively as and even better than the other experimental and control groups.

In 1991, VanPatten first introduced his theory of Input processing (IP). He states that L2 learners usually process the linguistic input by making form-meaning connections. He defines IP as 'the process of making form-meaning connections from the linguistics data in the input for the purpose of constructing a second language linguistic system' (Lee and VanPatten 1995 p.96). VanPatten argues that instruction should force learners to process the input, i.e.to make form-meaning connections. This is particularly important if the target forms are redundant for communicative purposes. For example, learners do not normally need to process the *–ed* in the following sentence to know that it happened in the past since the adverb yesterday conveys this meaning *He cleaned the car yesterday*. IP involves deleting *yesterday* and providing a context which conveys the meaning of the sentence, thus forcing learners to process the *–ed* form in the input.

VanPatten points out that learners usually process input for meaning before form; he refers to this as the *primacy of meaning principle* (VanPatten, 2004, Lee and Benati, 2007). VanPatten argues that in most cases, L2 learners are able to derive meaning from the input without attending to form unless they are pushed to encode the input for linguistic features, they will not do so.

Both VanPatten and Truscott and Sharwood Smith base their models in generative SLA and assume a central role for linguistic mechanisms/UG. Schmidt (1990), among others, examines processing not from a linguistic but from a general cognitive perspective. Schmidt (1990) identified two types of processes, conscious and unconscious. Unconscious processes 'are not under voluntary control and are difficult to modify, but are fast, efficient and accurate' (p138). The same definition is echoed in what has usually been referred to as automatic processing (McLaughlin, 1987). Conscious processes are defined by Schmidt as 'the experiential manifestation of a limited capacity central processor' (p.138). In this sense, they are similar to controlled processing which is slow and often deliberate. The discussion on conscious and unconscious processes is often linked to the type of knowledge that learners utilize during processing. In brief, conscious processes are typically associated with explicit knowledge while unconscious processes are linked to implicit knowledge. As a central concept in the understanding of input processing, the implicit/ explicit knowledge/learning distinction is discussed in the next section when reviewing the literature on the role of instruction (see N. Ellis, 1994, 2005; Paradis, 2004, 2009; Hulstijn 2002, 2005 for further discussion). Briefly, the preliminary picture that emerges from the literature is that drawing on implicit knowledge is assumed to be fast and automatic while drawing on explicit knowledge is slower and more controlled (Segalowitz and Hulstijn, 2005).

The discussion above reveals one of the perplexing questions in SLA: Why do learners process certain linguistic features which might not be salient (in terms of frequency or perceptual enhancement) in the input but fail to process others that are supposedly more salient? Factors that have been identified to play a role in whether items are processed or not are consciousness, noticing, type of input, priming of the input (e.g. MOGUL; see above), modality of the input and instruction (see below) . However, Clahsen and Felser point out that 'further investigation of grammatical processing in language learners is necessary before any firm conclusions can be drawn' (Clahsen and Felser, 2006, p.35).

With respect to the application of processing research to instructed SLA, Ringbom and Jarvis (2009) note that 'much more knowledge is needed about the mechanism through which language learning proceeds before the field is justified in pronouncing definitive statements about how languages can be taught most effectively' (Ringbom and Jarvis, 2009, p. 114).

The present study uses user behaviour tracking technology to examine how learners deal with linguistic input in a Technology-Enhanced Language Learning (TELL) environment and it aims to add to our understanding of the mechanisms language learners employ to process linguistic information in the input. As we shall see, an important feature of a TELL environment and one which motivates the present study is that input can be presented using different media, video, audio, graphics, etc., factors that considerably affect the processing of input as will be shown in section 2.2.6 .

2.2.6 Modality of input

Most of the accounts of L2 input processing do not say much about the medium in which input should be presented to the learner in an instructed context. However, evidence from cognitive psychology strongly suggests that input modality affects processing, comprehension and memory (Treisman, 1969; Shaffer, 1975; Rollins and Hendricks, 1980; Penney, 1989; Beaman, 2002; Beaman and Morton, 2000; Bird and Williams, 2002). Research investigating the modality of input in instructed language learning is based largely in language pedagogy and Computer-Assisted Language Learning (CALL) research. The picture that emerges from studies on input modality from a methodological or pedagogical perspective is that the use of different modes has many advantages. These include more exposure to authentic input, increased learner motivation and better fit for different learning styles (Brinton, 2001). Pedagogically-based research on input modality is not reviewed here. Most of these studies focus on looking at the advantages of presenting linguistic information in more than one medium, but this is not the focus of the present study; rather the focus is on what elements of the input provided in different modes – not in combination- affect performance and processing.

Studies that have looked at input modality in L2 research include Lund (1991); Johnson (1992); Leow (1993, 1995); Murphy (1997); Brinton (2001); Wong (2001); Nassaji (2004); Plass and Jones (2005); and Sydorenko (2010). As the modality of input is one of the components of the learning process the present study focuses on, I review some of these studies in detail here.

Johnson's (1992) study is one of the earliest experiments on input modality in L2 research. She examined the effects of modality on ESL learners' performance on grammaticality judgment tasks which were used in an earlier study (Johnson and Newport, 1989) in aural mode. The study revealed that adult L2 learners scored much higher in the same tasks when they were presented in the written mode than in the aural mode. Similarly Murphy (1997) looked at the effect of modality on the performance on grammaticality judgment tasks and results showed that the aural group were slower and less accurate. Murphy (1997) pointed out that studies that present tasks in only one mode do not provide a complete picture of the processes and knowledge that are available to learners. Findings from Johnson (1992) and Murphy (1997) are compatible with those that emerged from Lund (1991) as they all show superiority of visual written input over the aural input in terms of recall scores and performance.

In 1995, Leow claimed that the role of modality surprisingly had not attracted much attention in SLA research. Learners are usually exposed to either written or aural input and even when sometimes both modes are used in a study, the role of modality was rarely investigated. Replicating his 1993 study, which focused on written input, Leow (1995) examined learners' intake of linguistic items, this time using aural input. He specifically looked at the effects of simplification of input, type of linguistic item and language experience. In his 1995 study, Leow found that learners attended more to the Spanish present perfect form than to the subjunctive form. This was not the case in his 1993 study where there was no difference amongst learners in relation to the linguistic items. His explanation was that in the aural mode, learners attend more to the phonological aspect of linguistic forms. He argued that learners' intake depends on cognitive constraints and strategies which operate and are employed differently in different modes. Based on this, he states that 'readers are generally regarded to be less cognitively constrained by their exposure to L2 data than are listeners' (Leow, 1995, p. 85). His findings highlighted the 'need for research to consider seriously the role of modality while addressing cognitive processes in SLA' (Leow, 1995, p.79). He also stressed that instructional approaches should look at the type of input, the cognitive processes and the modality in which input is made available to ensure learners' maximum benefit from their L2 exposure within the constraints of the classroom.

A more recent study on the modality of input in SLA was conducted by Wong in 2001. Wong argues that evidence from psychological research indicates that attentional constraints during input processing are not the same for the aural and written modes. Wong (2001) decided to replicate VanPatten's (1990) study where he concluded that learners find it difficult to attend to both form and meaning in aural input, more so when the linguistic form has no communicative value and is not essential for the understanding of the meaning. VanPatten (1990) also found that attending to both meaning and form in the input affects comprehension and that learners recall scores were significantly lower when they were asked to attend to both. Wong intended to see if these findings also hold for other modes of input, the written mode in particular. Results showed that performance was significantly better in the written task mode than in the aural mode, and that attending to both meaning and form in the aural mode affected comprehension while doing so in the written mode did not. This led Wong to conclude that attentional resources are not constrained in the same way during input processing in written and aural modes.

Nassaji (2004) examined the effect of different modalities on remembering name-referent associations. He divided his subjects into three groups exposed to input in either auditory only mode or visual mode only or both modes. The results showed significant differences between the single mode groups (whether visual or auditory) and the dual one, with learners being more accurate when exposed to input in both modes. Nassaji's findings are compatible with Plass and Jones' (2005) findings that a combination of print, audio, and imagery enhance input by making it more comprehensible.

Similar to Nassaji (2004), Sydorenko (2010) examines the effect of input modality on vocabulary acquisition but unlike Nassaji, she goes further than the effect of modality on recall and recognition to look at its effect on attention to input as well. Sydorenko's (2010) study involved three different modality groups: visual and audio, visual and written and visual, audio and written. The results indicate that the two groups that got written input scored higher on written recognition while the groups that received audio and visual input scored higher on aural recognition. The results also revealed that learners paid more attention to the written input, followed by the visual and finally the aural input.

To summarise the discussion, it is obvious from the studies above that the modality of input plays a central role in processing input: it affects learners' performance during testing and also during learning. The review of these studies also shows that the findings are inconclusive and that further research on the effects of input modality is required.

I now turn to another aspect of the learning process that has been argued to considerably affect input processing in SLA, namely, instruction itself.

2.3 The role of instruction in ISLA

As I have established from the discussion in the previous sections, despite the myriad possible factors affecting SLA, the need for sufficient exposure to input is central for language development. Although there is a consensus among SLA researchers about the importance of input, there is less clear agreement on the nature of that input and on how learners deal with it. Moreover, different SLA theories assign different degrees of importance to the role of instruction. These degrees range between the two extremes of having no role at all to being the most important element. Some researchers, mainly generative ones, have argued that L2 acquisition occurs in the same way L1 acquisition does so, in addition to learners' use of linguistic mechanisms, comprehensible input in the form of positive evidence alone is sufficient for both child and adult L2 acquisition (e.g., Krashen, 1977, 1994, Schwartz, 1993). Although Schwartz assigns a role for negative evidence, she questions the extent to which it engages linguistic mechanisms instead of general cognitive mechanisms.[7] This, however, is not the case among ISLA researchers, where there is a consensus that instruction facilitates language development (Spada, 1997; Norris and Ortega, 2000; Ellis, 2001; Spada and Tomita, 2010 among others). There is, though, less agreement on the *type* of instruction that best facilitates language development. In this section, the central issues surrounding the role of instruction in ISLA are discussed and some of the theories proposed are reviewed. This is, however, not meant to be an exhaustive review of theories but only the ones that are most relevant to the present study.

The recurring argument on the role of instruction is the basis of many studies in ISLA and is usually discussed in terms of implicit and explicit knowledge. Williams (2013)

[7] It is important to note here that the non-generativists do not distinguish between acquisition/linguistic competence and learning but instead between procedural and declarative knowledge, via implicit and explicit processing. This is not the same as the acquisition-learning distinction because modularized linguistic mechanisms are not assumed. The explicit-implicit distinction is discussed in section 2.3.1.

asserts that the answer to which type of instruction is most effective depends on how a given theory handles the explicit–implicit knowledge interface. Therefore, it is necessary to start with delineating the implicit-explicit distinction.

2.3.1 Implicit vs. explicit knowledge

Gass (1997) has argued that the first stage of input utilization is for researchers to acknowledge the fact that there is a gap between what the learners know already and what they have to know. Gass' statement and some researchers' beliefs that 'learners often seem to draw on knowledge they cannot articulate, and, conversely, are able to verbalize knowledge they cannot reliably use in communication' (Williams, 2013) are rooted in studies investigating the implicit-explicit interface. However, the explicit – implicit knowledge distinction is not easily defined as it is often mixed with discussion of implicit-explicit learning, as highlighted by Hulstijn (2002, 2005) and Ellis et al. (2009). Moreover, as noted above, many terms have also been used to describe the two types of knowledge: declarative-procedural; explicit-implicit; and controlled-automatic, among others. One of the earliest definitions of the implicit-explicit distinction is by Bialystok (1979). She defines implicit knowledge as what is 'used without attention to the rules or with inability to verbalise the rule' (p.82). In contrast, explicit knowledge involves the attention to the rules and ability to verbalize them. The ability to describe the rules is present in most definitions of *explicit* and *implicit* knowledge. Ellis (2008) provides more comprehensive definitions that summarise the findings of the research as follows:

Implicit knowledge is intuitive, procedural, systematically variable, automatic and thus available for use in fluent, unplanned language use. It is not verbalizable. According to some theorists it is only learnable before learners reach a critical age (e.g. puberty).
Explicit knowledge is conscious, declarative, anomalous and inconsistent (i.e. it takes the form of 'fuzzy' rules inconsistently applied), and generally only accessible through controlled processing in planned language use. It is verbalizable, in which case it entails semi-technical or technical metalanguage. Like any type of factual knowledge, it is potentially learnable at any age. (Ellis, 2008, pp. 6-7)

Ellis' definitions reflect the widely-accepted association of explicit knowledge with slow controlled processing and implicit knowledge with fast automatic processing (Hulstijn, 2005; Segalowitz and Hulstijn, 2005). For the purpose of the present study, implicit and explicit knowledge are identified in their broadest sense as knowledge *about* language (explicit) as opposed to knowledge *of* the language (implicit) (Ellis, 2004).

The most important questions in relation to instruction are what type of knowledge is affected by instruction and what type of knowledge learners draw from when dealing with different types of instruction. In the following section, I review some of the studies that tackle both questions. The body of research on implicit and explicit knowledge and learning is substantial so only studies that are closely related to the present research will be detailed.[8]

2.3.2 Implicit-Explicit learning

As mentioned earlier, a large and growing body of literature has investigated the role of instruction on language learning (R. Ellis, 1994; Doughty 2001, 2003; Norris and Ortega 2000; Robinson 2002; Sanz and Morgan Short 2005; Spada, 2005, 2009, 2013; R. Ellis et al. 2008; de Graaff and Housen, 2009; Spada and Tomita, 2010; Shintani et al., 2013; Williams, 2013 among many more). The studies on the role of instruction vary in their focus; some investigate the effect of instruction on certain features, others on noticing, while some studies examine the effectiveness of explicit and implicit instruction. I start by further defining what implicit and explicit learning is and move on to identify the main issues in the field before reviewing some studies in more detail.

Reber (1976) defined implicit learning as ‘a primitive process of apprehending structure by attending to frequency cues’ (Reber, 1976, p93). He contrasted this with a ‘more explicit process whereby various mnemonics, heuristics, and strategies are engaged to induce a representational system’ (ibid). Schmidt (1993) similarly differentiates between implicit and explicit learning in terms of intention: if something is learned unintentionally, then it results in implicit knowledge, and if it is learned intentionally, then it results in explicit knowledge. He adds that implicit learning occurs without awareness or understanding of what has been learnt. It is worth mentioning here that the implicit-explicit distinction should not be confused with the learning-acquisition distinction proposed by Krashen (1984). (See section 1.2.) Krashen distinguishes between acquiring an L2 language in a naturalistic setting and learning it in a formal

[8] It is important to note that as the main focus of the present study is not on the nature of explicit and implicit knowledge, studies that investigate this distinction will therefore not be reviewed here. For more information, please see N. Ellis (2005) and Paradis (2004, 2009).

setting. In contrast, both implicit and explicit learning can occur in the classroom, even simultaneously sometimes. So in a classroom where learners 'intend' to learn the present tense and are instructed explicitly about it, they might 'implicitly learn the article system.

Many studies have investigated the effectiveness of explicit and implicit learning in laboratory and classroom environments. However, after years of research on the efficacy of instruction, there is still a lack of consensus among researchers about whether and how instruction is effective. As noted above, some SLA researchers have claimed that instruction has no effect on how an L2 is acquired, i.e. on the acquisition of linguistic competence (Krashen 1985, 1993; Prabhu 1987; Schwarz 1993). There was even a call by Krashen and Terrell (1983) to abandon all instruction. Krashen argues that the acquisition of a language cannot be taught. Languages can only be learned (in Krashen's sense) in the classroom. But students acquire the language in spite of what goes on in the classroom (Krashen, 1985). Ellis, on the other hand, argues that 'while instruction may not always be necessary to achieve competence in the L2, it undoubtedly helps' (Ellis, 2005, p.725). This view is adopted by other researchers who see a beneficial role for instruction but either do not distinguish between acquisition and learning or do not support the interface hypothesis, that learning cannot become acquisition (Long 1983, 1988; Rutherford and Sharwood-Smith 1985; Ellis 1991, 2008; Norris and Ortega, 2000, DeKeyser, 2000; Doughty, 2003; White, 2003; de Graff and Housen, 2009). De Graff and Housen (2009) review the literature on the effectiveness of L2 instruction. They argue that there is evidence to suggest that instruction makes a difference in the following ways:

- Provides critical L2 exposure (Long 1988; Doughty 2003; Ellis 2008)
- Influences L2 propensity (Dörnyei, 1998, 2003; Platt and Brooks 2002; Ellis, 2003)
- Activates cognitive learning mechanisms (Skehan 1998; Housen and Pierrard 2005)
- Enables internalisation of new L2 knowledge, although researchers admit that findings on the effect of instruction on noticing are mixed and inconsistent (Ellis, 2001, 2002, 2005; Doughty 2003; Williams, 2005)
- Enables modification of L2 knowledge through gap noticing and corrective feedback (Swain 2005; Williams, 2005; Russel and Spada 2006)

- Enables consolidation of L2 knowledge (MacLaughlin and Heredia 1996; Gatbonton and Segalowitz 1998; Gass et al. 1999; Robinson 2001; Segalowitz 2003; DeKeyser and Juffs 2005; Swain 2005)

De Graff and Housen (2009) point out, though, that these effects are mediated by factors including the type of instruction, the type of L2 feature and the type of learner.
Having presented the arguments about the effectiveness of instruction in L2 learning, I now review some studies in more detail.

In one of the earlier studies, Hulstijn and Hulstijn (1984) investigated the effect of time pressure and focus on attention (information or grammar) on correct use of grammar. Their findings revealed that attention had a significant effect while time did not.
Another earlier study was by N. Ellis (1993) who examined the effects of implicit and explicit learning on the acquisition of grammatical features (soft mutations) in Welsh. He divided learners into three groups: exposure only, accompanied by rule presentation or with both rule presentation and exposure/examples. The results showed that the 'exposure only' group learned faster but showed little implicit knowledge and poor acquisition of the explicit knowledge as well. The 'rule presentation' group took more trials to arrive at an understanding of the structure, and the learners were able to verbalize the rules. However, they failed to apply these rules in practice. Finally, the rule-presentation-plus examples were the slowest learners; however, they were the only ones able to arrive at a working knowledge of the structure. They were able both to verbalize their knowledge explicit rules as well as implicitly generalize these rules to new structures.
Hulstijn and de Graff (1994) showed that instruction is effective if the target structures are complex while simple structures could be learned under implicit conditions. Their findings are consistent with those of Krashen (1982, 1994) and Reber (1989) that complex rules are best learned implicitly in meaning-based activities. Krashen and others base their argument on the observation that complex features are hard to notice in naturally occurring input, thus explicit instruction is necessary. On the other hand, Robinson's (1996) study revealed that explicit instruction was more effective in learning simple rules but implicit instruction was not more effective for complex rules. Robinson (1997), in fact, provided evidence that explicit instruction was effective for both complex and simple structures.

Dekeyser (1995) showed benefits for explicit rule presentation inn immediate and delayed post-tests, thus suggesting that the effect of such presentation is long lasting. He reported that explicit instruction was much better for learning categorical rules but there was no clear evidence of whether implicit instruction was effective for prototypical rules. In a later study, de Graff (1997) questions her earlier findings (Hulstijn and de Graff, 1994) that instruction is effective and argues that there is no relationship between type of instruction and type of language feature. Williams and Evans (1998) concluded that explicit instruction was best for simple rules but for complex rules, explicit and implicit effects were equal. On the other hand, Housen, Pierrard and Van Daele (2005) found that explicit instruction was effective for both simple and complex rules. Housen et al.'s (2005) findings corroborate those of earlier studies that showed that instruction is beneficial for language learning (Spada, 1997; Norris and Ortega, 2000; Ellis, 2001). In addressing the contradictory and unclear findings presented here, Norris and Ortega (2000) and Doughty (2003) argue that most studies on the explicit-implicit distinction are biased. Norris and Ortega's meta-analysis reported advantages of explicit instruction over implicit instruction in most studies reviewed, but they pointed out that measures of explicit knowledge rather than implicit knowledge are used to test knowledge after the treatment in most studies. According to Doughty (2003), this has led to an overstatement of the effects of explicit instruction. The problem with using explicit knowledge measures is that there is no evidence that the unanalysed, unconscious L2 competence necessary for rapid spontaneous communication is affected by explicit instruction (Doughty, 2003). Ellis (2005) emphasises that it is 'impossible to construct tasks that would provide pure measures of the two types of knowledge' (Ellis, 2005, p.153), but he asserts the importance of using multiple measures that tap into different types of knowledge (Ellis, 2002; 2004). To overcome this bias, Ellis (2006) conducted an experiment where he used explicit and implicit knowledge measures to test learners after the treatment. His findings revealed that students who were tested for their explicit knowledge on certain features and scored high failed to score the same when implicit measures were used. This led him to conclude that explicit knowledge does not necessarily correspond with better, more automatic use.

Spada and Tomita (2010) point out that it is not only the lack of implicit measures that lead to bias in stating the effects of explicit and implicit learning but that there is also lack of consensus on what constitutes simple vs. complex structures. In their meta-analysis study, they employ statistical tests to examine the effect size of the studies that looked at explicit and implicit instruction in relation to complex and simple structures and conclude that there is evidence that explicit instruction is more effective for both simple and complex structures than implicit instruction. There is also evidence from this meta-analysis that explicit instruction effectiveness is not short lived, as some have argued. The conclusions from Spada and Tomita's meta-analysis are consistent with other meta-analyses and reviews which reveal more effects for explicit instruction than implicit instruction (DeKeyser, 1995; Robinson, 1996; de Graff, 1997; Williams and Evans, 1998; Norris and Ortega, 2000; Housen et al., 2005).

However, there are still researchers who have not found any effect for explicit instruction on learning (Rosa and O‟Neill, 1999) or who argue that its effects are limited (VanPatten and Oikkenon, 1996; Benati, 2004; Sanz and Morgan-Short, 2004) or claim that instruction has no effect on how L2s are acquired (Krashen 1985, 1993; Prabhu 1987; Schwarz 1993).

As with the research on the role of instruction, research on the relationship between explicit-implicit knowledge and explicit-implicit learning is contentious. Alderson, Clapham and Steel (1998); Green and Hecht (1992); Brumfit, Mitchell and Hooper (1996); Metcalfe (1997); Terrell, Baycroft and Perrone (1987) and Scott (1989) all argue that learners' ability to describe rules of language accurately does not reflect their ability to use those rules. VanPatten and Oikkenon (1996) Sanz and Morgan-Short (2004) and Benati (2004) show that explicit knowledge alone does not lead to significant gains. Although, they acknowledge the role of instruction in resulting in learned knowledge, Schwartz (1993) and Truscott (1998) point out that the effects of explicit knowledge are short-lived.

Similarly, research on effects of instruction on noticing has so far produced mixed and inconsistent results (see Doughty 2003; Ellis, 2001, 2002, 2005; Williams, 2005). Researchers who advocate the superiority of instruction assume that learning is a conscious process, this means it involves higher levels of conscious awareness where instruction could increase the chance of learners' noticing the form (DeKeyser, 1998; Ellis, 1995, 2001; Sharwood Smith 1991,1993, 2008, 2009; Doughty and Long, 2003).

We can conclude from the above discussion that both the definition and effectiveness of implicit and explicit instruction is not straightforward. Nonetheless, the working definitions which will be used in the present study are modified versions of those proposed by Norris and Ortega (2000) in their meta-analysis and then voiced by Spada and Tomita (2010): instruction is considered explicit if it comprises rule explanation or learners are required to attend to form or 'to try to arrive at metalinguistic generalizations on their own' (Norris and Ortega, 2000, p. 437). Implicit instruction does not include rule presentation or attending to form as part of a study treatment. In the present study, instruction that does not involve explicit rule presentation but instead leaves it up to learners to arrive at metalinguistic generalisations by themselves is not considered explicit. This decision is based on the fact that even children acquiring their first language in naturalistic settings are observed to make metalinguistic generalisations (see e.g. Gombert, 1992).

The present study investigates how the use of user-behaviour tracking technology could help us investigate the effectiveness of different types of instruction in a technology-enhanced learning environment to shed new light on these issues. The study looks at the factors that contribute to differences in learners' performance in relation to input processing and type of input or instruction. In the next section, the reasons behind choosing three types of instruction in this study are discussed and relevant literature reviewed in detail.

2.4 Type of instruction

As suggested above and argued by De Graff and Housen (2009), the effectiveness of instruction is moderated by at least three factors: the type of instruction, the type of language feature and the type of learner. The focus in this section and in the present study is on the type of instruction rather than feature type or learner type. In reviewing the ISLA literature, it is obvious that many terms have been used to refer to the different approaches of providing instruction. In some cases the different terms are used to refer to the same concept (for further discussion, see Spada, 1997; Doughty and Williams; 1998, Ellis 2001). Therefore, before turning to the role of instruction in the classroom, it is important to start with defining the terminology used. In the next sections, I present the main terms and types of instruction then turn to examine types of instruction, focusing closely on the three types used in the study: focus on meaning (FoM), focus on forms (FoS) and focus on form (FoF).

2.4.1 Defining terminology

L2 instruction is defined as 'any deliberate attempt to promote language learning by manipulating the mechanisms of learning and/or the conditions under which these operate' (de Graff and Housen, 2009, p.726). Ellis (1997) distinguishes two types of instruction available to the language learner: Communication-Focused Instruction and Form-Focused Instruction. According to Ellis, the learners' attention in Communication-Focused Instruction is directed to meaning; in contrast, learners' attention is directed to language form in Form-Focused Instruction. Spada (1997) defines Form-Focused Instruction as 'any pedagogical effort used to draw the learners' attention to language form either implicitly or explicitly' (Spada, 1997, p. 73). As is clear from Spada's definition, Form-Focused Instruction is used in the literature as a cover term to refer to the general concept of providing form for the learner. As it does not specify how or when the form is presented, it embodies a multitude of concepts. Doughty and Williams (1998) and Long (2000), on the other hand, differentiate between three types of instruction and divide FFI into two categories. For Long (2000), there are three type of instruction: Focus on Meaning (FoM, henceforth), Focus on Form (FoF) and Focus on Forms (FoS). Long (2000) summarizes the three basic options associated with instruction in the classroom in Figure 2.3 below.

Focus on meaning	*Focus on form*	*Focus on forms*
Natural Approach	TBLT	GT, Silent Way, ALM, TPR, etc.
Immersion	Some content-based LT	
Procedural syllabus	Process syllabus	Lexical, structural, N-F syllabus, etc.
Etc.	Etc.	Etc.

Figure 2.3: Three options for the design of a second language course (Long, 2000, p. 180)

(TBLT = Task-Based Language Teaching LT = Language Teaching GT = Grammar Translation

ALM = Audio-Lingual Method TPR = Total Physical Response NF = Non-functional)

Long (1991) coined the term *focus on form* and used it to refer to the incidental focus on linguistic form when the learner's or teacher's goal is not metalinguistic knowledge. He defined focus on form as input which 'overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication' (Long, 1991, pp.45-46). Here he did not differentiate between the two types of instruction based on explicitness per se, but rather based on the main focus of instruction. Just like Focus on Forms/FoS, Focus on Form/FoF can be explicit but unlike in FoS, in FoF, the mastery of linguistic features is not the focus of instruction; rather, the focus is on communication and meaning. Norris and Ortega (2000), on the other hand, distinguish between explicit and implicit presentation of form but their ideas are similar to Long's: they use *focus on form* to refer to the implicit embedding of form in communicative tasks and *focus on formS* to refer to the explicit abstract teaching of form. Ellis et al.'s (2002) definition is similar: focus on form is 'the treatment of form in the context of performing a communicative task' (Ellis, et al., 2002, p. 419). Ellis does not note whether the treatment of form is explicit or implicit, but he distinguishes between 'planned' and 'incidental 'focus on form and also between 'pre-emptive' and 'reactive' focus on form (Ellis 2001). With the increasing interest in focus on form/FoF over the years, the term *focus on form* has extended beyond Long's original definition. As pointed out early on by Doughty and Williams (1998); 'there is considerable variation in how the term *focus on form* is understood and used' (Doughty and Williams, 1998, p. 5). The confusion does not stop at the definitions, and even when researchers distinguish between explicit and implicit instruction, they sometimes use different acronyms to indicate the same type of instruction. For example, Long and Robinson (1998) used FonF to refer to focus on form instead of FoF while de Graff and Housen (2009) used FFI (form-focused instruction) to refer to both types and they used the terms with adjectives (implicit FFI and explicit FFI) to differentiate between focus on form and focus on formS, respectively. The terminology and acronyms used carry different connotations, but most of them share basic dimensions. Focus on meaning is similar to Communicative-Focused Instruction in the sense that learners' attention is only directed to meaning without any reference to linguistic forms. FoS, on the other hand is used as a cover term to refer to the explicit abstract presentation of a form where the main goal of instruction is the mastery of linguistic features, while FoF is used to refer to the incidental and most likely implicit and communicatively-based presentation of form. FoF is only supposed to be what is provided when the learners struggle to complete a task.

The implications of the different terms and definitions are critically revisited in the sections below when reviewing the literature on the different types of instruction. For now, the three acronyms: FoM, FoS and FoF, are used throughout this thesis to refer to the three types of instruction as outlined earlier. That is, FoM is used to refer to instruction when the focus is mainly on using language in communicative tasks where the primary focus is completion of a task. FoS is used when the attention of the learner is explicitly directed to linguistic features and pedagogic grammar rules are presented as part of the instruction (see section 2.5 below about pedagogic grammar). FoF is used for instruction where the main focus is on meaning but learners are given the chance to focus on linguistic forms as and when they decide to.

The term *form* has also been extensively used in the literature to refer to morphsyntactic features of the language, or in more pedagogic terms, grammatical features. However and as rightly argued by Ellis, Basturkmen and Loewen (2001), *form* can be linked to phonology, discourse, grammar, vocabulary and spelling. The present study looks at a specific grammatical construction and the discussion here will therefore not extend to these other possibilities.

2.4.2 Focus on forms (FoS)

As discussed earlier, FoS entails the teaching of linguistic forms in isolation. This approach has long been used in language learning environments. As described by Long and Robinson (1998), FoS teaching involves breaking the L2 into 'words, and collocations, grammar rules, phonemes, intonation and stress patterns, structures, notions or functions', that is, traditional foreign language teaching (Long and Robinson, 1998, p. 15). The linguistic features are then presented to learners in a linear order based on for example frequency or assumed difficulty. The effectiveness of FoS has been investigated in many studies, some of which were reviewed in the previous section when examining the difference between explicit and implicit learning.

After reviewing the literature on the effectiveness of instruction, De Graff and Housen (2009) neatly list the features that distinguish the two extremes of implicit and explicit Form-Focused Instruction or FoF and FoS respectively, as shown in Figure 2.4. They do admit that there is a continuum along the two extremes. What they do not note, however, is that there is considerable difficulty in deciding where one extreme ends and the other starts. FoS involves direct attention to language forms which is predetermined and planned. It also comprises controlled practice of target forms through the use of metalinguistic terminology to present target forms in isolation.

Implicit FFI	Explicit FFI
• attracts attention to language form	• directs attention to language form
• language serves primarily as a tool for communication	• language serves as an object of study
• delivered spontaneously and incidentally (e.g., in an otherwise communication-oriented activity)	• predetermined and planned (e.g., as the main focus and goal of a teaching activity)
• unobtrusive (minimal interruption of communication of meaning)	• obtrusive (interruption of communication of meaning)
• presents target forms in context	• presents target forms in isolation
• no rule explanation or directions to attend to forms to discover rules; no use of metalanguage	• use of rule explanation or directions to attend to forms to discover rules; use of metalinguistic terminology
• encourages free use of target form	• involves controlled practice of target form

Figure 2.4 : Implicit and explicit form-focused instruction (de Graff and Housen, 2009, p.737)

As mentioned earlier, several studies have shown that FoS makes a difference in learners' performance. Norris and Ortega's (2000) meta-analysis revealed that FoS effects are larger and more durable. They do, however, point out methodological bias in the studies in terms of treatment length and assessment measures, as also noted above. Other studies indicate that FoS is suitable for certain features or under certain conditions (see section 2.3.2 above). As also pointed out earlier, one of the problems when classifying the type of instruction as FoS and FoS is that one cannot draw boundaries between the two, and where one study might consider *incidental* explicit focus on form as FoF, another would consider it FoS. Studies and meta-analyses that have shown an effect for any sort of form-focused instruction include Hulstijn and Hulstijn (1984); Long (1991, 1996, 1997, 1998, 2000); Ellis (1994, 2001); Spada (1997); Doughty and Williams (1998); Long and Robinson (1998); Norris and Ortega (2000); Doughty (2003) and Nassaji and Fotos (2004).

In 2000 through his seminal work, Long raised doubts about the effectiveness of FoS and proposed FoF as a balanced approach to language learning. According to Long (2000), FoS suffers from major problems. It is a 'one-size-fits-all' approach which means that it always results in either teaching too much of what is not needed or teaching too little of what is actually needed. Another problem with FoS, as noted by many researchers, is that simplified input is used to provide linguistic forms in isolation rather than in context. To put it in Wilkins' words 'parts of the language are taught and step by step so that acquisition is a process of gradual accumulation of parts until the whole structure of a language has been built up...At any one time the learner is being exposed to a deliberately limited sample of language (Wilkins, 1976, p.2). There are two problems with this sort of approach: simplification of input and order of acquisition. The simplification of input involves stripping it from the new or difficult forms that learners usually encounter when using the language. Materials used in FoS approaches are usually artificial / simplified dialogues. This in turn results in unrealistic language usage rather than language use (Widdowson, 1972). Furthermore, the findings of SLA research indicate that learning a language is not a process of accumulating and memorizing the features of the language (see Selinker 1972 up to White 1989 and beyond). Widdowson (2008) stresses that learners do not learn by adding items of linguistic knowledge but rather by a process of continual revision and reconstruction; that is, learning is ‘ continual cognitive adaptation as the learner passes through different transitional stages’ (Widdowson, 2008, p. 211). In addition and as pointed out by Cook (2008), the order of presentation in the classroom and language teaching materials does not comply with the findings of SLA research. For example, beginner learners are presented with full sentences where they need to mark tense and agreement such as *He goes to school every day* although evidence from SLA research suggest that these forms are acquired later (see Dulay and Burt, 1973, 1974).

Further criticism of an FoS approach comes from processing-based research. FoS does not presume any importance for language learning processes. In this sense, it suggests that learners acquire linguistic forms in the same order they were exposed to in the syllabi. Ellis (1989) and Pienemann (1984), among others, argue that teachability is not the same as learnability, i.e. teaching a form does not mean that learners have learnt it, and definitely it does not mean that they have learnt it at the same time it was taught. It has also been argued that an extreme model of ‘getting it right’ from the beginning, i.e. teaching grammar explicitly from the very start so that students get things right, does not benefit effective communication (Lightbown and Spada, 1993).

In response to the problems found in FoS, a new type of instruction has emerged: focus on meaning (FoM). In the next section, I examine the literature on FoM and its effectiveness in the classroom before I return to FoF.

2.4.3 Focus on meaning

According to Willis and Willis (2007), meaning-focused instruction is an essential element of the language teaching pedagogy. Ellis (1999) has argued that any deliberate effort to engage the L2 learner in the communicative exchange of relevant meanings and authentic messages could come under the umbrella of FoM. The main assumption behind this relatively new approach to instruction is that second language learning is very similar to first language learning and so by providing the same setting, learners will be able to acquire the second language successfully. This assumption implies that learning a second language is 'incidental' in the sense that learners acquire linguistic forms while they are doing something else, for example, communicating (Long, 2000). Comprehensible Input (Krashen, 1984, 1985), the Natural approach (Krashen and Terrell, 1983), input flood, Content-Based Teaching (Wesche and Skehan 2002) and Communicative Language teaching (Wilkins, 1972; Widdowson, 1978; Brumfit, 1979; Littlewood, 1981; Savignon, 1997) are all examples of FoM instruction. Materials used in this approach mainly focus on communication and meaning and enriched input believed to be enough to acquire the grammatical rules of the target language subconsciously.

According to Long (2000), FoM has problems. As in FoS, there is no needs analysis involved in this approach so in one way it is still a one-size-fits-all model. In addition, it ignores research on maturational constraints, for example, the Critical Period Hypothesis.[9] If adult L2 learners cannot achieve native-like proficiency for biological reasons, FoM will not be sufficient since recreation of the L1 acquisition environment will not enable L2 acquisition to successfully take place. This implication is supported by research that confirms that adult L2 learners cannot achieve native-like grammatical competence even after prolonged periods of exposure (Swain, 1991, but see Herschensohn for an overview of more recent research). In addition, and taking into account the point that languages do not share the same grammatical systems, certain target language forms will be more difficult than others to acquire. Some argue that this is where explicit learning plays a role. For example, White (1991) argued that the ungrammaticality of placing an adverb between a verb and a direct object such as in the example below is unlearnable from comprehensible input, i.e. from FoM, only.
Example:

---

[9] The hypothesis was introduced by Penfield and Roberts (1959) and Lenneberg (1967) and proposes that after certain age, e.g. puberty, the language learner will not be successful.

1. *She drinks always milk
2. She always drinks milk

Apart from (1), adverb placement is fairly free in English and in a FoM classroom; L1 French learners of English will hear utterances with adverbs in a number of positions. They need to notice that structures such as in (1) are missing from the input in order not to produce ungrammatical forms based on what is possible in their L1 French. Producing structures similar to (1) will not, however, break communication making it less salient and recognizable.

The argument so far, is that FoM instruction totally lacking any focus on form does not lead to effective learning. Solutions proposed by researchers to overcome the inadequacy of a meaning-focused instruction include: focus on form/FoF (Long, 1991; Doughty and Williams, 1998; Fotos and Nassaji, 2007), Input Enhancement (Sharwood Smith, 1993), Input Processing (VanPatten, 1996, 2002, 2004) and conscious raising (Ellis, 2003; Cook, 2008).

In the next section, I evaluate the effectiveness of FoF and review some of the studies that have looked at FoF under different conditions.

2.4.4 Focus on form

By the late 1990s, voices of doubt were raised about the effectiveness of FoM and FoS instruction. So far, the argument is that neither FoS nor FoM is sufficient for acquiring grammatical competence in the target language. FoS provides learners with information on what is grammatical or ungrammatical but is not enough to show them how to use the grammatical forms. On the other hand, FoM provides learners with input rich in grammatical structures but does not help them eliminate ungrammatical structures and it presumes that learners will notice the presence of grammatical structures and the absence of ungrammatical ones. As mentioned earlier, among the solutions proposed to enhance the effectiveness of instruction is Long's *focus on form* (FoF).

In their summary (see Fig 2.4 above), de Graff and Housen (2009) argued that implicit form-focused instruction/FFI, referred to as FoF in the present study, involves directing the learner's attention to form in an unobtrusive way. Similar to FoM, language serves as a tool for communication and attention to form is spontaneous and incidental. Blake

(2008) defines focus on form from a cognitive-pedagogic perspective. He states that FoF is a 'task based methodology that calls on L2 learners to solve specific tasks' (p.19). While doing so, the students will need to negotiate meaning of forms to solve the problems entailed in completing the tasks and by doing so, they focus on the gaps in their linguistic knowledge and analyse their own interlanguage. One of the reasons for the empirical and theoretical support FoF has received as a type of instruction is that it complies with the findings of psycholinguistics-based language learning research. For example, as FoF requires learners to pay attention to form under certain conditions, it has support from Schmidt (1990) who has in his Noticing Hypothesis long argued for the importance of attention in language learning (see above). According to Long (1998) and Doughty (2003), the beneficial effect of FoF instruction is accelerating the passage through the sequences of L2 development and extending the scope of application of grammatical rules. Thus, it results in improving accuracy, rate of learning and level of ultimate attainment.

In contrast to FoS and FoM, FoF assumes an important role for language learning processes. Learners focus on forms while processing the input for meaning. Long (1996) proposed that learning takes place via negotiation of meaning in interaction. In the same way, when there is a breakdown in communication during interaction, learners resort to focus on forms to negotiate the meaning and convey the message. VanPatten (2002, 2004) also suggested that learners find difficulties in processing the input for meaning and forms at the same time. He argues that learners process input for meaning first as their primary intention is to maintain communication and comprehend the message. VanPatten proposed his Processing Instruction model as a way of directing learners' attention to forms without providing explicit rule presentation which would result in a communication breakdown. The role of output is also highlighted in FoS. Learners produce output which might contain indications of problems which will result in the learner's noticing aspects of linguistic forms, making new hypothesis and producing more output; this is what is called negotiation of meaning process (Long, 1996).

In seminal work, Ellis, Basturkmen, and Loewen (2002) identified two types of FoF: pre-emptive and reactive. Reactive focus on form can be classified based on whether it is conversational or didactic and whether it is implicit or explicit. For example, corrective feedback, an example of focus on form, can be implicit or explicit: implicit by means of recasts or explicit by telling the learner what was wrong. Pre-emptive focus on from, on the other hand, could be teacher initiated or student initiated. First the authors advise having teachers trained on the use of focus on form before they start using it. The two types of FoF are distinguished: planned focus on form, which involves designing communicative tasks to elicit the use of specific grammatical forms in the context of meaning focused use and incidental focus on form which involves designing communicative tasks to generate general samples of the language rather than specific forms. In such a case, it is expected that many forms will be attended to briefly rather than focusing on one form. They also proposed how FoS and FoF should be presented in the classroom. For FoS, they suggest the PPP technique (presenting a grammatical form, practicing it in controlled exercise then producing it freely). For FoF, they suggest giving the students an information-gap exercise and in the course of doing so drawing their attention to one or more grammatical forms which are needed to complete the task. In this sense, the focus on linguistic forms is raised through communicative need. In their discussion, they also argued that typically learners will not achieve high levels of competence from entirely meaning-focused instruction. At the same time, however, learners will not acquire the forms through entirely FoS instruction. Although learners will get high marks or pass a grammar test, they will not be able to use the forms in spontaneous, free conversation. Ellis et al. also found that pre-emptive FonF was as common as reactive FonF and concluded that pre-emptive FoF is most likely student-initiated and the new language forms encountered during pre-emptive FoF were more likely to be taken up and used by them subsequently.

A number of empirical studies have investigated the frequency of these different types of FoF and their contribution to learner uptake. Loewen (2003) found great variability both in the frequency of FoF episodes, and also in the extent to which individual students took part in these. Similarly, Mackey, Polio and McDonough (2004) compared novice and expert teachers. Their findings indicate that experienced teachers used significantly more pre-emptive FoF than the inexperienced teachers. The inconsistencies in the findings of studies that looked at the different types of FoF or the different types of instruction meant that firm conclusions were not possible (see Norris and Ortega, 2000, and Mackey and Goo, 2007). There have been calls for experimental work that uses mixed methods and longer treatments to yield more valid conclusions (see section 2.3.2 for discussion and Ellis et al. 2006 for an example).

The interest in FoF and its implementation is obvious in the growing body of research on the topic and not all of it supports FoF. One of the criticisms addressed to Long is that he conducted his research outside the classroom so his research is flawed. For example, Poole (2005) evaluated FoF instruction and concluded that FoF is only effective in particular settings and highlighted the need for research that gauges the appropriateness of FoF in different settings and for different learners. Norris and Ortega (2000, 2006) and de Graff and Housen (2009) highlighted the difficulty of drawing firm conclusions on the effectiveness of any type of instruction due to methodological and experimental bias. Along the same lines, Doughty (2003) pointed out that the effectiveness of instruction is overestimated in many studies.

Although this will not be considered in the present study, it should be noted that many factors are believed to affect the efficacy of instruction, including individual learner variables, type of linguistic features, settings, among others. The type of linguistic feature has particularly been the focus of a growing body of research. Researchers have looked at whether L2 instruction is more effective for some L2 features than for others and whether some L2 features are more susceptible or responsive to instruction (more 'teachable') than others. Early on, there were studies that postulated that not all linguistic items in the input are attended to equally by L2 learners (e.g. McLaughlin, Rossman and McLeod, 1983). The observation is that learners appear to attend to linguistic forms based on their communicative value (Klein, 1986; VanPatten, 1985, 1990, Leow, 1995). For example, Williams (1995) concluded that when the target structure is complex (e.g. relative clauses and passives) enriched input, i.e. FoM, may be as or more effective than explicit instruction with feedback, FoS. However, when the target feature is simple (participial adjectives), FoS may be more effective. Long and Robinson (1998) state that if the forms are ‘rare and/or semantically lightweight, and/or perceptually non-salient, and/or cause little or no communicative distress’ (p.23), they are less likely to be acquired without instruction. Other studies have shown that instruction is more effective for, and should focus on, easy/simple L2 features/forms (DeKeyser 1995; Krashen 1994; Pica 1985, Robinson 1996). But agreement is lacking here. Hulstijn and de Graaff (1994), de Graaff (1997) and Housen et al. (2005) conclude the opposite: in general instruction is more effective for, and should focus on, difficult/complex L2 features. The verdict is therefore still out.

Taking into consideration Ellis et al.'s (2002) distinction between two types of focus on form, the present study includes pre-emptive, planned focus on form/FoF along with focus on meaning/FoM.

2.4.5 Summary of the role of instruction literature

There is no space in this thesis to review all studies and experiments that have dealt with the role of instruction in the classroom. The main findings, however, can be summarised in the following points, as highlighted by reviews and meta-analyses (Norris and Ortega, 2000, 2006; Mackey and Goo, 2007; de Graff and Housen, 2009; Li, 2010; Lyster and Saito, 2010; Spada and Tomita, 2010; Plonsky, 2011). First, instruction can have a positive and durable effect on L2 learning when compared to achievements in conditions of just naturalistic exposure to the L2 (de Graff and Housen, 2009). It seems to be most effective if it includes both comprehension-based and grammar-based activities (Shintani, Li and Ellis, 2013).

- FoF is characterised by allocation of attentional resources to language forms raised by communicative demand. It is believed to speed up the rate of learning
- FOS is characterised by explicit focus on the elements of the grammar; it is common in a structural syllabus where repetition of models, memorization of short dialogs and error correction are required. FoS instruction alone results in high accuracy but low fluency.
- FoM is based on authentic language use. It is used in a meaning-based syllabus where implicit grammar teaching might be present but it is not the focus of the learning. FoM results in low accuracy but high fluency

Regardless of all the criticism directed to teaching based on a purely FoS syllabus, it is often the main teaching approach in many countries (Long and Robinson, 1998, Nunn, 2011), and research findings, with their limitations, indicate that explicit instruction is beneficial (Norris and Ortega, 2000; Spada and Tomita, 2010). However, the distinction between FoF and FoS is not always clear-cut (see Sheen 2002); the present study attempts to clarify the distinction in a TELL environment.

The discussion in this section focused on the role of instruction in language learning from a psycholinguistic and cognitive perspective. As this study is conducted in a pedagogic setting, it deemed necessary to look at the role of instruction from a pedagogic perspective. One of the cornerstones of the discussion of the role of instruction from the pedagogic side is the debate on whether, how, and when to integrate grammar instruction in the classroom. This is the focus of the next section.

2.5 Grammar Instruction in the classroom: research and methodologies

Although SLA theories do not directly mention language teaching methods, they do form the underlying principles for some of these theories by identifying the most favourable conditions for L2 learning to take place. As discussed earlier in the chapter, theories of second language acquisition have changed considerably over the years, and new theories are still being developed. The different theories imply different approaches to language learning and also different implications for the classroom (Richards and Rodgers, 2001; Piske and Young-Scholten, 2009; Whong, 2011). In the next section, I look at the links between ISLA research and grammar teaching in the classroom.[10]. I then move on to look at Task-Based Language Learning and its suitability for the present study.

2.5.1 ISLA research and grammar teaching

Cook (2009) correctly points out that 'many of the changes in thinking about the language teaching over the last decades can be traced back to the overall ideas about the nature of the learner developed in SLA research' (Cook, 2009, pp. 141-142). Just as the debate on the type of instruction that best facilitates learning has been going on for decades now in ISLA, the question of whether and how to teach grammar in the classroom has been and still is at the centre of a large body of language teaching research. The two debates are interrelated, and findings of ISLA studies often refer to pedagogical implications. In fact, Cook (2009) states that the start of SLA research around the 1960s was an attempt to underpin language teaching but unfortunately for teachers only a fraction of SLA research has been applied to the classroom. One of the language teaching principles that could be tested or justified from SLA/ISLA research that Cook (2009) lists is the assumption that 'teachers should avoid explicit discussion of grammar' (Cook, 2009, p. 149), see also Lightbown (1985). SLA researchers who have proposed certain techniques on how and when to teach grammar include among others Spada (1997); Long and Robinson (1998); Doughty and Williams (1998); Lightbown (2000); Norris and Ortega (2000, 2006).

[10] Grammar is used in this thesis to refer to pedagogical grammar, as presented in language learning materials and classroom teaching.

Grammar instruction has always been at the centre of most theories and research in the field of second language learning. For many years, it was thought that it was sufficient to consciously know the grammar to acquire/learn a language (Rutherford, 1988). This is still the thinking in many countries (Nunan, 2006; Nunn, 2011). The origin of this belief partly has its roots in Skinner's Behaviourism (1957). Under the behaviourist-based methods such as the Grammar-Translation method and the Audio-Lingual Method, [11] language learning was seen as formation of habits. These methods, referred to in the literature as 'traditional' can also be classified as language-centred methods since the goal is the mastery of the language forms. Linguistic competence (including under generative linguistics) is believed to be the core of learning as learning a second language meant learning the grammar of that language.

Cook and Singleton (2014) state that the Grammar-Translation method which was universally popular in schools until the 1950s is still used in universities today. Grammar-Translation involves rote-learning of grammar rules and the translation of texts. The focus in the Grammar-Translation method was on 'the conscious memorization of grammatical paradigms and rules, as well as lexical items and expressions' (Cook and Singleton, 2014, p. 112). With the decline of interest in Grammar-Translation in schools and EFL classrooms, the Audio-Lingual Method appeared as an alternative. In the Audio-Lingual Method, parts of the language are presented as chunks and structure rules which are practiced again and again through memorisation of dialogues and teacher-directed instruction. According to Rodgers (2009), until the mid-years of the 20th century, Audio-Lingual Method dominated classroom practice, particularly in the USA. Though in Europe and other parts of the world, it was not as dominant. Cook (2008, p.17) states that although the Audio-Lingual method reached its 'peak of popularity' in the 1960s, its use was not prevalent in British-influenced EFL. What is referred to as 'mainstream' EFL (Jin and Cortazzi, 2011, p.563) was common in the classroom. This approach appeared around the 1930s and it was marked by its eclectic nature. Jin and Cortazzi describe eclecticism here as not indicating a random combination of techniques and methods but involving a focus on analysis and a rationale for sequences and choices of strategies and structures. Cook

[11] Other methods, but which are outside of mainstream language teaching practices such as the Silent Way, the Direct Method, Total Physical Response, Suggestopedia, are not discussed here as the aim of this section is to explore the links between mainstream teaching methods and ISLA, and these methods were not widespread, their influence is marginal (see Richards and Rodgers, 2001 and Rodgers, 2009 for details).

(2008) states that the Audio-Lingual method arrived in Europe from the USA when language laboratories were popular, which meant that some of the techniques involved in the Audio-Lingual Method, such as repetition and drilling, worked well. Cook and Singelton (2014, p. 114) argue that Audiolingualism in the US 'was paralleled in Europe by the 'audio-visual' method'. They explain that the Audio-Visual method was similar to the Audio-Lingual method but differed in its focus on the creation of meaning through the use of pictures and sentences to help create associations in the learners' mind.   In all the aforementioned teaching methods, learners are viewed as passive recipients of the knowledge provided to them by the teachers. The most common technique is the PPP, (presentation, practice, production) and the focus of instruction is mastery of grammatical rules, i.e. linguistic competence. Although these traditional methods had fallen out of favour by the end of the last century, they still represent 'poles of thinking about teaching and about language that are still highly relevant today' (Cook, 2014, p. 117). For example, as was established earlier in the chapter, when discussing explicit and implicit learning and the role of instruction, the focus on grammatical accuracy which was a main aspect of these approaches and methods has not faded.

By the 1970s, generative-linguistics based SLA theories had been introduced with the work of Noam Chomsky. Chomsky's main impact on L2 classroom teaching can be summarized in the assumption that children acquire languages without being instructed or explicitly knowing the grammar, so grammar instruction is not actually necessary. To put it in Cook's words, 'as the universal grammar in the student's mind is so powerful, there is comparatively little for the teacher to do' (Cook, 2001p. 183). This idea, though not expressed in reference to UG, dates back to at least Krashen (1985), with his Natural Order hypothesis. Moreover, there was growing body of evidence from classroom research that the traditional methods were not yielding the desired outcome as even after years of exposure to the language in the classroom, learners were not able to use the language when communicating in real life. The Communicative Approach and meaning-focused methods emerged out of this dissatisfaction with the traditional methods, the influence of SLA theories and the ideas introduced by Hymes (see below). This shift in perspective was also spurred on by the belief that L2 learners can acquire the language in the same way as children, through exposure to the language. Comprehensible Input (Krashen, 1984), the Natural Method and the Natural Approach (Krashen and Terrell, 1983, 1989) were all proposed as the best approaches to language learning. These approaches are similar in that they make no or little provision of formal instruction of grammar. The focus is on the comprehension of language as used in everyday situations/contexts. Krashen's Comprehensible Input model no doubt had a crucial role in the rise of communicative and meaning-focused methods. However, Communicative Language Teaching (CLT) emerged separately, as a mainstream teaching method which gained support fairly quickly. It came into existence with the works of applied linguists such as Hymes and Wilkins. In 1979, Hymes coined the term 'communicative competence' which he used to refer to the 'overall underlying knowledge and ability for language use which the speaker-listener possesses' (1979, p.13). Hymes states that communicative competence does not depend only on grammatical knowledge but also on knowledge of the sociocultural norms governing day to day communication. CLT's main goal is to provide opportunities for learners to practice language use through meaning-focused activities, assuming that learners will learn the linguistic forms as a by-product while fulfilling their communicative needs (Kumaravadivelu, 2006). According to Cook and Singleton (2014), the communicative method 'sought to connect language teaching and learning in the classroom as transparently as possible to learners' likely uses of the target language' (p.118) Kumaravadevilu (2006) and Cook and Singleton (2014) emphasise that CLT, with its different manifestations, is still the most teaching approach in use worldwide.   Unlike the Grammar Translation and the Audio-Lingual Method which were language centred, CLT is learner-centred. Finocchiaro and Brumfit (1983) contrast the main features of both approaches in detail which were then summarized by Rodgers (2009) (see Table 2.1 below). Typical classroom activities are those that encourage the negotiation of meaning such as information-gap exercises where one learner has part of the information and the other learner the other part and they need to negotiate to complete the activity. Unlike the traditional methods where learners are passive recipients, in learner-centred approaches, learners are viewed as active participants in the learning process.

Table 2.1 Features of the Audio-Lingual and communicative approaches (Rodgers, 2001, p.351)

AUDIO-LINGUALISM	COMMUNICATIVE APPROACH
1. Attends to structure and form more than meaning	Meaning is paramount
4. Language learning is learning structures, sounds, or words	Language learning is learning to communicate
7. Native-speaker-like pronunciation is sought	Comprehensible pronunciation is sought
9. Communicative activities only come after a long process of rigid drills and exercises	Attempts to communicate encouraged from the beginning
13. The target linguistic system will be learned best through the overt teaching of the patterns of the system	The target linguistic system will be learned through the process of struggling to communicate
14. Linguistic competence is the desired goal	Communicative competence is the desired goal
16. The sequence of units is determined solely by principles of linguistic complexity	Sequencing is determined by any consideration of content, function or meaning that maintains interest
18. "Language is habit" so errors must be prevented at all costs	Language is created by the individual, often through trial and error
19. Accuracy, in terms of formal correctness, is a primary goal	Fluency and acceptable language is the primary goal; Accuracy is judged not in the abstract but in context
21. Students are expected to interact with the language system, embodied in machines or controlled materials	Students are expected to interact with other people, either in the flesh, through pair and group work, or in their writings
22. Intrinsic motivation will spring from an interest in the structure of the language	Intrinsic motivation will spring from an interest in what is being communicated by the language

However, after the novelty of the new methods faded, criticism started to rise. Teachers were found to be using techniques from the traditional methods to deliver a supposedly communication-focused syllabus. Instead of the original PPP, they introduce the linguistic form and practice it but then give learners freedom in how they will produce it. This led some researchers to claim that nothing had changed: CLT 'has not been significantly different from or demonstrably better than the language-centred pedagogy it sought to replace' (Kumaravadivelu , 2006, p.132). In the 1980s, Howatt (1984) had already introduced the term 'weak CLT' to refer to the CLT where learning was still viewed as linear and involved presentation and practice of the target language. As a result of the criticisms addressed to weak CLT, strong CLT emerged. In the weak form of CLT, learners focus on meaning as well as form while in the strong form they focus primarily on meaning while trying to solve a problem (e.g. in an information gap activity).

In strong CLT, the underlying assumption is that learners will eventually master the target language (Howatt, 1984; Savignon, 2002, Howatt and Widdowson, 2004). As strong CLT started to attract interest, there were renewed calls for exclusion in the classroom of previous methods which relied heavily on explicit rule presentation such as the Audio-Lingual Method and the Grammar Translation method and an absolute abandonment of all kinds of grammar instruction. Lightbown and Spada (1993), criticising grammar instruction in the classroom, argue that an approach 'which is aimed at helping the learner get everything right from the beginning does not benefit effective communication' (1993, pp.79-83). The change in the principles and implementation of CLT also led to a change in the focus of research and experiments from the product of learning to the process of learning. While weak CLT is described as learner-centred, strong CLT is learning-centred. However, focusing on meaning only and excluding any provision of grammar instruction – in other words FoM - has not been as effective as was predicted. As early as 1981, L2 researchers argued for the importance of formal grammar instruction. For example, Sharwood Smith (1981) argued that grammar teaching plays a central role in raising learners' awareness of the language forms. More and more scholars are now firmly convinced that formal grammar instruction should not be swept out of language teaching classes (Nassaji and Fotos, 2004; Ellis, 2006; Cullen, 2008) and the result on FoF discussed above support this. It seems that strong Communicative Language Teaching widely adopted in the 1990s is giving way now to a more balanced approach. Unlike weak CLT, instead of being the focus of learning, grammar is now viewed as a facilitator to communication rather than abstract rules to be memorized. Byram and Mendez Garcia (2009) encapsulated the state of CLT in the late 2000s as follows;

> In spite of the originally heated debate on the exclusion of grammar from the communicative paradigm, nowadays there seems to be an agreement on the need to incorporate a focus on form, although always integrated with a parallel focus on meaning, for the individual's development of a global communicative competence in which discourse features, appropriacy, and communication are also an integral part. (p.505)

What is noteworthy here is that in the early days of CLT, the debate on the role of grammar instruction was mainly about *whether* to teach grammar or not. However, the debate has shifted in recent years to the question of *how* to teach grammar and *when.* Grammar instruction is viewed as an indispensable resource for communication, not simply the focus of study by the learner. The question was now how grammar instruction should be presented to learners; this includes the question of what type of grammatical input is most effective.

Earlier in this chapter, I discussed the two main approaches proposed by applied linguists, namely, Focus on FormS (FoS) and Focus on Form (FoF) (Long, 1988, 1991). Recall that FoS encourages the explicit teaching of grammar rules as separate units while FoF claims that grammar instruction should be limited and should happen only when it facilitates communication. FoF was claimed to be a balanced approach that seeks to overcome the deficiencies of an absolute focus on grammar/form, associated with traditional methods and an absolute abandonment of grammar/form associated with strong CLT. Ellis (2006) emphasized the idea that FoF implies grammar teaching integrated into a curriculum consisting of communicative tasks and Poole (2005) critically evaluated Focus on Form instruction in the classroom along these lines. Poole concluded that FoF instruction can meet its instructional objectives only if the following elements are present: principles of CLT are accepted in activities and assessments; classes are sufficiently small for teachers to be able to work individually with students, and teachers and students are proficient enough in English so they do not switch to their native language when communicative difficulties are encountered. He emphasized the need for future research to determine whether FoF instruction is appropriate for different groups of learners and different settings.

From the 1990s onwards, other scholars have highlighted the need for research that investigates innovative and effective approaches to how to incorporate grammar instruction into broadly communicative teaching. In the field of materials design there were calls for research that investigated the features of instruction that promote learners to notice important aspects of the language. Dating back to the seminal work of Candlin and Murphy (1987), this led to one of the offshoots of CLT, namely; Task-Based Language Teaching (TBLT), to increasingly gain interest. TBLT is, as we will see, a way to implement FoF.

TBLT considers the type of activities that promote language use in meaning-focused settings. Its main emphasis is on providing learners with meaningful tasks that help them use the language to communicate effectively in real life situations. TBLT is a balanced approach where although the focus is still on meaning, attention to form is permissible under certain conditions. In the next section, I explore TBLT and the central questions associated with its use in different language learning environments. In the course of this thesis, I use the term Task-Based Language Learning and the acronym TBLL instead of Task-Based Language Teaching (TBLT) as the present study was conducted in a self-accessed learning environment and no ‘teaching’, in the traditional sense of the word, takes place.

2.5.2 Task-based Language Learning

As noted in the previous section, recent years have seen a growing interest in Task-Based Language Learning and the role of tasks in L2 learning (Fotos and Ellis 1991; Skehan 1996; Bygate, Skehan and Swain 2001; Laufer and Hulstijn 2001; Ellis 2003, 2005; Nunan 2005, 2006; Reinders 2010; Robinson 2011; East 2013). In this section, I look at the definitions of a ‘task’. I then move on to look at the cognitive and pedagogical underpinnings of TBLL before I examine the factors that contribute to task effectiveness.

2.5.2.1 What is a task?

Since the emergence of TBLL, researchers have provided various definitions of *task*, the unit at the heart of a TBLL approach (Long 1985; Crookes 1986; Richards, 1986; Breen 1987; Prabhu, 1987; Nunan, 1989, 2004, 2006; Carroll, 1993; Willis, 1996; Skehan, 1996, 1998; Lee 2000; Bygate, Skehan and Swain, 2001; Ellis, 2000, 2003). A thorough look at the definitions reveals that these vary according to the focus of research and the perspective adopted by the researcher: cognitive, linguistic, or pedagogic. Early on, most definitions were relatively general and did not list features or components of a task. Since then, definitions have become more detailed and have highlighted particular features of a task.

One of the earliest definitions of a *task* is by Long (1985). Long argued that a task is

> a piece of work undertaken for oneself or for others, freely or for some reward. Thus examples of tasks include painting a fence, dressing a child….In other words, by ‘task’ is meant the hundred and one things people do in every-day life, at work, at play and in between. (Long, 1985, p.89)

What is interesting about Long’s definition is that it does not relate tasks to events that take place (or might take place) in the classroom or even to the use of language itself. Tasks are more related to real-world events in Long’s perspective. In contrast, Crookes (1986) and Richards (1986) defined tasks in a more specified context. Crookes stated that a task is ’a piece of work or an activity, usually with a specified objective, undertaken as part of an educational course, at work, or used to elicit data for research’(Crookes, 1986, p. 1). Even though Crookes was more precise in his definition, he did not detail the features of a task. Similarly, Richards' (1986) definition was ambiguous and generalised the use of the task to any activity that involved the processing or understanding of language. So, any activity in an educational context would be considered a task under Crookes or Richards' definitions. At that point, tasks did not necessarily involve the production of language and none of the previous definitions linked tasks to an outcome. It was not until 1987 that tasks were associated with the necessity to have an outcome. Prabhu (1987) and Breen (1987) were the first to point this out:

> Any structured activity which required learners to arrive at an outcome form given information through some process of thought and which allowed teachers to control and regulate that process was regarded as a 'task'' (Prabhu, 1987, p. 24)
> 'Any structured language learning endeavour which has a particular objective, appropriate content, a specified working procedure and a range of outcomes for those who undertake the task. (Breen, 1987, p.23)

Yet both definitions did not specify the focus of the task or the primary goal of completing the task. In the mid 1960s, Skehan provided a concise account of what a task is in one of the most used definitions in TBLL research. Skehan (1996, 1998) was very specific in his definition and outlined the features that constitute any activity as a task. For him, an activity is a task only when its main focus is on meaning, it is related to real world, task completion is essential and the assessment of performance is in terms of the task outcome. The same features are echoed in Lee's definition (2000). Lee argued that a task is

1. A classroom activity or exercise that has
    a. an objective obtainable only by the interaction among participants,
    b. a mechanism for structuring and sequencing interaction, and
    c. a focus on meaning exchange
2. A language learning endeavour that requires learners to comprehend, manipulate, and/or produce the target language as they perform some set of work plans

Nunan (2004) added that learners' attention during a task should be focused on conveying meaning rather than manipulating form. According to Nunan (2004), tasks should also have a beginning, middle and an end.   He divided tasks into two kinds: real world tasks and pedagogical tasks. The first one takes place in the real world while the second take place in the classroom.

Ellis 2009 makes a crucial and useful distinction between ‘task’ and ‘situational grammar activity’: a task needs to meet four conditions: the primary focus is meaning; there should be a gap, for example, to convey information; learners have to rely on their resources to complete the task and there is a clearly defined outcome (Ellis, 2009, p. 223). This is distinguished from a ‘situational grammar activity’ where the focus is not on meaning and the outcome is simply practicing correct language. Ellis (2009) further differentiates between focused and unfocused tasks where unfocused tasks aim at providing opportunities for language use in general while focused tasks provide these opportunities while focusing on specific linguistic features. Ellis (2009) warns that a focused task is not the same as situational grammar activity in the sense that the target linguistic feature is hidden in focused tasks while explicit in the situational grammar activity. What is important here is that ‘learners are expected to orient differently to a focused task and a situational grammar exercise’ (Ellis, 2009, p. 224). This distinction is very important as it indicates that even when the context is exactly the same and the only difference is that learners are made aware of the linguistic feature, there will be differences in the outcome and in how learners face the tasks.

Ellis’ types of task conform to the three theoretical underpinnings of the three main types of input available in the classroom and used in this study (see Table 2.2). In the course of the present thesis, Ellis’ distinction is taken as the basis for the design of the three different types of input.

Table 2.2 Task and Input types

Input Type	Focus on Meaning (FoM)	Focus on Form (FoF)	Focus on Forms (FoS)
Activity Type	Unfocused tasks	Focused tasks	Situational grammar activity

As it is obvious from the discussion, although the definitions vary in their focus, most of them share basic assumptions. Tasks are used to practice the target language for communicative purposes rather than for the sake of practice alone. Moreover, a task should be meaning focused, related to the real world and evaluated through outcome, where it is usually the teacher who decides what a successful completion or outcome is.

Having established the definition of task, I turn now to address TBLL from cognitive and pedagogical perspectives.

2.5.2.2 TBLL:Cognitive perspectives

Many researchers argue that tasks should be designed and sequenced in ways that increasingly approximate the demands of real-world target tasks (Long and Crookes, 1992; Robinson, 1998, 2001, 2002, 2005). As mentioned in the previous section, TBLL is promoted as a balanced approach where the main focus is meaning but attention to form is permitted or required under controlled conditions. Cognitive demands are signalled as a crucial factor that contributes to task effectiveness. Different types of task put different cognitive demands on learners and force them to direct attention to meaning or form alone or both. Robinson (2001) states that 'the greater the *cognitive demands* of a task, the more they engage *cognitive resources* (attention and memory), and so are likely to focus *attention* on input and output, which will have *performance effects*' (p. 305). It could be stated then that the more we understand the task's cognitive load, the better we will be able to interpret the learners' performance on it. Prominent scholars who have researched the cognitive demands of tasks include Skehan (1996), Skehan and Foster, 2001), Robinson (2001, 2002, 2005, 2011) and Ellis (2005, 2009). According to Skehan (1996), *demand* is directly related to the amount of processing required by a given task. Robinson (2001) identifies task demands as the 'attentional, memory and reasoning demands of tasks that increase the mental workload the learners engages in performing the task' (Robinson, 2001, p. 302). Cognitive load or demand could be ascribed to different elements of task design and implementation including task familiarity, task focus, task complexity, task planning and communicative stress/time pressure. I will look at some of these elements below.

Task familiarity is a crucial factor that contributes to cognitive load. If the tasks are familiar, the learners are able to draw on their existing knowledge and experience to complete the task which entails relatively light cognitive load. If the tasks are unfamiliar, learners require more cognitive processing to work out solutions. Skehan (1998) puts the debate about task familiarity in simple terms:

> Tasks based on familiar information with clear discourse structure, for example of a pair of students giving one another instructions to get to their respective home, will probably have low task demands, while a task requiring imagination and abstraction, and a complex outcome, such as agreeing on the solution to a moral problem, will probably make much higher ones. (Skehan, 1998, p.51)

The implication of task familiarity is strongly associated with the task focus. If processing has to be directed at the cognitive problem involved in the task, there is less attention left to focus on forms.

Another element that contributes to task demand is the task focus. In precise words, the degree to which the task requires learners to direct their attention to forms or meaning is a determining factor of the cognitive processing load. VanPatten (1996, 2002, 2004) argues that learners' priority is always to process the input for meaning unless they are pushed to process it for forms (see section 2.2.5). Similar to VanPatten, Schmidt (1993, 1995, 2012) also emphasizes the importance of attention to forms during a task as the basis of learning, i.e. what is learned is what is noticed. When the learner is under processing pressure, attention to forms is not possible. One way of redirecting the learner's attention to forms is that processing conditions need to be pedagogically manipulated to maximize the opportunities for focusing on form. Hypotheses and theories that have attempted to manipulate the input for this purpose include Input Enhancement (Sharwood Smith, 1991, 1993) and Input Processing (VanPatten, 2002, 2004). Input Enhancement focuses on the input and entails increasing the saliency of linguistic forms in the input to raise the learner's awareness of them. Rutherford and Sharwood Smith (1985) used consciousness raising in a task-based approach to provide input that is communicative but requires the learners to attend to language forms consciously. Input Processing (VanPatten), on the other hand, focuses on processing conditions and involves manipulating the input to increase the opportunities of making form-meaning connections by learners. The implications of both models is that focusing on both meaning and forms is more cognitively demanding than focusing on forms or meaning alone.

Ellis, on the other hand, argues that tasks are different in terms of the knowledge they measure or tap into. He maintains that a distinction needs to be made between tasks that measure or require application of explicit knowledge and those that measure implicit knowledge as they implicate unequal cognitive load. He operationalized the constructs of implicit and explicit knowledge and put forward the criteria that distinguish them, as shown in Figure 2.5 below.

Characteristics	Implicit knowledge	Explicit knowledge
Awareness	Intuitive awareness of linguistic norms	Conscious awareness of linguistic norms
Type of knowledge	Procedural knowledge of rules and fragments	Declarative knowledge of grammatical rules and fragments
Systematicity	Variable but systematic knowledge	Anomalous and inconsistent knowledge
Accessibility	Access to knowledge by means of automatic processing	Access to knowledge by means of controlled processing
Use of L2 knowledge	Access to knowledge during fluent performance	Access to knowledge during planning difficulty
Self-report	Nonverbalizable	Verbalizable
Learnability	Potentially only within critical period	Any age

Figure: 2.5 Key characteristics of implicit and explicit knowledge (Ellis, 2005, p.151)

Once Ellis established the distinction between the two types of knowledge, Ellis (2005) elaborated on the criteria and elements a task needs/has that determine if implicit or explicit knowledge is measured or tapped. The elements are summarized in Table 2.3.

Table 2.3 Operationalizing the constructs of implicit and explicit knowledge (based on Ellis, 2005)

Criterion	Implicit knowledge	Explicit Knowledge
Degree of awareness	Learners make use of *feel* to respond to the task	Learners make use of *rule* to respond to the task
Time available	The task is time pressured	The task is performed without time pressure
Focus of attention	The task's primary focus is on meaning and conveying information	The task's primary focus is on form as in traditional grammar exercise
Systematicity	The task results in consistent responses that tap into implicit knowledge	The task results in variable responses that elicit explicit knowledge
Certainty	the learners are confident about their responses to the task	Learners are less confident about their responses
Metalanguage	The task does not require the learner to use metalinguistic knowledge or terms	The task requires the learners to use metalinguistic knowledge
Learnability	The task is more suitable for learners who began learning as children	The task is more suitable for learners who have received formal instruction

Ellis (2005) has rightly pointed out that even when a task is designed, according to the seven features mentioned above to encourage learners to use one type of knowledge, it is 'impossible to construct tasks that would provide pure measures of the two types of knowledge' (p. 153).   He also highlights that when dealing with tasks, learners are likely to draw on other resources available to them at the time, not only on the ones provided by the task.

Task complexity is probably the most widely researched element of task design. Robinson (2001) differentiates between task *complexity* and task *difficulty*. He uses complexity to refer to the 'design features of tasks, which are proactively manipulative by the task designer, and can be used as the basis of sequencing decisions' (Robinson, 2001, p. 295). Difficulty, on the other hand, refers to 'learners' perceptions of the demands of the task' which are determined by affective (such as motivation) and ability (such as aptitude) factors (Robinson, 2001, p. 295). He also distinguishes between task demands that affect performance and those that affect development. He argues that performance and development task demands can be manipulated separately but 'they are often drawn on simultaneously during real-world performance' (Robinson, 2005, p. 2).

Task complexity (cognitive factors)	*Task conditions* (interactional factors)	*Task difficulty* (learner factors)
(a) resource-directing e.g., ±few elements ±Here-and-Now ±no reasoning demands	(a) participation variables e.g., open/closed one-way/two-way convergent/divergent	(a) affective variables e.g., motivation anxiety confidence
(b) resource-dispersing e.g., ±planning ±single task ±prior knowledge	(b) participant variables e.g., same/different gender familiar/unfamiliar power/solidarity	(b) ability variables e.g., working memory intelligence aptitude
Sequencing criteria	*Methodological influences*	
Prospective decisions about task units	On-line decisions about pairs and groups	

Figure: 2.6 A triad of task complexity, task condition and task difficulty factors (Robinson, 2005, p. 5)

What is interesting about Robinson's (2005, 2007, 2009) argument is that he assign a crucial role to perceived complexity, i.e. whether learners perceive the task as easy or difficult. Cognitive, perceived and design complexity all contribute to task complexity. I will look later at some studies that have investigated task complexity in relation to task type.

Another crucial element of task demand is related to communicative stress, particularly time pressure. Time pressure is linked to all the previous elements. For example, if the tasks are familiar, learners most often use 'analogical problem solving' strategies to complete the tasks. However, if the tasks are not unfamiliar, learners spend long time thinking before deciding on their moves. Learners need to use their old knowledge to interpret and understand the new problem. This means that with unfamiliar task, learners spend more time processing the input. Furthermore and as mentioned earlier, VanPatten's work has revealed that learners cannot attend to meaning and form in the input at the same time (see section 2.2.5 and earlier discussion in this section). Learners process the input for meaning first in order to verify the communicative message, and only when comprehension takes place, attention turns to form. The implications of VanPatten's findings are that learners who attend to form and meaning will spend more time processing the input compared to learners who process input for meaning only. The other implication of VanPatten's findings is that the urgency required for task-completion affects the learner's decision to process the input for form or meaning. This assumption complies with findings of earlier research on cognition conducted by Chaudron. Chaudron (1985) proposed a model to account for the dimensions of tasks. In his model, he used the 'degree of processing' to refer to the amount of time involved in an activity and the various stages of intake processing. Chaudron argues that activities/tasks that involve more processing require more time.

Several studies have been conducted to explore the extent to which design features of the task can manipulate learner attention and lead to noticing of linguistic forms; or can affect the learner's performance. Other studies have compared the effectiveness of different task types. I will look at some studies below, but the discussion will be limited to the studies that have focused on the acquisition of morphosyntactic and grammatical structures[12] as these structures are the focus of the current study.

For example, Tarone (1985) examined the effects of different tasks on the production of morphological and grammatical forms and found that participants' performance varied considerably across task types. Surprisingly, learners' performance in spontaneous production (such as an interview task) was more accurate than in grammaticality judgment tasks. At the end of her study, Tarone cautioned against interpreting results from learners' performance on different activities as it might not represent the learner's interlanguage state. Hulstijn and Hulstijn (1984) looked at the effects of time and focus (content or grammar). They interviewed learners after completion of the task to assess their explicit knowledge. Attention to grammar positively affected accuracy but it was not related to the learners' explicit knowledge. There was no effect for time.

[12] When it comes to vocabulary, there is a large body of research on the effectiveness of different task type. See Hulstijn et al., 1996; Hulstijn & Laufer, 2001; Laufer, 2001; 2003, 2005, 2006, 2011; Webb, 2005; Peters, 2006; Peters et al., 2009; Keating, 2008; Kim, 2008.

In a later study, Hulstijn (1989) investigated the effect of attention to form, meaning or both on the acquisition L2 Dutch and an artificial language. Learning took place implicitly and incidentally. The learners were divided into three groups: Form, Meaning and Form and Meaning and learners were post- tested using cued recall tests. The findings indicated that the Form group performed better in terms of the structural/grammatical aspects of the recall test while the Meaning and the Meaning and Form groups performed better on the content aspects of the test. The Form and Meaning group outperformed the other two groups when content and grammar were combined. The results indicate that attention during tasks affects performance: attention to form leads to better intake of form while attention to meaning leads to better intake of the meaning.

Swain and Lapkin (2001) examined attention by using two different tasks, a dictogloss, which was expected to encourage learners to focus on form, and a jigsaw, which was expected to offer the learners opportunities for negotiation of meaning. The learners' performance was more accurate and complex on the dictogloss task, but they did not produce many form-related episodes during interaction. There were no significant differences, however, across the tasks in relation to post-test scores or form-focused episodes.

In an interesting study, Hu (2002) used different tasks to examine the extent to which explicit knowledge is available for use in spontaneous writing. On the basis of the empirical findings, Hu asks whether the question of L2 knowledge use is not a matter of implicit versus explicit knowledge, but rather a matter of the extent of explicit knowledge use made possible by the interaction between different elements including automaticity, prototypicality, and task conditions. In Robinson's (2005, 2007) reports on findings of several studies he conducted. He notes that most show that task complexity affects the quantity and quality of interaction, the amount of uptake and intake and performance.

Following on from Sharwood Smith's ideas, Eckerth (2008) investigated the effect of conscious raising tasks on short and long term through the use of dyadic tasks and pre-test, post-test and delayed post-test. Eckerth conducted an experiment over five weeks using three different text repair tasks and two different text reconstruction tasks. The tasks were presented to two classes, with a frequency of one task each week. The results showed significant learning gains for the short and long term for the conscious-raising tasks. Furthermore, the results revealed that the learning gains are not limited to the target structures, but extend to non-targeted L2 elements.

In a more recent study, Reinders (2009) examined the effects of three task types: dictation, individual reconstruction and collaborative reconstruction on the acquisition and uptake[13] of linguistic forms. One major difference, Reinders listed between the three tasks is the extent to which they engage cognitive processing. The dictation task is viewed as less demanding as it only required learners to memorize the sentences for a very short time. The reconstruction tasks, however, required learners to remember longer texts for a longer time, so they are considered as more cognitively demanding. Reinders' results revealed that there were no differences between the three tasks on acquisition, but the dictation group outperformed the other groups on uptake.

Sasayama (2011) investigated the effects of cognitive task difficulty on ESL learners' written and oral performance. The results indicate that the difficult writing task elicited more complex production than did the simple task and the same was observed for the speaking tasks. Accuracy, though, remained at the same level across all tasks. Sasayama argued that the results show that learners are capable of directing their attention to complexity without sacrificing accuracy 'when tasks are designed to pose higher cognitive, functional and linguistic demands for their successful completion' (Sasayama, 2011, p. 123)

Other studies have investigated the effect of task planning, i.e. pre-task activity or planning time, on the use of explicit knowledge (For example, see Foster and Skehan, 1996,1999; Mehnert, 1998; Ortega, 1999; Skehan and Foster, 1997,1999). These are not discussed here as the current study does not employ task planning conditions in the design.

It is clear from the discussion and review above that the verdict is still out on which task type provides learners with the most opportunities for negotiation of input and modification of interlanguage which results in language development or whether better opportunities for language development result from the sort of task design which

[13] Uptake is measured in terms of the learners' suppliance of the target structure during the treatment while acquisition is measured in terms of the learners' performance on the post-tests.

triggers more cognitive processing, leads to noticing and allows opportunities for recycling information. Although there is less of a consensus on what constitutes tasks' cognitive demands, there is an agreement that the more cognitively demanding tasks, the more the learners will prioritize focusing on meaning and will draw more on implicit knowledge.

In this section, I looked at the cognitive underpinnings of tasks. In the next section, I focus on the pedagogical and methodological keystones of tasks.

2.5.2.3 TBLL: Pedagogic perspective

The pedagogical aspects of TBLL in particular and FoF in general have been at the heart of a large body of research (Long and Crookes, 1992; Skehan, 1996, 2003; Ellis, 2000, 2003, 2006; Bygate, Skehan and Swain, 2001; Willis and Willis, 2007; East, 2012). There is no doubt that the ways tasks are implemented in the classroom have a considerable impact on their effectiveness. As can be inferred from the discussion in the previous sections, within the framework of TBLL, tasks are used as pedagogical tools to direct learners' attention to form or meaning; to raise learners' awareness of linguistic forms and to increase their chances of making form-meaning connections. Task features and design that facilitate these objectives have been researched from a pedagogical perspective to inform the design of syllabus and curriculum and to create more favourable conditions that better aid learning. In this section, I examine the views and issues related to implementing TBLL in the classroom.[14]

Ellis (2006) states that various designs have been proposed regarding how tasks should be presented in a lesson (Willis, 1996; Skehan, 1996; Lee, 2000; Ellis, 2003, 2006; Klapper, 2003; East 2012). The main argument is that classroom tasks should imitate the same conditions of real life tasks in order to be effective and aid learning (Van den Branden, 2006; Willis and Willis, 2007). Van den Branden maintains that tasks should 'elicit the kinds of communicative behaviour (such as the negotiations of meaning) that naturally arise from performing real-life language tasks, because these are believed to foster language acquisition' (Van den Branden, 2006, pp. 8-9). Similarly, Willis and Willis claim that 'the most effective way to teach a language is by engaging learners in real language use in the classroom' (Willis and Willis, 2007, p. 1), which can be done

[14] Studies that investigated TBLL from the teachers' perspectives are not reviewed as they are out of the scope of the current study which is conducted in a self-accessed language learning environment. For further information, see Van den Branden (2006, 2009), Carless (2007, 2009) and East 2012b.

through the use of tasks. Learners are supposed to engage in real life like tasks which require them to use the language for communicative purposes to achieve an outcome. Klapper (2003) distinguishes between strong and weak CLT and argues that TBLL is more effective than both versions of CLT. As outlined above, the belief underlying strong CLT (Howatt, 1984) is that learning happens through natural processes in the learner's mind and the teacher has no knowledge of or control over these processes so his/her role is limited to providing activities that promote these processes. Thus, learners learn to communicate by communicating. The problem with strong CLT is that it does not lead to the full development of linguistic competence, i.e. to accuracy, as argued by Swain. Swain (1985) argues that 'simply getting one's message across can and does occur but with grammatically deviant forms and sociolinguistically inappropriate language' (Swain 1985, p. 248). On the other hand, in the weak version of CLT, the primary focus is still meaning and communication and attention to form is permitted to overcome communicative difficulty. As noted in section 2.5,weak CLT is usually applied through PPP which is a technique associated with traditional methods. Here, Klapper argues that TBLL resolves most of the limitations associated with strong and weak CLT in the sense that it provides sufficient comprehensible input and also opportunities for the learners to use the language in meaningful ways. Klapper maintains that TBLL is an offshoot of CLT but it differs in that it assigns a role for instruction in triggering learning. Klapper acknowledges that TBLL is based on sound principles and conforms better than many pedagogical models to what is known about SLA but points out some shortcomings which are discussed in the next section.

There is more than one proposal for task-based syllabi: the Procedural Syllabus (Prabhu, 1987), the Process Syllabus (Breen, 1984) and Long's Task Syllabus (Long and Crookes, 1992). All of these adopt the task as the unit for building up the syllabus. The Procedural Syllabus comprises a series of tasks divided into opinion-gap, information gap and reasoning-gap. The focus of this syllabus is task completion. It is criticised for its lack of evaluative components and specificity of the notion of task (Brumfit, 1984; Long and Crookes, 1992). The Process Syllabus (Breen, 1984, 1987) focuses more on learning than on language. Learners have control over the choice of tasks, objectives, and content. One of the criticisms addressed to the Process Syllabus is that it requires high levels of linguistic competence from learners to be able to negotiate the tasks and their contexts (White, 1988). The third syllabus is the Task Syllabus. Long and Crookes (1992) attempt to provide an integrated, internally coherent approach to TBLL which is compatible with SLA theory. They use tasks to present appropriate target language samples and to deliver comprehension and production opportunities of negotiable difficulty. They distinguish between *target* tasks, which are real tasks in everyday life and *pedagogic tasks* which are derived from target tasks and adapted for the classroom. Learners are assessed based on how they complete the tasks according to pre-set criteria established by experts in the field. Long's Task Syllabus seems to be the only one that was later subjected to empirical research, and is the closest to what is now known as TBLL.

Ellis (2009) argues that there is no 'single' way of doing TBLL and compares his own approach to TBLL (2003) with Long's (1985) and Skehan's (1998) as shown in the figure below.

Characteristic	Long (1985)	Skehan (1998a)	Ellis (2003)
Natural language use	Yes	Yes	Yes
Learner-centredness	Yes	Yes	Not necessarily
Focus on form	Yes – through corrective feedback	Yes – mainly through pre-task	Yes – in all phases of a TBLT lesson
Tasks	Yes – unfocused and focused	Yes – unfocused	Yes – unfocused and focused
Rejection of traditional approaches	Yes	Yes	No

Figure: 2.7A comparison of three approaches to TBLL (Ellis, 2009, p. 225)

Ellis points out that all three approaches are similar in their emphasis that TBLL should provide opportunities for natural language use and should include devices for focusing learners' attention on form. They differ, however, on how focus on form is achieved. . Rodgers (2009) concedes that one of the difficulties of TBLL as a methodology is the lack of consensus of what constitutes a *task*. As mentioned in section 2.5.2.1 above, there is a consensus that a task should be meaning –focused, outcome-evaluated, rand related to real world. But unfortunately, the consensus stops there. There is less agreement on what an outcome is and how it should be evaluated. Ellis (2006) argues that although differences exist as to what constitutes a task, all accounts of task design have three phases in common: pre-task, during task and post- task (e.g. outcome). Ellis also points out that the only obligatory phase is the 'during task' while the other two are optional. According to Rodgers (2009) a number of systems have been made to group tasks into categories as a basis for their design. Pica, Kanagy and Falodun (1993) talk about jigsaw, information-gap, problem-solving, decision-making and opinion exchange tasks. Willis (1996) classifies tasks into six types: listing, ordering, comparing, problem-solving, sharing personal experiences and creative tasks. Under all these classifications and definitions, it seems easy to consider any activity in the classroom a task. As stated in section 2.5.2.1 above, Ellis (2009) argues that although various definitions of tasks are provided, most of them approve that for a language teaching activity to be a 'task', it must meet the following criteria (Ellis, 2009, p. 223):

- The primary focus is on meaning
- There is a gap such as conveying information or inferring meaning
- Learners need to rely on their resources to complete the activity
- There is a define clear outcome

Nassaji and Fotos (2004) notes that meaning-focused tasks 'containing communicative instances of target forms are useful for developing learner awareness of grammar structures that are too complex to be understood through formal instruction alone' (Nassaji and Fotos, 2004, p. 135).

To summarise the argument in this section, it is obvious that there is neither one way of implementing TBLL nor one way of designing tasks. Evidence from the literature suggests that tasks with grammatical structures, whether presented implicitly or explicitly, and cognitively demanding tasks, seem to be more effective in promoting awareness of the target structures and providing learners with plentiful of opportunities to use the language in meaningful ways.

2.5.2.4 Concluding notes on TBLL

In their book *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*, Bygate, Skehan and Swain (2001) provide an extensive discussion of TBLL and its application and implications. Most of the discussion can be summarized by noting that there is no single way of implementing, designing or evaluating TBLL. However, the main debate and issues in TBLL research and methodology can be summarized in the following points:

- In task-based language learning, success is defined by communicative success not formal accuracy. Therefore, tests need to be modified to account for this fact. (East, 2012)

- Task based learning instruction must be ‘challenging’ and ‘engaging’ (Platt and Brooks, 2002; Ellis, 2003).
- Whether TBLL is ‘viewed in terms of syllabus or methodology, it is clearly incorrect to claim that it ‘outlaws grammar’. Grammar may not be central to TBLT, but it has an important place within it’ (Ellis, 2009, p. 232).
- Successful performance on a classroom task is not necessarily a good predictor of acquisition (Reinders, 2009).

The increasing interest in TBLL should not shelter the critical voices of applied linguists and educationalists that highlight its methodological and empirical limitations. For example, Widdowson (2003) highlights that TBLL overemphasizes the importance of authentic language use. He also points out that the defining features of tasks are overly loose. On the other hand, Seedhouse (1999, 2005) argues that a ‘task’ does not constitute a valid unit to create a language teaching curriculum. Voicing another criticism, Swan (2005) states that the essential argument against TBLL in instructed learning contexts is that ‘time and opportunities for language exposure are limited. Similar concerns are raised by Bruton (2002), who argues that the language used in interaction in TBLL leads to uneven oral development and is inappropriate for beginners. He also argues that coverage of all structures is not possible in TBLL, and the type of interaction in meaning-focused activities means that learners are exposed to large amount of non-native input which might not be helpful for their interlanguage development (Young-Scholten, 1995). Klapper (2003) claims that there is no guarantee that a purely task-based approach will provide information about all the target structures that learners need to encounter. He also claims that based on the evidence that there are many successful learners who had only been taught by traditional methods, TBLL ‘has a lot more work to do before it can provide a convincing alternative pedagogical model, still more before it can claim superiority over other approaches’ (Klapper, 2003, p. 40). Other researchers have also questioned the applicability of TBLL to different learning settings where traditional approaches are dominant or to typical secondary classrooms (Li, 1998; Carless, 2004; Brunton, 2005; Butler, 2005). However, Ellis (2009) reviewed most of the critical points discussed above and managed to refute most of them based on empirical and theoretical evidence.

Regardless of sceptical voices, there is no doubt that TBLL is increasingly gaining interest and strong support from applied linguists and educationalists (Willis 1996;

Long and Norris 2000; Bygate, Skehan and Swain, 2001; Skehan 2002; 2003; Ellis 2003; Nunan 2004; Edwards and Willis 2005; Van den Branden 2006; García Mayo 2007; Samuda and Bygate 2008). While there is still need for research that evaluates TBLL as a whole curriculum in real classrooms rather than more research that investigates the effects of a single task or a limited number of tasks, mostly in laboratory/classroom settings. It is also necessary to continue to develop our understanding of what TBLL is or ought to be in time-constrained FL contexts as pointed out by East (2012a). The current study serves this purpose; it is an attempt to contribute to the on-going consideration of the perceived suitability of TBLT in a specific context. Tasks are used to deliver the different types of input discussed above in a technology-enhanced learning environment, a setting under researched so far. User-behaviour tracking technology is used to record all learners' actions to obtain a picture of how learners deal with the tasks/input in real time. Such data is valuable in exploring the effectiveness of tasks as highlighted by East

In this section I discussed TBLL and the effects of pedagogical and cognitive task features. In the next section I look at individual differences as one of the variables that can explain variations in the process and outcome of language learning. In applying tasks in a technology-enhanced learning environment, it is possible to track such differences, as we shall see in subsequent chapters.

2.6 Individual differences in ISLA

There is no doubt that individual differences have great effects on the process and product of learning. Robinson (2001) argues that research into the effects of individual differences on learning under different exposure conditions is well established in the fields of cognitive psychology and psychotherapy. He highlights that individual differences research can strongly contribute to the understanding of SLA in general, and notes the following three issues in particular. Among other things, individual differences can:

- Explain variation in language learning success under particular instructional conditions.
- Explain differences between, implicit, incidental, and explicit learning processes
- (Robinson, 2001, pp. 368-369)

Even very early on, individual differences were credited as variables in ISLA. Individual variables were examined from different standpoints: cognitive, generative and pedagogical. Within the cognitive paradigm, it has been assumed that variation results from differences between 'learning characteristics and learning contexts' (Robinson, 2002, p. 2). Leow (1995) concedes that 'learners at different levels of language experience demonstrate a different pattern of performance while internalizing written input' (Leow, 1995, p.80). Although, Leow only focused on written input, his statement is evidently applicable to all types of input. On a more general basis, Lightbown and Spada (2006: 177–178) state that it is necessary to 'find the balance between meaning-based and form-focused activities,' and although they acknowledge that this is not as easy to do as it sounds, the right balance is likely to be influenced by the characteristics of the learners when decisions are made about the amount and type of form focus to offer.

Within the SLA field, individual differences are mainly discussed in terms of age, exposure to the language, type of exposure and type of instruction. The interaction between type of instruction and individual variables is the most interesting for the present study. However, de Graff and Housen (2009) argue that research on the interaction between individual learner variables and instruction has been too restricted to draw any firm conclusions. On the other hand, individual differences were extensively researched in language pedagogy in terms of motivation, aptitude, learning styles and learning strategies. As interesting as some of these variables are, they are not the focus of the present study as the focus here is on the patterns of decisions that learners exhibit while dealing with different types of exposure. In order to examine how language develops, it is very important to look closely at language use and learning processes from an individual perspective. Looking at individual differences under different exposure conditions might reveal how resource allocation and learning mechanisms work. Such research might help in explaining how information available to the learner is processed. The current study included qualitative analysis of selected learners to do so.

I turn now to a brief justification of why I chose to conduct the experiment in the current study in a Technology-Enhanced Language Learning environment before I move to Chapter Three where I revisit some of the issues discussed in this chapter from within a TELL perspective.

2.7 Why the technology-enhanced environment?

As can be seen from the previous sections, all areas of language learning and teaching are undergoing changes, in some cases drastic ones. At the same time, advances in technology have revolutionized how learning takes place. It is not surprising then that there is an increasing call for research that investigates the effects of the new technologies on the learning process and product; parallel with this is a call to utilise the findings of ISLA research to inform the design of technology-based materials. These two calls, among other reasons detailed later, drive my choice to conduct the present study in a TELL environment.

Although the use of technology for language learning has been at the heart of many research papers and books, there is a small number of SLA researchers who question whether the use of new technologies in language instruction furthers second language acquisition. In addition, Computer-Assisted Language Learning (CALL) researchers agree that the biggest problem in CALL systems that have been developed so far is the lack of underlying theory based on ISLA findings that could be used to design better courseware (Levy 1997; Chapelle 1997, 2001).

Chapelle (2009a) points out that when researching input, SLA has ignored the role of CALL from the beginning. She argues that in a CALL environment, CALL designers have the option of selecting, sequencing and modifying the input, and they base their design decisions on a theorized role for the various types of input. Moreover, learners also, when dealing with the material, have the same options; each sequence, selection or modification that learners or teachers make results in a different type of input and consequently, different processes and products. In other words, the individuality of the learning process that CALL materials offer has not been examined adequately from an applied linguistics perspective. However, SLA has started to incorporate different approaches that do assign a role for instruction and CALL (Ortega, 2005).

As Hulstijn (2000) put it 'overall, one could say that SLA data are seldom elicited with the use of computer-aided techniques' (p. 32). But the advantages of using technology-supported data collection instrument are vast. Schrooten (2006) summarizes the potential of technology for language learning in the following points (p. 129):

- It allows a high degree of differentiation where individual needs and abilities are met;

- The use of technology elicit a high degree of learner motivation and involvement;
- Technology offers enriched content in different modes;
- The use of technology frees the teacher's time so s/he can be more supportive.

Schrooten maintains that the use of technology in language teaching is surely gaining ground. One of the main advantages of using technology /CALL in SLA research is that instructional conditions can be controlled for in better and easier ways than in classroom research (Chapelle, 2007). Moreover, if technology is used effectively in the classroom, it can help in overcoming input-related inadequacies such as insufficient, impoverished or interlanguage input as it allows access to relatively large authentic samples of the target language that are suitable for different developmental levels. Poole's (2005) study detailed in section 2.5 above, concluded that for FoF to be effective in classroom, classes should be small enough for teachers to be able to work individually with students and teachers and students should be proficient enough in English so they do not switch to their native language when communicative difficulties are encountered. These limitations can be easily resolved by using technology. There are already TELL materials designed in a way that modifies input to correspond to individual responses and needs. Since the interaction is mainly learner-machine rather than learner-learner or teacher-learner, the chances of code-switching into the learner's native language are non-existent. Chapelle (2003, p. 55) claims that 'a useful theory of interaction in CALL needs to define broadly what interaction consists of, what kinds of interaction are believed to be important for SLA, and why. This general understanding provides an essential basis for conceptualizing and evaluating the new types of interaction made available through CALL.'

TBLL has been criticized on the basis that it might not provide sufficient opportunities for FoF and that the teacher will not be able to meet individual learners' needs and might not always have the resources to provide form-focused information. One advantage of using technology in a FoF classroom alongside tasks is that it allows access to a vast number of resources in different media, so when a learner's need arises, it can easily met by a click.
So far, most studies concerning input based approaches to grammar instruction and interaction have been carried in naturalistic, immersion, laboratory and classroom settings. Research on classroom interaction mainly focuses on analysing spoken

interaction to identify form-focused or meaning-focused events. But as Chapelle has pointed out, the nature of interaction in a CALL environment is different, and one should not assume that what apply in the classroom is applicable to a CALL context. There is also a widely accepted view now that the nature of communicative competence has changed dramatically over the last years with technology a major influence. In addition, the availability of online help and dictionaries options has resulted in different linguistic choices being made by learners.

Garrett (1998) points out that little research has been conducted to assess learners' behaviour to gauge whether learning goals have been met. In order to do this, all actions and choice made by the learners need to be recorded and analysed, a task made much easier by using the advances in technology as will be demonstrated in Chapter Three. On another note, it has been argued that one of the problems with the cognitive-based processing models studies that they are mainly based in laboratory settings which limit the applicability of the findings (Mitchell and Myles, 2004). Also, as studies often focus on the spoken interaction among teachers and learners, the focus is actually on communication rather than on cognitive and processing skills. Even studies that attempt to focus on processing and use think aloud protocols have limitations (see Chapter Four). Asking learners to report on their thinking while carrying out the treatment put extra demand on their cognitive abilities which then affects their performance and calls into question the validity of the data.

Most importantly and as rightly pointed out by Ellis (2010) in his forward to Thomas and Reinders (2010) *Task-Based Language Learning and Teaching with Technology*, the literature so far on TBLL has dealt 'almost exclusively with TBLT as practices in face-to-face (FTF) classrooms. There is still relatively little published about TBLT in technology-mediated contexts' (Ellis, 2010, p. xvi). Two years on from Ellis' (2010) remark and there is still a dearth of research on the application of ISLA theories and models in general and TBLL in particular to technology-enhanced learning settings.

The current study aimed at filling this gap in the literature. By conducting the experiment in a TELL setting, we are providing insights into the applicability of the findings of ISLA research to the different learning environments and also on the possibility of borrowing whole theories and models from ISLA to inform the design of technology-enhanced materials. The current study thus aimed to provide information on how learners process input, and what contexts or aspects of processing input lead to more efficient processing, as reflected in better performance. One of the main reasons behind conducting the study in a technology-enhanced environment is that it allows us to obtain reliable and detailed records of all the learners' behaviour at the same time with no interference from the researcher and no extra cognitive demands on the learners (as is the case with the think-aloud protocol). It has been pointed out in Chapter One that the study is not an acquisition-based study as the input provided during the experiment could have been influenced by many factors (see Chapter Three) and is in no way assumed to be sufficient for the learners to have acquired the target construction. In fact, the goal of the study was to replicate three classroom scenarios (FoF, FoS and FoM) but in a self-accessed technology-enhanced environment with textbook materials widely used in the EFL classroom while tracking learners' behaviour in real time. In order to be able to track all actions taken by the learners, it was important to limit the amount of input provided. As pointed out by Chun (2013), 'data documenting what learners actually do in CALL activities can provide valuable insights into both second language acquisition and pedagogical design' (p. 256). Among the benefits that Chun (2013) lists of collecting and analysing such data are: ascertaining precisely what learners do or do not do and determining whether there is a relationship to learning and documenting the learning process (Chun, 2013, p. 256). Thus, performance was assessed in relation to the factors (FoF, FoS, FoM) that led to un/successful learning. In brief, the main goal of the present research is to closely examine the learning process, in this sense the use of technology to deliver the input and collect the data is vital. This will be discussed in more details in Chapter Four. In the next Chapter, I will review the literature on Computer-Assisted Language learning revisiting some of the issues discussed in this chapter.

Chapter Three

Technology-Enhanced Language Learning

3.1 Introduction

As mentioned in Chapter Two, the present study is conducted in a Technology-Enhanced Language Learning (TELL) environment. This chapter provides a detailed review from an educational technology perspective of the studies and theoretical concepts underpinning the study. The review is not intended to report on studies where only generic uses of the computer or technology are involved, such as writing reports using Word, designing pictures to print out later and use or analyse data (using Excel). Rather, the review is concerned with the specific uses of technology to teach an L2, or to elicit or collect L2 data. The discussion is organised to answer the question proposed by Hubbard (2002): when using technology in the classroom, is it possible to simply borrow theories, frameworks and models wholesale from the field of Instructed Second Language Acquisition (ISLA) or do the differences in and relationships between language forms, nature of interaction, participants' roles and environment need to be accounted for differently in such an environment?

The chapter is organised as follows. Section 3.2 gives a historical background of Computer-Assisted Language Learning. Section 3.3 examines the relationship between TELL and ISLA. Section 3.4 looks at studies that have investigated the interaction between TELL and Task-Based Language Learning (TBLL). The discussion turns, then, to the advantages of using technology in section 3.5. The relationship between ISLA research, technology and materials design is discussed in section 3.6. Finally, section 3.7 presents the research questions of the present study and the hypotheses underlying these questions.

3.2 CALL and TELL

I start this section with an overview of the history of TELL in section 3.2.1. I then move on to look at the focus of TELL research in section 3.2.2 before I elaborate on the interface between TELL and other fields of research in section 3.3.

3.2.1 Historical background

Researchers have for several decades now investigated the history of technology use in language learning from different perspectives (O'Shea and Self 1983; Ahmad et. al. 1985; Warschauer, 1996; Levy 1997; Vallance 1998; Warschauer and Healey 1998; Delcloque, 2000; Warschauer, 2000; Bax, 2003; Beatty, 2003). In their 1998 paper, Warschauer and Healey could already note that computers had been used for language learning and teaching since the 1960s. This seems to have begun with one machine at the University of Illinois where PLATO, the first computer-assisted instruction system, was created in 1960. Now more than fifty years later, the use of computers and technology in education is not only taken for granted but unavoidable.

Warschauer (1996, 2000) and Warschauer and Healey (1998) divided CALL history (up to 2000) into three phases: Behaviourist, Communicative and Integrative. The Behaviourist period, from the mid 1950s to early 1970s, was dominated by practice and drill software that was based on the Behaviourist approaches to language learning (see section 2.5.1), such as the Audio-Lingual Method, that were dominant at the time. Materials reflected the prevailing views about language learning, particularly in the USA, and focused on the mastery of linguistic forms through repetition and imitation (see section 2.5.1). During the next ten years (later 1970s-1980s), Communicative CALL based on Communicative Language Teaching/CLT appeared. CLT-based CALL was intended to be more learner-centred, representing a shift of focus from the language itself to the use of language. Materials focused more on learners' use of the forms of language rather than on the forms themselves, which was the case during the Behaviourist period. During the CLT CALL period, materials were designed in a way that allowed more learner control and choice, in keeping with learner centeredness. Materials included practice activities (not in a drill format), simulation tasks, discussion and concordance programmes. The work of Underwood (1984) represents typical views at that time. Although the call to integrate the use of computers in communicative language learning was at its peak in the 1980s, the programmes produced did not meet expectations. Cook (1988) argues that the connection between CALL and language teaching was still 'tenuous'. The CALL materials used at the time were for drilling and helping the learners master grammatical rules/forms rather than to provide them with opportunities to use the language in real life situations in response to the CLT approach (see section 2.5). As an attempt to bridge the gap between teaching and technology, Cook reviews three programmes that were available at the time and highlights the fact that CALL materials could be improved by 'drawing more on the computer's ability to handle human language than CALL has already done' (1988). The same criticisms presented by Cook are echoed later in Bax' (2003) reclassification of the history of CALL (see below) and in research into the effectiveness of CALL (see section 3.2.2).

From the 1990s onwards, Integrative CALL has focused on utilising new advances in technology in the learning process. At the same time, interest in Communicative Language Teaching and Task-Based Language Learning has increased (as established in section 2.5.2). Specialist computer-based software has been designed, and similar to communicative language learning CALL materials, the focus of the materials has been more on language use in authentic contexts than on the language itself, and on giving learners more control and choice in their learning. However, examining the CALL materials almost two decades later, Schrooten (2006) revealed that Cook's (1998) criticism still hold and the bulk of available materials were still essentially Behavioursitc.

Although Warschauer and Healey's 1998 classification is one of the most cited references on the history of CALL, it has its pitfalls. Bax (2003) questions this analysis of the history of CALL on the grounds of apparent inconsistencies and problems with the notion of chronological phases and takes a different view and categorises CALL in terms of approaches rather than phases. For Bax, CALL is restricted, open or integrated. Bax refers here to the type of activities and software that are present in each approach. Table 3.1 summarizes Bax's discussion. While Bax avoids chronology, he states that the next stage in CALL is or will be normalization, defining this as a stage where technology is invisible and taken for granted in everyday life. This stage is the most important characteristic of the integrative CALL approach. Normalization allows teachers and learners to realise the full benefits of CALL (Chambers and Bax, 2006). In addition, when normalization takes place, links between CALL and the wider literature on educational and language learning research will be easier to establish. However, this stage poses difficulties for researchers in controlling variables.

Table 3.1 Three approaches to CALL (adopted from Bax 2003, p.21)

Content	Type of task	Type of students activity	Position in curriculum	Position in lesson	Physical position of computer
Restricted CALL **learning systems**	Closed drills Quizzes	Text reconstruction Answering closed questions Minimal interaction with other students	Not integrated into syllabus-optional extra Technology precedes syllabus and learner needs	Whole CALL lesson	Separate Computer lab
Open CALL **system and skills**	Simulations Games CMC	Interacting with the computer Occasional interaction with other students	Toy Not integrated into syllabus-optional extra Technology precedes syllabus and learner needs	Whole CALL lesson	Separate Computer lab-perhaps devoted to languages
Integrative CALL **Integrated language skills work Mixed skills and system**	CMC WP e-mail Any as appropriate to the immediate needs	Frequent interaction with other students Some interactions with computer through lesson	Tool for learning Normalised Integrated into syllabus, adapted to learners' needs Analysis of needs and context precedes decisions about technology	Smaller part of every lesson	In every classroom, on every desk, in every bag

One issue Chambers and Bax (2006) identify as an obstacle in the normalization of CALL is the integration of technology with learning aims, a problem exacerbated by the lack of 'authorable' CALL materials (materials that can be adapted to better fit the curriculum and student needs). This issue is discussed further in section 3.3.
What is worth noting here is that no matter how one approaches the history of CALL, there is a consensus among CALL researchers that new advances in technology define and shape language teaching (Garrett, 1998, Chapelle, 2009). There is also the inherent characteristic of technology to constantly change. Change can have positive and/or negative outcomes, and as I will show later, changes and advances in technology bring about new trends in language acquisition research.

Levy (1997) argued that the acronyms that had at the time been used over the years to refer to the utilisation of computers in learning provide clues to the role/s the computer plays in the process of teaching/learning. Where CAI (Computer-Assisted Instruction) and CAL (Computer-Assisted Learning) assumed a secondary role for computers in their use of the word 'assisted', CBE/CBI (Computer-Based Education, Computer-Based Instruction) suggested a more fundamental role. CALL (Computer-Assisted Language Learning) connotes an auxiliary role for the computer in language education.

CALL began to replace CAL, CBL, CAI or CALI in the early 1980s, through the work of people like Davies, Higgins and Johns in the UK with the first informal meeting of EUROCALL taking place in 1985. CALL is probably still the most widely used term, with terms such as ICT for LT never really having replaced it. With the rapid advances in technology, however, computers alone are no longer the centre of CALL research. The most recent studies are starting to investigate the use of Podcasts, MP3/4 players, virtual learning environments, interactive whiteboards, iPhones, and so on. In view of the advent of alternative means of delivering electronic materials, I would argue that the term CALL has outlived its usefulness and should be replaced with Technology-Enhanced Language Learning (TELL). Technology-Assisted Language Learning (TALL, Beatty, 2003) has also been used, but given the centrality of technology in today's learners' lives, 'enhanced' is more accurate than 'assisted'. In the context of this study, I, therefore, use the acronym *TELL* whenever I refer to the use of computers and technology in language learning. I use the term *CALL*, however, only when it is used in the original research referred to.

3.2.2 Research Overview

It could be argued that the relationship between technology and learning has gone through two stages. Since its very early use around the 1960s when computers where taking their first steps in the classroom, the technology drove the educational process. As teachers and other pedagogues were not familiar with the new technology, software developers and computer technicians were the ones who decided what to include in software packages and even how to use them. After recovering from the shock of the new technology, pedagogues and teachers started to take the initiative. They started to realise that it is not the new technology that is going to change the teaching/learning environment; rather it is the method underlying its use, or as Schrum (2000) puts it 'technology is a means, not an end'.

The interaction between technology and language teaching has been investigated extensively over the last thirty years. Research topics on CALL corresponded to the developments in technology and to the prevailing teaching approach at each period (see section 3.2.1). There is no doubt that the early stages of CALL research were more 'efficacy studies' concerned with the advantages of using technology in classroom in an attempt to overcome the barriers of integrating technology into instruction (Kenning and Kenning 1983; Higgins and Johns, 1984; Hope et al. ,1984; Stevens, 1984; Underwood, 1984; Ahmad et al., 1985; Barker and Yeates, 1985; Windeatt, 1986; Hainline, 1987; Higgins, 1988; White and Hubbard, 1988; Pennington, 1989). Most of these studies focused on the need for computers in the classroom and drew comparisons between CALL and traditional learning in terms of effectiveness. Although the results did not show good practice in relation to the integration of CALL materials in the classroom, most of the findings indicated that the use of CALL in the classroom was an inevitable next step. Later on, practitioners and academics started to understand that technology itself is not a method; rather it is a tool or means to deliver the method. As early as the mid-1980s, Jones pinpointed the problem. In his seminal article entitled 'It is not so much the program, more what you do with it: the importance of methodology in CALL', Jones (1986) argued that computers have to be used in the classroom like any other tool. Importantly, Jones stressed, the use of computers does not involve the removal of the teacher from the classroom. According to Garrett (1998), researchers and practitioners realised that the efficacy of the tool depends much more on the content than the delivery platform. During the 1980s, research by Davies, Higgins, Johns and Jones offered new ideas on how computers could be used in the classroom (Davies 1989, Davies and Higgins, 1982, 1985; Higgins and Johns (1984); Jones and Fortescue, 1987). In his review of the history of CALL, Davies (2002) listed some of the ideas at that time. His list included text manipulation, word games, action mazes, simulations, adventures, reading and listening comprehension and so on (for a review of the history of CALL in Europe during the 1980s and 1990s, see Chambers and Davies, 2001; Davies 1989, 1993 and 2002). By the late 1990s, other technological platforms had made their way into the classroom and consequently became the focus of research. For example, research began to focus more and more on the use of the internet and the web. However, research in the 1990s was still more concerned with the relationship between CALL and the method of teaching; the development of CALL activities and environments in the classroom; and the efficacy of CALL in relation to certain pedagogic issues rather than the theories and models of SLA that should, inform the use of CALL in the classroom despite the call for such research by prominent researchers such as Chapelle (1998).

Other contributions, with varied findings and conclusions, come from Hardisty and Windeatt (1989); Pennington (1991); Hagen (1993); Scrimshaw (1993); Mohan (1994); Higgins (1995); Chen (1996); Lasarenko (1996); Debski (1999) and Sinclair et al. (2004) among others. Some of these were research studies and some ideas about how computers might be used for teaching. The studies' findings, for example, did not always lead to advocating CALL; some of them revealed negative results on the use of computers. For example, Mohan's experimental study (1994) revealed that the computer's role as a stimulus for speaking was not entirely appropriate. Furthermore, Chen (1996) argued that teachers and students often demonstrate unrealistic expectations of CALL. Adding to this were Lasarenko's (1996) observations that teaching 'with' computers often means teaching 'about' computers, not language; of course this is no longer the case in 2014. Somewhat more recently, Sinclair's (2004) experimental study also concluded that students who used paper-based, less automated exercises were able to score higher than those who used computers to do the same exercises. Regardless of these results, the use of computers in education has continued to attract the attention of both teachers and researchers. This is due to the fact that technology is now an unavoidable part of our lives.

Since the 1990s, a number of empirical studies carried out on CALL have focused on the use of computers in particular classroom settings (Dunkel, 1991; Hulstijn, 1993; Kern, 1995; Chun and Plass,1996; Sullivan and Bratt, 1996; Grace, 1998; Warschauer, 1996; 1998; Meskill and Krassimira, 2000; Laufer and Hill, 2000; Kamhi-stein, 2001; Chapelle, 2001; Hill, 2003; Littlejohn, 2003). By the turn of the 21st century, research included studies on the use of mobile phones, PDAs (personal digital assistants), MP3/4 players, iPhones and so on. Recent research topics are more concerned with how synchronous (skype) and asynchronous (e-mail) communication is related to language learning, the use of multimedia (video, audio, graphics) in the classroom, the effect of social media and the individuality and autonomy of learning that CALL/technology offers (see for example, Ensslin and Krummes, 2013). In 2009, Hubbard listed three areas that looked promising for future CALL research: Web 2.0, mobile learning and virtual learning worlds. Since 2009, these technologies have been taken up, not necessarily with positive outcomes. The focus has turned now to other intelligent CALL tools such as speech recognition, instant translation and interpreting, Web 3.0 and tracking learners' behaviour to provide a more individualized, and consequently effective, learning experience and to better understand the learning process.

The integration of technology in language learning settings can be summarised in two respects: the technologies used and the reasons they are used. Reviewing the research on the application of technology in classrooms, three categories of TELL materials can be identified

1. TELL materials: software designed especially to be used in language teaching;
2. Non-TELL materials: materials not specially designed for language teaching but which are used for this purpose;
3. The Internet, including its various applications such as: e-mail exchange; WWW; e-learning, blogs (and more recently Facebook, Twitter, and so on).

Scanning TELL-related literature, Schrooten (2006) highlights the potential of technology for language learning in the following four points (also see section 2.6):

1. It allows a high degree of differentiation;
2. It elicits a high degree of learner motivation and involvement;
3. It offers enriched content and allows intense multisensory learning process;
4. It makes teaching more efficient as it frees the teacher's time. (Schrooten, 2006, p.129)

As is obvious from the discussion above, the rapid changes in information technology that have been experienced, particularly in the past decade, have placed great challenges as to how best to exploit the new media associated with these technologies. As Salaberry (2001) states, the most important aspect of TELL for researchers to study is the pedagogy, not the technology itself. Earlier, Carrier (1997) also called attention to the fact that all technological approaches should be essentially driven by the teacher's methodological design, i.e. by pedagogy. Thus, as TELL and its associated platforms and digital devices is now considered a taken-for-granted method in teaching languages, the question is no longer whether to use computers in education (Hubbard, 2009); as stated earlier, it is rather a question of how, when and where to use computers. In other words, what is questioned now is the role that computers should play in the educational process. The position I adopt in this thesis is that technology and computers are tools, which, if correctly used, can enhance the language learning curriculum and aid the learning and teaching process, yet it is the theory, visible or invisible, behind that technology that counts.

Having provided an overview of CALL/TELL research, I turn now to look at the pedagogic and research possibilities that are made available by technological developments. I focus in the next section on the interaction between TELL (using this more general term) and Instructed Second Language Acquisition (ISLA).

3.3 TELL, ISLA and Language Learning

Since its emergence in the 1950s, TELL has been closely associated with language learning and teaching and has been studied from various perspectives. More recently, there has been a call in the TELL field for more studies that connect it with other fields such as SLA, applied linguistics, psycholinguistics, and cognitive science (Chapelle, 2007, 2009). However, I focus in this section only on the interaction between TELL research and ISLA. I start by reviewing the main issues and then list some relevant studies. The studies vary in their focus: some researched the use/potential of technology in ISLA, others used ISLA models to inform the design of TELL materials, others employed TELL-based materials for data collection or elicitation with no theoretical framework underlying their use or no technology-related implications identified in the findings and finally some were conducted within the TBLL approach to examine task features and design. The review in section 3.3 does not cover studies within TBLL as those are covered in section 3.4 and it does not cover task features and design as these are reviewed in section 3.5.

3.3.1 TELL and ISLA: research focus

According to Chapelle (2009a), although not stated explicitly, CALL studies were strongly influenced by generative SLA theories during what Bax (2003) calls the *open CALL* period (see section 3.2 above). In typical studies at that time, CALL was viewed as a tool to provide comprehensible input (Krashen 1982; 1985) rather than as instruction (Chapelle, 2009a). This is mainly because generative SLA, at that time, viewed acquisition not only as a natural process but also one unaffected by instruction under, for example, Krashen's work and others' ideas (e.g. Schwartz, 1993; see Chapter Two). More recently, TELL has been examined for its potential in using the Communicative Language Teaching approach in general and in particular Task Based Language Learning/TBLL; this is discussed in the next section below. There is also a growing interest in the use of technology in ISLA research. This trend is best illustrated by the fact that most SLA handbooks now have a chapter or even a section on ISLA dedicated to the topic, although less than a decade ago, none of these handbooks would have had one. One of the problems when researching the use of technology in SLA is that in many cases technology is taken for granted and is not mentioned as a variable/factor when reporting on the results. This observation illustrates what Bax (2003) refers to as *normalization*, as discussed above (see section 3.2); that is, use of technology and computers in SLA nowadays is integrated in a way that makes them invisible. The observation also echoes Selber's (2004) remark that 'good tools become invisible once users understand their basic operation' (Selber, 2004, p. 36), and this is unarguably the case with technology in 21st century experimental ISLA studies.

In 1987, Doughty published her seminal article entitled 'Relating second language acquisition theory to CALL research and application'. It is one of the earliest publications connecting SLA and CALL. Doughty concluded that the design of CALL materials should be informed by SLA theory; consequently, the findings of CALL research will inform practice. She also argued that using technology to collect data from learners allows more precision and control. Later, in 1991, Garrett posed some questions about the use of technology in language learning in line with Doughty's conclusions. Three of these questions still stand today and are the centre of on-going debate: (1) What is the relationship between a theoretically and empirically based understanding of the language learning process and the design and implementation of technology-based materials? (2) What cognitive strategies or problems are implied? (3) What kinds of research does the use of technology for language learning demand or enable? (Garrett, 1991, p.74)

As important as these questions are, there are few studies that have tackled them. Chapelle, one of the leading researchers on the relationship between TELL and SLA, has published many articles and books detailing the persistent questions and issues in the two fields and calling for more interdisciplinary research (1998, 2001, 2003, 2005, 2007, 2009, 2009a). In her (2009a) paper, she reviews the SLA literature listing the main theoretical approaches to SLA with the aspects of SLA theory in which technology integration would be useful (Table 3.2); see Chapter Two for a review of some of these theoretical approaches in SLA. Chapelle's analysis of the literature clearly shows that the link between SLA theories and models and the use of technology cannot be ignored. These are shown in table 3.2.

Table 3.2 SLA approaches and CALL features (adapted from Chapelle, 2009a, p.744)

SLA Approach	Focus of Theory	Example Implications for CALL
	Cognitive Linguistic Approaches	
Universal Grammar	Internal linguistic mechanisms	May provide a basis for sequencing grammatical forms in a syllabus for individualized learning
Autonomous Induction Theory	Internal linguistic mechanisms	May provide a basis for sequencing grammatical forms in a syllabus for individualized learning
Concept-Oriented Approach	Linguistic mechanisms for making form–meaning connections	May provide a basis for sequencing the teaching of form–function mappings in individualized learning
	Psycholinguistic Approaches	
Processability Theory	Psycholinguistic processes for comprehension and production	Provides a basis for sequencing the teaching of grammatical structures in individualized learning
Input Processing	Psycholinguistic mechanisms for making and learning form–meaning mappings	Provides a basis for suggesting the format of instructional materials to draw learners' attention to target form–meaning mappings
Interactionist	Psycholinguistic processes through noticing linguistic forms during meaning-oriented tasks	Provides a basis for suggesting meaning-oriented activities that engage learners' attention to form.
	General Human Learning	
Associative–Cognitive CREED	Cognitive mechanisms for perception and learning of linguistic patterns	Provides suggestions for learning through repeated exposure
Skill Acquisition	Cognitive mechanisms for learning through practice	Provides suggestions for learning through practice and for assessment of successful learning
	Approaches to Language in Social Context	
Sociocultural	The context in which learners communicate	Points to contextual factors such as time, place, and mediating technologies that are relevant for communication
Language Socialization	Communities and their practices	Provides concepts and terms for analysis of how learners' identities as language users evolve through group participation
Conversation Analysis	Language used for social action	Provides methods for analysis of how learners accomplish social action through conversation
Systemic–Functional	The linguistic resources used to make meaning	Provides terms and concepts for analysis of how learners' language constructs meaning
Complexity Theory	The interplay among cognitive and contextual factors	Provides a conceptual framework for the integration of various facets within a system

Processability Theory (PT) (Pienemann, 1998, 2007) assumes that linguistic structures are acquired in a certain order determined by processing conditions. TELL researchers can use the theory to sequence the presentation of linguistic forms in TELL materials in

accordance with the order proposed for their acquisition. The results, if compatible, can then be used to strengthen the evidence on the validity of the Processability Theory. Van Patten's Input Processing (IP) (1991, 1993, 1996, and 2007) model identifies the best processing conditions to maximise the opportunities for learners to make form-meaning connections in meaning-focused communication. IP is one of the most theoretically important theories of SLA where TELL offers great promise in terms of allowing the manipulation of input in easy accessible ways. Interaction-based SLA models such as Long's (Long, 1983, 1985, 1996) are already the most researched in relation to TELL and materials design. Long argues that learners' attention is best directed to linguistic forms through learners' negotiation of meaning, when they communicate in the classroom and they focus on form to overcome a communicative difficulty.
N. Ellis's (2007) Associative-Cognitive framework assumes that learners will acquire forms through repeated exposure to input. In this sense, TELL can play a crucial role as it allows limitless opportunity for exposure to different types of input in different media. Finally, Skill Acquisition theory (DeKeyser, 2007) presumes that language learning is similar to any other cognitive skill so it applies the principles of cognitive psychology to the study of language learning; in order for language use to become automatic, knowledge has to be transferred from declarative to procedural one and this is best done through practice. Most TELL materials allow for repetition of tasks as often as learners need.

Chapelle (1998) argues that to maximize learners' engagement in interaction that is useful for them to draw form-meaning mappings, TELL can be used to direct learners' attention to form by providing opportunities for repetition and modifications; this is reinforced by giving learners control of *when, if* and *how* to modify, repeat or review a structure and also *when* and *how often* to request help using technology that is able to take their individual differences into account (see 2.6). Learners are most likely to notice gaps and correct errors when they initiate form-focused episodes.

As I demonstrate below when reviewing the relevant studies, it seems that both Long's Interaction Hypothesis and VanPatten's Input Processing (IP) are the most researched models when it comes to the application of technology in exploring SLA theory. This is mainly because the two focus on the ways learners interact with and process L2 input through modifying the input to make the linguistic forms more salient or through engaging learners in interaction involving form-meaning negotiations and the options offered by the advances in technology in relation to these two aspects are substantial.

The question here (repeated from above) is: when using technology in the classroom, is it possible to simply borrow theories, frameworks and models wholesale from the field of ISLA or do the differences in the language forms, nature of interaction, participants' roles and environment need to be taken into account? Egbert and Hanson Smith (1999, 2007) argue that 'educators do not need a discrete theory of CALL to understand the role of technology in the classroom; a clear theory of SLA and its implications for the learning environment serves this goal (2007, p. 3). I examine this statement now in relation to interaction and input types as these are the issues related to the current study. For example, in relation to the study of interaction in TELL environments, Chapelle (2003) states that

> a useful theory of interaction in CALL needs to define broadly what interaction consists of, what kinds of interaction are believed to be important for SLA, and why. This general understanding provides an essential basis for conceptualizing and evaluating the new types of interaction made available through CALL (Chapelle, 2003, p. 55).

Chapelle (2003) argues that there are three theoretical perspectives on interaction that are applied to CALL: the interaction hypothesis, sociocultural theory and depth of processing theory. However, one important difference in relation to interaction in a TELL environment is that face-to-face interaction is missing, so the traditional instances of learner-learner or teacher- learner negotiation of meaning are missing as well. Blake (2008) and Pellettieri (2000) disagree with this view and argue that such instances are always evident in synchronous network-based communication such as e-mail and chat. However, Chapelle (2009a) challenges Blake's (2008) argument and maintains that because most of the interaction that takes place in a TELL is written, it creates different processing conditions than those in face-to-face oral communication.

Chapelle (2009a) summarises how principles based on the interactionist theory have been used in the development of materials used in SLA studies. Most of these studies have shown better results for the group using the theory-based materials over other materials. These studies include

- modified input using L2 (Borras and Lafayette, 1994; Guillory, 1998)
- modified input using L1 (Grace, 1998)

- multiple forms of help that make the form more salient and prompt noticing (Chun and Plass, 1996; Plass et al. 1998; Jones and Plass, 2002)

In terms of the input, most ISLA theories are based on the assumption that the input learners receive in the classroom, or in their experimental group, is the same but their processing of it is different for each learner. Chapelle (2003) proposes three areas of input research where TELL could be beneficial: saliency, modification and elaboration. However, she argues that input delivered using technology is not the same as input in the classroom. Chapelle (2007) also argues that 'all approaches to SLA that theorize a role for input need to consider the way that technology changes linguistic input and how learners access to new forms of input might affect acquisition' (Chapelle 2007, p. 107). Broadly speaking, each choice a learner makes in a TELL environment results in a different type of input; if one learner decides to listen to optional audio input and another learner decides not to, this means the type of input they are exposed to is not the same and consequently, this might form a variable that affects their performance and ultimately their acquisition. There is a need for further research in order to understand how elements of a theory are reinterpreted in an environment where technology is mediating the learning process, a question that the present study focuses on. Few studies have tackled this question; I review some of them next, in section 3.3.2.

<u>3.3.2 TELL and ISLA: studies</u>

Researchers (Collentine, 1998; Schulze, 1998; Egbert and Hanson Smith, 1999, 2007; Hwu, 2004; Rosa and Leow, 2004; Heift and Schulze, 2007; Rousell, 2008; Chapelle, 2009, 2009a; De La Fuente, 2012) have called for more studies that look into whether L2 acquisition theories in terms of input type are useful in CALL environment and for CALL materials design. I review some studies that have tackled this question either directly or indirectly.

Manning's (1996) study is mainly about evaluating the merits of exploratory learning as opposed to testing explicit vs. implicit approaches in presenting input (see Chapter Two). Manning carried out her study on a group of 30 UK learners of French in secondary schools, on gender agreement rules using a specifically designed CALL program. The research mainly involved an intact language learning environment. Part of the study showed that an implicit approach to learning the gender agreement rules was inefficient. The first experiment was through traditional tests but not a CALL test at this stage. Then, computer software was developed to compare the merits of implicit, explicit and exploratory teaching. The design was as follows:

1. Implicit mode: examples, exercises and revision.
2. Explicit mode: explanations, examples, and exercises
3. Exploratory mode: a choice of access to either.

To measure progress, identical pre- and post- tests were given to each learner in each of the modules and each of the three modes. L The answers were recorded manually and electronically and also a record was kept of the learners' reactions to different features of the programme. The results showed that the exploratory mode was faster and more efficient, better in terms of rule acquisition and considered more flexible and enjoyable by learners. However, this mode created problems in navigation and decision making and was more suitable for learners with more developed learning strategies. Moreover, it did not work for the more difficult rules for any of the learners, where the explicit mode was more efficient. Manning therefore suggested training all learners to use the exploratory facilities and giving particular attention to helping learners develop more learning strategies.

Zhao (1997) investigated the interaction between speech rate and listening comprehension and the effect of giving control to learners over their input, particularly control over audio input. One group was given control over selecting the speech rate they preferred before they listened; the other group was given control while listening by clicking on a 'faster' or 'slower' button. Zhao concluded that when given control, the participants' listening comprehension improved.

A particularly interesting study, as it is relevant to the present study, is Schulze (1998). His study is mainly about interlanguage grammar and how grammar checking software (Textana) can help parsing an interlanguage variety. The software is a grammar checker which can be used in any writing process but is not intended to be used as a grammar teaching package. Schulze describes how the software helps to facilitate focus on form by providing a grammar checker at the post-editing stage of writing. In other words, grammar instruction is provided in meaningful communicative tasks; therefore, learners will concentrate more on the linguistic form. The research basically shows that L2 acquisition theories in terms of input type are useful when designing CALL software or using existing software.

Collentine (1998) outlines principles with which CALL educators can design effective input-oriented tasks targeting grammar instruction. He describes how cognitive principles can operate in a CALL environment by describing a software prototype that targeted the instruction of the Spanish subjunctive within a modified version of VanPatten's (1993, 1997) Input Processing/IP model. Collentine's discussion is about how input theories inspire CALL designers and IP includes some of the features which, in his view, should be included in an input-based CALL task. This conclusion is echoed in Chapelle's list (Table 3.2 above) and her call for more interdisplinary research. One of the main assumptions is that in input-oriented tasks requiring students to focus on the message conveyed by the language input, they will indeed engage such structures provided that items representing the target structure possess stimulus novelty. Another assumption is that highlighting a target structure is not only useful to enhance reading and listening skills but also productive skills. The aim of this is to increase the probability that learners will detect the target structure and also ensure that learners engage in meaning-form connections. A technique that is claimed by VanPatten to be effective in promoting the intake of target structures that are not redundancy-dependent is to present learners with two situations, typically in the form of illustrations or cartoons. After studying each situation learners are provided with a sentence containing the target structure and then asked to determine the situation which the sentence describes. For example to teach the simple past, one might give learners two cartoons of a woman playing baseball, under one cartoon is the caption 'last week' and under the other 'right now'. Learners then hear or read a sentence such as 'Mary played baseball' and are prompted to indicate to which cartoon the sentence refers. In such a context, CALL applications are highly meaningful as they can be used to present structured-input tasks. Collentine concludes by suggesting that CALL applications can be particularly effective at facilitating the intake of grammatical structures that normally have little communicative value in input.

Collentine (1998) proposed that CALL materials designers and educators explore mind-centred theories, such as those that recognize the importance of providing learners with comprehensible input. Collentine's study used computerized conscious-raising tasks to train and test 40 L2 learners of Spanish on indirect speech. He used user-behaviour tracking technologies to identify which aspects of the instruction contributed to improvement in the participants' performance. His findings indicated that learners underutilised some of the instructional features provided to them. The results also showed that audio events correlated highly with instructional benefit. His findings corroborate Chapelle's (1998) call at the time that SLA theory and research might be consulted to suggest approaches to multimedia CALL design.

Hulstijn (2000) reports on four studies conducted in the Netherlands which used computers for both data collection and language training. Technology-aided tools involved word and sentence recognition tasks, response time and speed measurements, data recording using log files and online dictionaries. The main aim was to investigate the automaticity of word recognition and reading processes of participants. Hulstijn highlights the fact that the use of software (on laptops) was the obvious choice to conduct such studies. The use of log files meant that the researchers obtained a full unobtrusive observation of the participants' behaviour. The way Hulstijn programmed the log files was invaluable in revealing what no other technique can easily reveal; that is, *which, when* and *in which order* the learners processed the lexical items.

Pellettieri's (2000) study was on negotiation of meaning, and her results show clear instances of this through interaction in network-based synchronous communication. She found that in text chats learners tend to correct themselves and others as well, and thus, engage in form-meaning negotiations that, she argues, help them notice the forms. Pellettieri did not, however, measure noticing or its effect. Similarly, Blake (2000) has found that written synchronous online exchanges contain episodes of meaning negotiation held to be essential for the enhancement of learners' interlanguage.

Individual learners' control of input through technology is a topic several researchers have taken up. Shea (2000) investigated how comprehension can be enhanced by learner control of L2 input which is offered in TELL environments. In a study on captioned interactive video and SLA, Shea found that being able to control the pace of language captions during a computer-based language activity was beneficial for those language learners classified as weaker. Fogg (2003) also discusses this, referring to it as the principle of 'tailoring' in educational technology. By this he means that technology helps in providing information tailored to each learner's individual needs, interests and personality, usage, context and other individual factors.

Similar to Collentine's (1998) findings referred to above, Hwu (2004) concluded that to create a potentially effective input application for grammar instruction, CALL designers need to consult other relevant areas in addition to considering SLA theory. One such example is Knutsson, Cerratto and Severinson (2003) who emphasize that NLP (Natural Language Processing) tools are particularly important and useful when trying to direct learners' attention to form through an FoF approach.

Another issue raised during this decade of research is that of exposure under different task conditions. Rosa and Leow (2004) examined whether exposure to L2 data under different computerized task conditions had an impact on learners' ability to recognize and produce the target structure immediately after exposure to the input and over time. They used recognition and controlled production tasks to assess learners' L2 development. They manipulated the degree of explicitness by combining three features: (a) a pre-task explicit grammatical information, (b) feedback concurrent with input processing, and (c) variation in the nature (i.e., implicit or explicit) of the feedback in those cases in which it was provided. The results showed advantages of processing input under explicit conditions; these advantages were more visible for production than recognition.

Heift and Schulze (2007) were more concerned with how technology can help in error recognition and analysis. They reported on the 'German Tutor', a programme based on intelligent parser-based systems. It prioritizes errors and provides single instances of feedback to the learner even when multiple errors occur in a task. By keeping a record of all learners' performances, the Tutor analyses errors and develops a model that informs subsequent feedback, assessment and remediation.

A study which addressed amount of input and how learners deal with input is Rousell's (2008). It compared the performance of listeners under three different conditions: the learner listened once, listened twice, and the learner controlled their own listening. The dependent variable was recall of idea units. The learners were able to recall more idea units when they controlled the listening section. In a later study, Rousell (2008) investigated the effects of the ability to control listening on processing. L2 learners of German had the ability to control listening input by using a computer mouse. Rousell recorded the physical movements of the mouse to measure metacognitive activity. The results showed that the ability to control information, i.e. audio input, improved all participants' information processing.

De la Fuente (2012) investigated a number of the issues considered in the studies described above, namely the effects of the medium of aural input during listening tasks on noticing and type of comprehension: top-down or bottom-up. L2 learners of Spanish were exposed to focus on form listening tasks in a technology-enhanced classroom using two media of delivery of input: learner-manipulated mobile assisted language learning and instructor- manipulated language learning. Immediate post-tests were used to measure participants' reported noticing and type of comprehension. De la Fuente operationalised noticing using think-aloud protocols. The findings of the study indicated that learners in the learner- manipulated group showed significantly higher levels of noticing, bottom-up comprehension, and top-down overall comprehension than learners in the instructor- manipulated group. At the end of the study de la Fuente recommended that instructors take a principle-oriented approach to the use of technology where technology serves a clear pedagogical purpose and does not interfere with - but rather facilitates - learners' attentional and language processing mechanisms.

From the above studies, it is apparent that the features of technology that are relevant to SLA research are timing, multimodality, access to help and feedback and directing attention in relation to input (Chapelle, 2009a). All the above studies highlight the potential of technology in ISLA research and stress the need for more research that investigates the applicability of existing ISLA theories and models to TELL environments. The previous discussion mainly focused on the integration of technology in ISLA and SLA research and studies. I move now to consider the integration of ISLA models and theories in CALL/TELL practice.

3.3.3 ISLA based TELL

Even after 50 years of research, TELL materials design still lacks a specific pedagogical framework. In her 2009a publication, Chapelle states that frameworks and guidelines are required when assessing technology-based materials in terms of the opportunities they provide for L2 learners. Yet there is a dearth of research that evaluates multimedia and TELL systems, or even typical materials, from a theoretical rather than a practical perspective. Here, a few researchers have attempted to set criteria for the evaluation of CALL materials, some of which are based on SLA research findings. For example, Egbert and Hanson Smith (1999, 2007) argue that SLA conditions can inform the design of technology-enhanced materials and suggest ways in which CALL can enhance and promote language learning. In her earlier 2001publication, Chapelle identified an

SLA-theory-based framework for evaluating CALL by drawing on the concepts and practices of CALL testing materials. In addition to what is shown in Table 3.3 below, she recommends that language materials should also embed the following features, regardless of theory of SLA:

1. Provide help opportunities for learners;
2. Require learners to focus on meaning;
3. Be at the right level ;
4. Tailor activities to fit the learners.

Table 3.3 Characteristics of materials and relation to SLA theories based on Chapelle (2001)

Material Characteristics	SLA theory	Aspect of the theory measured
Language learning potential	Interaction Input processing Skill acquisition	Quality of interaction Utility of selected input Quality of the practice
Meaning focus	All Theories	Availability of rich, interesting input that provide opportunity to produce and comprehend meaning
Learner fit	All cognitive and psycholinguistics theories	Level of language
Authenticity	Systemic linguistics	Relationship between input used in instruction and that learners will use in real life
Positive impact	Sociocultural theories	Benefits, linguistics and non-linguistics of the experience
Practicality	Skill acquisition Cognitive processing	Degree of access and skills needed to work on tasks

Tomlinson, (2003) also argues that materials should be based on principles derived from SLA theories. Blake (2008) maintains the view that teaching methodology and CALL design should be informed by what is known about the nature of the SLA process. In contrast, Garrett (2009) questions the need for relying heavily on SLA findings when designing CALL materials. She argues that SLA theory has been developed mainly from studies on English and points out that although there are many studies now that investigate other languages, SLA theory is based on and derived primarily from the study of ESL. Garrett (2009) states that what is applied to English does not necessarily apply to other languages. She does, however, acknowledge the role SLA theory could play in motivating and justifying some technological interference. She also points out that SLA theory is less concerned now with the acquisition of grammatical forms and argues that most SLA research now is related to communication rather than acquisition, that is to sociolinguistics, pragmatics and discourse analysis. In this sense, SLA theory and research can only inform CALL materials when the focus is to provide grammar in the context of a communicative approach.

When it comes to task-based language learning and technology, Doughty and Long (2003) argue that any language programme must be carefully planned based on a clear understanding of learners' needs. They identify the ten relevant methodological principles shown in Table 3.4 below. In their list, they try to integrate interpretations of SLA theory and research findings into a coherent design for the delivery of instruction using CALL. According to these criteria, the designer has to take into account the content of the activities and the type of interaction it results in in order to create a rich learning environment (Schrooten, 2006).

Table 3.4 Methodological principles for CALL (Doughty and Long, 2003, p.52)

	Principles (adapted from Long, in press a)	L2 Implementation	CALL Implementation
ACTIVITIES			
MP1	Use tasks, not texts, as the unit of analysis.	task-based language teaching (TBLT; target tasks, pedagogical tasks, task sequencing)	simulations; tutorials; worldware
MP2	Promote learning by doing.		
INPUT			
MP3	Elaborate input (do not simplify; do not rely solely on "authentic" texts).	negotiation of meaning; interactional modification; elaboration	computer-mediated communication / discussion; authoring
MP4	Provide rich (not impoverished) input.	exposure to varied input sources	corpora; concordancing
LEARNING PROCESSES			
MP5	Encourage inductive ("chunk") learning.	implicit instruction	design and coding features
MP6	Focus on form.	attention; form-function mapping	design and coding features
MP7	Provide negative feedback.	feedback on error (e.g., recasts); error "correction"	response feedback
MP8	Respect "learner syllabuses"/develop-mental processes.	timing of pedagogical intervention to developmental readiness	adaptivity
MP9	Promote cooperative/ collaborative learning.	negotiation of meaning; interactional modification	problem-solving; computer-mediated communication / discussion
LEARNERS			
MP10	Individualize instruction (according to communicative needs, and psycholinguistically).	needs analysis; consideration of individual differences (e.g., memory and aptitude) and learning strategies	branching; adaptivity; autonomous learning
(This table is adapted from Doughty 2000b, 2001b.)			

In line with Doughty and Long's (2003) call, Ellis (1998) highlights the need for empirical SLA theory-based evaluation of materials in general. The point here is that any learning activity, whether it is delivered using technology or not, will not be effective if is not pedagogically planned.

As can be noted from the above discussion, almost all researchers in TELL and SLA alike recommend that materials are designed in accordance with the findings of ISLA research. I turn now to investigate if this is actually the case with published materials. Since the current study is on different types of grammar instruction, the discussion is mainly focused on grammar instruction materials.

Although many new text books and materials in general claim to have progressed beyond the Grammar Translation Method and the Behaviourist Audio Lingual Method, they all seem to have roots in these methods, just with more attractive packaging (Wells, 2000) and pedagogical guides with communicative recommendations (Schrooten, 2006). Years of developing EFL/ESL textbooks and published materials have resulted in a heavy reliance on them in classrooms. These materials usually provide grammar rules as summaries at the end or as boxed information within sections of the book. By doing so, there is the assumption that learners' attention will be drawn to these parts. On the other hand, CD and DVD based teaching materials are often overloaded with graphics, audio and video items so it is sometimes the case that the medium (i.e. technology) dominates the message. Schrooten (2006) started a mission to find existing software that complies with Task-Based Language Learning/TBLL principles for his study. After conducting an extensive exploratory search, he determined that the bulk of available software packages were still essentially Behaviouristic and that no suitable software was available. Similarly, one of the obstacles Hulstijn (2000) pointed out in his review of the use of computers in experimental SLA was that no commercially available software was suitable for the requirements of the tasks. It has also been argued that although ISLA research findings highlight the importance of directing learners' attention to form and meaning, very few CALL grammar activities that require attention to both form and meaning had been produced (see Hubbard and Bradin Siskin, 2004). This highlights the more important issue of separation between the fields of TELL and ISLA (Thomas and Reinders, 2010)

In 2007, Levy also argued that although there was a wide scope for the design of sophisticated grammar programmes using technology. At the time, they were still not on the market and he pointed out that most of the available grammar programmes were still very basic in terms of input processing, error diagnosis and feedback provision. But Schrooten (2006) underscored the point that integrating the use of technology in TBLL is actually not 'self-evident' and that 'the principles underlying a lot of the currently available educational software seem to be flatly opposed to the principles of task-based language learning' (Schrooten, 2006, p.130). Chapelle (2009a), however, rejects previous criticism and points out that theory-based principles stemming from SLA theory had been used in the design of commercial and academic materials for the previous ten years but because material developers were less likely to talk about their materials, the connection was not observed directly.

It is important to note here that every ISLA theory or model has its limitations, as was made clear in the previous section. No single theory accounts for the cognitive processes, the nature of the input and the external and internal factors that are involved in L2 acquisition. Therefore, a single theoretical framework can also not provide a full account of what is needed in TELL software. A mixture of all these approaches is what is usually called for to promote acquisition in the classroom (as shown in the tables presented above). Furthermore, evaluating technology-based materials requires a holistic approach, looking at design, and implementation at the minimum without overlooking the theoretical framework and implications.

In the next section, I examine the research on TELL and Task-Based Language Learning (TBLL) in more detail, as one of the aims of the current study is to evaluate the effectiveness of different task-based types of input in a TELL environment and to investigate how the use of user-behaviour tracking technologies can help us understand the learning process.

3.4 TELL and TBLL

Thomas and Reinders (2010) argue that the separation between TELL and TBLL research is evident. They claim that the reason behind this separation is that SLA has marginalised CALL and TBLL and has mainly focused on face-to-face learning settings. They stress that research on the interface between TELL and TBLL approaches is long overdue. The few exceptions that have bridged the gap between the two include Doughty and Long (2003), Skehan (2003), Gonzalez-Lloret (2003, 2007), Schrooten (2006) and Chapelle (2007), and more recently Thomas and Reinders (2010); Heift and Rimrott (2012) and Thomas et al. (2013). I review the studies and the main issues arising from them in publications discussed below.

Chapelle (2001) states that ’the study of the features of computer-based tasks that promote learning should be a concern for teachers as well as for SLA researchers who wish to contribute to knowledge about instructed SLA’ (Chapelle, 2001, p. 2). In 2007, Chapelle described computer-based tasks and uses the acronym ‘CASLR’ to refer to Computer Assisted Second Language Research tasks. In this publication, she defines CASLR tasks as ‘tasks that require learners to work on the target language interactively with a computer program or with other people through the medium of computer’ (Chapelle 2007, p. 98). She states that these tasks can be viewed by learners as a regular part of instruction and she lists two main scenarios where such tasks are useful

- To ‘operationalize’ learning conditions in order to test hypotheses about SLA
- To gather data to make inferences about learners’ knowledge and strategies.

Chapelle (2007) argues that since the 1990s, CASLR tasks have proven to be a reliable source of data about specific aspects in instructed SLA. Studies that have used CASLR tasks include Doughty (1991) and Sanz and Morgan-Short (2004). What is worth noting here is that most of the studies that have used tasks in TELL have rarely referred to them as computerized tasks. Most of the research on TBLL using technology has focused on computer-mediated communication (CMC) tasks, such as chatting, and its role in promoting interaction similar to the interaction in typical classrooms (Hampel, 2006; Ortega, 2009; Smith, 2009). Researchers claim that there is evidence that the benefits of CMC tasks include increased participation, increased quantity and better quality of learner output, improved attention to linguistic form, more saliency of linguistic input and output and increased willingness to take risks among learners with their second language (Salaberry, 2000; Izumi, 2003; Smith, 2004). What they also point out is that more research is needed to evaluate the effectiveness of CMC tools for language learning.

Can tasks and CMC be integrated? Hampel (2006) takes this up, and discusses a framework for the development of tasks in a synchronous online environment used for language learning and teaching and demonstrates that a theoretical approach based on SLA principles, sociocultural and constructivist theories, and research on multimodality can influence the design and implementation of tasks for computer-mediated communication. Hampel reported on a study that investigated task design and implementation in CMC. The findings showed that tasks encouraged active participation and fostered interaction among students and between students and the teacher. She pointed out that the implementation of tasks also revealed that the computer medium cannot be used in the same way as a conventional classroom setting and that ‘both the design of the tasks and their implementation needed to reflect the affordances of the environment’ (Hampel, 2006, p.118).

Task-based computer-mediated communication was also addressed by Smith (2003, 2004, 2009) who conducted a series of studies to look for evidence of direct and indirect target language benefits of task-based computer-mediated communication. Results showed that learners engaged in negotiation of meaning in relation to lexical items but not to grammar. Smith argues that this result is not surprising since learners are more likely to focus on meaning over form particularly when their proficiency level is low (as was the case for these learners) and also because lexical items carry more weight in communicating the basic meaning than does grammar (Smith, 2009). The results also showed that learners decided to engage in ‘compensatory strategies and modify their own output when faced with instances of communicative breakdown or non-understanding often resulting in more target-like modifications’ (Smith, 2009). Smith notes that not all learners were able to notice moves by their interlocutor indicating problems with form. The study revealed that chat records may provide good diagnostic information for teachers about learners’ interlanguage stage, and such information can help teachers make better data-driven decisions regarding which grammatical features to target (and when) for each learner. Smith also points out that in this technology-mediated setting, learners were able to produce more of the target language than they normally do in typical face-to-face settings and that they also received rich input from their interlocutors - two findings that highlight the advantages of integrating technology in TBLL.

The effectiveness of a task-based CALL program designed to promote interaction is taken up by Gascoigne (2006) who reports on a study conducted to investigate this. The design of the software integrates methodological principles of language teaching that are well grounded in SLA. The tasks were designed to engage learners in the learning process by promoting interaction; input was modified as well to be salient and elaborated. The results indicated that learners were able to engage in L2 interaction negotiated in ways that facilitate comprehension and lead to language acquisition.

Fewer studies, however, have looked at input type, task type and task design all together in a TELL environment and at the elements of the medium that affect performance. Studies that have tackled these topics include Hill and Laufer (2003) and Heift and Rimrott (2012). Hill and Laufer (2003) compared the effect of three task types on incidental L2 vocabulary learning, on the task-induced amount of dictionary activity, and time-on-task. Participants read a text with unfamiliar target items and later performed one of these three tasks: a form-oriented production task, a form-oriented comprehension task or a meaning-oriented task. The computer was used to present the text and the words could be looked up in electronic glosses. The results of immediate and delayed post-tests showed that form-oriented production and comprehension tasks yielded better results than the meaning-oriented task. There was no significant difference in time-on-task, but there was a significant difference in the amount of dictionary activity the tasks generated. Hill and Laufer (2003) found that the amount of word-related activity that the task prompts determines task effectiveness.

Task-related variation in learner performance in a CALL environment was also investigated by Heift and Rimrott (2012) who used three types of activities: free composition, translation and sentence building and participants were L2 learners of German. The results revealed that grammatical accuracy with respect to German word order was significantly higher with the meaning-focused task type (i.e., free composition) for both the beginner and intermediate levels. Proficiency level was another factor that explained variation in L2 word order accuracy: beginner-level students performed significantly better than intermediate-level learners on the two form-focused task types (i.e., translation and sentence building). Heift and Rimrott concluded that learner performance varies according to the teaching objective of the instructional task: grammar or topic/meaning.

In short, and as highlighted by the findings of experimental studies, TELL environments bring new dimensions to Task Based Language Learning which are different from those in face-to-face communication. If TBLL is to expand its applicability to learning contexts beyond the typical classroom, these dimensions need to be explored. The current study focuses on one of these dimensions: the use of user-behaviour tracking technology to inform the learning process. Heift and Rimrott (2012) emphasise that many questions about the effects of task-related variation on learner performance in TELL environments are unanswered. Similarly, Thomas (2013, p. 335)) maintains that

> while a number of educators have developed language learning environments utilizing new digital media tools, it is necessary to acknowledge the extent to which such virtual environments are different from face-to face classrooms in order to develop materials and teaching and learning strategies than can fully take advantage of the opportunities [they present]

The current study makes a contribution in this direction by investigating the effectiveness of different types of task-based input in a TELL environment.

I turn now to look at the technology-enhanced research tools used in ISLA as the main research tool in the present study is real time log files. I review the literature here and evaluate the importance of log files in Chapter Four while describing the methodology.

3.5 Technology-Enhanced ISLA research tools

Reviewing the literature, it is evident that TELL-based tools have been mostly used within the field of cognitive SLA research (Hulstijn, 2000). Technology is used to elicit L2 data or record how L2 learners process L2 input as well as to teach the L2. The scope for the use of technology in SLA research is very wide. Typical uses of technology in SLA research have so far included

- The use of computers and the internet to provide input, authentic, modified, enhanced, etc.;
- Electronic dictionaries, corpora, search engines, etc.;
- Interaction among learners and native speakers in chat rooms, forums and websites;
- Cultural information about the L2 environment;
- Personalised learning opportunities that allow learners to repeat, sequence and control their learning choices;
- Using computers to elicit or record data.

TELL-based tools in SLA research can be classified according to whether they are specifically designed to conduct research in SLA or designed for another purpose but used in SLA research. Software packages that designed specifically to aid SLA research include:

- COALA, (Pienemann,1992) to analyse learners' interlanguage according to the processability theory;
- COMOLA (Jagtman and Bongaerts, 1994) to analyse learners' interlanguage;
- CHILDES (MacWhinney, 1995; Sokolov and Snow, 1994) to transcribe, analysis and store learners' utterances;
- SPELL (Spoken Electronic Language Learning), which combines automatic speech recognition with advanced graphics for virtual humans to create effective spoken natural language eLearning software. It integrates speaker-independent continuous speech recognition technology with virtual worlds and embodied virtual agents to create an environment in which learners can converse in the target language within meaningful contextualized scenarios (Morton, 2005, 2008; Anderson et al., 2008)

However, most of the software packages widely used in SLA research have been developed in other fields, particularly psychology. Programmes such as *e-prime* were developed by psychologists and are used by SLA researchers. E-prime is designed to administer and record psychological stimuli and is widely used in SLA to record and analyse response time in timed or untimed grammaticality judgment tasks (Marinis, 2010, Wright, 2010, 2012). Eye tracking technology, mainly developed in psychology to investigate brain-related events, is increasingly being used in investigating attention and noticing in SLA research (Kahoul, in press). Nerbonne (2003) also lists concordancing, speech recognition, syntactic processing and machine translation as technological tools used in language research. With the rapid advances in technology, more tools are being developed and used. In 2000, Hulstijn 2000 argued, however, that few areas within SLA research were using computer-assisted elicitation techniques and listed the following as the only types of tasks that were at the time administered using computers in SLA

- Grammaticality judgment (GJ) tasks where learners need to press one of two buttons to indicate their perception of the grammaticality of a sentence. Software usually records the responses and the reaction time.
- Sentence matching tasks where learners are asked to indicate if a sentence matches another one that appears on the screen within a timed period. The response and reaction time are recorded in such tasks.
- Oral production tasks where speech recognition technology is used.

- Word recognition tasks which are similar to GJ tasks as learners need to indicate whether a sequence of letters form a word by pressing ‘yes’ or ‘no’ buttons.
- Sentence and paragraph reading where learners are presented with text or sentences and they need to press an arrow to move to the next one. The time they spend on each session is recorded. Sometimes, they are asked to recall the story and their responses are analysed against the time they spent reading and other variables.
- Form-function mapping tasks where learners are presented with a number of words and are asked to give a response, for example, indicating which word is the subject. Responses and reaction time are recorded.
- Connectionist simulations which imitate learners’ production.
- L2 learning experiments where learners are asked to respond to tasks conducted under controlled conditions. Studies that have used such tasks require learners to learn an artificial language over a certain period of time and then to respond to tasks that measure different variables in accordance with the research interests. (Hulstijn, 2000, pp. 33-35). Although, things have moved on since 2000 and more technology-based tools are used now in SLA research such as eye tracking, more research is needed on the effectiveness of the tools and how best to use them to elicit and analyse L2 data .

Overall, the technology-enhanced tools mentioned above are used to test attainment of learners at a certain point in time rather than examine the learning process itself to shed light on how acquisition takes place. A recent review and study by Chun (2013) highlights the importance of technology-enhanced tools, such as user behaviour tracking technologies, in analysing the complexity and abundance of information that dynamic learning environments generate. More recently, and as stated in section 3.4, computerized tasks are increasingly used to investigate the multi-dimensional aspects of task design. Some recent studies have also used eye tracking and MRI techniques to track learners’ processing.

Research is now more directed towards the use of ICALL (Intelligent CALL) and NLP (Natural Language Processing) in SLA research. Such technologies provide better tools that allow researchers to get more insights into the processes of SLA. However, as ICALL and NLP rely on artificial intelligence, their use so far is limited as knowledge in programming is essential in the design and execution. Corpus linguistics is now also

increasingly used to investigate learners' interlanguage in an attempt to draw conclusions in accordance with SLA theories (Myles, 2005).
The relationship between technology and SLA is best described in Hulstijn's (2002) words which stand as true today as they then 'new theories might lead to the search for innovative methods and tools. And at the same time, new methods might result in new theoretical thinking and approaches'.

In the next section I examine the advantages of using technology in ISLA research before moving on to state the research questions and hypotheses of the current study.

3.6 Advantages of using technology

The advantages of using technology in ISLA research are viewed as two-way: technology-based materials should be informed by ISLA findings while SLA should utilize new advances in technology in data collection, input provision and data analysis.

One of the main advantages of using technology in ISLA research is that instructional conditions can be controlled for in better and easier ways than in classroom research (Garrett, 1998; Chapelle 2007). Moreover, Levy and Stockwell, 2006; Reinders and White (2010) and Thomas and Reinders (2010), all argue that CALL technologies make it possible to engage in learning and teaching across boundaries and to consider the multi-dimensional nature of task design. TELL resources can considerably increase L2 learners' contact with the target language and culture, for example, and the internet allows exposure to authentic and varied input. In addition, technology now makes it increasingly feasible to individualize and personalize the learning process, resulting in much desired self-empowerment and autonomy in learning (Warschauer et al. 1996; Murray 1999). Kessler and Ware (2013) concede that by 'pairing tasks with appropriate technology, or with a combination of technologies' (p. 103), the number of competencies addressed in the classroom can be increased.

Blake (2008) argues that the possibilities are vast for the integration of technology in language learning. He lists three technological platforms that are viewed as tools that assist the language learning process: the web, CD-ROM or hypermedia applications and network-based communication such as e-mail and chat rooms. In addition, new advances in software development and technology have allowed for analysis of large databases of learner language (corpora). Such analysis could be conducted in a 'bottom-up' process to identify patterns and systematic themes in the data or in a 'top-down' process to test pre-created hypothesis (Granger, 1998; Rutherford and Thomas, 2001; Myles, 2005). Furthermore, the internet provides limitless opportunities to enhance language learning[15]. Littlejohn (2003) argued that utilizing the internet in the field of learning and teaching offers new opportunities to increase flexibility

1. in time and location of study;
2. in forms of communication and types of interaction;
3. and in access to, and availability of, information and resources through the World Wide Web.

Some of the advantages of using web-based learning are listed by Jolliffe et al. (2001) as follows: feasibility of delivering learning in terms of time and location; ease of updating materials; smoothness of interaction and communication among learners and between learners and teachers; and the availability of an unlimited range of learning resources.

Computers are seen as tools for individual learning and many researchers have focused on the advantages that computers offer in terms of privacy, self-pacing and personal feedback (Curtin et al. 1972; Nelson et al. 1976; Higgins 1988; Kenning 1996; Jones 2001). In fact, as stated by Jones (2001), a common justification for the use of computers in language teaching and learning is that they promote learner autonomy. Levy (2009) also describes how new advances in Intelligent CALL (ICALL) and Natural Language Processing (NLP) could help in SLA research (see above). He gives the example of a tutoring system, designed using ICALL and corpora principles where learners' errors are recorded and analysed then used to create models for feedback. One of the advantages of such systems is that errors can be categorized and annotated according to certain criteria such as errors by learners of the same L1, or errors of a certain proficiency level. The system can then be programmed to present individualized feedback and training. There is still, however, a need for software that provides better informed analyses and feedback of learners' behaviour (Dodigovic, 2005; Heift and Schulze, 2007) which could be used to test ISLA models. After reporting on the findings of several studies, Hulstijn (2000) concluded that computer-aided tools

[15] Other publications reporting on the use of the internet in relation to teaching and learning are Teeler and Gray (2000); Warschauer (2000); Windeatt et al. (2000); Hoshi (2002); and Linder (2004).

available at the time indeed provided SLA researchers with means to get closer to the processes of language acquisition and use.

The advantages of the use of technology in ISLA can be encapsulated in three dimensions: input delivery, input processing, and data collection. However, the use of technology should not be seen as a separate competency, but rather as ‘a vehicle for accomplishing other tasks directly related to, and in support of, language teaching and learning’ (Kessler and Ware, 2013, p. 104).

Having now established the theoretical framework of the current study in the fields of ISLA, language pedagogy and TELL, I move on to state the research questions and hypotheses that the present study tackles.

3.7 Research Questions and Hypotheses

The best way to summarize the discussion so far in relation to input is through Gregg’s (2001) words:

> it is uncontroversial that learners need input in order to acquire a language … unfortunately the consensus stops about there. How much input is necessary? What kind of input? Under what conditions need it be provided? (Gregg, 2001, p.167).

Grammar instruction is one aspect of input where agreement is far from being reached and Gregg’s questions are far from being answered. In addition to the contradictory findings on the type of input that best promotes second language learning, instructional approaches predict what learners take in from the input, but without any empirical evidence to support this prediction. The role of modality of input in SLA has been under researched, and further research is not only essential for guiding the selection of input, but also for increasing language teachers' awareness of what learners do with the input that is made available to them. Further research is also needed on the applicability of TBLL to different contexts, particularly with the increasing interest in TBLL (see section 2.5.2). Most of the relevant research is conducted in classroom and laboratory settings where face-to-face communication is the norm. Ellis’ (2010) remark encapsulates the argument on TBLL and TELL and drives the current study. Ellis (2010) emphasises that

> we cannot assume that tasks work the same way in FTF [face-to-face] classrooms and in technology-mediated environments. Nor can we assume that they work in the same way in the highly varied environments that technology now affords. Given the current advocacy of TBLT and the increasing use of

> technology in language teaching it is important that we develop fuller understanding of how to design tasks for use with different technologies and how best to implement them in ways that will foster language learning (p. xviii).

The current study makes a contribution in this direction. It aimed to look at the effect of different types of input on the learners' intake in a TELL environment and the factors contributing to variability in attainment through the use of user-behaviour tracking technologies and in the context of a TBLL lesson. For the purpose of the present study and as intake itself is not the goal of the treatment in the present study, intake is used in its broadest form, to represent surface linguistic and metalinguistic gain. In this sense, intake is operationalized in terms of performance so learners' scores are taken to represent their intake of the linguistic items. Linguistic gain is identified through the differences in scores between time one and time two. Three different types of input are used as they represent the main approaches to grammar instruction (see Chapter Two): Focus on Forms (FoS), Focus on Form (FoF) and Focus on Meaning (FoM). Components of the learning process that have been identified to affect learners' intake and are part of the present study include modality of input, time/speed, trial numbers, access to grammar and task type.

The advantages of including TELL research tools in ISLA research have been established for a while now; however, studies employing such tools are still few. Garrett (1991) summarises what many researchers have been calling for in the field for over at least the last 20 years,

> A CALL lesson which creates an environment for some interesting language learning activity could be fitted with a program collecting data on how the learner make use of that environment, and those data can not only feed back into improving the pedagogy but can also contribute to the development of second language acquisition theory (Garret, 1991, p.94)

More recently, Chun (2013) notes that

> despite the fact that data documenting what learners actually do in CALL activities can provide valuable insights into both second language acquisition and pedagogical design, a surprisingly large proportion of CALL studies do not report on tracking data (Chun, 2013, p. 256)

Also MA (2013) points out that

> further research needs to focus on the identification of the key user actions related to learning outcome. Only with a good tracking system can CALL effectiveness be proven, useful design features be identified, and the appropriate applications be selected. (Ma, 2013, p. 230)

The present study accordingly employed log files to collect data about learners' performance and the learning processes they exhibited while dealing with different types of input. In this sense, log files help in evaluating the purported benefits of using such techniques in ISLA research. As was pointed out earlier, the present study's focus is not on the learning outcome itself as it is acknowledged that the amount of input is not sufficient to acquire the target structure. However, in order to be able to connect the learners' actions (process) to the learning outcome (product), it is important to measure the outcome before and after the intervention. So the goal of the study is to replicate a classroom in a TELL environment to investigate the ability of user-behaviour tracking technologies to inform the learning process.

Ellis (1990) identified two approaches for studying the relationship between L2 learning and instruction: linguistic/psycholinguistic and pedagogical/education. By the linguistic approach, he referred to the type of research where the attempt is to falsify or verify a certain theory of language or language acquisition through a hypothesis testing model. The pedagogical approach, on the other hand, aims at identifying the assumptions behind a specific pedagogical principle or technique and/or evaluating these assumptions in relation to language learning theories. The current study covers linguistic, pedagogic and methodological issues.
The following is a summary of the principal issues and suggestions which have arisen in the discussion in this and the preceding chapter:

- The role of instruction: after decades of research, agreement is lacking on which type of input is most effective for language learning. The case is worse in relation to technology-enhanced learning environments due to the dearth of research.
- The role of grammar instruction: from the pedagogic side, the debate is still on as to whether, how, and when to integrate grammar instruction in the classroom.

- ISLA and TELL: there are increasing calls for research on the application of ISLA theories and models in general and TBLL in particular to technology-enhanced learning settings
- Face-to-Face classrooms and TELL environments: without evidence, we cannot assume that the findings of classroom-based research apply in a TELL environment.
- User-behaviour tracking technology: data, collected through the use of electronic log files, documenting what learners actually do in CALL activities in real time can provide valuable insights into the process of learning and knowledge construction and can prove very useful in analysing the complexity and abundance of information that dynamic learning environments generate.

These issues are echoed in the presentation of the research questions and hypotheses below.

Research Question One: Which type of input, Focus on Forms (FoS), Focus on Form (FoF) and Focus on Meaning (FoM), is most effective in a TELL-based environment?

Null Hypothesis: There will be no differences in performance among the three groups who each receive a different type of input.

Alternative Hypothesis: There will be differences in performance among the three groups. Based on the literature review in Chapter Two and on the findings of the meta-analysis studies by Norris and Ortega (2000) and Spada and Tomita (2010), it is expected that the Focus on FormS (FoS) group, which receives explicit grammar instruction, will outperform the other two groups in relation to learning of the forms of the target items. It is also expected that the FoM group will outperform the other two groups in relation to overall performance as these learners will only focus on meaning. The first research question is related to the effectiveness of instruction in general and the three input types in particular. Three input groups are compared in terms of their performance. Two measures of performance are implemented: communicative and linguistic. Communicative performance is represented by task completion (also expressed as 'overall communicative performance score') while linguistic performance (also expressed as 'target linguistic performance score') is represented by the gain in relation to the target structure. Learners' performance is expected to vary according to the type of input they are exposed to.

Processing time is also expected to vary based on VanPatten's claim that learners find it difficult to process the input for meaning and form at the same time. Therefore, processing time is also measured and analysed in the present study as an indication of the cognitive demands of the different types of input. The following sub-hypotheses are formulated:

1.a There will be differences among the three groups in their overall *communicative performance*. The FoM group will outperform the other two groups.

1.b There will be differences among the three groups in their target *linguistic performance* score. The FoS group will outperform the other two groups.

1.c There will be differences among the three groups in terms of the processing time. Learners in the FoF group will take longer to process the input as they will need to focus on form and meaning.

Research Question Two: (a) What factors or decision processes exhibited by the learners while dealing with the different types of input contribute to differences in attainment? (b)How do these factors/processes map on to performance?

The second research question is related to the factors that affect performance in a TELL- environment. It is expected that because the three different groups of learners are exposed to different types of input, they will deal with the input differently and this will be reflected in their decisions. Also, the patterns of behaviour in terms of what the learners decide to do or access during the experiment will have an effect on their performance. Based on the literature and the discussion in the previous chapters, the following null hypothesis has been generated for the second research question:

Null Hypothesis: *There will be no differences among the three groups in how they deal with the input.*

Alternative Hypothesis: *There will be differences among the three groups in how they deal with the input.*

The following sub-hypotheses are formulated

2a There will be differences among the three groups in terms of the number of trials.

2b There will be differences among the three groups in terms of task types.

2c There will be differences among the three groups in terms of modality of input.

2d There will be differences among the three groups in terms of access to grammar help.

It is not possible at this stage to predict how the differences will work in relation to learners' decisions as this aspect of the research is quite novel and has not been investigated in previous studies in a TELL environment. Therefore, the question is approached in an exploratory mode.

Research Question Three: How can the use of user-behaviour tracking technologies (log files) help us explain the variability in performance among learners?

No hypotheses are stated as this is more of an exploratory question to evaluate the efficacy of using log files as a data collection instrument in ISLA research. However, based on previous research by Chun (2013) and Ma (2013), it is expected that the use of user-behaviour tracking technology will help us identify precisely what learners do and do not do when dealing with input in a TELL environment under three different input types. It is also expected that such information will help us identify whether there is a relationship between learners' behaviour and performance; in other words, it will help us link the product to the process in learning.

3.8 Conclusion

Chapters One, Two and Three have set the theoretical framework of the current study. Relevant issues and studies in the fields of SLA, ISLA, pedagogy and TELL were reviewed and central arguments were highlighted. The three research questions underpinning the experiment have been stated with accompanying hypotheses. In the next chapter, I turn to a description of the experiment used to answer the research questions and test the hypothesis.

Chapter Four

Methodology and Research Design

4.1. Introduction and Research Questions Revisited

The present study empirically examines the effectiveness of three different types of input and explores the aspects of Instructed Second Language Acquisition processes in a Technology Enhanced Language Learning environment in an attempt to categorize processing skills and to pinpoint definable patterns of processing that might inform ISLA research and the design of TELL materials.

What makes conducting this sort of study a challenge is that it draws from five different disciplines and sub-disciplines, namely Second Language Acquisition, Computer Assisted Language Learning, second language pedagogy, second language materials design and English Language Teaching. One of the challenges is that these operate under different paradigms and employ different methodological approaches as can be noted from the discussion in Chapters Two and Three. The following section is not meant to be a comprehensive review of methods used in each discipline or sub-discipline but rather a description of this study's research design, data collection and analysis with rationale for why such design and methods are used. The rationale will make it clear how different research paradigms have contributed to the present study. In the process, methodological approaches relevant to the current design will be reviewed.

In elucidating, examining and justifying the methods used in the study, the chapter starts by again presenting the research questions and hypotheses and elaborating on issues surrounding them. Then a detailed description of the research design, materials, participants, treatments and tests will be provided. There were two pilot studies conducted prior to the main study and these will be described in detail. Issues of reliability and validity of the tests and treatment will also be considered. Section 4.5.2 closes the chapter by providing a detailed description of the tasks used to elicit data as they form a crucial part of the study.

The present study aims to contribute some answers to the questions initiated by SLA researchers and TELL practitioners on the implications of ISLA theories for TELL material design and the processes learners go through when dealing with input, as discussed in Chapter 3..

The research questions are

1. Which type of input, Focus on Forms/FoS, Focus on Form/FoF and Focus on Meaning/FoM, is most effective in a TELL-based environment?
2. What factors or decision processes exhibited by the learners while dealing with the different types of input contribute to differences in attainment with respect to the construction selected? How do these factors/processes map on to performance?
3. How can the use of user-behaviour tracking technologies help us explain the variability in performance among learners?

Before we turn to the construction selected for the present study, the research design is discussed in detail.

4.2. Research design

In the following section, the justification for an empirical study is first provided, followed by presentation of the experimental design and the rationale for its suitability to the present study is provided. This includes a detailed description of the study design, the materials, the participants and procedures.

4.2.1 Hypothesis testing

The present study was carried out using a hypothesis testing experimental approach as this was most appropriate for the first and second research questions. Before we turn to the hypotheses formulated for the present study, it is useful to consider the sort of experimental methodology common in SLA. Loewen and Reinders (2011) define a hypothesis as an ‘idea or assumption that the researcher holds about a specific aspect of L2 learning’ (p.81) or, as Rasinger (2008) puts it, a statement ‘about the potential and /or suggested relationship between at least two variables’ (Rasinger, 2008, p.175). Hypotheses need to be measurable and phrased in such a way that evidence shows they are confirmed or falsified (Rasinger 2008). Hypothesis testing is common when studies involve the comparison between groups, aspects, and variables. As hypotheses are statements about the possible outcome of research, they can be stated as *null* or *alternative* hypotheses where it is assumed that a relationship will or will not exist among the variables, respectively (Hatch and Lazaraton, 1994). A null hypothesis always assumes that there will be no significant differences/relationships between variables, while alternative, i.e. directional hypotheses are formulated according to the previous findings in the literature or observations of the field/classroom. An alternative hypothesis can thus be formulated in a one-way or two-way direction (Dörnyei, 2007; Rasinger, 2008). If the researcher has reason to believe that the relationship is only from one variable to the other, a one-way or *one-tailed* hypothesis is stated. An example is the relationship between gender and language learning where the researcher assumes a one-tailed relationship in that gender affects the learning process; the reverse is unlikely. However, when the researcher has no good reason or evidence from previous research to assume a one-directional relationship, a *two-tailed hypothesis* is formulated (Dörnyei, 2007; Hatch and Lazaraton, 1994). For example, when investigating the relationship between social status and language proficiency, a researcher might assume a two-way influence so s/he will be looking at how the social status affects the learner's proficiency and also how language proficiency impacts on the social status of the learner.

Statistical tests are then used to reject or accept either the null or the alternative hypotheses. In such an approach, the validity of the hypothesis is tested by observation of the domain under investigation; then, the hypothesis is either confirmed if it is compatible with the observation or rejected if it is not. Hypothesis testing in SLA involves generating assumptions about specific aspects of L2 learning, then conducting empirical research to evaluate these assumptions, and finally drawing conclusions (Loewen and Reinders, 2011). ISLA in particular relies heavily on empirical and experimental studies (Gass and Mackey, 2007) and on intervention studies where the independent variable is some sort of instructional technique or method. Some reasons behind this will be discussed in the following sections.

For the first research question, which type of input, FoS, FoF, FoM is most effective in a TELL-based environment, a 1-tailed hypothesis was formulated to refer to the effectiveness of the type of input where direction is one-way, that is each type of input predicts the following

1.a There will be differences among the three groups in their overall communicative performance score. The FoM group will outperform the other two groups.
1.b There will be differences among the three groups in their target linguistic performance score. The FoS group will outperform the other two groups.
1.c There will be differences among the three groups in terms of the processing time. Learners in the FoF group will take longer to process the input as they will need to focus on form and meaning.

In other words, the assumption is that being in a certain experimental group (FoM, FoS or FOF) will affect the learner's performance.

For the second research question (What factors or decision processes exhibited by the learners while dealing with the different types of input contribute to differences in attainment with respect to the construction selected? How do these factors/processes map on to performance?) Assumptions are stated about the learner's processing patterns; however, as the investigation in the current study into the learner's processing is more exploratory than experimental in nature, the hypotheses were formulated in a non-directional way. Dörnyei (2007) emphasises that it is acceptable to start with non-directional hypotheses when the research question is an exploratory one and asks 'how can we start out by producing a specific research question when we know relatively little of the topic, which is exactly why we want to research into it?' (Dörnyei, 2007, p. 72). The hypotheses therefore predict only that there will be differences among the three groups (in number of trials, in terms of task types, in terms of modality of input and in terms of access to grammar help.

For the third research question, how can the use of user-behaviour tracking technologies help us explain the variability in performance among learners, As noted in Chapter 3, no hypotheses were stated due to the exploratory nature of this question.

4.2.2 Experimental

The current research fits into the category of empirical quasi-experimental research. A quasi experiment is very similar to a true experiment with the exception that the groups or the sample are not randomly assigned (Brown, 2011). Brown states that given the difficulty of collecting data from a true random sample of the population, most second language studies are quasi experimental. This lack of randomization could be due to ethical or practical reasons. In the current study, for ethical, practical and technical reasons, the groups were assigned to the three different types of input according to the classes they attended at a language centre. For ethical reasons, intact classes are typically used where no learners in a given group are excluded from participation. According to Gall, Borg and Gall (1996), using such classes does not affect the internal validity of the experiment as long as classes are randomly assigned to the groups in the experiment. Although the groups might be similar in their proficiency level, age, background, etc. and there is no reason for assigning a particular class to a specific type of input, this type of allocation is not considered a random one and hence the term 'quasi-experimental' is used. The advantage of using a quasi-experimental design is that it allows the comparison of different equivalent or non-equivalent variables/groups as in the current study, as we shall shortly see.

Commonly used study designs in experimental second language acquisition research are longitudinal, case and cross-sectional. Each design has its advantages and disadvantages and thus they suit some kind of research but not others. Dörnyei (2007) described longitudinal research as referring to a family of methods rather than one. However, these methods have one thing in common 'information is gathered about the target of the research (which can include a wide range of units such as people, households, institutions, nations, or conceptual issues during a series of points in time (Dörnyei, 2007, p. 79). In other words, longitudinal studies involve the collection of data at more than one point in time. Cross-sectional studies are used to study behavioural patterns at a single point in time and they study different groups or individuals from more than one perspective (e.g. proficiency level). Case studies are used to study one case (individual, institution, group) in details and do not allow for comparison across groups. On the other hand, longitudinal studies are non-interventionist in nature while cross-sectional research allows intervention (Loewen and Philp, 2011). This means if a longitudinal study design is adopted; the progress of learners is followed as it occurs in the normal course of actions without any intervention or manipulation. As this study manipulates the type of input the learners receive, longitudinal design is not viewed as appropriate. Longitudinal studies are very common when studying the development of a particular phenomenon over time, while case studies are used for in-depth examination of a particular phenomenon, learner, or any single aspect/element at a single point in time. In this sense, the present study is a cross-sectional one as it involves more than one variable and data was collected at one point in time from a relatively large number of cases.

Table 4.1 Summary of study design

	No. of variables	No. of cases	Time	Treatment
Case study	One or more	One case	Not restricted in time	Short in comparison with other designs
Longitudinal	Limited	Limited	Data collected over more than one point in time	Longer
Cross-sectional	More than one	Relatively more than other designs	Data collected at one point in time	Not necessarily included but short (if there is any)

Cross-sectional studies are a common and familiar type of study in classroom research and are used to examine a wide range of phenomena of interest to the researcher. The synchronic nature of the present study allows subjects' decisions to be traced as and when they are made and helps in examining consistency in performance. In addition, the main focus of the present study is not to draw generalisations about the interlanguage systems of the learners or to claim that all learners will go through the same processes. The focus is to identify different patterns of decision making and explore any associations between these patterns and learners' performance, if there are any. A cross sectional study design is a valid research design for providing reliable information about general patterns among large samples of learners. Its disadvantage is that since information is collected only at a single point in time, no claims can be made about time progression or across individuals.

4.2.3 Review of ISLA studies' design

Considering that ISLA refers to acquisition where instruction that target forms and/or meaning is part of the input learners receive, most studies in the field of learning and teaching could come under ISLA studies. When reviewing the literature, different researchers classify studies in SLA according to various criteria. When discussing data elicitation for second and foreign language acquisition research, Gass and Mackey (2007) divide studies into seven categories (1) Psycholinguistics-based research; (2) Cognitive processes and strategies-based research;(3) Linguistics-based research;(4) Interaction-based research; (5) Sociolinguistics and pragmatics-based research; (6)

Survey-based research; (7) Classroom-based research. Gass and Mackey use *research area* to categorise studies this is not completely clear. Take for example, category number (6) 'survey-based research'; it doesn't correspond to an area of research but rather a method of conducting the research while linguistics-based research need not correspond to any particular method.

Loewen and Philp's (2011) classification of studies according to the methodological approach is clearer. They summarize studies under four categories:

- Comparative methods studies;
- Observation studies;
- Non-interventionist quasi-experimental studies;
- Interventionist quasi-experimental studies.

Their classification is simple and allows for the comparison of different methods in ISLA regardless of the specific focus. For the purpose of this section, this approach will be used as it allows a more critical review of ISLA studies based on methods.

Early studies (1960-late 1970s) in the field were **comparative method studies** aimed at comparing different intervention methods and using a pre-/post-test design to assess intact classes (Chaudron, 2001). Any differences found among groups were attributed to the different types of intervention. Intervention is also sometimes referred to as treatment. It involves the manipulation of one or more variables of the experiment in order to allow the study of the effect of this manipulation on other variables. These studies were heavily criticised as there was typically no monitoring of whether a teacher has adopted a certain teaching method or that s/he has adopted it correctly (Long, 1980). Long therefore claims that there is no way of knowing to what to attribute the difference or gains in performance. Ammar and Spada, (2006) point out in addition to information about treatments not always accessible, there were difficulties in accessing classes, getting permission to record and ethics sensitivity in certain classes (those of children or vulnerable people).

Loewen and Philp (2011) stress that early (and some recent, in my view) ISLA studies have two additional major limitations. The first is the choice of methods used to evaluate effectiveness as instruments are recommended that do not always allow the researcher to reach better understanding of the learner's knowledge. The second is the availability of data on the nature of the intervention in practice.

Since the 1980s researchers have been more aware of the limitations of early studies and have focused more on researching the process rather than the product itself. Most ISLA studies were **observation studies** where the aim was not only to compare the different teaching/intervention methods but to research the processes and examine how they affect the outcome. Classroom activities were observed and most of the time these were audio or video recorded and subsequently analysed.

To overcome the limitation of lack of information on the nature of the intervention in observation studies, researchers developed new observation tools such as COLT (Communicative Orientation of Language Teaching, Spada and Frohlich, 1995). These tools are descriptive coding systems which a researcher or an observer uses to record and identify activities in the classroom and have been used in many ISLA studies (Ellis, Basturkmen and Loewen, 2001; Loewen, 2004; Sheen 2004; Lyster and Mori, 2006). There are many advantages of such unified coding systems; they are simple, quick to use in the classroom, it is easy to compare outcomes from different environments, their use doesn't require audio or video recording of the class, which overcomes the difficulties associated with getting permission to do so. Early observation studies were mainly descriptive where categories and items were pre-identified, then their use in the classroom was reported. Later studies took an interactionist perspective. Discourse analysis was used to analyse classroom interaction and draw conclusions. While observation might give the impression of being objective, such studies have been accused of imposing the researcher's point of view on the classroom. If this view is not shared with the students or teachers, problems arise. There is another disadvantage to such systems which can be applied to current classrooms: they cannot be used in blended classes where technology is used as part of the classroom to deliver input, to assess its effects, or to provide opportunities for interaction and negotiation or in e-learning. Although these studies reveal much about what happens in the classroom, most of them have been concerned with the actual learning that took place in these classrooms as a result of the activities in the classroom.

Non-interventionist quasi-experimental studies also tackle the issue of measuring L2 learning through specific types of instructional activities (Loewen and Philp, 2011). Such studies involve looking at the classroom process but unlike comparative methods studies, they measure the outcome, and the researcher has control of what and how much to measure or focus on. Some studies choose to manipulate nothing; these studies have high ecological validity and the classrooms are not affected by the research (Loewen and Philp, 2011). One of the disadvantages of such studies, however, is that the researcher has no control over variables that might affect the outcome. Also, as classes are used intact, many issues related to the linguistic ability of individual students or their personality and other relevant attributes made it hard to draw comparisons or arrive at generalizations.

To allow manipulation and control over variables, **interventionist quasi-experimental studies** are used. These studies are widely used in ISLA to explore a range of research questions. Loewen and Philp (2011) identify several strengths of the interventionist quasi-experimental design. The researcher has more control over the design and implementation of the study, in other words, over what to focus on, the type of instruments to use and the type of intervention. The researcher is also able to eliminate any variables that might affect the outcome; for example, if the research findings might be affected by the learner's L1, the researcher could use learners with one specific L1 or similar L1s. Interventionist quasi-experimental studies are sometimes criticised for being artificial and not of high internal validity as the sample is not random. Therefore, researchers always attempt to keep intervention and manipulation of variables to the minimum needed. In addition, statistical tests are used to ensure that there is no loss of internal validity due to nonrandomized samples. With the increasing number of powerful, parametric and non-parametric statistical tests, such worries can be easily eliminated. In this sense, Isaac and Michael (1995) highlight that quasi-experimental design is as close to a true experiment as an experiment with intact groups can be.

In view of the discussion above, this study is an interventionist quasi-experimental study in the sense that the researcher did manipulate input to address the features associated with different types of input. Also, control over certain learner variables such as proficiency level, amount of instruction, etc. was exercised. In terms of focus, the present study examines the way in which knowledge is constructed in a TELL environment. The difficulty of this type of research is, as Hmelo-Silver (2003, p.398) states, that 'one needs to use multiple methods to understand the interaction' when analysing knowledge construction. In the next sections, detailed description of the methods used to collect the data, materials and procedures will be provided and this point will be addressed.

4.2.4 Study design

According to Rasinger (2008), ‘traditionally, experimental designs are quantitative in nature and consist of the comparison between the experimental group (the one affected by the manipulation of variables) and a control group (not affected by manipulation)’ (Rasinger, 2008, p. 42). The initial design of the present study involved recruiting three experimental groups according to the three different types of input. According to Rasinger (2008), a control group should also be used. When thinking about the recruitment of a control group, a couple of options emerged. As the type of input was the variable that was manipulated in the three experimental groups, the option existed of having a control group which would receive no input at all. However, as the focus of the present research was not on the difference between a naturalistic and an instructed environment in terms of the acquisition of the target construction, a control group that received no input would not add much value to the findings or help in understanding the learning process. There was also the option of having a control group which receive the input in a non-technology environment. But again, such a group would only help us to identify the effectiveness of instruction in technological over a non-technological environment, and as the research focus was not on the efficacy of technology in comparison to traditional instruction. Therefore, a control group with traditional instruction was not used. A control group of native speakers could have been used, but based on research on multicompetence (Cook, 2002, Bassetti and Cook, 2011), there is evidence that native speakers proficient in their own language are different from bilingual or multilingual learners and this would have introduced a variable that was not under examination. In the present study, it was expected that the participants would use different processes or exhibit different patterns of decision making when dealing with the input. As the focus of the research was on how input is processed by learners in a technology-enhanced environment and how knowledge is constructed in such an environment, it was decided that a control group was not going to add value to the research design or help in interpreting the findings (see section 6.7 for the shortcomings of the research design and the experiment and their effects on the findings).
The three different types of input were chosen based on

- Feasibility of designing software
- Prominent effect on classroom practice as shown in previous studies (see sections 2.4 and 2.5 for other types of input and more details on the ones used in the study).

The aim was to programme three software packages to teach three different target constructions. The initial design was to teach each group each construction using the different types of input as shown in Table 4.2.

Table 4.2 Initial design

	FoM	**FoF**	**FoS**
Group A	structure 1	structure 2	structure 3
Group B	structure 2	structure3	structure1
Group C	structure 3	structure 1	structure 2

However, when the labour-intensive process of developing the software started (see section 4.5 below on the design of the materials), it was decided that it would not be possible due to financial, time and resource limitations of this study to develop extensive software packages. Also, when piloting the trial software, it was noted that there would be no way to control all the variables that might contribute to variation in rate and attainment. It was decided to introduce only one construction using three software packages that corresponded to the different types of input. Consequently, the initial design was modified so each of the groups was exposed to only one type of input.

Table 4.3 Modified design

Group A	FoM
Group B	FoF
Group C	FoS

A pre-test/post-test design was implemented. According to Loewen and Philp (2011), if the research question addresses the effectiveness of an instructional technique, a pre-/post-test design is necessary. The first part of the study involved the measurement of the effectiveness of the different input types; thus, a pre-test/post-test approach seemed the most valid method.

Pre-test Treatment Post-test

It was felt that a delayed post-test was not needed for the following reasons: this study was not concerned with the long term effect of linguistic input; there are too many variables that would have had to be strictly controlled, i.e. asking the learners to provide IDs they could remember to log in later or information about their university IDs and this was not permitted by the centre; it would not have been easy to keep track of participants. It is also not very standard in classroom research.

The same tasks were used for the pre- and post-test to increase the validity of the results obtained and to eliminate any other variables that might affect the findings such as familiarity with lexical or context-related elements. Using the same tasks for pre- and post-tests raises a concern about potential practice effects. Odlin (1994) defines practice effect as a 'mean gain score that is influenced by familiarity and/or practice in taking the post-test rather than as a result of the experimental treatment' (p. 327). Brown (1988) warns that when the same test is given repeatedly in a study to determine if there are changes in performance, the researcher needs to consider practice effect. The main issue with administering the same test is that learners can learn from their mistakes, i.e. if the test is re-administered in a short period of time, the learners will remember their responses and the feedback they received. Many techniques could be used to reduce this effect such as administering the post-test after a long period of time to ensure learners cannot remember the answers or withholding answers and feedback or producing more than one version of the test where items are presented in a different order. In the context of the present research, it was not possible, due to technical reasons, to administer the post test at a later session (see section 4.5 and 4.6 below). This is because the learner's data was recorded using a random login ID that was linked to the time and the packages that the learners used but not a real ID and if the post–test was administered at a later session, it would not be possible to match the learner's previous data to the new record. To reduce practice effect in the study, another solution was adopted: feedback was not offered to the learners at any stage of the testing or treatment. All the scores were withheld and given at the end of the post-tests (see section 6.7 for further discussion on practice effect). Also, the tasks were presented in a different order in the post-test.

As stated in Chapter Three, section 3.1, the first research question was about the efficacy of the three types of input in a TELL environment. Three software packages were developed to correspond to the types of input

Input English 1	FoM
Input English 2	FoF
Input English 3	FoS

In Input English 1, the focus was on meaning with the assumption that the target construction would be acquired incidentally while the learners were attending to meaning only.
In Input English 2, the focus was on meaning with occasional focus on form. The assumption was that learners would focus on meaning but also shift their attention to form when they face difficulty in completing the tasks. In Input English 3, the focus was on form and meaning so learners would receive explicit instruction about the form. It was assumed that learners would be able to attend to form and meaning simultaneously.

The study design aimed at replicating a classroom in a technology-enhanced environment to enable a closer look at the learning process. One of the main aims of the research was to identify how knowledge is constructed by looking at how learners deal with the different types of input they receive in real time, as they are learning. (See section 4.5 for a detailed description of the materials). In order to deal with variation across individuals, log files are used to track individual decision making and to help in drawing comparisons across individual patterns, as will be described later

4.3. Informant Recruitment: Interviews with Programme Leaders and Teachers

Recruitment interviews were conducted with three language centre programme leaders and four teachers. The aim was also to get an overview of students' needs and the forms they had problems with. Also, more information was needed about the programs the learners go through and the amount and content of the input they were exposed to. They were two male and five female teachers. Their mean age was 37 and they were all L1 speakers of English. Their teaching experience varied between 7 and 24 years. The interviews were audio recorded and were later analyzed to identify emerging themes and topics to apply to the design of the materials for the study. The programme leaders pointed out that the students to whom access could be provided were all on their 'foundation' year, i.e. they were all going to be studying at university level the next year. Their intended specialties varied across humanities, social sciences, science and medicine. They had English classes for 20 hours a week from 9am -1pm. They went to two-hour subject specific classes twice a week in the afternoon (see section 4.8 below about description of the participants). It was important to find a construction the participants had not yet acquired, and the following points highlighted by the teachers helped narrow this down:

- The learners' most common problems were tenses, subject-verb agreement, articles, prepositions, complex sentences, reporting verbs and structure, and modal verbs. The teachers reported certain problems to be common among all learners such as complex sentence structures and reported statements and others that were L1 related, for example, for Arabic learners: articles, subject-verb agreement, and for Chinese learners: plurals, double subjects.

In creating interesting materials, it was also important to find out what their current interests were and teachers mentioned personal relationships, social problems, technological development, and the environment.

After analysing the interviews, it was noticed that all teachers mentioned complex structures and reported statements/speech as one of the main problems all learners have, regardless of their native language. As most text books and language learning materials at the intermediate level have units to teach reported statements and questions and as one of the goals of the research is to explore the implications for pedagogy and materials design, the decision was made to use reported statements as the target construction in the software package.

4.4. Target Form: Reported Statements and Questions

One of the reasons for choosing the reported speech (RS henceforth) as the target construction in this study, other than that the teachers and programme directors highlighted it as one of the problematic structures is that it does not have a strong communicative value. In other words, learners do not necessarily need to attend to reported speech to understand the message conveyed in the input. However, it is one of the most taught linguistic constructions at the intermediate stage in traditional classrooms and one that learners have problems.

The focus of the present study was not on acquisition of this construction, but rather on learners' responses to input as they attempted to master something they had not yet acquired. Chapter 5 refers to learners' errors and it is therefore useful to provide a description of the phenomenon here. Jakobson (1971) stated that RS is a 'crucial linguistics and stylistic problem' (Jakobson, 1971, p.130). Comrie (1985) states that in English there is a clear distinction between direct and indirect RS. In direct RS, the speaker uses the original utterance of another speaker without any changes whatsoever;

see examples C and D below.   On the other hand, in indirect RS, there is a change in the tense of the verb and also in the deictic centre, see example A, B and E below.

A. She said she had been playing before.
B. She said I had been playing before.
C. She said: ‘I had been playing before
D. He said, ‘I’m the strongest man here.’
E. He said he was the strongest man there.

The semantic and syntactic differences between direct and indirect RS were the focus of many studies in the 1970s and 1980s, e.g.: Banfield (1973, 1982); Partee (1973); Wierzbicka (1974) and Li (1986). In brief, in morphosyntactic terms, the requirements of reported speech are that there are two subjects and there is tense agreement/marking. These studies are not reviewed here as the emphasis in this study is not on morphosyntactic features of reported speech but rather on how it is presented in traditional classrooms and language learning settings.

McCarthy (1998) argues that ‘it is hard to conceive of achieving any intermediate level of competence in a foreign language without needing to know how the speakers of that language make speech reports’ (McCarthy, 1998, p.150). However, many researchers (Carter and McCarthy, 1995; Carter, 1998) have emphasised that ESL textbooks and materials provide inadequate coverage of reported speech. The same could be argued about descriptive grammar books such as the *Oxford English Grammar* (Greenbaum, 1996). Barbieri and Eckhardt (2007) point out that there is a lack of fit between textbook grammar descriptions and real language use. This lack of fit might be linked to many factors including the fact that textbooks often simplify language use for pedagogic purposes; usually present grammar rules as generalizable; are usually based on written norms; and are not informed by empirical evidence about the relative frequency of occurrence of linguistics features (Barbieri and Eckhardt, 2007, p.321). The case of reported speech is not an exception to these observations. It is obvious through the corpus analysis and the ESL textbooks review by Barbieri and Eckhardt (2007) that RS grammar rules presented in pedagogic materials do not account for the variation across spoken and written registers ‘the complexity of RS revealed by these corpus-based analyses contrasts with the descriptions of RS found in many popular

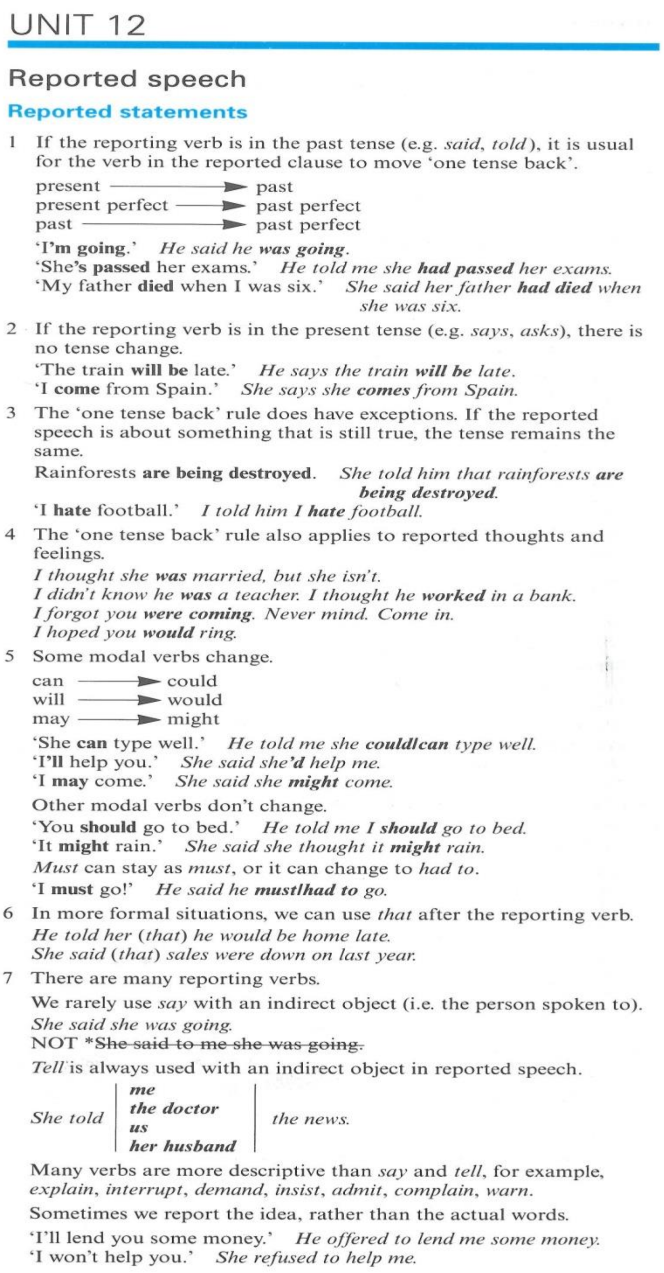
ESL/EFL textbooks, which typically focus on grammatical transformations and backshifting, and neglect register variation in the use of this structure' (p. 338).

In most textbooks the focus is on the tense shift in reported speech. Barbieri and Eckhardt (2007) report on a study by Eckhardt (2001) in which she reviews RS rules in seven widely used grammar textbooks (Murphy et al., 1989; Fuchs and Bonner, 1995; Bland, 1996; Raimes, 1998; Eastwood, 1999; Elbaum, 2001; Thewlis, 2001); this includes the textbook from which lessons were drawn for the present study (see below). Eckhardt's review (as reported in Barbieri and Eckhardt, 2007) indicates that most textbooks focus only on indirect reported speech rather than direct speech. It also revealed that *say* and *tell* are the main two verbs that introduce reported speech. Most importantly, the review shows that in pedagogic materials, reported speech is taught and represented in terms of the following points verbs used to introduce the RS, mainly *say* and *tell*, verb tense combination; and the use of indirect reported speech or what is sometimes referred to as *backshifting*.

Comrie (1985) introduced the term backshifting to refer to the change in the original tense of the direct speech utterance triggered by a past tense reporting verb. Barbieri and Eckhardt (2007) argue that the tense backshift is the main aspect that all ESL textbooks focus on by presenting 'examples almost exclusively in the past tense'. The same observation is stated in Charkova and Halliday (2011) where 'a central issue in the construction of English indirect reported speech is the phenomenon of tense backshifting' (p. 6).

The focus in the present study was not on the adequacy of the ESL materials nor on the acquisition of RS but rather on replicating what is typically found in a traditional classroom in a TELL environment. Therefore, the target construction was introduced in the software along the lines it is presented in ESL textbooks in general and the very commonly used *Headway Intermediate* in particular. This means that the focus of grammatical instruction was on the tense shift and the past tense in particular. The study materials did not include examples where the subject of the main and reported clause was the first person singular, where grammatical differences were neutralised and reference must be either retrieved from context or marked phonologically. Nor did the materials include examples of direct reported speech where there was no tense shift (see examples B and C above).

In *Headway Intermediate* (2003), the textbook used in this study, reported speech is presented in the student's textbook through tense shift, using reporting verbs (mainly *say* and *tell*) and use of *that* in reported statements. Figure 4.1 below shows the relevant page.



UNIT 12

Reported speech

Reported statements

1 If the reporting verb is in the past tense (e.g. *said*, *told*), it is usual for the verb in the reported clause to move 'one tense back'.

present → past
present perfect → past perfect
past → past perfect

'**I'm going**.' *He said he **was going**.*
'She**'s passed** her exams.' *He told me she **had passed** her exams.*
'My father **died** when I was six.' *She said her father **had died** when she was six.*

2 If the reporting verb is in the present tense (e.g. *says*, *asks*), there is no tense change.

'The train **will be** late.' *He says the train **will be** late.*
'I **come** from Spain.' *She says she **comes** from Spain.*

3 The 'one tense back' rule does have exceptions. If the reported speech is about something that is still true, the tense remains the same.

Rainforests **are being destroyed**. *She told him that rainforests **are being destroyed**.*
'I **hate** football.' *I told him I **hate** football.*

4 The 'one tense back' rule also applies to reported thoughts and feelings.

*I thought she **was** married, but she isn't.*
*I didn't know he **was** a teacher. I thought he **worked** in a bank.*
*I forgot you **were coming**. Never mind. Come in.*
*I hoped you **would** ring.*

5 Some modal verbs change.

can → could
will → would
may → might

'She **can** type well.' *He told me she **could/can** type well.*
'**I'll** help you.' *She said she**'d** help me.*
'I **may** come.' *She said she **might** come.*

Other modal verbs don't change.

'You **should** go to bed.' *He told me I **should** go to bed.*
'It **might** rain.' *She said she thought it **might** rain.*
Must can stay as *must*, or it can change to *had to*.
'I **must** go!' *He said he **must/had to** go.*

6 In more formal situations, we can use *that* after the reporting verb.

He told her (that) he would be home late.
She said (that) sales were down on last year.

7 There are many reporting verbs.

We rarely use *say* with an indirect object (i.e. the person spoken to).
She said she was going.
NOT *~~She said to me she was going.~~
Tell is always used with an indirect object in reported speech.

She told	***me*** / ***the doctor*** / ***us*** / ***her husband***	*the news.*

Many verbs are more descriptive than *say* and *tell*, for example, *explain*, *interrupt*, *demand*, *insist*, *admit*, *complain*, *warn*.

Sometimes we report the idea, rather than the actual words.

'I'll lend you some money.' *He offered to lend me some money.*
'I won't help you.' *She refused to help me.*

Figure 4.1 Grammar reference: reported speech (Headway Intermediate, 2003, p.155)

These rules can be summarised in the following points:

- A verb in a reported statement moves one tense back if the reporting verb is in the past tense. e.g.

Table 4.4 Tense change in reported statements

Present	Past
'I'm going'.	He said he was going.
Present Perfect	**Past Perfect**
Mary has passed away.	She told him Mary had passed away.
Past	**Past Perfect**
'My father died'.	He said his father had died.

- There is no tense change;

1. if the reporting verb is in the present tense

'I'm going'. He says he is going.

2. the reported speech is about something that is still true

'I hate football'. She told him she hates football.

- Reporting verbs include: *tell*, *say*, *ask*, *explain*, *admit*, etc. *Tell* is always used with an indirect object while say is used without one.

He said he was going He told me he was going.

- You can use *that* after *say* and *tell*.

He said that he was going. He told her that he was going.

- When you report a question, the word order changes. There is no inversion of subject and auxiliary verb. You can use *if* and *whether* when you are reporting a question. For example,

Are you married? He asked if I was married.

Where have you been? He wants to know where I have been.

These explicit points were all used in the materials for grammar instruction in the two software packages, FoF and FoS, but not for FoM (see section 4.5 below). What is worth emphasising here is that the target construction itself was not the focus of the study. Rather the study focused on what learners attend to and notice, i.e. the processes involved in dealing with the input that is rich with a target construction they have not yet acquired, as presented in the three different manners.

4.5. Materials

After choosing the target construction, the researcher started developing the materials that were used for the treatment and research. Content and design needed to comply with the rules of TELL (Technology-Enhanced Language Learning) and the features for each type of input. Three options were available for the content (1) inventing it from scratch; (2) using existing materials or (3) adopting existing materials.

The advantage of the first option was that the researcher would have control over all the variables. The disadvantage was that many pilot studies would have been needed to ensure validity and reliability of the material. The second option was the best to save time and effort but the problem was that there was no control over the variables. There was no existing software package that appeared to represent any SLA research. Also most of the software packages use audio lingual methods of drill and repetition, which violates the principles of communicative and task based learning. The third option was the best fit for this research as it allowed for control of variables, eliminated extra pilot studies to validate the content and most importantly, allowed application of SLA research. The other advantage of this option was that one of the ultimate goals of the research was to explore the implications for pedagogic practices and material design, so it seemed more useful to use materials that were already used by teachers and software developers. This meant that findings could be related more easily to existing practices and modifications could be carried out smoothly. Why is this software different from existing marketable materials? The answer is that input is controlled so the focus is on one specific construction. The content for the materials was adapted from *Headway Intermediate.*

It was very important when designing the three versions to keep the context as similar as possible in the three packages. It has already been argued in the previous chapters that the difficulties centred on contextual factors push the learners to a higher number of reformulations and repairs in order to overcome a wider range of difficulties. This, in turn creates a complex set of variables that should be considered when analysing the outcome. Using the same tasks and input for the three different versions eliminated the possibility of the learners performing differently because of cognitive and contextual task demands.

The materials included three versions of a software package designed in accordance with the three types of input. Materials were designed using Macromedia Authorware. Authorware is a visual authoring tool that is used for creating rich-media e-learning applications. The applications can then be delivered through CD/DVD, local network or the web. The advantage of using Authorware was that although sophisticated software could be created using scripts and coding, basic programs and activities could be created without the need for any programming experience, just training on the actual software charts and interface design. This fact was critical when the pedagogic value of the software was considered. However, as one of the main aims of the present study was to investigate the use of user-behaviour tracking technology to provide insights into the learning process, programming was crucial for all the collection of such data. Therefore, when working on the design of the software, two aspects needed to be considered: developing the content and implementing built-in scripts to allow data collection.

Authorware is a flowchart-based package used to create interactive programs. Figure 4.2 below shows the software structure for the log-in screen from the designer's perspective and the learner's perspective. Each icon contains further sub icons and charts as can be seen from figure 4.3 further below.

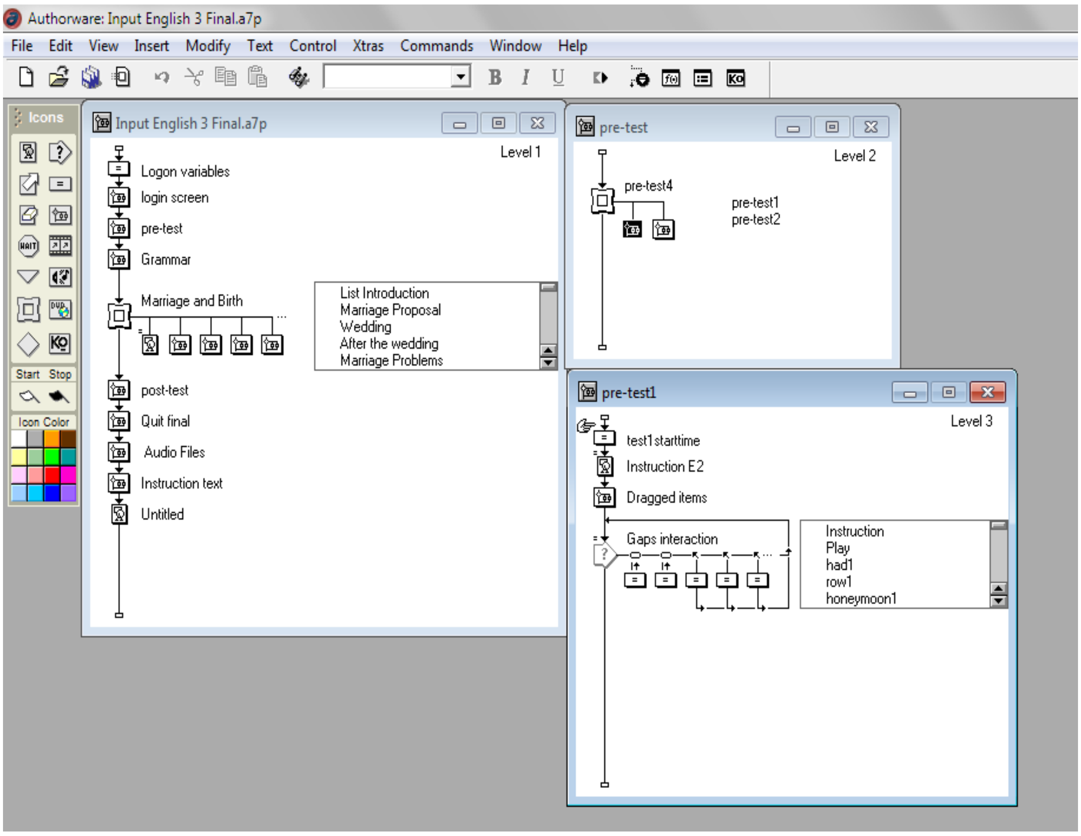


Figure 4.2 Screenshot of the flow chart

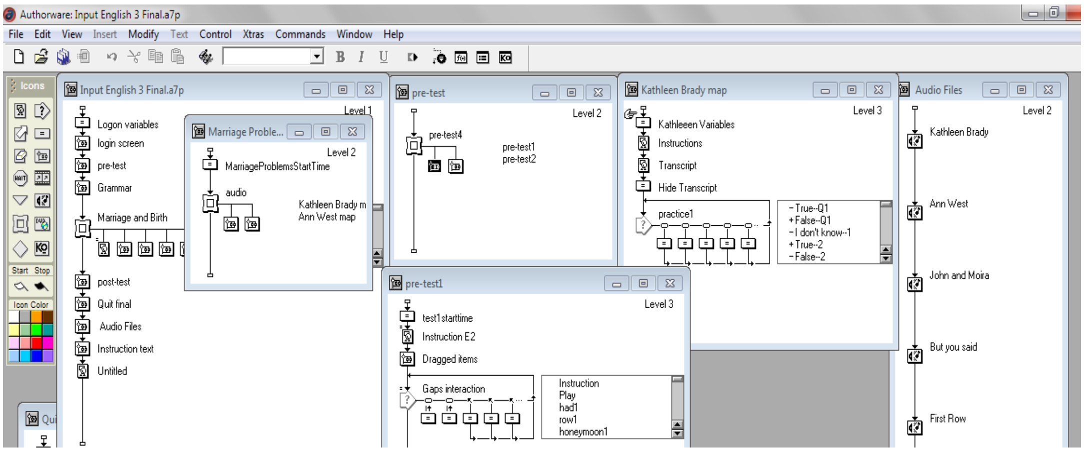


Figure 4.3 Screenshot of the sub-charts

While developing the content for the lessons, Java –based script was used to write the codes for the log files. This meant that a script was needed for every action that learners were predicted to make. Scripts were used to record the time, order, and outcome of every action. For example, if the learner dragged an item, information would be recorded about the time he dragged the time, which item he dragged, whether it was dragged to the correct gap and how many times was it dragged. To clarify how the tracking script and data worked, I will use the word ‘had’ from the first task as an example. The following script was used to record the response to dragging this word to any gap and to calculate the score for that item.

Example one: sample script used in the log files

```
if ObjectMatched = ‘had2’ then
   had:=‘1’
else
   had:=‘2’
end if
if had=1 then
   hadscore:=1
else
   hadscore:=0
end if
AppendExtFile(PathUser^UserFile, Return^Return^’48 = ‘^had)
```

The above script means that if the word ‘had’ has been dragged to the gap ‘had2’, then write 1 to the data file; otherwise write 2. Then if 1 is written to the data file, add 1 to the score of the learner, otherwise add 0. Then write all the information to the data file. In order to enable an accurate measure of whether the item was dragged to the right gap, it was important to first set the X and Y axes for the item itself and the gap, as can be seen from Figure 4.4 below. Another script was included to record the number of times the item was dragged (referred to as trial number) and also the actual time the item was dragged. Similar scripts were included for every action that learners could make.

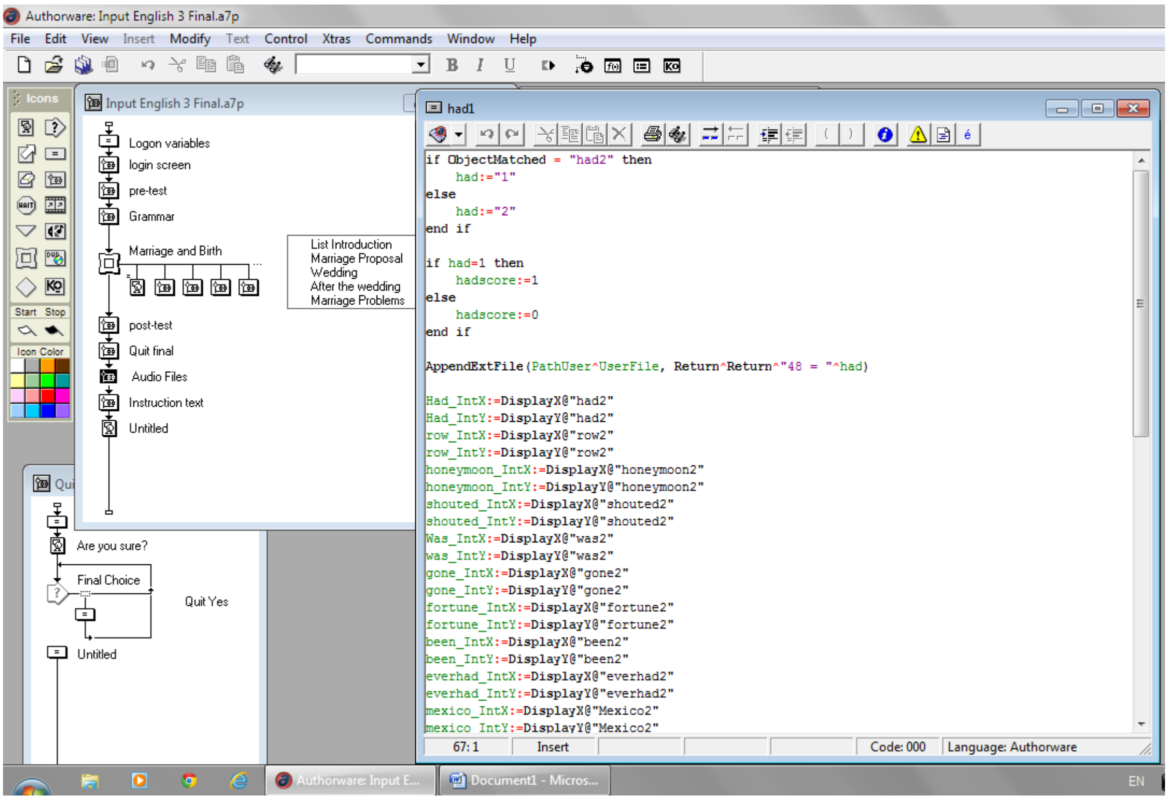


Figure 4.4 Screenshot of a sample of the script used for data collection

Data from the previous scripts were automatically written to a text file with the user name. For example, the data from the above script for 'had', will be written to the individual text file as

had response = 2
had time= 09:18:29
had score=0
had trial =1
had response = 2
had time= 09:18:35
had score=0
had trial=2
had response =1
had time = 09:18:57
had score =1
had trial =3

This would indicate that the learners dragged 'had' three times, the first two to the wrong gap then to the right one, and his final score was 1. A script was included in the design to allow the automatic creation of an individual data file once the learner logged in then all the information was written to the user file as they were using the software. The scripts were adjusted after the pilot study to include only number. In other words, it was clear after the pilot study that there would be a problem with the data as it had text and most statistical programmes do not support text analysis. Therefore, the scripts were modified to replace action explanation with numbers so the script above changed to

had response=	71	=2
had time=	72	=09:18:29
had score=	73	=0
had trial=	74	=1
had response=	71	=2
had time=	72	=09:18:35
had score=	73	=0
had trial=	74	=2
had response=	71	=1
had time=	72	=09:18:57
had score=	73	=1
had trial=	74	=3

A script was also included in the *Quit* button at the end of the software to enable data to be written to an *index* file (see Figure 4.5 below), where data from all learners were written to one file and only the last record for each action was recorded. So for example, in the above example only the following information was written for each learner.

User ID	had response	had time	had score	had trial
User1	1	09:18:57	1	3
User 2	2	10:10:23	0	1

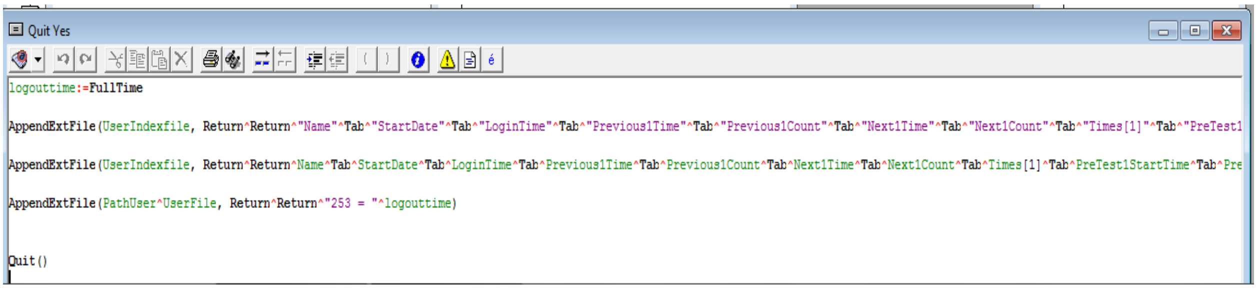


Figure 4.5 Script for index file data

Having the two files proved very useful later for analysis. As some of the learners did not complete all the tasks or click the *Quit* button, the individual files provided information about what they did during the time they were using the software. The index file data, on the other hand, was crucial for drawing comparisons among the learners to identify patterns of behaviour.

This section detailed the technical side of designing the software; the next sections will focus on the pedagogic and linguistic aspects of the design.

4.5.1 Input features

The input types under investigation here were focus on meaning (FoM), focus on form (FoF) and focus on forms (FoS) (see Chapter 2 for detailed discussion). The main distinction between the three types is the way grammar is presented in the input. The software package had to be designed according to the following (see Table 4.5);

- For FoM, no grammar is provided; all tasks should be meaning focused.
- For FoF, grammar should only be referred to incidentally and only when there is a breakdown in communication or task completion.
- For FoS, grammar is essential and it should be referred to explicitly and available throughout the tasks.

Table 4.5 Types of input and focus of instruction

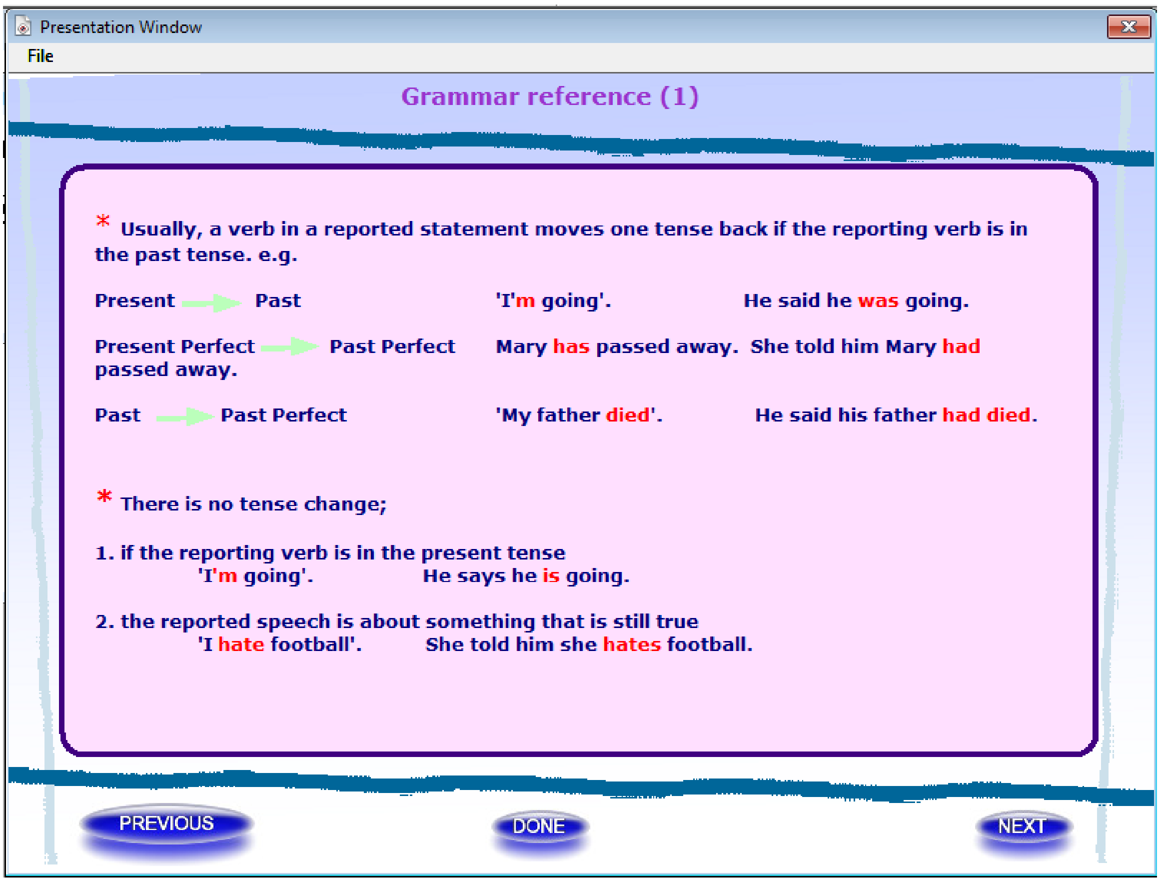
	Software package	meaning	form
FoM	Input English 1	Y	N
FoF	Input English 2	Y	incidental
FoS	Input English 3	Not necessarily	Y

Three packages were developed to correspond to the three types, i.e. Input English 1/FoM; Input English 2/FoF; and Input English 3/FoS. In the next section I describe the operationalization of the three types of input in the treatment packages.

In Input English 1, FoM, the focus was mainly on meaning so no grammar instruction at all was provided; the learners only received meaning-focused tasks where the text and audio files were rich with the target input. When a learner logged on to the application, they went through several screens that introduced them to the application and the navigation buttons, they then moved to the test tasks which were introduced as practice tasks. Once they were finished, , they moved straightaway to the six treatment tasks and when they finished and click *done*, they went to the post-test tasks. There was no focus on form at all. The aim was to flood the learners with input rich with the target construction. Care was taken even not to use words that might indicate any formal references such as, verbs, sentences, grammar, etc.

In Input English 2, FoF, the aim was to provide the learners with meaning focused tasks and also to provide them with incidental focus on form. To achieve this in the application, grammar instruction was provided as a help icon which learners could click on as and when they wanted (see Figure 4.4 below for the actual grammar instruction included in the software package). After the learner logged in and once they were done with the test tasks, they went to a screen which introduced the treatment tasks and also informed them that there was a new button added which was the *grammar help* button. They were informed that they could click on the button if they were stuck or whenever they wanted and for as many times as they wanted. This meant that the learners had control over when and for how long to focus on the forms involved in the construction. Learners were not provided with any grammar instruction other than the optional *grammar help* button. This was a main difference from a classroom environment where usually a teacher is in control of when to focus on form and for how long.

In Input English 3, FoS, the focus was on providing the learners with explicit grammar instruction. After the learner logged in and finished the test task, s/he was directed to a grammar explanation about RS (see Figure 4.4 below for the actual grammar instruction included in the software package). The grammar instruction was the same as the one provided in the optional *grammar help* button. In this package, the learners could not move to the treatment tasks until they went through all the screens with the grammar instruction. The learners were given control over how much time they wanted to spend on the explicit explanation and this time was measured for analysis later. The learners then needed to click on a *done* button to move to the treatment tasks. During the treatment tasks, the learners had access to the optional grammar button. In brief, to allow an explicit focus on formS, learners were provided with explicit grammar information on how to form reported speech in English before going into the main treatment section. They were also provided with the grammar help icon as part of the main treatment.



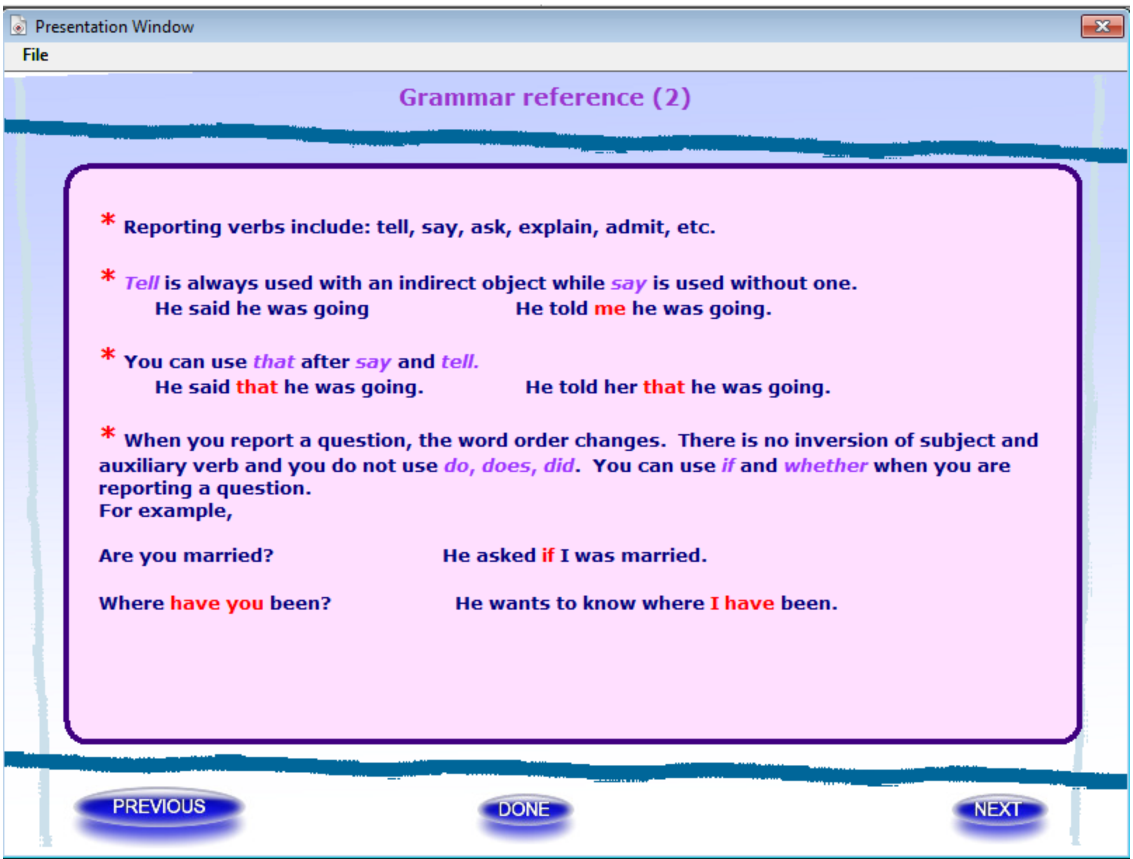


Figure 4.6 Screenshot of the grammar instruction provided to the FoF and FoS groups

For each version of the software there was

1. Pre-test: two tasks, fill in the gap and drag and drop.
2. Treatment: six tasks, two fill in the gap, two drag and drop and two multiple choice.
3. Post-test: two tasks the same as the pre-test but in different order (see section 4.5.2 below)

The tasks in the three versions were exactly the same apart from the extra grammar information. This was meant to control for the number of tokens, quality of the materials and any other variables that might affect the outcome.

	2 tasks	6 tasks	2 tasks
Input 1	Pre-test	Treatment	Post test
Input 2	Pre-test	Treatment with Grammar	Post test
Input 3	Pre-test	Grammar Info+Treatment with grammar	Post test

In addition, the same tasks were used for the pre- and post-tests but were presented in a different order; the drag and drop was first in the pre-test while fill-in-gaps was the first in the post test. This was meant to reveal any differences between time 1 and time 2 and eliminate any other test effects. Another reason for using the same tasks for pre- and post-tests was to eliminate variables such as quality and quantity of the input, familiarity with the topics, etc. This design was intended to keep the experiment within similar parameters.

4.5.2 Tasks

Different criteria were applied when selecting tasks. In the following sections, I will discuss the features of the tasks from different perspectives.

4.5.2.1 ISLA perspective

Chaudron's (2003) exposition of data is widely used in SLA. He classified data into naturalistic, prompted production and prompted response data. This classification reflects the common tendency in SLA to distinguish between naturalistic data and elicited data. Clearly, researchers in SLA always prefer to obtain naturalistic and spontaneous data whenever possible. This, however, is usually not the case in ISLA as by definition, it involves learning through instruction. Many studies in ISLA focus on identifying the features in the input that prompt or hinder learning (Ellis, 2001, 2006). The most common method used in ISLA to draw conclusions about learners' underlying grammar is through the use of quantitative empirical data that has been collected from tasks to elicit production and test comprehension of the target language. Heavily used techniques in ISLA include grammaticality judgment tasks, interpretation tasks and elicited production tasks (Hulstijn, 2000; Mackey and Gass, 2012).

Interpretation tasks are mainly used when the researcher is concerned about the link between form and meaning. Typically, learners are asked to listen to or read a pair of sentences or a short passage and then answer questions related to the meaning. For example; learners read the following sentence in a story: *John was watching TV when the door rang* and then they are asked whether John usually watches TV. The assumption is that learners will realize that a simple present form is what indicates a habitual action and as the previous sentence is in the present progressive, they will answer no or I don't know.

Although the current study is concerned with form-meaning connection, its focus is not on whether learners are able to demonstrate their understanding of how target grammatical structures affect the meaning. The focus is rather on how learners respond to different types of input when the focus of the input is either on form or meaning or both. Therefore such tasks were not used.

Grammaticality judgment tasks (GJ) are commonly used in generative SLA. Generative SLA researchers (White, 2003) argue that grammaticality judgment tasks help in investigating the learner's internal grammar, i.e. their implicit knowledge of the target language structure. GJ tasks have been used so far to examine linguistic competence, which can be accessed only indirectly and under controlled conditions (Mitchell and Myles, 2004). Although they are and continued to be heavily used in generative SLA, GJ tasks have been criticized widely. Bley-Vroman and Masterson, 1989 and Sorace, 1996 have raised concerns about the suitability of GJ tasks for beginners and intermediate students and Schutze (1996, 2005) questions the type of knowledge that GJ task really test. The current study did not aim at testing or falsifying the nature of the knowledge acquired by learners; it is concerned with the processes that affect this knowledge. GJ tasks were not required for this study. Data obtained through elicited production tasks was deemed more reliable.

Elicited production tasks are used heavily in ISLA to collect written and spoken data or to elicit responses to prompts. Elicited production tasks are the most common when investigating the effectiveness of different types of intervention or the processes involved. For example, Ortega (2009) used story retelling to investigate the effect of pre-task planning on the focus on form episodes that learners engage in during the actual performance of the task. De Jong (2005) used self-paced reading and sentence matching to examine whether the comprehension of target structure leads to implicit or explicit knowledge of that structure and whether this knowledge is available for production later.

Stimulated recall (immediate or delayed) and **think aloud protocols** are used by SLA researchers to explore learners' processes and strategies. Learners are usually prompted to recall thoughts and feelings that they had while or after dealing with the intervention or participating in an activity. Gass and Mackey (2000) and Bowles (2010) extensively discuss the use of stimulated recall and think-aloud techniques in second language research. As the most popular techniques to investigate internal processes, many studies have employed stimulate recall to determine the effectiveness of feedback (Swain and Lapkin, 2002; Mackey, Gass and McDonough, 2000). Think-aloud protocols are used to examine the relationship between input and intake (Leow, 1998) or to identify the relationship between types of exposure and levels of awareness (Rosa and O'Neill, 1999). As popular as they are in processing studies, think-aloud protocols and stimulated recall have many pitfalls, as will be discussed below in section 4.7.1.

In reviewing the literature on tasks used to measure input effectiveness and processes in ISLA, it was noted that different tasks are used depending on the nature of knowledge tested (implicit, explicit), the focus of the input (communicative, meaning-focused, form-focused), the perspective of the researcher (processing, efficiency, etc.). The main focus of the present was not on the nature of knowledge acquired through different types of input but on the effectiveness of the three different input types and the processes that correlate or lead to effective production. In this sense, the tasks adopted just needed to measure effectiveness of the input and the processes learners used. Elicited production tasks were therefore used in the study along with user-behaviour tracking technologies (see section 4.7.1 for discussion of these technologies).

4.5.2.2 TBL perspective

The tasks created for the materials in the present study met the conditions known to allow task implementation in a way that optimizes output and helps in assessing learners' performance (as stated in Pica at al., 1993; Crookes, 1986, Skehan, 1996). Skehan (2001) lists the following as basic task features

- The focus is on the meaning;
- There is a problem to solve;
- Performance is outcome evaluated;
- There is a real world relationship.

The study tasks had a real world context; '*relationships' problems*', a topic that was recommended by all five teacher interviewees (see section 4.3 above). The tasks were also presented in a very close-to-real-life replica of a negotiation situation in order to motivate interaction and help learners' activation of the sort of communicative skills required for interlanguage development. Keeping some of these elements under control is basic to accomplishing the task in an efficient way.

As discussed in Chapter 2, in section 2.5.2.1, R. Ellis (2009) differentiates between focused tasks, unfocused tasks and situational grammar activity. In a situational grammar activity, the focus is not primarily on meaning and the outcome is simply practicing correct language. Unfocused tasks aim at providing opportunities for language use in general, while focused tasks provide these opportunities while focusing on specific linguistic features. Regardless of whether the tasks are focused or unfocused, they still need to meet the four conditions similar to those proposed by Skehan (2001) and shown above and explained by Ellis (the primary focus is meaning; there should be a gap, for example, to convey information; learners have to rely on their resources to complete the task and there is a clearly defined outcome; Ellis, 2009, p. 223). Ellis (2009) warns that a focused task is not the same as situational grammar activity as the target linguistic feature is hidden in focused tasks while explicit in the situational grammar activity. This framework is used in the current study to inform the design of tasks for the three different versions of the software, that is:

Table 4.6: Type of tasks according to the type of input

Type of input	**FoS**	**FoF**	**FoM**
Type of task	Situational grammar activity	Focused tasks	Unfocused tasks

As mentioned above, very common task types in classroom research are story re/telling and picture description but these were not used in the current study due to the difficulty of obtaining data in terms of decision making; they are not common in TELL as they are not easy to assess; they can't be analysed in a computerised quantitative way.

Another important issue here is that most studies of TBLL are conducted in a classroom environment. Typically, the interaction taking place between the learners and the teachers or among learners in the classroom is recorded and then analysed and compared with performance results. In this context, story re/telling and picture description are very useful as they potentially increase the quantity of produced utterances. The difference in the current study is the medium in which the learning took place. In a TELL environment, the interaction is between the learner and the technology which is different to the human-human interaction (teacher –students and student-student) in traditional classroom studies. Consequently, different tasks are needed to explore task processing and task features.

Important attempts to improve elicitation procedures in language production within the issue of more authentic communicative goals derive from the application of the principles of a 'task-based methodology'. Therefore, when designing the tasks for the present study, two points were taken into consideration: avoid activities that restrict language input to only the grammatical forms or are decontextualized and at the same time elicit input in a format analysable electronically to create a model.

Different task types have varied impact on the opportunities for the negotiation of meaning, and different task types activate different processing skills. The relationship between task type and task demand is critical as the task demand impacts on the learners' attentional resources. Some tasks require word-level processing while others require sentence or discourse processing. In brief, different task types make different demands on the learners' noticing behavior and therefore could trigger attention to form at certain points to overcome difficulty in completing the task. As a result of the different task demands, learners decide to choose to focus on meaning or form or both.

4.5.2.3 Information processing perspective

To account for the factor of attention, different scores were calculated.

- Overall communicative performance score: score of all correct answers to the target and non-target items;
- Target linguistic performance score: score of the target items only, i.e. the items that needed attention to form to supply the correct answer. For example, in the information-gap test task, learners needed to attend to form to supply had instead of has in the following gap; it _________large tropical gardens.

In addition, for the FoS and FoF groups, access to the grammar help was monitored and checked against each task type. As the task processing demand was expected to vary according to type, results should indicate that learners' access patterns correlate with certain task types,

Of course, if the results reveal that certain tasks trigger attention to form or result in more accuracy, pedagogical claims for classroom and material design could be made and, claims could be drawn regarding facilitation of language acquisition or use. In brief, different demands are associated with different types of tasks and consequently activated processing skills vary according to the task type. The outcome of a task is directly related to high or low processing demands required by each task type. By tracing performance behaviour, i.e., decision processes, when learners are coping with a similar task in different input-based contexts, patterns can be revealed of how learners react in the face of a breakdown in (simulated) communication. Assessing these patterns against learners' performance scores can shed light on what works and ultimately might lead to effective acquisition (as measured by learners' performance scores).

4.5.2.4 Pedagogical perspective

Different tasks were used for the treatment and tests (see below for explanation). The main types of tasks used were drag and drop, gap-filling and multiple choices, as shown in Table 4.7.

Table 4.7 Summary of task types and testing

Task	Type of task
Pre-test task one	Drag and drop
Pre-test task two	Fill in the gap
Treatment T1	Fill in the gap
TT2	Drag and drop
TT3	Drag and drop
TT4	Multiple choice
TT5	Multiple choice
TT6	Fill in the gap
Post test T1	Fill in the gap
Post test T2	Drag and drop

The reasons for choosing these three task types were:

1. They are relatively easy to construct and score;
2. They are compact and efficient;
3. They are used to gather a large amount of information in a short period of time and space;
4. The effects of guessing are less when compared to other types such as true/false.

Multiple choice questions, used as one of the tasks, are believed to be best used to assess simple and complex learning while matching (not used) is best used to assess association and relationships (Jolliffe et al. 2001). The main focus of the present study was to create a purpose for communication that would promote spontaneous reactions. This was best achieved through information-gap activities when learners are asked to complete a communicative task where they are required to attend to the meaning. The tasks were designed so that they would trigger a genuine context for negotiation of meaning and also provide opportunities for learners to stretch their interlanguage resources to produce probably more accurate forms. This was achieved by presenting learners with different choices for the drag and drop task and the learners needed to notice the difference in grammatical forms to supply the correct answer. Although, from a TBLL perspective, the learners could have completed the task successfully without attending to form, they needed to attend to form to supply the accurate verb form. As argued by Pica and Doughty (1986) and Bygate (1987, 1988), the key is to find a purpose for communication. In this study, this purpose comes via opportunities for interaction created through an information-gap and drag and drop tasks.

4.5.3 Pre-/post test tasks

The following section will describe the two tasks used in the pre-and post-tests. As was noted earlier, a drag and drop and an information gap task were used in the pre-and post-test, but they were presented in a different order to reduce test effect. Information gap is used here as fill-in-the gap as the completion of the task required the learners to attend to the information provided in a brochure learners saw during the testing and treatment.[16]

4.5.3.1 Task 1: Drag and Drop

The task required the learners to drag items from a box to the correct space. The text was an entry in the diary of a newly-wed woman called Karen who had just had her first row with her husband, Tom. An audio file was provided for the actual row and learners were supposed to listen to the file and choose the right answer.

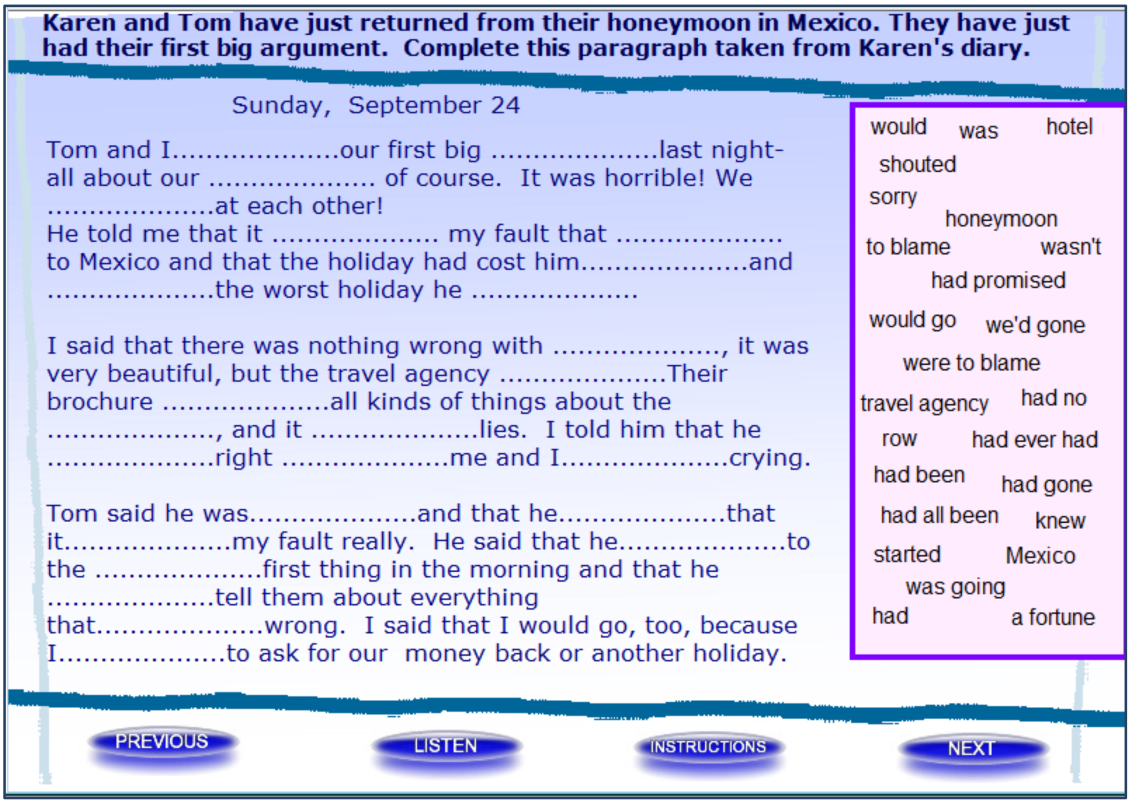


Figure 4.7 Screenshot of the drag and drop task in the pre-test

[16] This is different from traditional information-gap tasks used in TBLT where learners are usually given part of the information and are asked to talk to each other to produce a complete version of the information.

4.5.3.2 Task 2: Information gap

The task required the learners to write responses in specific gaps. Karen and Tom are talking to a travel agent about the hotel they stayed in. A brochure is provided about the hotel and learners needed to use the information in it to complete the conversation.

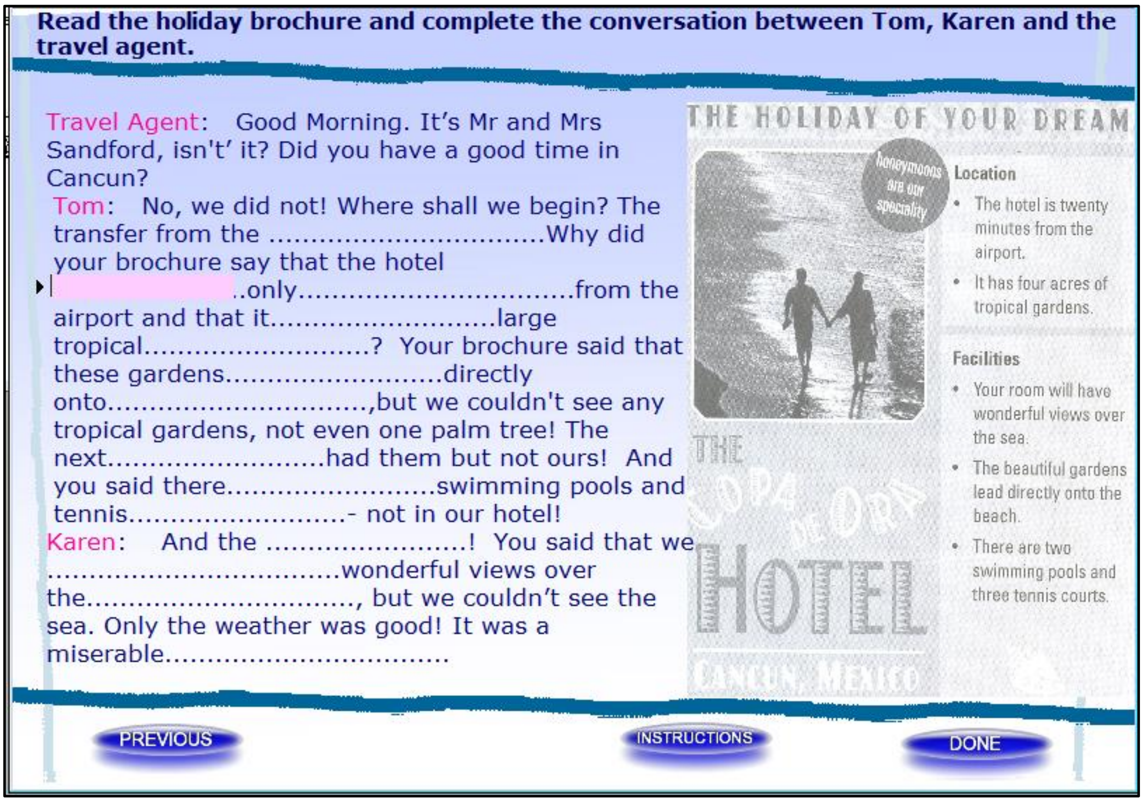


Figure 4.8 Screenshot of the information-gap task in the pre-test

In total, there were 25 items in the first task, 15 target items and 10 distractors. In the second task, there was a total of 14 items, 5 target and 9 distractors. Table 4.8 below presents a summary of all the target items and associated verb forms. As mentioned earlier in section 4.4 above, the main reporting verbs used in pedagogic materials are *say* and *tell*; they were both used in the tasks.

Table 4.8: Target items and forms

	Item	Form
Pre-Test Task 1	Would	*Will* past
	Was	*Be* past
	Wasn't	*Be* past negative
	Had promised	*Have* past
	Would go	*Will* past
	We'd gone	Past perfect
	Were to blame	*Be* past
	Had no	*Have* past
	Had ever had	Past perfect
	Had been	Past perfect
	Had gone	Past perfect
	Had all been	Past perfect
	Knew	Past
	Was going	Past continuous
	Had	Past had
Pre-Test Task 2	Was	*Be* past
	Had	*Have* past
	Led	Past
	Were	*Be* past
	Would have	*Will* have past
Total	20 target items	6 forms

It is important to mention again here that it was not assumed in the present study that the number of token items or forms provided was sufficient for the forms to be acquired, and in any case, consideration of learners' acquisition was not the aim of the study. Two main factors affected the amount and quality of the input provided. The first was that the aim was to replicate what is used in the classroom so it was important to adhere to what is in the textbook. The second was that it was essential to control the amount of input provided as the method used for data collection generated a considerable amount of data as all learners' actions were recorded. Even with the limited amount of items provided, some learners' data included over one thousand entries.

4.5.4 Treatment tasks

Treatment tasks were exactly the same for all learners in terms of content with the important difference that the FoS group had a grammar explanation before they were exposed to the treatment tasks and a ‘grammar help’ button on each screen, the FoF group only had access to the ‘grammar help’ button and the FoM group had no access to either the grammar explanation or the ‘grammar’ help.

4.5.4.1 TT1: Marriage proposal

The first treatment task was an information gap task. Learners were asked to complete the conversation between a male and female friend talking about their trip when the man proposed to his friend. A summary retell was provided at the top of the screen. The learners also had the option of listening to the actual conversation. The aim was to focus on meaning but at the same time provide learners with written input rich with reported statements and questions. The learners had to change the form of verbs in the retell before using it to complete the actual conversation.

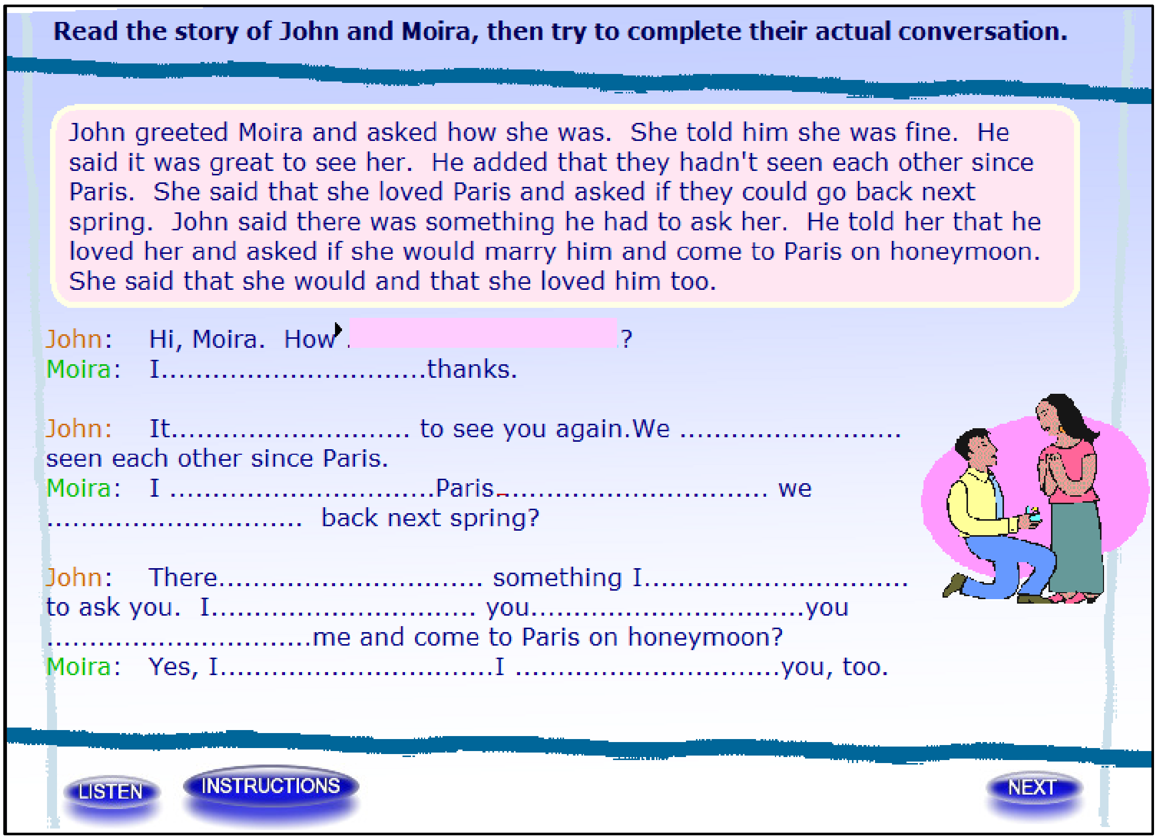


Figure 4.9 Screenshot of treatment task one

4.5.4.2 TT2: Wedding

The next task was a sequencing task. Learners were asked to sequence the conversation between two people who met at a mutual friend's wedding. Learners also had the choice of listening to the actual conversation. The focus of the task was mainly on meaning and completing the task successfully. The audio file again provided oral input rich with the target structure.

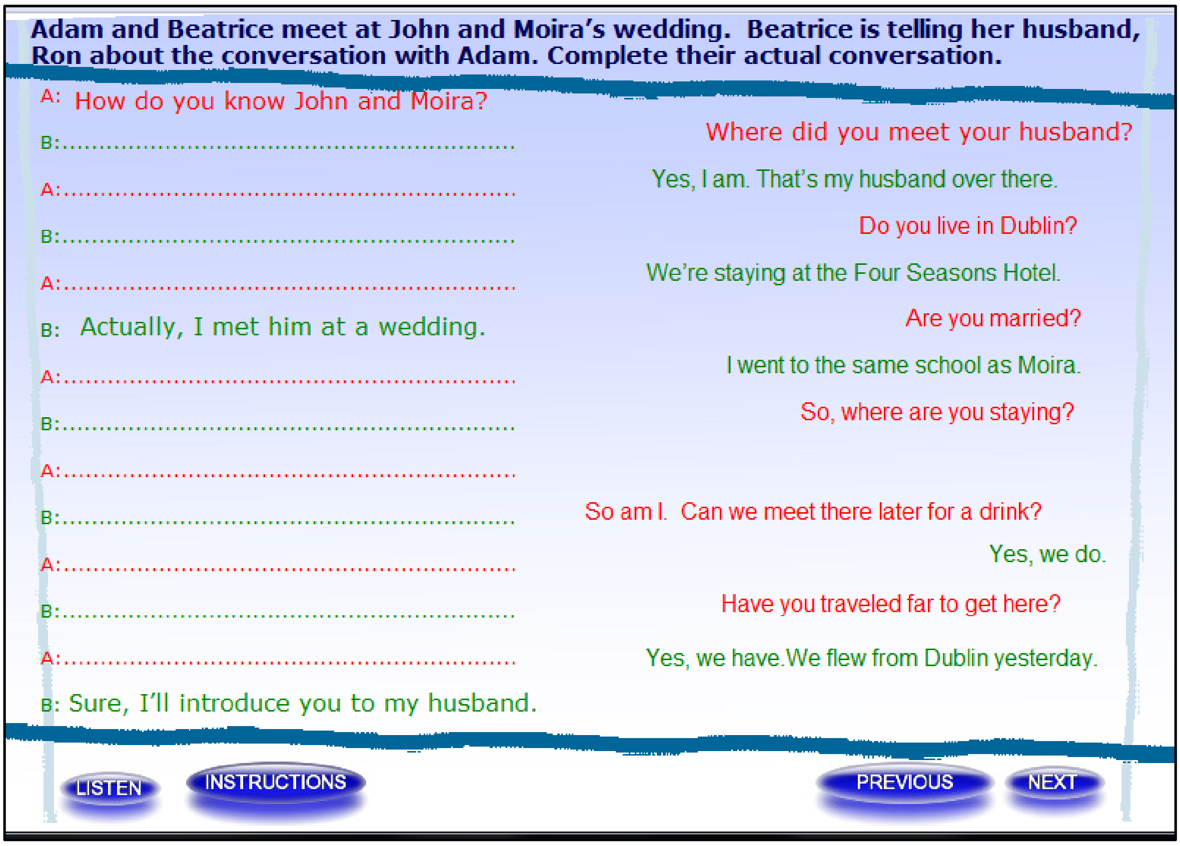


Figure 4.10 Screenshot of treatment task two

4.5.4.3 TT3:After the wedding

This task was a drag and drop task. Learners were asked to complete a conversation between Beatrice and Ron about Adam, whom they'd just met at John and Moira's wedding. The focus of this task was meaning with attention to form needed to complete the task successfully. The learners had the option of listening to the conversation which was rich with reported statements. The items they needed to choose from had different statements with different verb forms which they needed to attend to.

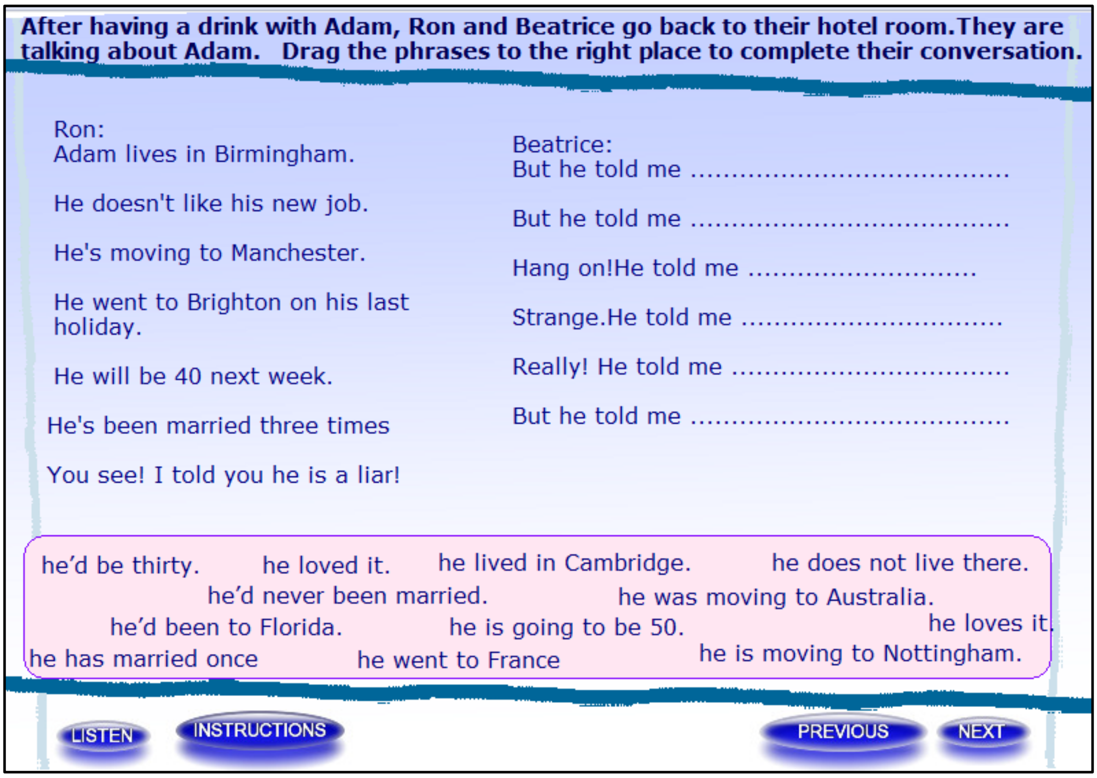
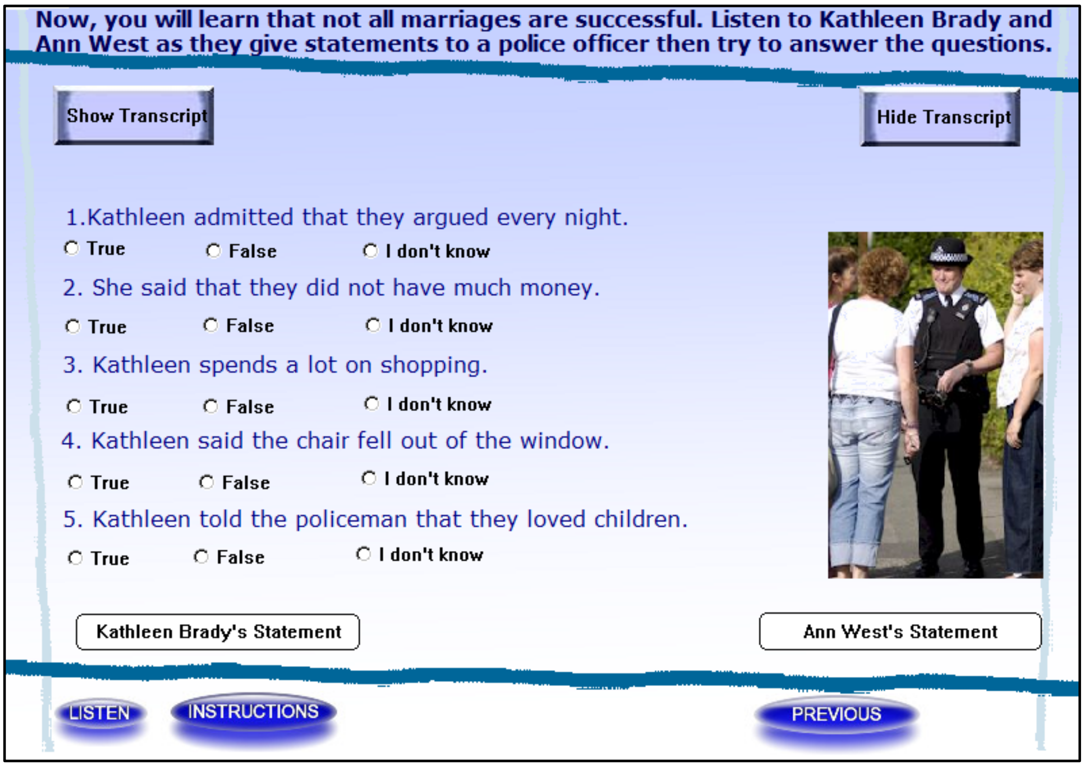


Figure 4.11 Screenshot of treatment task three

4.5.4.4 TT4: Marriage problems

This task was a multiple choice task which had two parts. Learners were asked to listen to statements from two women who are neighbours giving information to the police. The first statement is from Kathleen Brady who had a big fight with her husband the night before and the neighbours had called the police. Kathleen is being interviewed by the police officer. The second statement is from one of the neighbours, Ann West. Audio recording and written transcripts of the statements were provided. The focus of the task was mainly on meaning and completing the task successfully. Target structures were included in the questions and the statements.



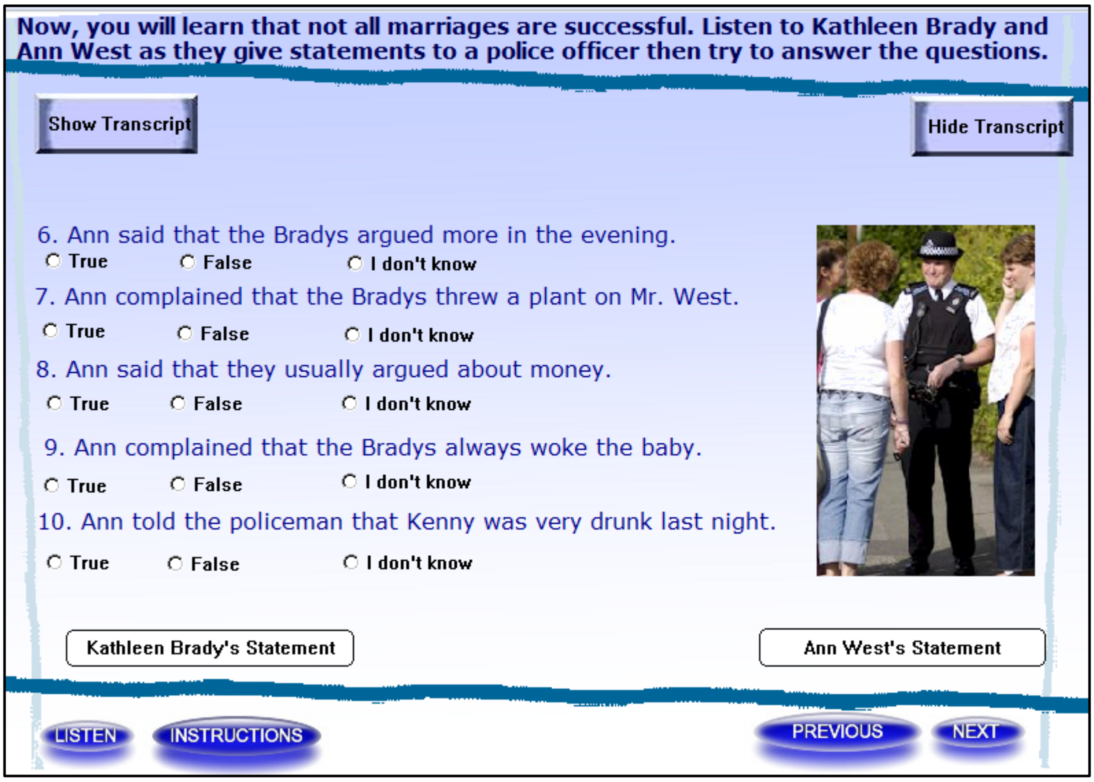


Figure 4.12 Screenshots of treatment task four

4.5.4.5 TT5: A Birth story

The task is a multiple choice task. Learners are asked to listen to Catherine telling the unusual story of how her sister gave birth. The focus is on meaning. No attention to form was required to complete the task successfully. Learners needed only to attend to meaning to complete the statements.

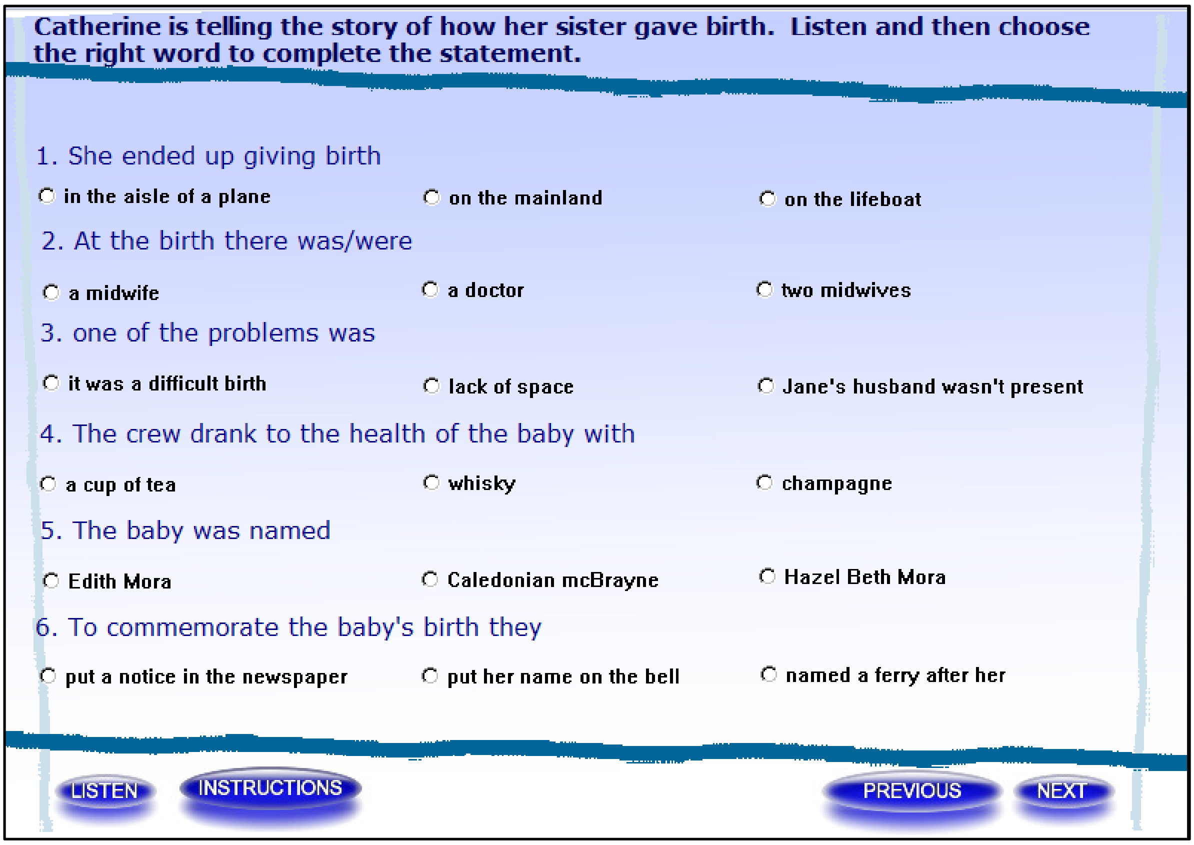


Figure 4.13 Screenshot of treatment task five

4.5.4.6 TT6: An unusual birth

This task was an information gap task. It was based on the same story as the previous task, the unusual birth. Learners needed to attend both to meaning and form to complete the task. They had the choice of listening to the original story told by Catherine but were asked to complete the story as it was retold by Catherine's friend. Written input was rich with the target structure.

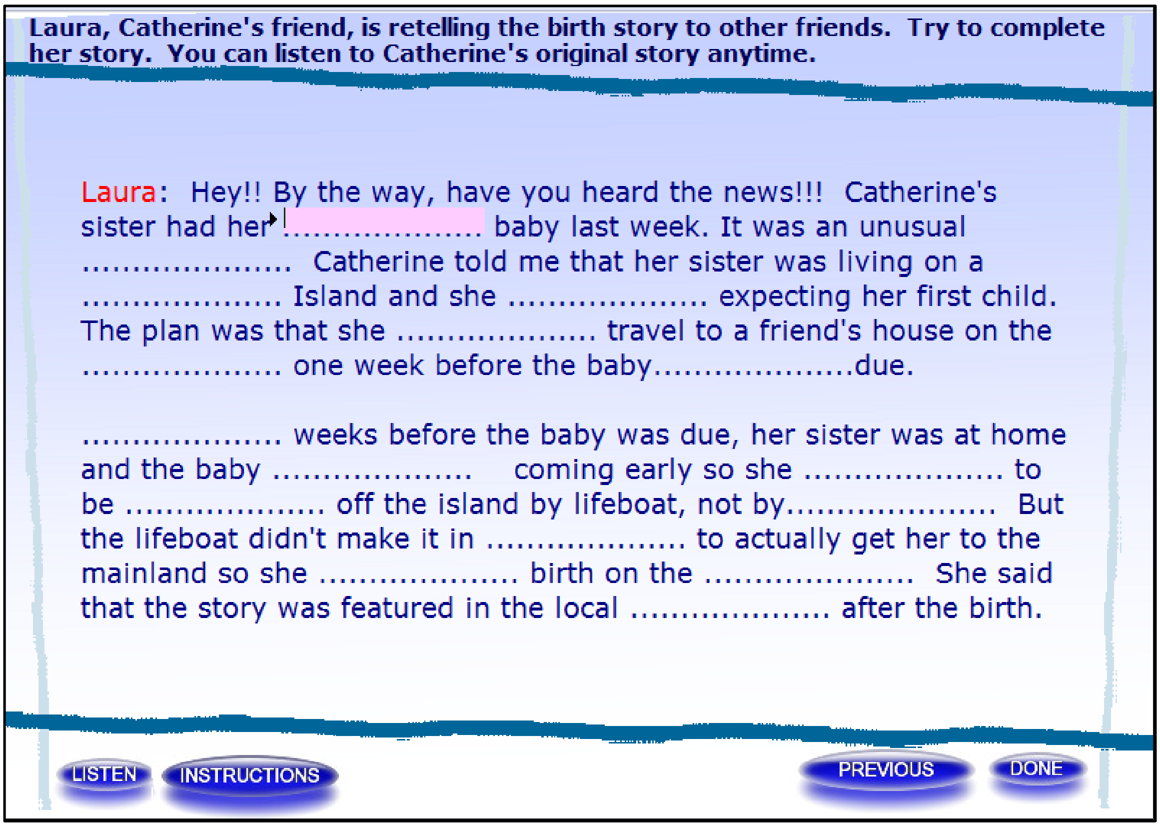


Figure 4.14 Screenshot of treatment task six

Two the pilot studies were conducted on the above materials prior to the main study.

4.6. Pilot Studies

The first pilot study was to establish user-software interface characteristics and the second to use the software among the target learners to confirm that the language level and content was appropriate and to decide on the time needed to complete the tasks.

4.6.1 Pilot study one: user-software interface

An initial pilot study was carried out to ensure that the software was user-friendly. The study involved 10 students and two language teachers. Participants were undergraduate and postgraduate students, studying different disciplines. The students were seven females and three males. Their age range was 20-29 with a mean of 26. Two students were English native speakers and eight were non-native speakers of English with different L1s including Arabic, Chinese and Korean. The trials were conducted individually and participants were asked to comment on presentation, navigation and design while going through the software. The researcher observed and took notes. The modifications were applied in accordance with the responses received as detailed in Table 4.9 below.

Table 4.9: Pilot Study Modification

Problem	Solution
Some learners didn't figure out how to do the first pre-test	Instructions were added
Didn't know where the tab+shift are	Added 'use tab+shift on your keyboard'
Didn't know how to hide the instruction pop out	Added info to the instruction itself asking them to click again on the button to hide the instruction pop out.
Drag and drop task background will move if they didn't click on the right word	Changed the properties of the background to 'unmovable'
When the students press the listen button in 'A birth Story', they could answer the questions until the audio file is paused or stopped	Changed code so they can do both simultaneously
Some learners clicked on the 'done' button at the end of the pre-test by mistake and had to exit and log in to complete the other tasks	A confirmatory question was used before the last screen on the pre-test to avoid clicking on the 'done' button by mistake
Students going back and forth between exercises to be able to listen more than once to the audio files, this was messing up the data	Changed code so they can pause and listen to the audio file more than once without having to restart the activity
Students struggled with the tab movement in the gap-information activities.	Navigation between gaps was modified; instead of just using tab to navigate which means they will go through all gaps between 2-5 to move from gap 1 to 6. A new design is used where they could just click on the gap or move forward using tab and backwards using Tab+shift
Could not identify gaps easily	Use boxes to highlight gaps
Was difficult to control the audio navigation bar	Removed the audio bar and replaced it with a button where they could listen and stop by one click
Learners kept going back and forth between the treatment exercises	Removed the previous button so they can't go back to a previous exercise (was messing the data so much) and included a message about this at the welcome page
Audio file will start automatically when they move to a new activity	Changed code so they had to click on the 'listen' button to start the audio file
Images were distracting	Wash out the images
A list was provided on the third screen listing the pre-test, treatment and post test activities but was confusing to the students as there was a 'done' button	Change the order of pages to go through the pre –test then show the list of treatment activities

after the pre-test and some thought that they have missed the treatment	
Click on the 'done' button for the pre-test before moving to the second pre-test activity either by mistake or intentionally	'Done' button only works after they move to the second exercise in the pre-test
Learners start the activities and they are not familiar with the navigation bar and button so they spend long time on the first activity trying to figure out the navigation which increase session time for that activity and give a false interpretation of complexity	Include an information page before the test to explain navigation and to inform learners about how many activities. Include the sentence 'have you done this?' at the end of the information page to increase the possibility learners taking the time to explore the buttons
The 'Wedding' activity's instruction read 'listen and complete' but there was no need to get the learners to listen first	Changed the instruction, just used 'complete' so learners have more control over their choices.
Grammar reference was put on a different navigation bar so there was two bars on the screen	Grammar reference included also on the navigation bar
Long pauses and silence in the original audio activity	Audacity software was used to edit the audio files to reduce pauses and silent periods
Learners clicking on the 'Restart' and 'Quit' buttons by mistake which means they have to go through the materials from beginning again. This posed two problems, they will be more familiar with the materials the next time. As the log in details were anonymous, there was no way of identifying the previous work of individual students	Remove the 'Restart' and position the 'Quit' only on the last screen
Some learners didn't like the font	San Serif fonts are used
Display was different according to screen size which messed some of the activity	Screen size was adjusted to fit different options
Clicking on different buttons by mistake	Deactivate buttons when not in use
When learners were going back and forth, the 'session time' data was messed as it was only recorded for the final time they accessed the activity	By solving the listening problems, learners didn't have to go back and forth all the time. Also, a new code was added so 'session time' was recorded every time the learners accessed an activity and then all logs were added manually
Some learners wanted more personalised messages	Personalise the software by using ID to welcome and give feedback

The modifications shown in the right-hand column were all applied. The second pilot study was then conducted to pilot the first version of the software among target subjects and to identify the best way to implement the main study.

4.6.2 Pilot study two: target subjects and implementation plan

For this pilot study, data was collected between October and December of that year. 51 participants took part in this study. All participants were ESL learners at an intermediate level, as determined by the language centre. They had all scored 5.5 in IELTS prior to joining the class on a foundation course. They came from 10 different L1 backgrounds: 31 Chinese, 1 Vietnamese, 9 Arabic, 1 Turkish, 4 Cantonese, 1 Nigerian, 1 Urdu, 1 Malay, 1 Russian, and 1 Azerbaijani. Their age range was 17-23 with a mean age of 19.

The main aim was to make sure that the lexicon and construction were at the right level and to find out the time needed to complete the software lessons. Log files, observation and questionnaires were used to collect data. Questionnaires were used to collect biographical information about the learners and also to check their IT knowledge. Observations were used to examine any problems with the software in terms of navigation or presentation and also to look for any problems encountered by the learners when browsing through the software. Log files were used to collect data regarding the time spent on each task, scores, and decision processes.
The following was noted and modified as a result of the pilot study:

- The sequencing exercise seemed difficult to complete; to make it easier different colours were used to indicate different speakers;
- Clearer, more detailed instructions were recommended, so instructions were divided into two parts: navigation and task instruction. In addition, an instruction button was added on each screen for ease of navigation. Navigation instructions were included as part of the navigation menu as a button and were displayed only when the button was pressed. They were highlighted and put inside a box and larger text was used while task instructions were positioned at the top of the screen and were present all the time.
- Codes were added to record data regarding how many times the instruction button was pressed and for how long.

- More distracters were added as some learners were concerned about the focus of the materials. .

4.7. Main Study: Data Collection Methods

As in the second pilot study, and after applying all the modifications according to the findings of pilot study two, data for the main study was collected using log files, questionnaires and observation. Questionnaires were used to collect biographical information about the learners and also to check their IT knowledge. Observations were used to examine any problems with the software in terms of navigation or presentation and also to look for any problems encountered by the learners when browsing through the software. Log files were used to collect data regarding the speed, accuracy, task efficacy and decision processes. Methods used in ISLA were reviewed in the previous sections (see section 4.2.3) so the following section will discuss log files, the main method of data collection in this study and its suitability.

4.7.1 Log files: user behaviour tracking records

Log files, also referred to as user-behaviour tracking technologies by Collentine (2000), is a non-invasive technique for collecting behavioural data. To put it simply, log files are an electronic register of events in real time. Researchers or designers use programming scripts to collect information about events that take place when participants are exposed to technology-based materials.

4.7.2.1 Use of log files

For years now, this technique has been used to study consumers' online behaviour by recording information on which pages are visited and which links are clicked. Recently (since the 2000s), it has been increasingly used in language and learning research. Its most common use is in web-based and online learning, where information is stored about learners' actions such as buttons pressed, links visited and words looked up. In ISLA research, log files have been increasingly used in grammaticality judgment tasks to record reaction time and responses (Marinis, 2010; Wright, 2010, 2012) and also in eye tracking experiments. Eye-tracking technologies mainly look at one aspect of the behaviour, which is visual attention. Information is usually obtained about what learners looked at and for how long, and in some cases the order of areas looked at. As important as that is for ISLA, it does not provide insights into learners' actual behaviour when dealing with input for the purpose of the present study. Therefore, although eye-tracking technologies utilize log files, they will not be discussed within the scope of this study.

From an ISLA perspective, experimental techniques such as think-aloud protocols, eye-tracking and brain screening, mostly borrowed from the psychology field, are applied to the study of language processing by adults and the call to use such tracking technologies in language research started in the 1990s. Cubillos (1998), discussing tracking features, stated that students' logs and students' records of all kinds can provide teachers with unprecedented insights into their students' SLA processes (p.45). The same argument is made by Garrett (1998) who suggested that the only way we could make claims about the efficacy of technology use in language teaching is by using tracking software. What is needed is data collected 'on what students do with technology-based language learning materials while they are in the act of working with them' (Garrett, 1998, pp.9-10). Highlighting the importance of the use of log files in experimental language research, Hulstijn (2000) notes how it allows 'an unobtrusive observation' of participants' 'look-up' behaviour, 'with these computer-aided tools, however, researchers have the means to get closer to the processes of language acquisition and use' (Hulstijn, 2000). Chun's (2013) recent review into the contributions of tracking user behaviour to SLA research concludes that such data can provide 'valuable insights into both second language acquisition and pedagogical design' (p.256) and that collecting data on learners' actions through log files can help in identifying underlying strategies and navigation behaviour or document which features or components of the input are most frequently used and how this maps onto learning outcomes (Fischer, 2007, 2012; Chun, 2013).

From a TELL perspective, Garrett (1998) points out that in CALL research, data collected using log files could be used to address three areas:,

- Evaluation of the software (features used, outcomes achieved, examination of individual differences in terms of strategies and attainment)
- Investigation of non-technology based teaching (evaluating the pedagogical principles, assessing knowledge absorbed)
- Students' strategies and behaviours in the new learning environments (links between technology and learning and teaching processes).

Garrett's claim that by tracking and monitoring what language learners do in 'the whole range of language learning environments and materials' (Garrett, 1998, p.10), researchers can inform and define the ways technology shapes language learning. These points still hold in 2014.

Although the call for the use of tracking technologies started in the 1990s, ISLA studies using them did not appear until the year 2000. The first study in ISLA that reported the use of log files was Collentine (2000). He used what he refers to as 'user-behaviour tracking technologies' to record all the events that learners generated while using a specially-designed software application. Collentine concluded that learners did not exploit opportunities to engage in exploratory strategies even when they were available, most learners provided very short answers and those who did so did not benefit from instruction, learners kept moving backwards and forwards between slides to explore the information before answering any questions relating to conscious-raising activities. He also concluded that the use of tracking technologies provided insights into the processes of L2 acquisition that are difficult to observe using other techniques.

Other, later studies include Bruckman (2000, 2006). Bruckman (2006) used log file data to analyse reading and learning behaviour in children aged eight and above. Results showed that the use of log files enables researchers to gain significant insights into user behaviour and learning. Fischer (2007, 2012) and Ma (2013) also emphasise the importance of tracking data in revealing how learners actually interact with the learning system. As Ma (2013) puts it 'only with a good tracking system can CALL effectiveness be proven, useful design features identified, and the appropriate applications be selected' (p. 230).

4.7.2.2 Advantages of using log files

Log file data is believed to be the best method when examining user behavior in different disciplines (Bruckman, 2006). This is mainly because it allows the collection of a comprehensive record of all interactions and activities. Another reason is that log files are non-invasive, so there is no risk of altering the participants' behavior. In fact, it has been argued that even telling the students that their responses are recorded does not result in altering their behaviour or their level of participation (Black, 2008, Pellettieri, 2000). In brief, log files are considered a reliable method of data collection as they provide a comprehensive and complete record of all activities with a guarantee of the accuracy of information. They are easy to use for data collection once the coding is set and are not biased by the subjectivity of the researcher or any external variables.

In traditional classroom research, think-aloud protocols and observations have been used to collect the kind of information obtained by log files. Think-aloud protocols are usually administered in two ways: retrospectively, after the learners finish the treatment and concurrently, while they are doing the treatment. For the purpose of the present study, using retrospective think aloud protocols did not seem suitable. One of the many problems is that the treatment lasted for about 90 minutes and it was not possible for learners to precisely recall what they were thinking or their actions. Ellis (1997) claims that asking learners to describe how they learnt is limited as learners are sometimes unaware or cannot remember the actual learning processes they engaged in. On the other hand, concurrent think-aloud protocols are believed to slow down processes (Sanz et al., 2009) which would affect study results. Most studies that have used think-aloud protocols involved training the participants through pre-task activities to familiarize themselves with the method (see Bowles, 2010 for a full account of the use of think–aloud in L2 research). Think-aloud protocols were not suitable for the current study as they would have put extra load on the learners' cognitive system which would have affected processing time and patterns. This was pointed out by researchers such as Seliger (1983) who warned that verbalisation adds another task for the participant and thus affects performance. Observations, on the other hand, which are widely used in classroom process research (Ellis, 1990), were used in the present study as a support tool of data collection rather than the main data collection instrument. The data needed for the current experiment involved recording all actions taken by learners while dealing with the input. Apart from using log files, there was no way that the researcher would be able to record all actions taken, even if the experiment was administered one learner at a time.

Ma (2013) points out that one of the key benefits of collecting data using log files is that it can record 'an array of user actions for learners to manage and monitor their learning as well as provide an intelligent learning system for evaluation purposes based on which the system may recommend appropriate learning paths for learners' (p. 231). Fischer (2007) also asserts that tracking user-behaviour using log files has the advantage of 'unobtrusively observing students' behaviour' (p. 411). Despite increasing interest in log file data, very few studies have been carried out to investigate the actual relationship between user actions and the actual learning outcome.

As the focus of the present study is on how learners deal with the input and construct knowledge in real time, log files were considered the most valid and reliable method for data collection. One crucial reason is monitoring consistency of use. This is directly associated with whether certain types of behaviour can further or hinder completion of tasks. When the communication demands made upon the learners are exceed their current linguistic competence, they are forced to adopt alternative paths to overcome communicative problems. There is no other method available that can provide such information in a TELL environment. The feasibility of collecting information about all events that take place during the experiment and linking each event to a specific response or time was only possible through log files. The importance of this kind of information is that it can provide an interesting profile of individual patterns of behaviour/strategies which can reveal specific patterns underlying performance.

In addition to tracking the decision processes exhibited by learners when using different types of input, in the present study, log files were used to measure speed and accuracy. Information was recorded about when learners moved an item/ filled a gap/ made a choice and how many trials they took each time. A trial is taken here to mean any attempt a learner makes towards completing the task. For example, in the drag and drop task, a trial is when the learner drags an item. If he or she drags an item once, it is recorded as one trial, when s/he goes back and drags the same item; this is considered another trial and so on. In the information gap task, a trial is when a learner fills in the gap, if s/he changes his/er mind and changes the answer, this is considered as a second trial. In the multiple choice task, every time the learner changes his/her mind and clicks on a different choice is considered a trial. The number of trial for individual items was recorded along with an overall number of trials for the whole task. Also information was recorded on how long learners spent on each activity.

As was explained earlier in section 4.5, information was written to two different files, the 'index' file and the 'individual ID' file. The files were in a simple text format. The 'index' file had summary information from all learners in each group. The script in the software package was encrypted so information was written in a grid format where rows represent learners and columns represent events recorded. For example, the following table is from one of the index files; information written here includes the learners' ID used to login, the date, the exact time they logged in to the software, the last time the learner pressed the previous button and the number of times s/he pressed it. Columns were delimited by tab for ease of transfer to statistical package or database software to analyse.

Name	StartDate	LoginTime	Previous1Time	Previous1Count
Wong	Wed, June 18	11:22:03	11:41:08	4
Lucy	Wed, June 18	12:56:12	13:10:30	1
Dais1	Wed, June 18	09:16:44	09:21:12	1

In addition to the 'index' file, each learner had a file titled according to their unique ID. To follow the same example above, three files are created titled *Wong*, *Lucy* and *Dais1*. The 'individual ID' file was a more detailed file. Events were recorded in order and as they happened. Each event was assigned a number, for example, Name =1, Date=2, Time previous button is pressed =46, number of times previous button is pressed =47 and so on. This was done to make it easier to analyse the data as many statistical programs do not allow tests on text data. So every time the previous button is pressed, information is recorded about the exact time it was pressed and the number of times it was pressed so far. In the example below, the learner 'PEI' has pressed the previous button two times at 15:59:47 and at 15:59:48.

1 =	PEI
2 =	19 June 2008
3 =	15:59:20
9 =	15:59:35
46 =	15:59:47
47 =	1
46 =	15:59:48
47 =	2

The information was also in the format of a grid where rows represented events and the two columns represented the event number and the value of the event retained.

It is assumed in the context of this study that using such techniques will help in answering questions about the approaches learners exhibit when attempting to grasp the relationship between meaning and form.

4.7.2.3 Problems with collecting data using log files

To put it in Bruckman's (2006) words, one of the major problems with data collected using log files is that 'log file data is more often collected than analysed' (2006, p.1449). This statement encompasses the main problem of collecting data through log files. Very often, the researcher ends up with a considerable amount of data that s/he will find difficult to analyse or does not know how to analyse. This is echoed by Chun (2013) who notes that some of the common reasons for not incorporating user-behaviour tracking techniques are the sheer quantity of available data and the time required to process and analyse the data This problem is exacerbated by the lack of research that addresses the problem or models that could be referred to. Even studies that have used log files rarely detail how data was extracted from the coded information. Thus, although researchers have rich data, they struggle to make meaning out of it. In other words, the difficulty is in the identification of the key user actions.

Another problem when collecting data using log files is producing the scripts/codes to encrypt information. The scripts are written in widely-used programming languages such as Java or HTML (for web-based materials) or specific authoring languages created by authorware designers. In most cases, a programmer writes the script according to guidelines from the researcher or the teacher. The difficulty with this is that there is not a one-size-fits-all design available so it is more a trial and error procedure which is both time and effort consuming.

4.7.2.4 Issues related to electronic data collection

One of the important issues I had to consider when collecting data electronically is data storage and security. Options for data storage could be encompassed as follows;

- Local disc: needs simple programming and no access permissions required as the data is saved to a space on the temporary local drive. The problem with this option is that the data will be lost when the user logs off so data need to be transferred to another location straight after it is written to the space. This poses a problem particularly when collecting data from groups rather than individuals.
- Shared network drive: needs programming and institution's permission to access the drive to write to and to collect the files written to it. But the advantage is that it can be used to collect all the information at once and it is not time consuming

or demanding. Also data can be accessed and transferred at any time as long as access to the drive is permitted.

- Students' file space: needs institution and individual access permission to write to the space which is difficult to get. Although, the data will not be lost when the students log off, it still needs to be collected from individual students.
- Data e-mailed automatically: it needs an internet connection and an e-mail with large capacity. This option is also not suitable for sensitive information as it poses security risks.

The researcher managed to obtain permission from the institution to access a shared drive and all information from all learners was written to this space. The researcher copied the files from the drive after each data collection session to a privately owned drive and deleted the ones on the shared drive.

4.7.3 Questionnaire

The questionnaire used was divided into three sections: background information, IT Skills and language background (see Appendix I). In the background section, participants were asked for their name, age, nationality, qualification and gender. In the IT skills section, there was a total of 16 questions and sub questions designed to check participants' knowledge of the basic IT skills needed to complete the tasks. In the language background, participants were asked about their L1, duration of exposure to L2 in the UK and years of instruction.

4.7.4 Observation

Observations were used to eliminate any problems in the software or those encountered by students during the experiment and to examine any issues that might relate to the outcome of the treatment.

4.7.5 Researcher's informal diary

I kept a diary with notes on all issues related to the experiment. The notes included problems with the design and how they were resolved, observation notes from the two pilot studies and the main study. The diary was used to refine the software packages after the pilot study. Most importantly, it was very useful when writing this section of the thesis.

4.8. Participants

All participants in the main study were ESL learners at a university language centre in the UK. They were all studying at the foundation stage. There were 41 males (% 44) and 52 females (%56). Their age range was 17-23 with a mean of 19. All of them had scored at least 5.5 in IELTS before joining the class. They had all taken the centre's own placement test. The centre released anonymised language placement scores and these ranged from 49 to 69 with a mean score of 58. The placement test included listening and writing sections only. Participants were from a range of different linguistic backgrounds, similar to the second pilot study. As there is no evidence that L1 will have an effect on learners' decision processes, this variable was not controlled for (although a further study that examines this will be interesting).

Data was collected between May and July of that academic year. Participants' and institute's consent was obtained before any data collection took place and participants were assured of the anonymity and confidentiality of all data collected. Also, they were told that they could withdraw from the experiment at any time without giving reasons (Appendix II).

Participants were informed that the whole experiment was meant to be extra-curricular work, that it would not be part of their course and that it was in no way going to affect the final outcome. This was necessary as in the briefing sessions that took place before the experiment, most learners expressed concern that the experiment could be used to assess their performance in the current course and consequently could affect their final grade. This type of reaction is echoed in Gass and Mackey (2005) where they argued that stress, anxiety and demotivation are affective factors that might influence the outcome of any experiment. By keeping the learners informed and reassuring them of their anonymity and confidentiality before each session, it was hoped that such variables were eliminated.

As required by their programme, the learners attended 20 hours of formal instruction per week, from 9-1.00 every day. They also had some afternoon classes depending on the degree they had chosen for their university study in humanities, arts, business, science or medical studies. They had all received three sessions on IT skills and two sessions on computer-related research skills.

The learners were randomly assigned to the three experimental groups in their intact classes, though not all learners were included in the final analysis. After the pre-test, some learners were identified as outliers and their data was removed from the sample. Some studies use the class level as a measure of proficiency (Thomas, 2006) to assign experimental groups. One problem with such an approach is that class level is usually based on the institution's placement test which might not target the linguistic structure under examination or might not be appropriate. Other studies used developmental stages (Philp, 2003; Mackey, 1999; Mackey and Philp, 1998). The advantage of the pre-test post-test approach is that it gives an up-to-date measure of the learners' linguistic ability with regards to the target construction. The present study used both approaches; first data was collected from learners in their intact classes as determined by the institution's placement test and their IELTS score, but then pre-test scores were used to eliminate any outliers, those who had scored at ceiling levels or a very low. This served to maintain homogeneous groups.

4.9. Procedures

Data used for the main study were collected from 93 L2 learners of English who were randomly assigned to one of the three different types of input groups: FoF, FoS and FoM in their intact classes. The target form was the construction of reported speech in English and the independent variable was the way the associated grammar was represented in each software package.

The software packages were uploaded to PCs in a computer cluster where learners usually had their classes. The intervention was integrated within the learners' timetable and took place during a regular slot in the learners' programme. The intervention was carried out in a very low-stakes atmosphere to lower affective filters. Learners were assured that teachers would not have access to their scores and that their participation would in no way affect the assessment for their programme of study. Learners were also told that they could withdraw from the experiment at any point. Their anonymity was assured and learners were not asked to provide their real names or university ID when logging into the programme. Once the learners were provided with the materials, there was no time limit, i.e. they could take as long as they wanted to complete the activities. Help was only offered when learners had technical problems, for example, if a student was not sure how to move between tasks.

4.10 Data Analysis Approaches

In traditional TBLL studies, analysis usually focuses on the learner's performance and interaction during the experiment. Interaction is studied by examining the oral conversation between learners while completing the tasks and recording the incidents where they attend to form or meaning. As there is not an oral outcome here, which is the common standard in most TELL materials, interaction was analysed by looking at the decisions and processes that learners exhibited while dealing with the input. Also, the focus of traditional studies that deal with input is the difference between accuracy and fluency. However, as there is not an oral element in most TELL materials and as the focus of the study is not on which skills are acquired best, the analysis focused on the difference between communicative performance and linguistic performance. An 'overall communicative performance score' was calculated by adding the scores of the all individual items in each of the test tasks together, then an overall score of the two tasks was calculated; see example below. This score is taken to represent task completion, a performance measure used in Task-Based Language Learning to evaluate success. In total, there were 25 items in the first task, 15 target items and 10 distractors. In the second task, there was a total of 14 items, 5 target and 9 distractors. Therefore, the maximum overall score for communicative performance is 39.

Task 1	**Task 2**	**Total**	
25 items	14 items	overall score	39
15 target items	5 target items	target score	20

The overall communicative performance score, however, does not indicate whether learners have attended to form or meaning or both when dealing with the input. A 'target linguistic performance score' was calculated by adding all the scores of the target items only in both tasks. The target score is taken to represent the learners' linguistic performance as learners needed to attend to form to provide the correct responses. Although the target linguistic performance score is a subset of the overall communicative performance score, it did not reflect the same performance as the overall score. In other words, the reasons for having the two scores is that it was important to distinguish between a learner (A) who scores 20 by getting all the target items correct and none of the meaning-focused items right (the non-target items) and another learner (B) who scores 20 by getting 15 of the meaning-focused items and 5 of the target items right. The difference between the two measures indicates that learner A might have attended to form while learner B might have attended mainly to meaning when completing the tasks. As we shall see in the following chapter, this distinction is crucial. Furthermore, one of the crucial assumptions in TBLL is that learners attend to meaning and tasks should therefore be meaning-focused. However, it is an empirical question as to whether learners attend to form while they are attending to meaning.

The scoring system used for the two test tasks was the following:

- Drag and Drop: each item was scored automatically as was explained in section 4.5. If the learner dragged the item to the correct gap, s/he scores 1, if not s/he scored 0. If the learner dragged the item more than once, the last score was used. The score of all items was then added together to get the overall communicative performance score. The scores for the 15 target items were added together to get the target linguistic performance score.
- Information gap: at first each item response was marked as correct and scores 1 if the meaning and form were correct. So, if the learner wrote '*lead*' instead of '*led*', this was scored as 0. The individual items' scores were then added to calculate the overall communicative performance score which was used to analyse the learners' communicative performance. A second stage of the analysis was calculating the target linguistic performance score by adding the score of the individual target items together. This target score was then used to examine the learners' linguistic performance.

Further analysis examined processing time, trial numbers during pre- and post-tests, access to audio files, access to grammar help, and task type, namely:

- Processing time was calculated by measuring the overall time learners spend on the experiment. For the FoS and FoF groups, the time spent on the grammar instruction whether before the treatment or during the treatment was deducted to maintain the validity of the measure.
- Trial numbers were only be calculated for the pre- and post-tests' tasks. This was done by recording all the attempts that learners made during the pre- and post-tests whether the responses were correct or not. A sum was calculated for the pre- and post-tests separately and the results were compared to examine treatment effect, input effect and other performance related variables.

- Access to audio files was measured by recording the number of times, learners decided to access the audio file during the pre- and post-tests. The cumulative sum for each test was calculated and results compared.
- Grammar access was measured for the FoF and FoS groups. For the FoF group, who only had access to grammar during the treatment, the number of times the grammar help was accessed, was summed for each learner. Also, the time spent on the grammar help was calculated by adding the time spent on each access. For the FoS group, in addition to the measures used for the FoF group, the time spent on the main grammar instruction was recorded and used later for within group comparisons.

Descriptive and inferential statistics were used to verify the hypotheses. For the quantitative analysis, out of the 93 learners, only 71 were used. Some learners were excluded as they had not completed all the tasks, or their data was messy or they scored so high or so low on the pre-test. The number of learners in each group was as follows; FoS (20), FoF (25), FoM (26).

In terms of the qualitative analysis, detailed analysis of one learner from each group was carried out by looking at the individual log files. The focus of the qualitative analysis was on identifying patterns of behaviour that were similar or different among the three learners.

4.11 Research Ethics

Consent from institutions was obtained prior to data collection. Consent from learners was also given before any data collection took place and learners were informed that they could withdraw from the experiment without giving reasons. No sensitive information was collected and no personal ID's or names were asked for. Confidentiality and anonymity were taken into consideration during the collection of data and the writing of the thesis.

4.12 Conclusion

Research methods are strongly dependent on context, theoretical framework and the questions under investigation (Mackey and Gass, 2012). This chapter described the data collection methods used in this study with explanations of why these methods were chosen. As stated in section 4.10, the choice of methods informed the data analysis approaches. In the next chapter, the analysis of the data will be presented and the results discussed.

Chapter Five

Data Analysis

5.1 Introduction

Data generated from interventionist quasi-experimental studies are usually quantitative in nature (Leowen and Philp, 2011) and this study is no exception in this regard. In the present thesis, quantifying the relationship between measures of performance at given points in time provides empirical information concerning the individual changes in learners' interlanguage production at those given points. Note again that the focus of the present study is not on learners' interlanguage competence regarding the target construction. Rather it is on their performance on this construction in relation to provision of three different input types. Hypotheses are stated and then statistical tests are used to verify or reject the hypotheses. The validity of hypotheses is tested by observation of the learning domain, then the hypotheses are confirmed if they are compatible with the observation or rejected if they are not.

In this chapter, the analysis is organized in sections according to the three main research questions, which are restated below for the reader's convenience. In each section, the hypothesis and the sub-hypotheses are stated then the results of tests using descriptive and inferential statistical tests are given.

Research Question One*: Which type of input, Focus on Forms (FoS), Focus on Form (FoF) and Focus on Meaning (FoM), is most effective in a Technology Enhanced Language Learning environment?*

Research Question Two*: (a) What factors or decision processes exhibited by learners while dealing with the different types of input contribute to differences in attainment? (b) How do these factors/processes map on to learners' performance?*

Research Question Three*: How does the use of user-behaviour tracking technologies (log files) help us explain the variability in performance among learners?*

5.2 Framework for the analysis

A mixed methods approach to data analysis was used. First, the hypotheses were tested using quantitative statistical tests. It has been argued that finding a significant p value is not enough; rather, p values are seen here as not an end, but just the beginning of subsequent analyses that can help researchers better understand their significant results (Brown 2011). Therefore, a detailed qualitative analysis of three learners' actions was used to verify the outcome of quantitative analysis.

In addition to quantitative methods, the present study used qualitative methods to reinforce and cross-validate the results. Brown (2011) claims that mixed methods research (among other issues) is now playing a big role in quantitative study and mixed methods research is also supported by researchers such as Chaudron (1986, 2000), Lazaraton (2000, 2005) and Brown (2004a). Such an approach is argued to provide the most informative, complete, balanced and useful research results (Burke Johnson et al., 2007, p.129).

In terms of analysis, the hypotheses given below are informed by these three questions: Is there a quantitative difference among group 1, 2 and 3/ FoS, FoF, and FoM in terms of production (group differences)? Is there a quantitative difference between the three groups in terms of their patterns of behaviour (group differences)? ; Are there qualitative differences among all learners in terms of their behaviour (individual differences)?

In light of this, the following seven variables were used to analyse the data:

1. Overall communicative performance scores in the pre- and post-tests
2. Target linguistic performance scores in the pre- and post –tests

As was explained in Chapter 4, the 'overall communicative performance score' was calculated by adding the scores of the all individual items in each of the test tasks together, then an overall score of the two tasks was calculated. On the other hand, the 'target linguistic performance score' was calculated by adding all the scores of the target items only in both tasks. The overall communicative performance score represents task completion while the target linguistic performance score is taken to represent the learners' linguistic performance as learners needed to attend to form to provide the correct responses. (See section 4.10 for more details).

In addition, these five variables were used in data analysis:

3. Time needed to complete the treatment (processing time)
4. Number of trials in the pre- and post-tests (confidence level)
5. Scores on individual tasks in the pre- and post-tests (task effect)
6. Accessing audio input (modality of input)
7. Accessing the grammar help (grammar effect)

Information on how variables 3-7 are measured will be detailed in the following sections.

5.3 Effectiveness of input in a TELL-environment

Research question 1: *Which type of input, FoS, FoF and FoM, is most effective in a TELL-based environment?*

The first research question is related to the effectiveness of input in a TELL environment. As was discussed in the previous chapters, it has been argued by many ISLA researchers that different types of input lead to different outcomes. Based on the literature and the discussion in the previous chapters, the following null hypothesis was generated for the first research question:

Null Hypothesis (H0) *There will be no differences in performance among the three groups.*

Alternative Hypothesis *There will be differences in performance among the three groups.*

The assumption is that the different types of input place different demands upon the learner's language system; this affects their performance and is reflected in learners' scores. An underlying assumption here is that the learners have homogeneous language experiences as they come from matched backgrounds. However, the further assumption is that they do not have homogenous cognitive systems in terms of learning styles. So any differences in attainment are due to either different types of input (see chapter two) or to decisions made by the learners as a result of the exposure to different types of input, i.e. their individually varying cognitive predispositions. Based on the literature review and the discussion in the previous chapters, the following sub-hypotheses, based on the alternative hypothesis above, are formulated and are concerned with the differences among the different input groups across all tasks (pre- and post-test tasks).

Hypothesis 1.a There will be differences among the three groups in their overall *communicative performance* score. The FoM group will outperform the other two groups.

Hypothesis 1.b There will be differences among the three groups in their target *linguistic performance* score. The FoS group will outperform the other two groups.

Hypothesis 1.c There will be differences among the three groups in terms of the processing time. Learners in the FoF group will take longer to process the input as they will need to focus on both form and meaning.

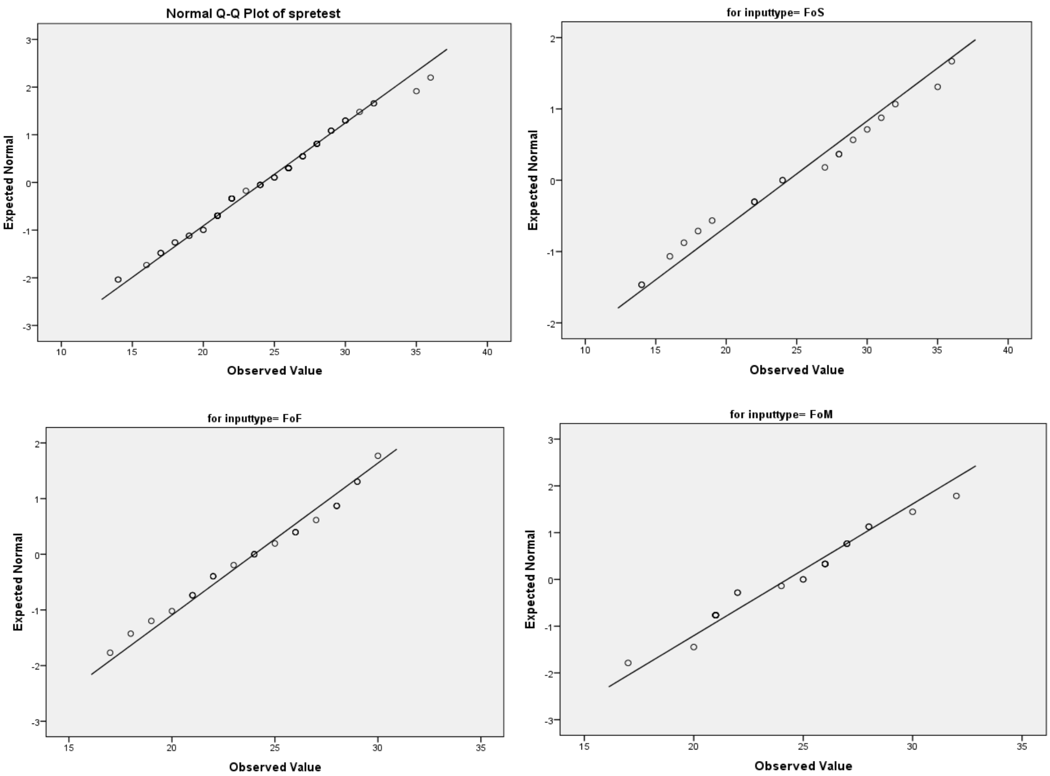
These sub-hypotheses are related to group differences rather than individual ones. In order to verify the alternative sub-hypotheses about the efficacy of the input, scores were calculated for all learners for each task. Statistical tests were then carried out to describe the results and look for any significant differences. The results of the analysis will be presented as follows: first normality tests were used to decide on the use of parametric or non-parametric tests. The descriptive results are then presented to show the comparison of mean scores across the input types. Then inferential statistical test results were used to check for any significant differences between time 1 and time 2 and among the three input groups. The independent variable for all the sub-hypotheses above is the type of input. The dependent variables are: overall communicative performance score, target linguistic performance score and processing time. Recall that overall communicative performance score is not a measure of accuracy; rather target linguistic performance score is. The results will be presented in the same order as the sub-hypotheses. This means that I will look first at the overall communicative performance score, which gives a comprehensive view of the learners' performance in regards to meaning and form. It also shows if learners within the three groups were able to complete the tasks and benefit from instruction to the same degree. I will then move on to look at the target linguistic performance score, which demonstrates whether the learners in the three groups were able to attend to form at the same level. I will finally look at whether learners varied with regards to the time they took to complete all the tasks and if the processing time affected their performance. Each section will begin by statistically confirming the normal distribution and equality of variance and then move to present the results.

5.3.1. Overall communicative performance score

An overall communicative performance score of each group for pre- and post-test is used here as it gives a comprehensive idea of the performance of each group. The pre-test score was calculated by adding the scores from tasks 1 and 2 in the pre-test. The same was done for the post-test. The maximum possible score is 39 in each test. The overall communicative performance score reflects task completion. As stated in the literature review, in Task-Based Language Learning (TBLL), success is measured by the completeness of the task, i.e. by *communicative performance*, rather than by accuracy or *meta/linguistic performance* (see Chapter Two and Chapter Four, in section 4.10). The analysis of the pre-test scores is presented first then the post-test analysis. The results will then be compared to explore any significant differences between the pre- and post-tests (treatment effect) and among the three groups (input effect). The dependent variables in this instance are the overall pre- and post-test scores, while the independent is the input type. (See Appendix III for full analysis.)

5.3.1.1. Pre-test overall communicative performance scores

Plot graphs for pre-test scores for all cases and for each group show that all cases cluster around the normal distribution line (Figure 5.0).

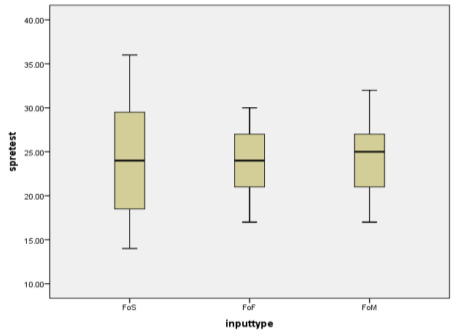


F

In order to verify the normality of distribution and decide on the appropriate statistical test, normal distribution tests for all cases and across all groups for the pre-test overall communicative performance scores were carried out. First, the Shapiro Wilk[17] test including all cases was run, and it revealed that the sample is normally distributed in relation to pre-test overall communicative performance scores, $p>0.56$.

Table 5.1 Shapiro-Wilk test: Pre-test overall communicative performance score

Test	Pre-test score		
Input type/Group	FoS	FoF	FoM
Sig	0.551	0.531	0.182



The Shapiro-Wilk test was also used to explore the normality of distribution across all groups before the use of inferential tests. The results indicated that the groups come from a normally distributed sample as all values are more than 0.05 (Table 5.1 above).

Mean and standard deviation for pre-test overall communicative performance score were calculated for each group (see Table 5.2 below). Results indicate that the three groups scored close to each other at the beginning of the intervention, with a mean pre-test score of 24.21 (SD 4.6).

Table 5.2 Mean and SD: pre-test overall communicative performance score

Groups	Pre-test overall communicative performance score
All groups	24.21 (4.6)
FoS (N=20)	24.404 (6.7)
FoF (N=25)	24.00 (3.6)
FoM (N=26)	24.26 (3.5)

17 The Kolmogorov-Smirnov test, which could have been used, is not sensitive to problems in the tails and it works reasonably well with data sets < 50. The Shapiro-Wilk test does not work well if several values in the data set are the same; it works best for data sets between 50-2000. However, it can be used with smaller data sets. It is, therefore, the better of the two SPSS tests for the present study. Thus, the Shapiro-Wilk test was used as it is more powerful and the sample size is acceptable.

To check whether there were any differences among the three groups at pre-test for overall communicative performance score, an ANOVA test was used. The ANOVA test showed no significant differences among the three groups, $F(2, 68) = .04$, $p = .9$. However, the homogeneity of variance test was used to check for equality of variances, an assumption needed for the ANOVA test, and it was significant ($P < .001$). This means that the equal variances hypothesis was violated; therefore, non-parametric tests were used to check for variances among the three groups. The Kruskal Wallis test was used, and it showed that there are no significant differences among the three groups in their pre-test score, $H(2) = .10$, $p > 0.9$. This means that all groups were at the same level at the beginning of the experiment and any differences in their post-test scores are not the result of their pre-test scores. The assumption, based on participants' language learning experiences, that the three groups are homogenous, is supported.

5.3.1.2. Post-test overall communicative performance score

The same procedure was carried out to check for normal distribution across all cases and treatment groups in regards to the post test overall communicative performance score. Plot graphs for post-test scores showed that all cases are close to the normal distribution line (Fig. 5.2).

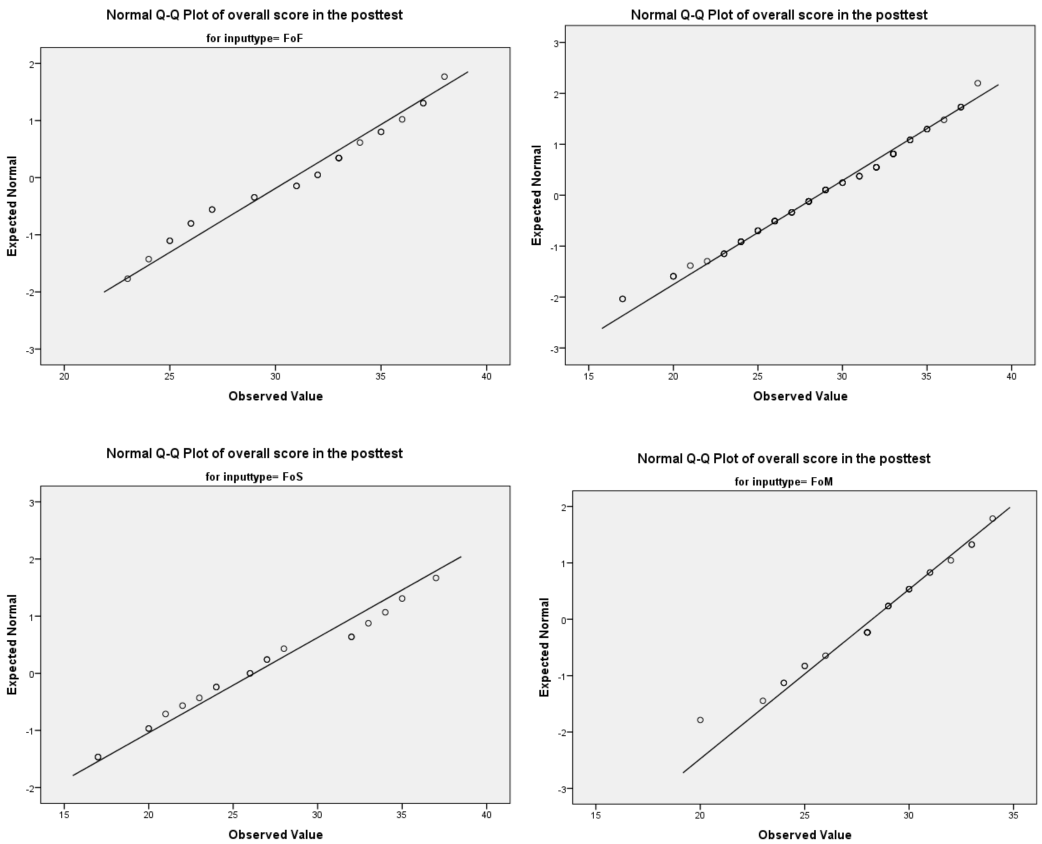


Figure 5.2: Normal distribution plot for all groups for pre- and post-tests

The Shapiro Wilk test including all cases revealed that the sample is normally distributed in relation to post-test overall communicative performance scores, $p > .39$. The results also indicated that all treatment groups come from a normally-distributed sample (Table 5.3).

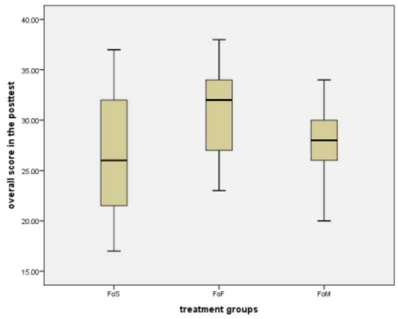
Table 5.3 Shapiro-Wilk test: post-test overall communicative performance score

Test	**Post test score**		
Input type/Group	FoS	FoF	FoM
Sig	0.49	0.21	0.44

Mean and standard deviation for overall communicative performance score were calculated for post-test for each group (see Table 5.4 below). Overall, all groups scored higher at the post test with a mean score of 28.94 (SD 4.8). However, the groups varied in their mean score: the FoF group outperformed the other two groups (M = 30.84, SD = 4.4). In comparison, the FoM group scored (M = 28.59, SD = 4.8) and FoS scored (M = 26.25, SD = 5.9).

Table 5.4: Mean and SD: post-test overall communicative performance score

Groups	**Post-test overall communicative performance score**
All groups	28.59 (4.8)
FoS (N=20)	26.25 (5.9)
FoF (N=25)	30.84 (4.4)
FoM (N=26)	28.59 (4.8)



To check whether the differences in mean overall communicative performance score were significant, an ANOVA test was used. ANOVA revealed that the differences among the groups are highly significant, $F(2, 68) = 5.65$, $p < .005$; however, the homogeneity of variance test was significant (p=0.009), indicating violation of the equal variances hypothesis and pointing to the use of non-parametric tests. Although equal variances among the three groups were not expected after the treatment but because it is an assumption of the use of ANOVA, it was decided to use the Kruskal Wallis test. Kruskal Wallis showed that there are significant differences among the three groups in regards to their post test score $H(2)= 8.56$, $p < 0.01$.[18] To check for which groups the

[18] Significant at the %99 level.

scores were significantly different (input effect), a post-hoc test was used. The Games-Howell post-hoc test was used and it confirmed that the differences in scores shown above were significant: the FoF group was significantly different from the FoS ($p<.02$) and FoM group ($p< .05$). The FoS was not different from the FoM ($p>0.3$) (see Appendix III).

Analyses thus far show that the three groups started at the same level but were different after the intervention. The results so far, demonstrate that the FoF group outperformed the FoM and the FoS groups with regards to overall communicative performance score.

5.3.1.3. Overall communicative performance score: treatment effect

Statistical tests showed that there was improvement over the period of intervention across all groups. The Kruskal Wallis results only showed that the group differed from each other at post –test but do not indicate whether the post-test scores were significantly different from the pre-test scores (treatment effect). To check whether the difference between pre- and post-test scores was a result of the intervention, a t-test was used. The t-test results indicated that on average, learners scored 4.38 points higher in the post-test and the difference between pre-test and post-test overall communicative performance scores was highly significant, $t(70) = 12.15$, $p < .001$. The results also indicated that there was a highly significant positive correlation between pre- and post-test scores $r(71) = .8$, $p < .001$. Thus, the learners who scored higher in the pre-test scored higher in the post-test. The correlation here does not indicate that there were differences among the groups in the pre-test scores, it just reveals that within each group, learners who scored slightly higher in the pre-test also scored higher in the post-test.

To sum up the results in relation to the learners' communicative performance,

Treatment effect: statistically highly significant. All groups improved after the treatment.

Input Effect: The FoF group significantly outperformed the FoM and FoS groups.

Although the FoM group scored higher in the post-test than did the FoS group, the difference was not statistically significant. This means that learners whose input included meaning-focused tasks but were advised that they could attend to form if they wanted to (FoF) outperformed those who were asked to focus only on meaning (FoM) or those who were explicitly asked to focus on form and meaning (FoS).

5.3.2. Target linguistic performance score

In addition to looking at the groups' overall communicative performance score which shows how well the learners completed the tasks, the target linguistic performance score was used to give a better idea of the learners' accuracy in relation to the target structure. Their linguistic performance in turn is an indication of whether they attended to the linguistic characteristics of the reported speech construction to form while they were doing the task. The dependent variable in this instance is the target linguistic performance score in pre- and post-tests while the independent is the input type. It was predicted that the groups that had access to formal instruction (FoF and FoS) would outperform the group which focused on meaning only (FoM) in their accuracy (target linguistic performance score) in particular. The underlying assumption is that while the FoF and FoS groups complete the task, they will attend to form as well as meaning. (See Appendix IV for the full analysis.)

5.3.2.1. Pre-test target linguistic performance score

The same procedure used for overall communicative performance score were used and plot graphs for pre-test target linguistic performance scores for each group showed that all cases were very close to the normal distribution line (Figure 5.3).

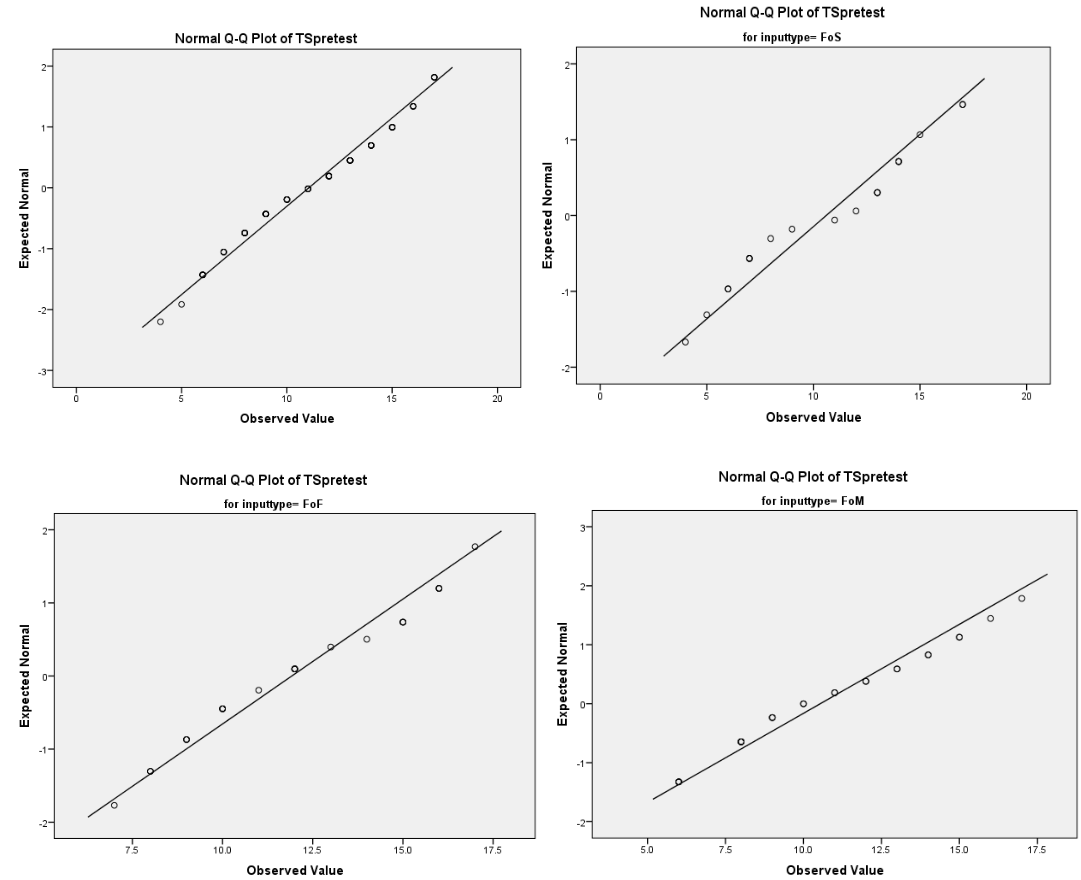


Figure 5.3 Normal distribution plots for all groups for pre- test target linguistic performance scores

The Shapiro Wilk test revealed that the sample was not normally distributed as a whole, $p<.03$. However, running the same test for individual groups showed that all groups were normally distributed. (Table 5.5)

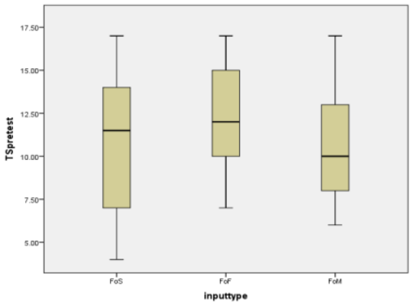
Table 5.5 Shapiro-Wilk test: pre-test target linguistic performance score

Test	**Pre-test target linguistic performance score**		
Input type	FoS	FoF	FoM
Sig	0.136	0.181	0.143

Mean and standard deviation for pre-test target linguistic performance score were calculated for each group. (See Table 5.6 below.) Results indicated that the three groups scored close to each other at the beginning of the intervention with a mean pre-test target linguistic performance score of 11.04 and SD of 3.44.

Table 5.6 Mean and SD: pre-test target linguistic performance score

Groups	Pre-test target linguistic performance score
All groups	11.04 (3.44)
FoS (N=20)	10.60 (4.10)
FoF (N=25)	11.92 (2.92)
FoM (N=26)	10.53(3.31)



In checking if the small differences among the three groups were significant, an ANOVA test was used and it revealed that the differences among the groups were not significant, F (2, 68) = 1.26, p >.2, but the homogeneity of variance test was significant (P< 0.04) so the Kruskal Wallis test was used. It showed that there are no significant differences among the three groups in their pre-test score H (2) = 2.52, p >0.2. This means that the three groups, FoF, FoM and FoS, attended to form at the same level at the beginning of the experiment.

5.3.2.2. Post-test target linguistic performance score

Plot graphs for post-test target linguistic performance scores showed normal distribution line (Fig 5. 4).

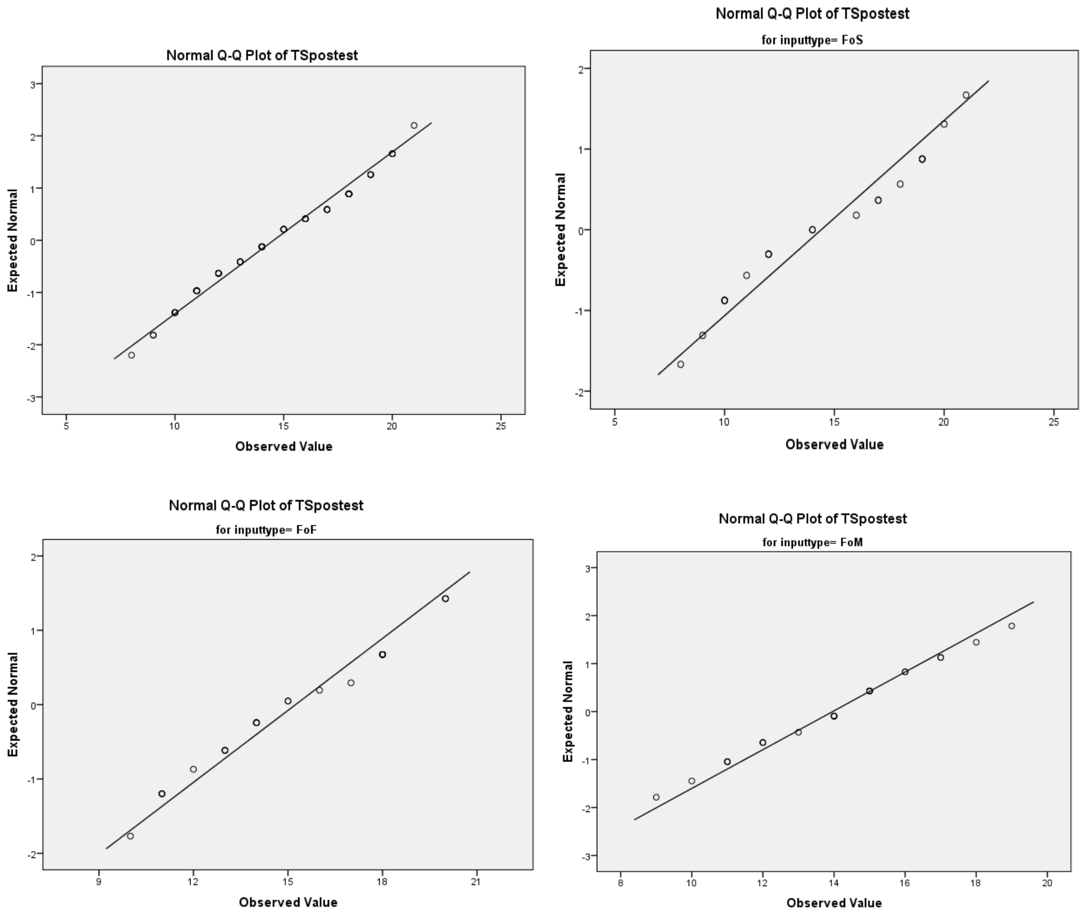


Figure 5.4 Normal distribution plots for all groups for post-test target linguistic performance score

Shapiro-Wilk confirmed the normal distribution, *p* >.06 of the whole sample. The results also indicated that all treatment groups were also normally distributed (Table 5.7).

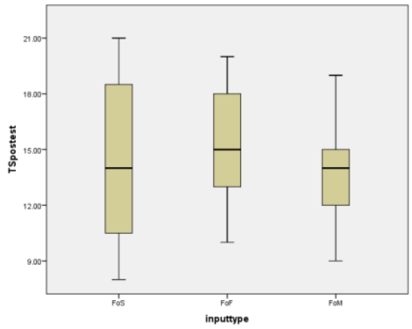
Table 5.7 Shapiro-Wilk test: post-test target linguistic performance score

Test	**Post-test target linguistic performance score**		
Input type	FoS	FoF	FoM
Sig	0.13	0.09	0.75

Mean and standard deviation for post-test target linguistic performance score were calculated for each group (See table 5.8 below). Overall, all groups scored higher at post-test with a mean score of 14.53 (SD 3.22). However, the groups varied in their mean score, with the FoF group scoring highest (*M* = 15.24, *SD* = 3.09) followed by the FoS (*M* = 14.40, *SD* = 4.13) and the FoM group came last (*M* = 13.96, *SD* = 2.47)

Table 5.8 Mean and SD: post-test target linguistic performance score

Groups	Post-test target linguistic performance score
All groups	14.53 (3.22)
FoS (N= 20)	14.40 (4.13)
FoF (N= 25)	15.24 (3.09)
FoM (N= 26)	13.96 (2.47)



ANOVA revealed that the differences among the groups were not statistically significant, $F(2, 68) = 1.02$, $p > 0.3$, but as the homogeneity of variance test was significant ($p<0.002$), non-parametric tests were used instead. The Kruskal Wallis test showed that there are not significant differences among the three groups in regards to their post-test target linguistic performance scores $H(2)= 1.71$, $p > 0.4$. There was no need to carry out a post-hoc test. Thus, although all the groups improved after the treatment, there were no differences across the FoF, FoS and FoM groups' improvement, i.e. no input effect. This means that three groups attended to form at the same level before and after the treatment, and the higher scores by the FoF and FoS groups are not related to the type of input they received.

A further analysis looked at the scores of individual target items in the post-test across all learners. Each item score was calculated by adding the score of all learners on that item so the maximum possible score for each item was 71 (number of participants). Looking at the scores, it was obvious that most learners struggled with the past perfect tense form, as their scores on these items were low. The two exceptions were the one where the subject was provided: *we'd gone* and *had promised* which they could have inferred from the context. Their scores were better on the past tense form, as can be seen in Table 5.9 below. In addition, comparing the forms across the two tasks indicates that the learners scored better in the drag and drop task than they did in the information gap task.

Table 5.9 Individual items' scores

Task	Score (out of 71)	Item	Form
Post-Test task 1	22	Had been	Past perfect
	31	Had gone	Past perfect
	33	Was going	Past continuous
	35	Had	Past had
	37	Would go	*Will* past
	40	Had ever had	Past perfect
	46	Had all been	Past perfect
	48	Were to blame	*Be* past
	51	Was	*Be* past
	53	Would	*Will* past
	53	We'd gone	Past perfect
	57	Knew	Past
	58	Had no	*Have* past
	59	Wasn't	*Be* past negative
	61	Had promised	*Have* past perfect
Post-Test task 2	25	Would have	*Will* have past
	32	Had	*Have* past
	35	Was	*Be* past
	41	Were	*Be* past
	43	Led	Past

5.3.2.3. Target linguistic performance score: treatment effect

Thus, statistical tests showed that the groups improved between pre- and post-tests in relation to their target linguistic performance score but did not indicate whether this was the effect of the treatment effect. So to check whether the difference between pre- and post-test scores was a result of the intervention (treatment effect), a t-test was used. The results indicated that on average learners scored 3.49 points higher in the post-test and this improvement was not due to chance, $t(70) = 12.80$, $p < .001$. The results also indicate that there is a highly significant positive correlation between pre- and post-test target linguistic performance scores $r(71) = .7$, $p < .001$. Thus, the learners who scored higher in the pre-test scored higher in the post-test as was the case with the overall communicative performance scores. So, this can be interpreted as all groups attended to form more after the treatment but they all did so at the same level. Individual learners, however, who attended to form more during the pre-test continued to do so during the post-test. It was not possible then to run individual items' analysis with respect to target forms (e.g. past vs. past perfect) across the groups as the test results above showed no significant difference among the three groups in relation to the target items.

To sum up the results in relation to learners' target linguistic performance,

Treatment effect: statistically highly significant. All learners were able to attend to form better after the treatment.

Input Effect: not significant, so regardless of which group they belonged to, learners attended to form at the same level.

So far, the results indicate that the treatment had an effect on both the communicative and linguistic performance of the learners. Moreover, type of input affected the learners' communicative performance but not linguistic performance. FoF learners were able to complete the tasks more successfully followed by FoM and finally FoS. All learners, however, were equally able to attend to form better after the treatment.
In order to account for the possibility of overestimating the significance of the statistical tests which results from running multiple ANOVA tests, a Mixed ANOVA analysis was conducted and the results are reported in the following section.

5.3.3. Mixed ANOVA: within and between group analysis

A typical repeated measures analysis was not suitable for the design of the present study as it involves measuring the same group over more than two points of time. This would be, for example, a group of learners who received pre-, post- and delayed post-tests. In the present study, there were three experimental groups which were tested twice; therefore, a Mixed ANOVA test was more appropriate. A Mixed ANOVA is a mixture of between-groups and repeated measures variables (Field, 2013). The main difference between a typical repeated measures ANOVA and a mixed repeated measures ANOVA is that the latter allows for between-group analysis. There is one within-subjects variable with two levels, (pre- and post-tests) and one between-subject variable with three levels (FoF, FoM and FoS). As can be seen from the plot chart below (Figure 5.5), all groups improved between the pre- and post-tests with the FoF group making the most improvement.

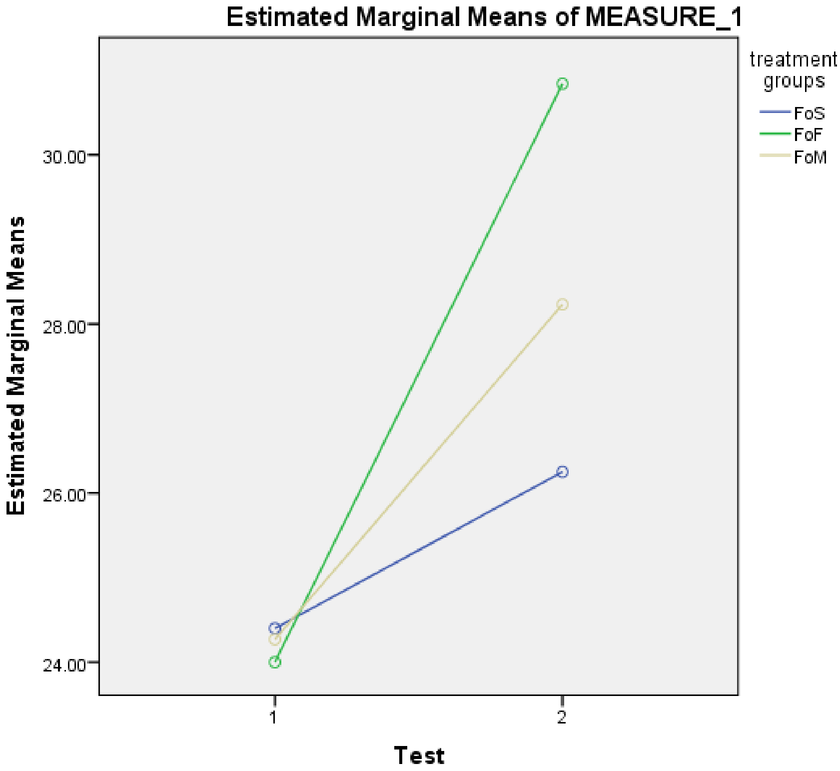


Figure 5.5 Plot graph Mixed repeated measures ANOVA: overall communicative performance score

A Mixed ANOVA test reveals a significant effect of the treatment F(1,68)= 260.8 , p <, 0.001. This confirms the results of the separate ANOVA tests. The test also showed that there was a significant relationship between the type of input and the learners' overall communicative performance score, F(2,68)= 29.7 , p <, 0.001. The results indicate a highly significant association between the type of input the learners received and how they performed in the post-test (see Appendix V for the full analysis tables).

The same Mixed ANOVA test was conducted to identify the significance of within and between subject differences with regard to the target linguistic performance score.

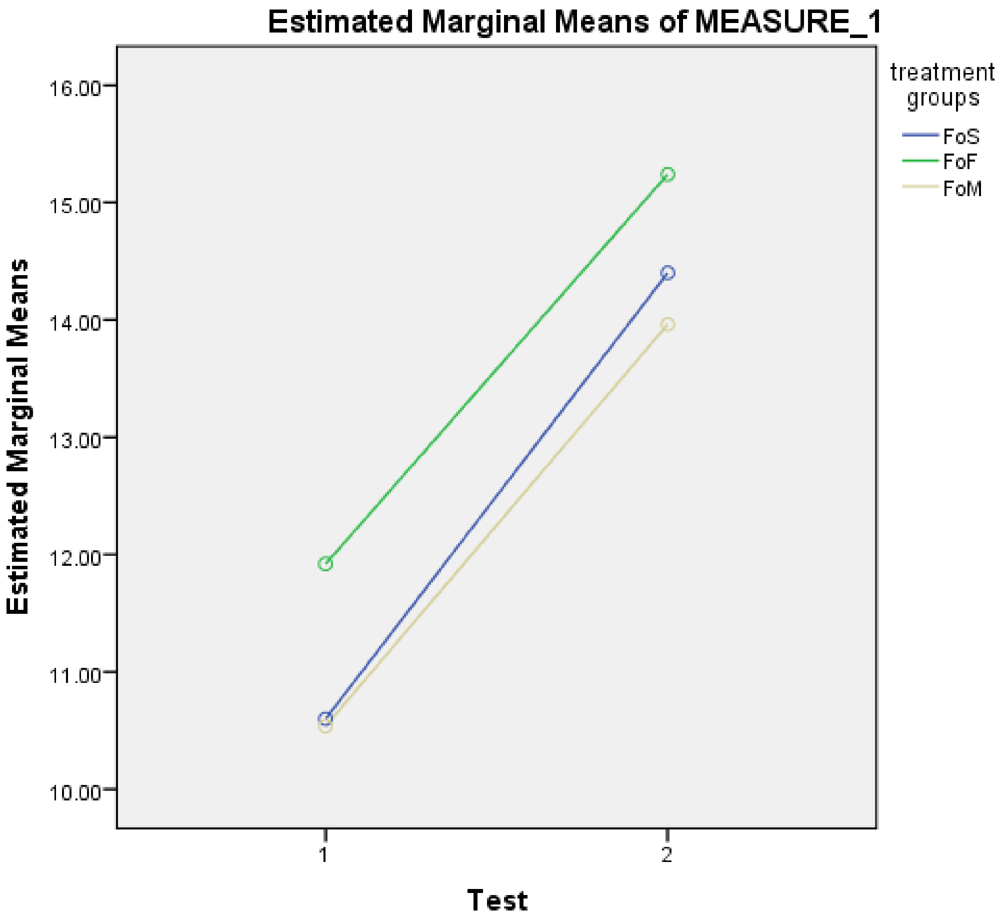


Figure 5.6 Plot graph Mixed repeated measures ANOVA: target linguistic performance score

Looking at the plot graph, it is obvious from the slopes that the three groups improved between the pre- and post-tests; however, the slopes indicate that all groups improved more or less in the same way. This observation was confirmed by the test output. The analysis revealed that the improvement between pre- and post-tests was highly significant, F(1,68)= 160.33 , p <, 0.001. However, the differences among the three groups were not significant, F(2,68)= 0.25 , p =, 0.77. The Mixed ANOVA results confirm what has been found earlier: that all groups improved regardless of the type of input. In relation to target linguistic performance, input type did not affect the scores. (See Appendix V for the full analysis tables.)

I will now move to look at the processing time and whether this had an effect on the learners' communicative and linguistic performance and also whether the type of input had an effect on the time taken to complete the tasks.

5.3.4. Processing time

This factor is more closely related to language development where it is assumed that the different types of input will result in variation in terms of the processing load. This is reflected in the time learners spend doing the tasks. The overall time spent on the software was used for analysis. This is due to the fact that the pre-test and post-test were short and immediate so no considerable variation was predicted among the groups in their processing time. However, the intervention varied and therefore learners were expected to spend more time processing the input if they had attended to form and meaning simultaneously. (See Appendix V for full analysis.)

In order to calculate the accurate amount of time spent processing the target input, the following was used:

> FoM: total time spent on the software.
>
> FoF: total time spent on the software minus the time spent on the grammar help. Time spent on the grammar help was calculated by recording the time the grammar help button was pressed and adding these times together for each learner.
>
> FoS: total time spent on the software minus the time spent on the formal grammar instruction and grammar help; this was calculated as above for FoF.

5.3.4.1 Processing time: mean and SD

Plot graphs for the time variable across all groups showed that the values are very close to the normal distribution line, with one exception. Further analysis of the time variable according to the input type showed that normal distribution was maintained within the groups as well.

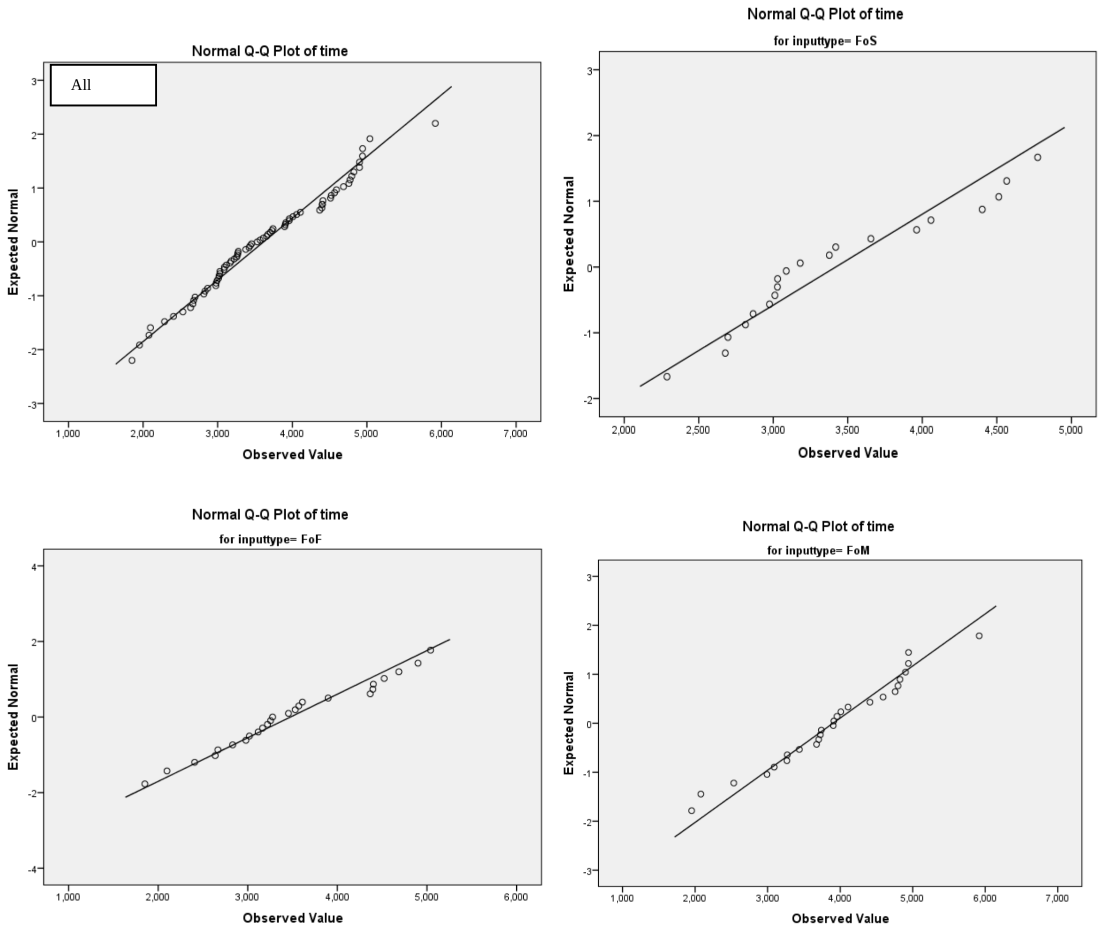


Figure.5.7 Normality distribution: processing time

Shapiro Wilk indicated that all the groups were normally distributed in regards to the time processing variable as $p < 0.05$ (see table 5.9 below).

Table 5.10 Shapiro Wilk test result for the time variable

Input type	**FoS**	**FoF**	**FoM**
Sig.	0.09	0.57	0.68

Learners varied on how long they took to complete the session. Overall, time spent on the software ranged between 1851 and 5917 seconds with a mean of 3615 seconds and SD of 873.56 seconds. On average, FoM spent more time (*M* = 3901.23, *SD* = 939.23) followed by FoF (*M* = 3475.68, *SD* = 866.81) and FoS (*M* = 3418.8, *SD* = 723.19).

Table 5.11 Processing time descriptives

Processing time	N	Mean	Std. Dev	Min	Max
FoS	20	3418.8	723.19	2287	4774
FoF	25	3475.68	866.81	1851	5040
FoM	26	3901.23	939.41	1952	5917
Total	71	3615.49	873.56	1851	5917

This means learners who were required to attend only to meaning spent a longer time processing the input while those who were required to attend explicitly to form and meaning spent less time.

5.3.4.2 Processing time: input effect

To check whether the differences in mean time spent processing the input were significant among the groups, inferential statistics were used. The test of homogeneity of variance was insignificant so one way ANOVA was used. The results showed no significant differences among the groups $F(2, 68) = 2.30$, $p > 0.1$. Thus, although the mean time spent by each group was different, these differences were not due to the input type. I will now look at whether the groups differed in their processing time in relation to the sub-categories of performance.

5.3.4.3 Processing time: overall communicative performance score

First, to check whether the processing time had an effect on the overall communicative performance score, a bivariate correlation is used. Pearson correlation revealed that there was a highly significant negative weak correlation between processing time and overall communicative performance score, $r(71) = 0.3$, $p < .001$. This means that across the whole sample, those learners who scored higher in their post-test overall communicative performance score spent less time processing the input.

5.3.4.4 Processing time: target linguistic performance score

Second, processing time was tested against target linguistic performance score to explore any effects, and a bivariate correlation is used. Pearson correlation revealed that there was a highly significant negative weak correlation between processing time and target linguistic performance score, $r(71) = 0.3$, $p < .001$. This means that learners who scored higher in their post-test target linguistic score spent less time processing the input. In other words, across the whole sample learners who attended to the linguistic characteristics of reported speech during their post-test were likely to spend less time processing the input.

To sum up the results in relation to processing time, the following could be identified:

> **Input Effect**: not significant, so the type of input the learners received did not influence the time they spent processing the input.
>
> **Performance effect**: highly significant association between processing time and both overall communicative score and target linguistic performance score. Learners who spent less time processing input scored higher on both measures.

The results so far indicate that the type of input learners received did not affect how long they spent processing the input. However, individual learners who were able to attend to both meaning and the linguistic characteristics of reported speech spent less time processing the input.

5.4 Factors that contribute to differences in attainment

Research question two: *What factors or decision processes exhibited by the learners while dealing with the different types of input contribute to differences in attainment? How do these factors/processes map on to performance?*

The second research question is related to the aspects of treatment that affect attainment in a TELL-based environment for each of the groups. It is expected that because learners in different groups are exposed to different types of input, they will deal with the input differently and this will be reflected in their decisions. Also, the patterns of behaviour in terms of what the learners decide to do or access during the experiment is predicted to have an effect on their performance. Based on the literature and the discussion in the previous chapters, the following null hypothesis was generated for the second research question;

Null Hypothesis (H0) *There will be no differences among the three groups on how they deal with the input.*

Alternative Hypothesis *There will be differences among the three groups on how they deal with the input.*

The assumption is that the different types of input will place different demands on the learners' cognitive systems; they will respond differently. It is not assumed here that all the learners share the same cognitive system, rather as a consequence of the input, their systems will respond in the same way. The hypothesis is concerned with how learners react upon perception of pressing communicative demands. The following sub-hypotheses were formulated and are concerned with the factors and decision processes that might be different among the three input groups:

Hypothesis 2.a There will be differences among the three groups in the number of trials.

Hypothesis 2.b There will be differences among the three groups in terms of task types.

Hypothesis 2.c There will be differences among the three groups modality of input.

Hypothesis 2.d There will be differences among the three groups in terms of access to grammar help.

As was explained in Chapter Three, section 3.7, the sub hypotheses were formulated in a non-directional way at this stage as the evidence from previous research either does not exist (e.g. trial number) or is inconclusive (e.g. task types) to support a direction-specific hypotheses. To answer the second research question, information was collected about all the decisions learners took when dealing with the input (see Chapter Four about data collection methods). It was predicted that learners using different types of input would exhibit different patterns of processing or decision making.

5.4.1 Trial numbers

Trial numbers will reflect the processing patterns of learners. If the learners went through many trials, it means they were less confident about their answers. First, the analysis aimed at looking for patterns, and then these patterns were compared among treatment groups using the performance measures already used above. (See Appendix VI.)

5.4.1.1 Trial numbers: pre-test

Plot graphs for trial numbers during the pre-test for all cases and for each group showed that all cases cluster around the normal distribution line (Figure 5.8).

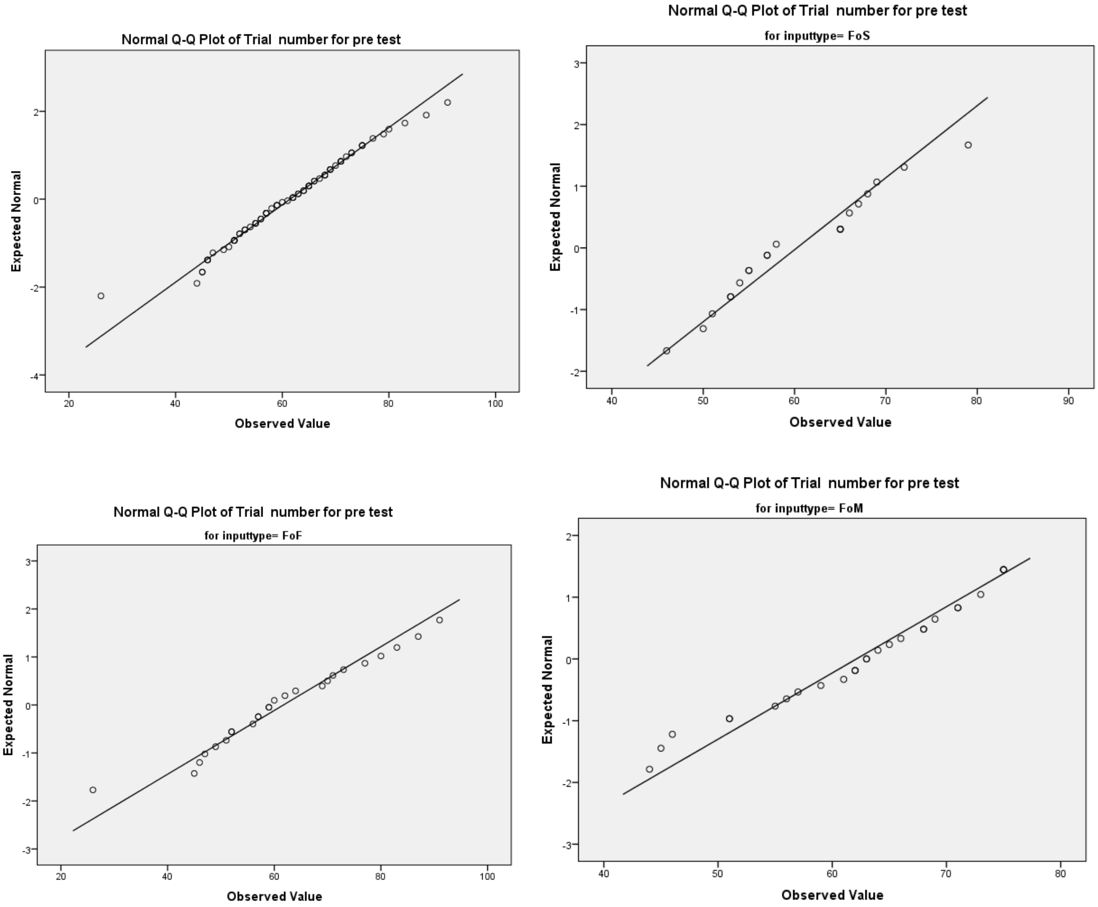


Figure 5.8: Normal distribution plot: trial number pre-test

The Shapiro-Wilk test including all cases revealed that the trial number variable was normally distributed $p > 0.65$. The results also indicated that all individual groups were normally distributed in their trial numbers (Table 5.12).

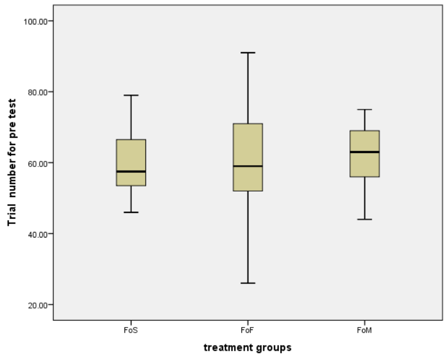
Table 5.12 Shapiro-Wilk test: trial no. pre-test

Test	**Trial number pre-test**		
Input type	FoS	FoF	FoM
Sig	0.462	0.792	0.186

Mean and standard deviation for trial numbers during the pre-test were calculated for each group (See Table 5.13 below). Results indicated that the three groups had the same number of trials during the pre-test at the beginning of the intervention with a mean trial number of 61.4, *SD*=11.3. On average, the FoS learners had fewer trials (*M* = 60.2, *SD* = 8.5), then the FoF group (*M* = 61.7, *SD* = 15), and FoM learners had the highest number of trials (*M* = 62.1, *SD* = 9.3).

Table 5.13 Mean and SD: trial number pre-test

Groups	**Trial number pre-test**
All groups	61.4 (11.3)
FoS (N=20)	60.2 (8.5)
FoF (N=25)	61.7 (15)
FoM (N=26)	62.1 (9.3)



ANOVA was used to check if the differences in trial numbers among the three groups were significant, and it showed no significant differences among the three groups, $F(2, 68) = .15$, $p = .8$. However, the homogeneity of variance test was significant ($p < 0.03$) which indicated that a non-parametric tests should be used. Kruskal Wallis showed that there are no significant differences among the three groups in their pre-test scores with respect to number of trials on the pre-test, $H(2)= .52$, $p > 0.7$. This means that all learners had pretty much the same number of trials during the pre-test.

5.4.1.2 Trial numbers: post –test

The plot graph for trial numbers during the post-test for all cases showed normal distribution. In addition, individual plot graphs showed that all groups were normally distributed (Figure 5.9).

Table 5.14 Shapiro-Wilk test: trial no. post-test

Test	**Trial number post-test**		
Input type	FoS	FoF	FoM
Sig	0.531	0.123	0.092

These observations were confirmed by Shapiro-Wilk test, $p > 0.17$. Also, the results of the test for individual groups indicated that all groups were normally distributed (Table 5.14).

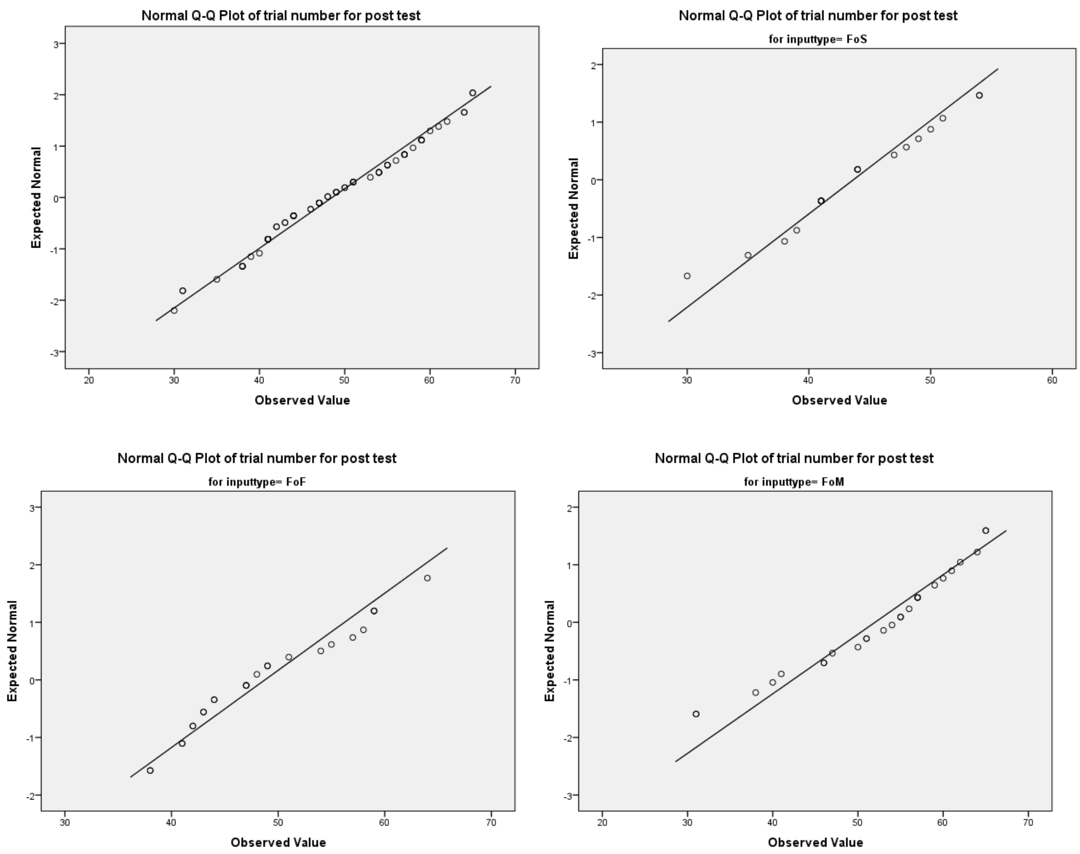
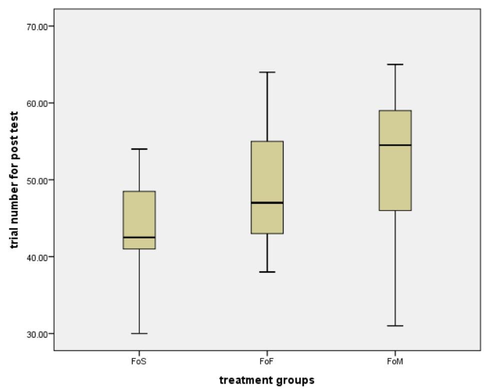


Figure 5.9: Normal distribution plot: trial number post-test

Mean and standard deviation for trial numbers during the post-test were calculated for each group (see table 5.14 below). Results indicated that the three groups had different numbers of trials during the post-test with a mean trial number of 48.5, *SD*=8.6. The FoS learners had fewer trials (*M* = 43.6, *SD* = 6.1), then the FoF group (*M* = 48.7, *SD* = 7.4) and FoM learners had the highest number of trials (*M* = 52, *SD* = 9.6).

Table 5.15 Mean and SD: trial number post-test

Groups	**Trial number post-test**
All groups	48.5 (9.02)
FoS (N=20)	43.65 (6.1)
FoF (N=25)	48.7 (7.4)
FoM (N=26)	52(9.6)



5.4.1.3 Trial numbers: input effect

ANOVA was used to check if the differences in trial numbers among the three groups were significant and it showed that there was a highly significant association among the three groups in relation to the number of trials during the post-test, $F(2, 68) = 6.11$, $p < 0.004$. This means that at the beginning of the experiment there was no difference in the number of trials, but there were significant differences among the three groups in their trial number during the post–test, after being exposed to different types of input. Post hoc tests were used to explore the nature of these differences. The Tukey HSD test was used and it showed that there was a highly significant difference between the FoS and FoM group and a significant difference between the FoS and FoF. However, the FoF and FoM group were not different (Table 5.15).

Table 5.16 Tukey HSD significance: trial number post-test vs treatment groups

Treatment groups	**FoS**		**FoF**		**FoM**	
	FoF	FoM	FoS	FoM	FoS	FoF
Sig	.094**	.002*	.094**	0.327	0.002*	0.327

*significant at the 99% level, **significant at the 90% level

So far, the results indicate that the learners who were required to explicitly focus on form and meaning during the treatment had significantly fewer trials than those who focused on meaning only or those who had the choice of focusing on form.

5.4.1.4 Trial numbers: treatment effect

To check if the treatment had any effect on the trial numbers, a paired samples T-test was used to compare trial numbers between the pre- and post-tests. The results showed that on average the learners had 12.9 fewer trials in the post-test and that there was a highly significant association between the trial numbers during the pre- and post-tests, $t(70) = 8.55$, $p < .001$. Pearson correlation also indicated that there was a significant (at the 90% level) association between pre- and post-tests trial number, $r(71) = .2$, $p < .08$. This means that all learners had fewer trials after the treatment and that some of the learners who had fewer trials in the pre-test continued to do so in the post-test.

5.4.1.5 Trial number: communicative performance score

To check if the number of trials during the pre- and post-tests had affected the learners' overall communicative performance scores, Pearson correlation was used. Correlating the trial number during the pre-test with these scores revealed that there was no association between the two, $r(71) = -0.1$, $p > .2$. Similarly, correlating the trial number during the post-test with overall communicative performance score showed no effect, $r(71) = .05$, $p > .6$.

5.4.1.6 Trial number: target linguistic performance score

To check if the number of trials during the pre- and post-tests had affected the learners' target linguistic performance scores, Pearson correlation was used. Correlating the trial number during the pre-test with these scores revealed that there was a weak negative association between the two, $r(71) = -0.2$, $p < .06$. On the other hand, correlating the trial number during the post-test with target linguistic performance score showed no effect, $r(71) = .04$, $p > 0.7$. This means that learners who had fewer trials during the pre-test scored higher in the post-test in relation to accuracy on the target structure.

To sum up the results in terms of the number of trials, the following can be stated:

> **Input effect**: There were significant differences between the FoS and FoM group and the FoS and FoF. However, the FoF and FoM group were not statistically different. This means learners who were required to explicitly focus on the linguistic characteristics of reported speech and meaning during the treatment were more confident about their decisions and had significantly fewer trials than those who focused on meaning only or those who had the choice of focusing on linguistic characteristics.
>
> **Treatment effect**: There was a highly significant association between the trial numbers during the pre- and post-tests, with learners going through fewer trials in the post-test.
>
> **Performance effect**: there was a significant association between the number of trials during the pre-test and target linguistic performance score. This indicates that learners who were more confident in their pre-test, i.e. had fewer trials (not necessarily those who scored higher in the pre-test), scored higher in the post-test.

5.4.2 Task type

The effect of task will be dealt with in detail in this section. There were two task types used in the tests: drag and drop and information gap. In the pre-test, the learners had to do the drag and drop task first and then the information gap. The order was reversed in the post-test as discussed in Chapter Four. Below, a descriptive summary of each task will first be presented and then the results of a comparison between the two tasks in relation to other variables will be presented. The descriptive statistics were based on the actual values of the task-related variables. However, in order to make direct comparisons between the two tasks and draw conclusions, the individual task variables were standardised. Z-scores,[19] on each task, overall communicative performance score) and target linguistic performance score, were used instead of raw scores. Also, z-trial numbers on each task were used instead of the actual number of trials. This procedure eliminated any variable effect as the two tasks had different distribution, namely: drag and drop (25 items) and information-gap (14 items). (See Appendix VII for full analysis and descriptive summary of the z values for all variables.)

5.4.2.1 Task type: treatment effect

As can be seen from table below, overall communicative and target linguistic performance improved across all groups in both tasks as the mean scores in the post-test were higher. Also, learners had fewer trials in the post-test in both tasks.

Table 5.17: Task type: mean and SD, pre- and post-tests

Task	**Drag and drop**						**Information gap**					
	Com		**Ling**		**Trial**		**Com**		**Ling**		**Trial**	
Test	**post**	**pre**	**post**	**pre**	**post**	**pre**	**post**	**pre**	**post**	**pre**	**post**	**pre**
Mean	19.45	16.70	12.66	9.99	30.61	39.71	9.14	7.51	1.87	1.06	17.88	21.73
SD	3.52	3.28	2.54	3.02	6.5	8.4	2.36	2.37	1.54	1.12	5.5	7.9

In order to check if these differences between time pre- and post-tests were significant across tasks, each task's variables were analysed individually. Paired samples t-tests were used.

Drag and drop vs. overall communicative performance score: differences between pre- and post-tests were statistically highly significant, $t(70) = 10.93$, $p < .001$. Also, there

[19] Standardized scores allow us to make comparisons of raw scores that come from different tasks. A z-score tells how many standard deviations someone is above or below the mean.

was a highly significant strong positive correlation between pre- and post-test scores in the drag and drop task, $r(71) = 0.8$, $p < .001$.

<u>Drag and drop vs. target linguistic performance score</u>: differences between pre- and post-tests were statistically highly significant, $t(70) = 11.36$, $p < .001$. There was a highly significant strong positive correlation between pre- and post-test scores in the drag and drop task, $r(71) = 0.7$, $p < .001$.
<u>Drag and drop vs. trial numbers</u>: on average learners had nine fewer trials in the post –test; the differences between pre-and post-tests were statistically highly significant, $t(70) = 8.57$, $p < .001$. There was a highly significant weak positive correlation between the number of trials during pre- and post-tests in the drag and drop task, $r(71) = 0.3$, $p < .008$.

<u>Information gap vs. overall communicative performance score</u>: There was statistically highly significant differences, $t(70) = 8.43$, $p < .001$ and a highly significant strong positive correlation between pre- and post-test scores in the information gap task, $r(71) = 0.7$, $p < .001$.

<u>Information gap vs. overall communicative performance score</u>: differences between pre- and post-tests were statistically highly significant, $t(70) = 8.09$, $p < .001$. There was a highly significant strong positive correlation between pre- and post-test scores in the information gap task, $r(71) = 0.8$, $p < .001$.

<u>Drag and drop vs. trial numbers</u>: on average learners had 3 fewer trials in the post –test; the differences between time pre- and post-tests were statistically highly significant, $t(70) = 3.67$, $p < .001$. However, there was not any correlation between the number of trials in the pre- and post-tests, $r(71) = 0.1$, $p > 0.1$.

This section investigated how learners performed in each task across all groups and looked closely at the treatment effect in each task. The results indicated that in the post test learners scored higher in overall communicative performance and in the target linguistic items and had fewer trials in the post-test drag and drop task and the information-gap task. The following section will detail the performance of individual treatment groups in each task across the experimental variables: task type, modality of input and grammar effect.

5.4.2.2 Task type: Drag and drop

On average, learners scored 16.7, *SD*=3.2 (out of a possible 25) in relation to overall communicative performance, and 9.98, *SD*= 3 (out of a possible 16) in relation to target linguistic performance and they went through 39.7, *SD*= 8.4 trials in the pre-test. In the post test, learners scored 19.45 out of 25, *SD*= 3.5 in relation to overall communicative performance, 12.66, *SD*=2.5 out of 16 in relation to target linguistic performance and went through 30.61, *SD*=6.5 trials.

Overall communicative performance score: As can be seen from Table 5.18 below, the groups' mean overall communicative scores in the drag and drop task (max score 25) were close in the pre-test, FoS (*M* = 16.95, *SD* = 4.8), FoF (*M* = 16.88, *SD* = 2.3) and FoM (*M* = 16.35, *SD* = 2.5). However, in the post-test, the groups varied, with the FoF group scoring higher (*M* = 21.04, *SD* = 3.8) followed by FoM (*M* = 19.27, *SD* = 2.1) and then FoS (*M* = 17.70, *SD* = 4.4).

Table 5.18: Summary of the descriptives for the drag and drop task

Drag and Drop		**Pre-test**			**Post-test**		
		Com	**Ling**	**Trial**	**Com**	**Ling**	**Trial**
FoS	*Mean*	16.95	9.60	38.00	17.70	12.25	27.95
	SD	4.88	3.53	5.45	4.47	3.31	4.80
FoF	*Mean*	16.88	11.04	40.80	21.04	13.56	29.60
	SD	2.39	2.42	11.04	3.18	2.14	4.46
FoM	*Mean*	16.35	9.27	40.00	19.27	12.12	33.65
	SD	2.51	2.93	7.56	2.16	2.03	8.26
All	*Mean*	16.7042	9.9859	39.7183	19.4507	12.6620	30.6197
	SD	3.27935	3.01659	8.46030	3.51645	2.54077	6.57781

Target linguistic performance score: the groups' mean target linguistic scores in the drag and drop task (max score 16) varied marginally in the pre-test, FoS (*M* = 9.60, *SD* = 3.53), FoF (*M* = 11.04, *SD* = 2.4) and FoM (*M* = 9.27, *SD* = 2.9). But these differences were not significant ($p > 0.5$). However, in the post-test, the groups varied, with the FoF group scoring higher (*M* = 13.56, *SD* = 2.1) followed by FoS (*M* = 12.25, *SD* = 3.3) and then FoM (*M* = 12.12, *SD* = 2).

Trial numbers: the groups' mean number of trials in the drag and drop task in the pre-test were slightly different but the differences were not significant; FoS (*M* = 38, SD = 5.4), FoF (*M* = 40.80, *SD* = 11) and FoM (*M* = 40, *SD* = 7.5). However, in the post-test, the groups varied more, with the FoS (*M* = 27.9, *SD* = 4.8) going through the least number of trials followed by the FoF group (*M* = 29.6, *SD* = 4.4) and then FoM with the most number of trials (*M* = 33.65, *SD* = 8.2).

The results can be summarized as follows;

	Communicative Score	**Linguistic Score**	**Trials**
Drag and drop	**FoF > FoM >FoS**	**FoF >FoS > FoM**	**FoS > FoF > FoM**

The descriptive summary above revealed that the FoF group outperformed the other two groups in the drag and drop task in both target linguistic and overall communicative performance scores while the FoS group had fewer trials than the other two groups. In order to check if these differences were significant, ANOVA was used.

Drag and drop vs. overall communicative performance score: differences were significant, $F(2, 68) = 5.75$, $p < 0.005$ but the homogeneity of variance criteria was not met so the Kruskal-Wallis test was used. It showed that the difference were significant, $H(2)= 8.33$, $p < 0.01$. The Games Howell post hoc test revealed that the difference between the FoF group's overall communicative performance score and the FoS ($p< 0.02$) and FoM group's overall communicative performance score ($p<0.06$) was statistically highly significant while the FoS and FoM overall communicative performance score was not ($p>0.3$).

Drag and drop vs. target linguistic performance score: differences were significant, $F(2, 68) = 2.53$, $p < 0.08$[20] but also non-parametric tests had to be used as the homogeneity of variance was not met. Kruskal-Wallis showed that the differences were significant, $H(2)= 4.7$, $p < 0.09$[21]. Thus, the groups varied in their target linguistic performance score.

Drag and drop vs. trial numbers: differences were highly significant, $F(2, 68) = 5.29$, $p < 0.007$ but non-parametric Kruskal-Wallis had to be used to overcome the significance of the homogeneity of test and it showed that the differences were significant, $H(2)= 6.5$, $p < 0.03$. The Games Howell post hoc test revealed that the significant differences

[20] Significant at the 90% level.
[21] Significant at the 90% level.

were between the FoM group and the FoS ($p < 0.01$) and FoF ($p < 0.08$[22]). The FoS and FoF were not different ($p > 0.4$).

Thus, going back to the results summarized above, the following differences were verified:

	Communicative	**Linguistic**	**Trials**
Drag and drop	**FoF > FoM >FoS**	**FoF >FoS > FoM**	**FoS > FoF > FoM**

In brief, the FoF group, who were asked to focus on the linguistic characteristics of reported speech if and whenever they wanted while completing the tasks, outperformed the two other groups in the drag and drop task in the overall communicative performance score and the target linguistic performance score. On the other hand, the learners in the FoS group who were required to explicitly focus on the linguistic characteristics of reported speech were more confident about their choices had fewer trials than the other two groups.

5.4.2.3 Task type: information gap

On average, learners scored 7.5, *SD*=2.3 (out of a possible 14) in relation to overall communicative performance score, and 1.05, *SD*=1.1 (out of a possible 5) in relation to target linguistic performance score, and they went through 21.73, *SD*=7.9 trials in the pre-test. In the post- test, learners scored 9.14, *SD*=2.3 in relation to overall communicative performance score, 1.8, *SD*=1.5 in relation to target linguistic performance score and they went through 17.88, *SD*=5.5 trials.

<u>Overall communicative performance score</u>: the groups' mean overall communicative performance scores in the information gap task (max score 14) was close in the pre-test, FoS (*M* = 7.45, *SD* = 2.7), FoF (*M* = 7.12, *SD* = 2.2) and FoM (*M* = 7.92, *SD* = 2.1). However, in the post-test, the groups varied, with the FoF group scoring higher (*M* = 9.8, *SD* = 2.1) followed by FoM (*M* = 8.96, *SD* = 2.1) and then FoS (*M* = 8.55, *SD* = 2.7).

[22] Significant at the 90% level.

Table 5.19: Summary of the descriptive statistics for the information gap task

Information gap		Pre-test			Post-test		
		Com	**Ling**	**Trial**	**Com**	**Ling**	**Trial**
FoS	*Mean*	7.45	1.00	22.25	8.55	2.15	15.70
	SD	2.76	1.03	6.21	2.74	1.42	3.11
FoF	*Mean*	7.12	0.88	20.92	9.80	1.68	19.16
	SD	2.24	1.09	9.99	2.18	1.65	5.34
FoM	*Mean*	7.92	1.27	22.12	8.96	1.85	18.35
	SD	2.19	1.22	7.08	2.13	1.54	6.76
All	*Mean*	7.5070	1.0563	21.7324	9.1408	1.8732	17.8873
	SD	2.37170	1.11979	7.93178	2.35612	1.53929	5.54604

Target linguistic performance score: the groups' mean linguistic scores in the information gap task (max score 5) were very close in the pre-test, FoS (M = 1, SD = 1.03), FoF (M = 0.8, SD = 1.09) and FoM (M = 1.27, SD = 1.22). The variances within each group were high considering the high SD. However, in the post-test, the groups varied, with the FoS group scoring higher (M = 2.15, SD = 1.4) followed by FoM (M = 1.85, SD = 1.5) and then FoF (M = 1.6, SD = 1.6).

Trial numbers: the groups' mean number of trials in the drag and drop task was close in the pre-test, FoS (M = 22.25, SD = 6.2), FoF (M = 22.92, SD = 9.9) and FoM (M = 22.12, SD = 7). But these differences were not significant (p >). However, in the post-test, the groups varied, with the FoS (M = 15.7, SD = 3.11) going through the least number of trials followed by the FoM group (M = 18.35, SD = 6.7) and then FoF with the most number of trials (M = 19.16, SD = 3.1).

The results can be summarised as follows;

	Communicative	**Linguistic**	**Trials**
Information gap	**FoF>FoM>FoS**	**FoS>FoM>FoF**	**FoS >FoM > FoF**

To summarise, learners who explicitly focused on the linguistic characteristics of reported speech scored less than the other two groups in their overall communicative performance score but higher than the two groups in their target linguistic performance score in the information-gap task. Learners in the FoS group had fewer trials in the information-gap task. In order to check if the differences among the groups were significant, ANOVA tests were used.

Information gap vs. overall communicative performance score: differences were not significant, $F(2, 68) = 1.71$, $p > 0.18$.

Information gap vs. target linguistic performance score: differences were not significant, $F(2, 68) = 0.5$, $p > 0.59$.

Information gap vs. trial numbers: differences were significant, $F(2, 68) = 2.39$, $p < 0.09$[23]. However, the homogeneity of variance test was also significant (p <0.001 so non-parametric Kruskal-Wallis was used. It showed that the differences were not significant, $H(2)= 4.2$, $p > 0.1$.

Therefore, none of these following differences were verified by the statistical tests.

	Communicative	**Linguistic**	**Trials**
Information gap	~~**FoF > FoM > FoS**~~	~~**FoS>FoM>FoF**~~	~~**FoS >FoM > FoF**~~

To recap the results with respect to the task type, input type had an effect on the learners' performance on the drag and drop task but not on the information-gap task.

5.4.3 Modality of input (i.e. access to audio file)

The modality of the input (e.g. written vs. aural) is another factor that has been found to play a role in the performance of learners and their processing of input, as established in Chapter Two. To investigate the effect of modality of the input on the learners' performance, information was collected on whether learners listened to the audio input. This input was provided only in the drag and drop task on the pre- and post-tests and during treatment. The following analysis will present an overview of the patterns of access to the audio file across all the groups and within each group for the pre- and post-tests. These patterns are then analysed in relation to the learners' communicative and linguistic performance. (See Appendix VIII.)

5.4.3.1 Modality of the input: pre-test

To investigate the effect of modality of input in the pre-test, I tracked whether each learner listened to the audio input at all and also the number of times they listened to the audio input. For the first stage of analysis, the number of times a learner listened to the audio file was treated as 1 so each learner was marked 0 (didn't listen) or 1 (listened).

[23] Significant at the 90% level.

Running descriptive tests for the access to audio file in pre-test revealed that most learners, (63.4%, i.e. 45 out of 71 learners), chose to listen to the audio file. Learners within input groups also varied. The majority of learners in the FoM group (73.1%) decided to listen the audio file, while only half of the learners in the FoF group (56%) and around 60% of the learners in the FoS did so (see Table 5. 20 below).

Table 5.20 Access to audio files: pre-test

Audio Access	Didn't listen	Listened
FoS (N=20)	(8) 40.0%	(12) 60.0%
FoF (N=25)	(11) 44.0%	(14) 56.0%
FoM (N=26)	(7) 26.9%	(19) 73.1%
All groups	(26) 36.6%	(45) 63.4%

In order to explore whether these differences were statistically significant, a chi square test was used. The modality of input did not differ by input groups $\chi^2(2, N = 71) = 1.73$, $p = 0.4$. This means that being in a specific treatment group did not affect whether learners decided to listen to audio input or not at pre-test

5.4.3.2 Modality of the input: post-test

To investigate the effect of modality of input at post-test, the same analysis was adapted as for the pre-test. I tracked whether the learner had listened to the audio, and the number of times a learner listened to the audio file was treated as 1, so each learner was marked 0 (didn't listen) or 1 (listened). Running descriptive tests for the access to audio file in post-test revealed that this time most learners, (66.19%, 47 out of 71 learners), chose not to listen to the audio file. Learners within input groups also varied. The majority of learners in the FoM group (61.53%) decided not to listen to the audio file while 72% of learners in the FoF group and around 65% in the FoS did not listen (see Table 5.21).

Table 5.21 Access to audio files: post-test

Audio Access	Didn't listen	Listened
FoS (N=20)	(13) 65%	(7) 35%
FoF (N=25)	(18) 72%	(7) 28%
FoM (N=26)	(16) 61.53%	(10) 38.46%
All groups	(47) 66.19%	(24)33.80%

5.4.3.3 Modality of the input: input effect

In order to explore whether these differences were statistically significant, a chi square test was used. The modality of input did not differ by input groups $\chi^2(2, N = 71) = 0.64$, $p = 0.7$. This means that being in a specific treatment group did not affect whether learners decided to listen to the audio input or not while completing the post-test.

5.4.3.4 Modality of the input: treatment effect

To check if the treatment per se had any effect on the learners' choice to access the audio file, a paired samples t-test was used. It showed that fewer learners decided to listen to the audio file during the post-test and that this difference was highly significant, $t(70) = 4.36$, $p < .001$.

5.4.3.5 Modality of the input: overall communicative performance score

To investigate the effect of modality of input on overall communicative performance (post-test score), an independent samples t-test was used. First the effect of listening to the audio file during the pre-test was tested against the overall communicative performance score. On average, learners who listened to the audio file during the pre-test scored 8% (3.12 points) higher (30.08) than the ones who did not listen (26.96). Independent sample t-test results indicated that this difference was statistically highly significant, $t(69) = 2.71$, $p < .008$.

Second, the effect of listening to the audio file during the post-test was investigated. Descriptive tests indicated that, contrary to the previous results, learners who did not listen to the audio input scored 1% higher in the post-test (29.08) than those who listened (28.66). However, the Independent samples t-test results revealed that this difference was not statistically significant, $t(69) = 0.34$, $p > 0.7$ so whether learners listened to the audio input during the post –test or did not do so did not affect their overall communicative performance score.

Table 5.22 Overall communicative performance scores according to modality of input groups

Modality of input group	Pre-test	Post-test
Listened	30.08	28.66
Didn't listen	26.96	29.08

To sum up, although there were differences in the number of learners in each group who accessed the audio files, these differences were not statistically significant. So, regardless of which group they were in, the learners exhibited the same pattern of access to the audio files However, there were significant differences in terms of overall communicative performance scores among those who listened to the audio input during the pre-test and those who did not.

5.4.3.6 Modality of the input: target linguistic performance score

To investigate the effect of modality of input on target linguistic performance (post-test scores), an independent samples t-test was used. First the effect of listening to the audio file during the pre-test was tested against target linguistic performance scores. On average learners who listened to the audio file during the pre-test scored 3.25% (1.27 points) higher than the ones who did not listen, 15 and 13.73 respectively. However, independent sample t-test results indicated that this difference was not significant, $t(69) = 1.6$, $p > 0.11$.

Second, the effect of listening to the audio file during the post-test was investigated. Descriptive tests indicated that learners who did not listen to the audio input scored 2.41% higher in the post-test (14.85) than those who listened (13.91). Independent samples t-test results revealed that this difference was not significant, t (69) = 1.15, $p > 0.2$ so whether or not learners listened to the audio input during the post-test did not affect their target linguistic performance score.

Table 5.23 Target linguistic performance scores according to modality of input groups

Modality of input group	Pre-test	Post-test
Listened	15.00	13.91
Didn't listen	13.73	14.85

To sum up the results in terms of modality of input effect, the following can be stated:

Input effect: there were no significant differences among the input groups. Thus, being in a particular group did not affect the learner's decision to listen to audio input.

Treatment effect: there was a highly significant association, with fewer learners listening to the audio input during the post-test.

Performance effect: There were significant differences in terms of overall communicative performance scores among those who listened to the audio input during the pre-test but not the post-test: learners who listened during the pre-test scored higher in the post-test. On the other hand, listening to the audio input did not have any effect on the learners' target linguistic performance scores.

5.4.4. Grammar effect

The different types of input demand different levels of attention to form and this will be reflected in the learners' decision to focus on the linguistic characteristics of reported speech, i.e. to access grammar help. FoF and FoS groups had access to grammar help but the FoM did not. Moreover, the FoS group had to go through grammar instruction after the pre-test. The following results show an analysis of whether the learners accessed the grammar help and if they did, whether there was any association between access of grammar help and performance scores and behaviour variables. It also examined the effect of grammar instruction, which the FoS group received, on performance and trial number during the post-test. (See Appendix IX.)

5.4.4.1 Grammar access: Input effect (FoS and FoF)

At the first stage of the analysis, there was no differentiation between the numbers of times the learners accessed the help. Learners were marked as (0) for no access and (1) for any access. If the access to grammar help was associated with performance scores, further analysis checked the significance of the number of times the grammar help was accessed. As it is clear from Table 5.24 below, 40% of the learners in the FoS group decided not to access grammar help at all during the treatment while 60% did. Similarly, 44% of the learners in the FoF group decided not to access grammatical help while 56% did.

In order to check whether there was any association between access to grammar and the treatment groups, a chi-square was used. The results showed no association between the treatment groups and access to grammar, $\chi^2(1, N = 45) = 0.07, p > 0.7$. This means the learners' decision to access grammar was not related to the treatment group they belonged to.

Table 5.24 Grammar access according to groups

Grammar access	No access	Access
FoS (N= 20)	40% (8)	60% (12)
FoF (N= 25)	44% (11)	56% (14)
Total (N=45)	42.2 % (19)	57.8% (26)

5.4.4.2 Grammar access: overall communicative performance score

An independent samples t-test was used to investigate the association between access to grammar and overall communicative performance score and it revealed that learners who accessed the grammar help scored very closely to those who did not 28.80 (SD= 5.5) and 28.78 (SD=5.8) respectively. It also showed that there was no association between accessing the grammar help during treatment and overall communicative performance score, t (43) = 0.01, $p > 0.9$.

5.4.4.3 Grammar access: target linguistic performance score

An independent samples t-test was also used to investigate the association between access to grammar and target linguistic performance score and revealed that learners who accessed the grammar help scored marginally different to those who did not 15.23 (*SD*= 3.5) and 14.36 (*SD*=3.7) respectively. It also showed that there was no association between accessing the grammar help during treatment and overall communicative performance score, t (43) = 0.7, $p > 0.9$.

5.4.4.4 Grammar access: trial number

Looking at the relationship between the number of trials in post-test and access to the grammar help, it is clear that those who accessed the grammar help had fewer trials 45.73 (*SD*=7.1) in comparison to those who did not 47.52 (*SD*= 7.6). However, t-test showed that the difference was not statistically significant, t (43) = 0.8, $p > 0.5$.

5.4.4.5 Grammar instruction: FoS group

In order to check whether the grammar instruction section had any effect on the performance of the FoS group, the time spent on the main grammar instruction was compared with the performance and trial variables. The plot graph (Fig 5.10) indicates that the time spent on main grammar is normally distributed. This was confirmed with the Shapiro Wilk test, $p> 0.1$. On average learners in the FoS group spent 95 seconds on the grammar instruction. Learners, though, varied considerably with 8 seconds as the least time spent and 283 seconds as the most, Mean=95 , *SD*= 75.2.

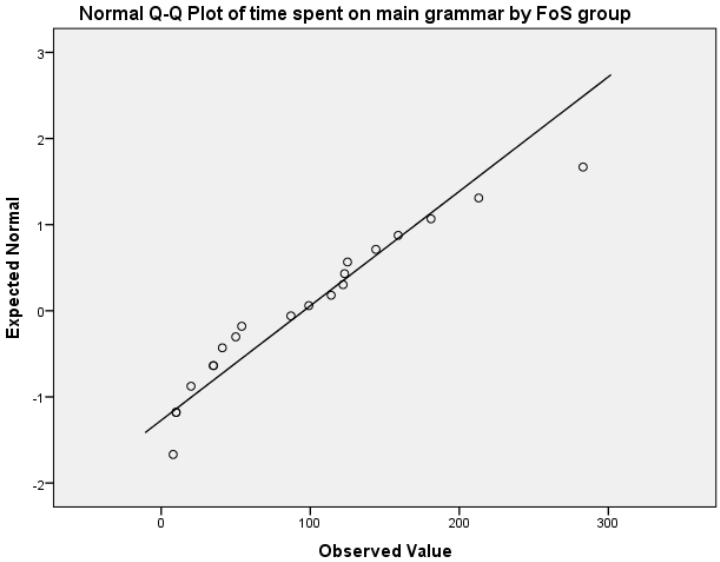


Figure 5.10: Normal distribution plot: time spent on grammar instruction by FoS group

In order to examine the effect of the time spent on grammar instruction on performance and trial variables, Pearson correlation test was used. The results indicated that the time spent on grammar instruction negatively correlated with overall communicative performance score ($r(20) = -0.4$, $p < 0.05$) and target linguistic performance score ($r(20) = -0.4$, $p < 0.03$). However, there was no significant association between the time spent on grammar instruction and trial numbers during the post-test ($r(20) = -0.08$, $p > 0.7$).

To sum up the results in relation to the effect of grammar instruction and help, the following can be stated:

Input effect: there was no difference between learners in the FoS and FoF groups in their access to the contextualized grammar help.

Performance effect: there were no differences between learners who accessed the contextualized grammar help and those who did not in terms of communicative or linguistic performance or the number of trials. However, there were significant differences among learners in the FoS group: the more time the learners spent on grammar, the lower their scores were in the communicative and linguistic performance.

5.5 Profiling learners

The analyses in the previous sections were based on data collected using quantitative methods to which statistical tests were applied in the aim of verifying or falsifying the hypotheses. Accordingly, the above sections looked mainly at group differences and the features of treatment in relation to the research questions. Regarding group differences, the data was analysed in terms of treatment (input), modality of input and grammar help and instruction groups. It also looked at the factors that affected attainment/behaviour, namely: task type, access to audio files and trial numbers during the pre- and post-tests. Turning now to the qualitative part of the analysis, the data of individual learners will be analysed in depth. The aim is to look at one learner from each of the three input groups and describe their actions and decisions to create a better picture of individual behaviour. Hence, this section is mainly concerned with individual similarities and differences. The three learners were chosen randomly from the 71 learners based on one criterion: their total scores from the tests and the treatment were as close as possible. The main reason behind this choice is that learners, whether in the classroom or in research studies, are always assessed based on their performance and an assumption is implied in ISLA and pedagogic based research that learners who score the same are at the same level or have acquired the target structure equally. Therefore, the aim of the qualitative analysis below was to examine closely how learners who are perceived as similar, based on their scores, react in real time when processing input.

As it is clear from Table 5.25, the three learners were very close in their total scores, tests and treatment. They were also relatively close in the number of trials they had during the treatment. This helps in examining their behaviour to see if they followed the same route to get the same results. The learning route is identified by examining the individual log files (see Chapter Four, section 4.7.1 for more details).

Table 5.25 Summary of the three learners' performance

	Type of input	FOF learner	FOM learner	FOS learner
	real time	2979	2991	2696
Drag and drop pre-test	Overall communicative performance score	17	20	23
	Target linguistic performance score	10	10	14
	Trial	28	28	30
Drag and Drop post-test	Overall communicative performance score	24	21	23
	Target linguistic performance score	16	11	16
	Trial	28	40	28
Information gap pre-test	Overall communicative performance score	12	12	12
	Target linguistic performance score	3	3	3
	Trial	28	16	16
Information gap post-test	Overall communicative performance score	14	13	12
	Target linguistic performance score	5	4	4
	Trial	15	16	13
Pre-test	Overall communicative performance score	30	32	35
	Target linguistic performance score	12	13	17
	Trial	57	44	46
Post-test	Overall communicative performance score	38	34	35
	Target linguistic performance score	20	15	20
	Trial	43	56	41
Treatment task 1	Score	12	14	12
	Trial	39	50	17
Treatment task 2	Score	11	11	11
	Trial	11	20	16
Treatment task 3	Score	5	5	6
	Trial	6	8	6
Treatment task 4	Score	8	8	8
	Trial	11	11	13
Treatment task 5	Score	6	6	5
	Trial	7	6	14
Treatment task 6	Score	16	16	15
	Trial	29	35	20
Total	Score	126	126	127
	Trial	203	230	173

5.5.1 FoS

The FoS learner spent 2696 seconds on the software package. He went through 173 trials and scored 127 overall.

The learner logged in and went through the information screens and then started the pre-test. Once the pre-test started, the learner pressed the 'listen' button and started listening to the audio file and dragging the items at the same time. He did not start in order; it seemed that he started with the items he was more confident about as he dragged each item only once. After the learner finished all the items he knew, he listened to the audio file again. This time he did not keep the file running while answering. He listened then tried some items then listened again and tried. In total he spent 11 minutes 46 seconds on the drag and drop task, scored 23 (out of 25) in that task and went through 30 trials. The next task was the information-gap task. There was no audio file here. Again, the learner started answering the gaps he seemed to know and then moved to the ones he did not know. This is evident in the order and trial number of each item. The learner went through one trial for the items he answered first while he had two or three goes at the items answered later. The learner then moved to the next section, which was the grammar instruction. He spent 9 seconds on the first screen and 13 seconds on the second. He then went back to the first screen and spent 13 seconds; in total he spent 35 seconds on the main grammar instruction. He then moved to the first task in the treatment, which was an information-gap task. The learner started filling the gaps in order and then he went back and refilled the three gaps he was not sure about. He spent 3 minutes and 7 seconds on the task, scored 12 (out of 14) and went through 17 trials. The learner then moved to the next treatment task, which was sequencing (treatment task 2). He started by answering the first item then decided to listen to the audio file. The same pattern of switching between items was followed. The learner decided to listen again to the audio file halfway through the task. He spent 4 minutes and twenty-one seconds on the task, achieved a perfect score of 11 (out of 11) and had 16 trials. The learner moved to the next task, drag and drop. He had a go at two items, which were answered correctly, and then started listening to the audio file. He stopped the file more than once but did not take action or move any items. After he finished listening, he answered the four other items correctly and then moved to the next task. The learner spent 1 minute, 41 seconds on the task, scored 6 (out of 6) and had 6 trials. The next task was multiple choice. The task had two sections each with a separate audio file and 5 questions to answer each. There was a choice of listening to the audio file or opening a transcript. Here the learner decided to listen to the audio file and started answering the questions in order. He then moved to the next section and started listening to the audio file. He answered four items (not in order) and then decided to listen to the audio file again and after that had 3 goes at the last item. The learner spent 4 minutes and 28 seconds on the task, scored 8 (out of 10) and went through 13 trials. The next task was also multiple choice again. The learner started listening to the audio file and answering the questions. He went through three trials for the first item then moved to the second. He answered the items in order but went through more than one trial for most items. The learner spent 2 minutes and 27 seconds on the task, scored 5 (out of 6) and went through 14 trials. The final task was an information gap. The learner started filling the gaps straight after he moved to the task. He filled them in order, leaving the ones he was not sure about blank. Once done, he decided to listen to the audio file and had a go at four more items. Then he decided to open the grammar help and spent 6 seconds on it and then closed it and answered the last four items. The learner spent 5 minutes and 18 seconds on the task, scored 15 (out of 16) and went through 20 trials. The screen after that was the information on post-test and then the post-test started. The first task was the information gap. The learner filled the gaps in order with one trial each and left one blank. The learner spent 1 minute and 38 seconds on the first post-test task, scored 12 and went through 13 trials. The final part of the treatment was the post-test drag and drop task. The learner started by dragging the items he was more confident about and then moved to the ones he was not sure about, as shown by the number of trials. After he dragged all the items, he decided to listen to the audio file and took no action. Once the audio file stopped, he moved to the last screen. The learner spent 3 minutes and 11 seconds on the second post-test task, scored 23 (out of 25) and went through 28 trials. The last screen was a summary of performance. He spent some time reading it, and then logged off.

<u>5.5.2 FoF</u>

The FoF learner spent 2979 seconds on the software package. He went through 203 trials and scored 126 overall.

The learner logged in and went through the information screens then started the pre-test. Once the pre-test started, the learner pressed the ‘listen’ button and started listening to the audio file. Unlike the FoS learner, who started dragging the items as soon as the audio file started, the FoF learner finished listening to the audio file, spent 10 seconds looking at the screen then pressed on the ‘listen’ button again. He listened for 2 seconds and then paused the audio file and pressed on the ‘instruction button’, spent 16 seconds on the instruction then pressed the ‘listen’ button again. He listened for 2 seconds then paused and pressed on the ‘instruction’ to close it and after that he pressed the ‘listen’ button. While listening, he started dragging the items in order, leaving the ones he was not sure about. Halfway through, he decided to listen to the audio file again. He dragged two more items while listening for 25 seconds and then paused the audio file and dragged the rest of items. He then pressed on the ‘listen’ button again to play the 10 seconds left on the audio file and went back and dragged the items he left before. When the audio file finished, he pressed again to listen and dragged the last three items to their destination. In total he spent 8 minutes and 37 seconds on the drag and drop task, scored 17 (out of 25) and went through 28 trials. A different pattern emerged in the information gap task. The learner went through all the gaps in order, one trial each and filled them with answers. Once done, he went back to the first gap and went through all the gaps in order again, correcting and changing the tense. The learner spent 5 minutes and 11 seconds on the task, scored 12 (out of 14) and went through 28 trials. He then moved to the first task in the treatment which was an information-gap one. The learner started the task and within 2 seconds he decided to move to the next task without taking any action on the first task. He spent 3 seconds on the second task and then went back to the first one. He filled three gaps then pressed the ‘listen’ button, listened for 8 seconds then paused and went back to the first two gaps and changed them. He pressed the ‘listen’ button again and listened for 16 seconds then paused and started filling more gaps. The same pattern - listening, pausing and then answering - continued until he finished the task. He spent 4 minutes and 15 seconds on the task, scored 12 (out of 14) and went through 39 trials. The learner then moved to the next treatment task, which was sequencing. He decided to listen to the audio file from the start and dragged all the items to their destination in one trial. He did not do this in order rather he dragged the ones he was sure about first. He spent 2 minutes and 58 seconds on the task, scored 11 (out of 11) and had 11 trials. He spent 2 seconds switching between the first two tasks then decided to move to the next task, drag and drop. He decided to listen to the audio file from the start and started dragging the items to their destination in order, going through one trial for each item. He finished dragging all the items by the time the audio file ended. He then decided to listen to the audio file again and check his responses.

He listened to the complete audio file, took no action and then moved to the next task. The learner spent 1 minute, 31 seconds on the task, scored 5 (out of 6) and had 6 trials. The next task was multiple choice. The task had two sections each with a separate audio file and 5 questions to answer each. There is a choice of listening to the audio file or opening a transcript. Here he decided to listen to the audio file from the start. He listened for 1 second then paused for 3 seconds and then started listening again and answering the questions, starting with the second item. He then answered the first and rest in order. He then moved to the next section. He started listening to the audio file and answered all the questions in order. The learner spent 3 minutes and 45 seconds on the task, scored 8 (out of 10) and went through 11 trials. The next task was also multiple choice. The learner spent 9 minutes looking at the screen and then started listening to the audio file and answering the questions in order, leaving the ones he did not know. Once the audio file finished, he pressed again to listen and answered another item within 30 seconds then moved to the next task. The learner spent 2 seconds at the next task and then went back to the multiple choice one and then again to the next task and then back to the multiple choice one. He then answered the last question and moved to the next task then back to the multiple choices then decided to start the next task. The learner spent 7 minutes and 27 seconds on the task, scored 6 (out of 6) and went through 7 trials. The final task was an information gap. The learner spent 5 seconds looking at the screen then started listening to the audio file and filled the first gap. He only listened for 20 seconds then paused the audio file. He then filled the gaps in order, leaving the ones he was not sure about blank. Once done, he pressed the 'listen' button to resume the audio file. While listening, he went back to the blank gaps and filled some in order, leaving some blank. When done he went back to the blank ones and filled them. He then decided to listen to the audio file for 10 seconds before moving to the next task. The learner spent 4 minutes and 4 seconds on the task, scored 16 (out of 16) and went through 29 trials. The screen after that was the information on post-test and then the post-test started. The first task was the information gap. The learner first listened to the audio file for 5 seconds and then paused it and started filling the gaps in order. He then decided to move to the next task then without any action on the task decided to go back to the information gap. He changed his answers on one gap then moved to the next task. The learner spent 2 minutes and 17 seconds on the first post-test task, scored 14 (out of 14) and went through 15 trials. The next task was a drag and drop task. After he started the second post-test task, he dragged one item and then within 2 seconds he went back to the information gap task and no action was taken, and then moved to the second post-test task again. The learner dragged all the items in order, with more than one trial for some. After he dragged 19 items, he decided to go to the previous task and then back to the current task. He then continued dragging items in order and when done, he moved to the last screen. The learner spent 3 minutes on the second post-test task, scored 24 (out of 25) and went through 28 trials. The last screen was a summary of performance. He spent some time reading it, and then logged off.

5.5.3 FoM

The FoM learner spent 2991seconds on the software package. He went through 230 trials and scored 126 overall.

The learner logged in and went through the information screens then started the pre-test. Once the pre-test started, the learner pressed the 'listen' button and started listening to the audio file. Unlike the FoS learner, who started dragging items straight away, and also unlike the FoF learner, who listened to the complete audio file before he started to drag items, the FoM learner listened for 4 seconds then paused the audio file and dragged one item. He switched then between the first and second task four times and then decided to work on the first task again. He started by clicking on the 'instruction' button, spent three seconds reading the instruction then closed it. He started dragging the items. He dragged four and then decided to listen to the audio file for two seconds then paused and dragged one item then listened for 3 seconds then paused and dragged another item. The same pattern continued throughout the task till all the items were dragged. Once, done, he decided to listen to the audio file again, this time pausing every second. He listened for a total of 16 seconds then paused, changed three items then clicked on the 'instruction' button again. He spent 15 seconds reading the instruction and then moved to the next task. The learner spent 6 minutes and 38 seconds on the task, scored 20 (out of 25) and went through 28 trials. The learner started the next task by clicking on the 'instruction' button, spent 8 seconds reading it and then closed it. He then started going through the first five gaps more than once leaving some blank and filling others. He reclicked on the 'instruction' button again and spent two seconds reading it, closed and filled more gaps. He accessed the instruction for the third time, spent five seconds reading it and then closed and went through all the gaps in order, changing the answers in some. Once done, he moved to the next task. The learner spent 6 minutes and 49 seconds on the second post-test task, scored 12 (out of 14) and went through 16 trials. The first treatment task was an information gap. First, the learner spent 4 minutes looking at the screen and then clicked on the ‘instruction’, spent two seconds on it and then closed and clicked on the ‘listen’ button. He listened for 5 seconds, then paused and filled the first three gaps, then listened for a couple of seconds, then filled more gaps and continued in the same pattern until he had a go at all the gaps more than once and listened to the audio file two times. He then decided to go through all the gaps in order and check them, changing the answers in some before he moved to the next task. The learner spent 5 minutes and 1 second on the task, scored 14 (out of 14) and went through 50 trials. The next treatment task was sequencing. The learner spent 3 seconds looking at the screen then clicked on the ‘instruction’ button, spent 13 seconds reading it and then closed it. He then clicked on the ‘listen’ button, listened for 20 seconds pausing two times. He then dragged the first item then resumed the audio file. He listened for 2 seconds then paused and dragged the second item. He continued in the same pattern, and once all items were dragged, he went through them all, not in order, and checked them. He spent 3 minutes and 10 seconds on the task, scored 11 (out of 11) and had 20 trials. The learner moved to the next task, drag and drop. He followed the same pattern as in the previous tasks, listen, pause, and then action. The only difference was that he did not check the instruction at the start of the task, but rather halfway through. The learner spent 1 minute, 4 seconds on the task, scored 5 (out of 6) and had 8 trials. The next task was multiple choice, with two sections each with a separate audio file and 5 questions and with a choice of listening to the audio file or opening a transcript. The learner decided to listen to the audio file for 2 seconds then paused and opened the ‘instruction’, spent 5 seconds reading it, then closed it and resumed the audio file. He listened for 4 seconds only and then paused and answered all the questions, not in order. He then moved to the next section and started listening to the audio file. He listened for 14 seconds and answered two questions, then paused and answered more questions. The learner spent 6 minutes and 47 seconds on the task, scored 8 (out of 10) and went through 11 trials. The next task was also multiple choice. The learner started listening to the audio file; he listened for 28 seconds then paused and dragged one item. He then resumed the audio file and answered the rest of the questions, leaving one blank. He moved to the next task, and within 2 seconds returned to the multiple choice one, answered the remaining question. He then switched three times to the next task and back to the current task and each time he clicked on the ‘listen’ button, listened for 2-5 seconds and then moved to the next task. The learner spent 7 minutes and 10 seconds on the task, scored 6 (out of 6) and went through 6 trials. The final task was an information gap. The learner looked at the screen for 17 seconds, then pressed the 'listen' button. He listened for 4 seconds, then paused and filled two gaps. He then resumed the audio file, listened for 34 seconds and filled 5 gaps (some more than one trial) while listening. He then paused and decided to go back to the previous task. He spent 2 seconds on the previous task then switched to the current task. He started to repeat the same pattern as before: listen, pause, then action. The learner spent 4 minutes and 46 seconds on the task, scored 16 (out of 16) and went through 35 trials. The learner then spent 1 minute and 4 seconds switching between all the treatment tasks before he decided to exit the treatment. The screen after that was the information on post-test and then the post-test started. The first task was the information gap. The learner filled the gaps in order, had more than one try at some. The learner spent 2 minutes and 56 seconds on the first post-test task, scored 13 and went through 16 trials. The final part of the treatment was the post-test drag and drop task. The learner started dragging items randomly and then after 1 minute and 33 seconds, he decided to listen to the audio file. He then listened, paused and then dragged a few items around. He continued in this pattern till he finished the task and moved to the last screen. The learner spent 4 minutes on the second post-test task, scored 21 (out of 25) and went through 40 trials. The last screen was a summary of performance. He spent some time reading it, and then logged off.

It is tempting to claim that all the learners went through the same processes while dealing with the input as the three learners' scores and trial number were very similar. However, the qualitative analysis revealed that although there were shared patterns of behaviour across the three learners, differences were evident. Even when the learners spent the same time, scored the same and had the same number of trials, the way they dealt with the input was different. These similarities and differences will be discussed further below.

5.6 Discussion

Before turning in the following chapter to an extensive discussion of the results in relation to literature on the issues addressed in this thesis, this section considers the results of the quantitative and qualitative analysis in terms of hypotheses and sub-hypotheses stated earlier in relation to group and individual differences.

5.6.1 Group similarities and differences

Research question one: *Which type of input, FoS, FoF and FoM, is most effective in a TELL-based environment?*

Hypothesis 1.a There will be differences among the three groups in their overall communicative performance score.

Input Effect: FoF significantly outperformed the FoM and FoS groups but although the FoM group scored higher in the post test than the FoS group, the difference was not statistically significant.

Treatment effect: statistically highly significant. All groups improved after the treatment

Hypothesis 1.b There will be differences among the three groups in their target linguistic performance score.

Input Effect: not significant so regardless of which group they belong to, learners attended to form at the same level.

Treatment effect: statistically highly significant. Learners were able to attend to form better after the treatment.

Hypothesis 1.c There will be differences among the three groups in terms of the processing time.

Input Effect: not significant, so the type of input the learners received did not influence the time they spent processing the input.

Performance effect: highly significant association between processing time and communicative and linguistic performance. Learners who spent less time processing input scored higher on overall communicative performance and in the target linguistic performance scores.

Research question two: *What factors or decision processes exhibited by the learners while dealing with the different types of input contribute to differences in attainment? How do these factors/processes map on to performance?*

Hypothesis 2.a There will be differences among the three groups in terms of the number of trials

Input effect: There were significant differences between the FoS and FoM group and the FoS and FoF. However, the FoF and FoM group were not statistically different

Treatment effect: There was a highly significant association between the trial numbers during the pre- and post-tests, with learners going through fewer trials in the post-test.

Performance effect: there was a significant association between the number of trials during the pre-test and the linguistic performance. Learners who had fewer trials scored higher.

Hypothesis 2.b There will be differences among the three groups in terms of task types There were significant differences among the three groups in their overall communicative performance score, target linguistic performance score and trial number in relation to the drag and drop task but not the information-gap. The differences are summarized as follows;

	Communicative	Linguistic	Trials
Drag and drop	**FoF > FoM >FoS**	**FoF >FoS > FoM**	**FoS > FoF > FoM**

Hypothesis 2.c There will be differences among the three groups in terms of modality of input

Input effect: there were no significant differences among the input groups. Thus, being in a particular group did not affect the learner decision to listen to audio input.

Treatment effect: there was a highly significant association with less learners listening to the audio input during the post-test.

Performance effect: There were significant differences in terms of overall communicative performance among those who listened to the audio input during the pre-test but not the post-test. Learners who listened during the pre-test scored higher in the post-test. On the other hand, listening to the audio input did not have any effect on the learners' target linguistic performance.

Hypothesis 2.d There will be differences among the three groups in terms of access to grammar help

Input effect: there was no difference between learners in the FoS and FoF groups in their access to the contextualized grammar help.

Performance effect: There were no differences between learners who accessed the contextualized grammar help and those who did not in terms of overall communicative or target linguistic performance or the number of trials. However, there were significant differences among learners in the FoS group: the more the learners spent on grammar, the lower their scores were in the overall communicative and target linguistic performance.

5.6.2 Individual similarities and differences

The section above summarized the findings on group differences and similarities. This section will summarize the findings on individual differences. Based on the detailed description in section 5.5 above, the following patterns could be recognized.

Similar Patterns

Learners start dealing with the items they are most familiar with or most confident about. This is evident from the number of trials and the scores. The first items answered were most likely to be correct with only one trial each.

Learners answer the questions in order when they perceive the task as easy, sometimes leaving difficult items blank. However, they answer randomly when they perceive the task as difficult.

Different patterns

Different patterns emerged on how learners deal with the audio input;

- listen to the whole audio file then deal with the task
- listen to it bit by bit and deal with the task while listening
- listen to it and deal with the task simultaneously
- deal with the task and then use the audio input to recheck answers

Overall, the patterns exhibited by each learner could be recapitulated as follows;

The FoS learner was more confident about his answers as he went through fewer trials and spent less time on the tasks. He listened to the audio input when there was a difficulty completing the task as he usually started dealing with the task correctly then decided to listen when he was not sure. The learner listened to the audio input from the start only when he perceived the whole task as difficult, this was evident in more trial numbers from the beginning. He only accessed the grammar help when the audio file did not help. When he perceived the task easy, he did not access the audio input or the grammar instruction.

The FoF learner exhibited two distinct patterns of behaviour. In the drag and drop task, he listened to the audio file extensively and answered the questions randomly while in the information gap, he filled all the gaps in order and even rechecked them in order. He followed a similar pattern when listening to the audio input; listen, pause and answer. He also used the audio input as a way of checking his answers. He switched sometimes between tasks even when he did not take any action on either.

The FoM learner displayed a clear pattern of dealing with the audio input as an essential part of the task. He listened to the audio input wherever available and always used it in the pattern, listen, pause and answer. He seemed the less confident about his answers as he kept going back and changing them all the time even when they were correct. He checked all the answers at the end of each task. He also showed a clear pattern of switching between tasks all the times, more than once in most cases. He was the only one among the three who accessed the instruction.

Based on the detailed description in section 5.5 above, the following patterns could be recognized.

Similar Patterns are that Learners start dealing with the items they are most familiar with or most confident about. This is evident from the number of trials and the scores. The first items answered were most likely to be correct with only one trial each. Learners answer the questions in order when they perceive the task as easy, sometimes leaving difficult items blank. However, they answer randomly when they perceive the task as difficult.

Different patterns emerged on how learners deal with the audio input:

- listen to the whole audio file then deal with the task
- listen to it bit by bit and deal with the task while listening
- listen to it and deal with the task simultaneously
- deal with the task and then use the audio input to recheck answers
- Overall, the patterns exhibited by each learner could be recapitulated as follows;

The FoS learner was more confident about his answers as he went through fewer trials and spent less time on the tasks. He listened to the audio input when there was a difficulty completing the task as he usually started dealing with the task correctly then decided to listen when he was not sure. This learner listened to the audio input from the start only when he perceived the whole task as difficult, this was evident in more trial numbers from the beginning. He only accessed the grammar help when the audio file did not help. When he perceived the task easy, he did not access the audio input or the grammar instruction.

The FoF learner exhibited two distinct patterns of behaviour. In the drag and drop task, he listened to the audio file extensively and answered the questions randomly while in the information gap, he filled all the gaps in order and even rechecked them in order. He followed a similar pattern when listening to the audio input; listen, pause and answer. He also used the audio input as a way of checking his answers. He switched sometimes between tasks even when he did not take any action on either.

The FoM learner displayed a clear pattern of dealing with the audio input as an essential part of the task. He listened to the audio input wherever available and always used it in the pattern, listen, pause and answer. He seemed the less confident about his answers as he kept going back and changing them all the time even when they were correct. He checked all the answers at the end of each task. He also showed a clear pattern of switching between tasks all the times, more than once in most cases. He was the only one among the three who accessed the instruction.

It is evident from the summary above and the description in section 5.5 that even when the learners scored the same in the task, they dealt with it differently. The implication of these results will be discussed in the next chapter. This now brings us to the third research question

5.7 User-behaviour tracking technologies

Research question three: *How can the use of user-behaviour tracking technologies (log files) help us explain the variability in performance among learners?*

How did log files facilitate the data collection and analysis? This research question has no accompanying hypotheses. This section is therefore more of an evaluation of the use of log files and whether it was beneficial in answering the questions about the effectiveness of input and the factors affecting performance than a question about the learners' performance per se.

In this study, the use of log-files was crucial, as it was the main method of data collection. Scripts were encrypted in the software package to log every event or action the learner took, the time it took place and the outcome of it (correct, incorrect). The data needed to answer the research questions included information on the performance of individual learners and groups, detailed information about the decision and processes taken during the experiment.

Tools used in similar studies that focus on input performance and learners' behaviour include; observation, think-aloud verbal protocol, retrospective verbal protocol, focus groups, questionnaires and more recently, screen recording. As discussed in Chapter Four, these tools were not suitable for the current study for the following reasons. Observations were not suitable because there is no way the observer will be able to capture all activities. Think-aloud or retrospective verbal protocols were unsuitable because the learner will not be able to remember everything he did during the experiment; even simultaneous verbal protocols were not suitable as they put extra load on the learners' cognitive system which would affect processing time and patterns. A focus group was ruled out because these data are not useful for the kind of data needed for the experiment as detailed record of the actions won't be accessible through focus groups. Focus groups would have been useful if we were interested in why the learner behaved in a certain way, as important as this is, it is outside the scope of the current study. Questionnaires would have been useful if we were surveying learners' views on the actions they are taking but, similar to focus groups, they will not generate the data needed. Finally, screen-recording or screen-casting, while increasingly used to monitor learners' behaviour in TELL environment and extremely useful to reconstruct the learner's actions is very time-consuming to code and analyse.

If log files had not been used in this study, more than one method should have been used to collect the data that was collected using log files alone. Log files were the most efficient and reliable method to collect all the detailed information needed in a format that would allow quantitative and qualitative analysis and would not interfere with the learners' performance or processing. The choice of methods was strongly dependent on feasibility and exclusion of unrelated variables as it is always recommended to control for as many variables as possible to have valid results and eliminate the errors. The use of log files will be evaluated further in Chapter Six by integrating the above observation with the literature review on log files presented in Chapter 4.

5.8 Conclusion

The main focus of this chapter was the quantitative and qualitative analysis of the data. Group and individual similarities and differences were examined and answers to the hypotheses were proposed. In the next chapter, I will turn to discuss further the results presented here in light of the theories and models presented in Chapters Two and Three.

Chapter Six

Discussion and Implications

6.1 Introduction

The present study was designed to examine how knowledge is constructed in a technology-enhanced language learning environment. The study employed user-behaviour tracking technology in the form of log files to investigate the relationship between the learning product and the processes learners engaged in. Data were analysed to look closely at the effectiveness of three different task-based types of input in a TELL environment along with the factors that contributed to attainment. These analyses were presented in Chapter 5.The use of user-behaviour tracking technology was also briefly evaluated in that chapter. In Chapters 2, 3 and 4, I reviewed the relevant literature, set the theoretical framework and described the methodology. In the present chapter, I consider the results presented in Chapter 5 in light of the literature discussed earlier and discuss the implications of these results from theoretical and practical perspectives offered in the literature. In section 6.2, I restate the research questions and the hypotheses the study tested and reject or accept them based on the results. I then move on to a more detailed discussion of the effectiveness of instruction in a TELL environment in section 6.3, and the effectiveness of the three different types of instruction in section 6.4. I next look at the factors that contributed to variation in the learners' performance in section 6.5. I examine individual differences in section 6.6. I turn then to consider the use of log files as a method for data collection in section 6.7. Finally, in section 6.8, I summarise the discussion.

6.2 Hypotheses: verified or rejected

The present study set out to test two hypotheses, restated here, and find answers to one exploratory question. I start the discussion by providing brief answers based on the results reported in Chapter Five

<u>Research Question One:</u>

Which type of input, Focus on Forms/FoS, Focus on Form/FoF or Focus on Meaning/FoM, is most effective in a TELL-based environment?

The aim of the first research question was to determine the effectiveness of type of instruction in a TELL environment. The following hypotheses were formulated:

Null Hypothesis: There will be no differences in performance among the three groups.
Alternative Hypothesis: There will be differences in performance among the three groups.
1.a There will be differences among the three groups in their overall communicative performance score. The FoM group will outperform the other two groups.
1.b There will be differences among the three groups in their target linguistic performance score. The FoS group will outperform the other two groups.

Results indicate that FoF instruction, which involved focus on meaning with incidental focus on form, led to significantly better gains in relation to task completion/communicative performance as operationalised in participants' total scores; focus just on meaning, with no resource to explicit grammar, namely FoM instruction, was less effective while the FoS instruction, which involved explicit grammar instruction, was the least effective. Type of instruction did not lead to any differences in relation to linguistic performance operationalised as 'target linguistic performance score'.

Table 6.1 Effectiveness of the type of instruction on performance

Variable	Group Differences
Overall communicative performance score	FoF vs. FoM and FoS
Target linguistic performance score	No significant differences

Therefore, the Null Hypothesis that there are no differences among the three groups is partially accepted. Sub-hypothesis 1a was accepted while sub-hypothesis 1b was rejected. The Alternative Hypothesis is accepted for overall communicative performance: Focus on Form, i.e. focusing on both forms and meaning, is most effective in a TELL environment.

Research Question Two:

What factors or decision processes exhibited by the learners while dealing with the different types of input contribute to differences in attainment? How do these factors/processes map on to performance?

The two parts of the second research question are related to the aspects of treatment that affect learners' attainment in a TELL-environment for each of the groups. Trial number, task type, modality of input and access to grammar help were identified as factors that lead to differences in performance. The following hypotheses were stated:

Null Hypothesis: There will be no differences among the three groups on how they deal with the input.

Alternative Hypothesis: There will be differences among the three groups on how they deal with the input.

As discussed in Chapters Three and Five, sub-hypotheses were formulated for this question based on literature of what might affect performance in a technology-enhanced environment; the following experimental variables were identified: task type, modality of input, trial number and grammar access.

The results were first compared across the three input groups to identify the factors contributing to group differences. Trial numbers and task type were significantly associated with type of instruction (Table 6.2). The FoS group had significantly fewer trials than the FoM and the FoF. Group differences were also found in relation to task type with the drag and drop task leading to group differences in communicative performance, linguistic performance and trial number.

Table 6.2 Factors contributing to group differences

Variable	Group Differences
Processing Time	No significant differences
Number of trials	FoS vs. FoM and FoF
Task Type (Drag and Drop)	(FoF >FoM >FoS) communicative performance
	(FoF >FoS > FoM) linguistic performance
	(FoS > FoF > FoM) trial numbers
Modality of input	No significant differences
Grammar access	No significant differences

The results were then analysed for all learners to identify the factors that contributed to variation in performance among learners regardless of the type of input they received. Processing time, number of trials, task type and modality of input have all contributed to differences in at least one measure of performance among learners, as can be seen from Table 6.3.

Table 6.3 Factors contributing to variation in performance across all the groups

Factor	Communicative	Linguistic
Processing Time	Yes	Yes
Number of trials	No	Yes
Task Type	Drag and Drop	Drag and Drop
Modality of Input	Yes	No
Grammar Access	No	No

In light of these results, the Null Hypothesis that there were no differences among learners on how they deal with input was rejected and the Alternative Hypothesis was accepted. As can be seen from Tables 6.2 and 6.3, differences exist among groups and among learners.

The third research question was an exploratory one about the use of user behaviour tracking technologies as a research tool. I discuss this in section 6.6 below. I now discuss the results in turn, in terms of (1) Effectiveness of instruction; (2) Effectiveness of the three types of instruction; (3) Factors contributing to performance.

6.3 Effectiveness of instruction in a TELL environment

As has been highlighted earlier, the aim of the present study was not to investigate the acquisition of reported speech but to look closely at the learning process by tracking what learners do or do not do when dealing with different types of input in a TELL environment. In order to map learners' behaviour to their attainment, intake (the processed input as identified by Corder, 1967, 1978; Krashen, 1981; Gass et al. 1998; VanPatten 1993, 1996) was taken to represent performance. This was operationalised through two scores, overall communicative performance scores and target linguistic performance score. In this section, I report on the results in relation to the effectiveness of instruction on learners' performance in a TELL environment. All learners improved after the treatment in relation to their overall communicative and target linguistic performance scores, and the improvement was highly significant. Instruction, whether meaning- or form-focused, also affected learners' confidence, with most learners going through fewer trials in the post-test. Instruction also had an effect on learners' decision to listen to the audio file, with fewer learners listening to the audio input during the post-test.

The effect of instruction found in the present study questions the claims in Krashen (1985, 1993) that instruction does not have an effect on learners' development of language.[24] However, the study's findings corroborate those of a great deal of the previous work in this field that has shown that instruction, mostly in laboratory and classroom settings, does facilitate language development, e.g. Spada (1997); Norris & Ortega (2000); Ellis (2001); de Graff and Housen, 2009; and Spada and Tomita (2010), among others. All types of instruction, both form- and meaning-focused, in the study led to improvement of learners' overall performance and linguistic performance which can be taken as an indication of facilitation of language development. That is, learners showed they attended to form and meaning better after the treatment. This supports Ellis' (2005) claim that 'while instruction may not always be necessary to achieve performance in the L2, it undoubtedly helps' (Ellis, 2005, p.725). The findings also support the argument by other researchers who see a beneficial role for instruction. In most cases these researchers do not make the distinction between acquisition and learning which Krashen (1985) and Schwartz (1993) make, or they do not support the interface hypothesis, that learning cannot become acquisition (Long 1983, 1988; Rutherford and Sharwood-Smith 1985; Ellis 1991, 2008; Norris and Ortega, 2000, DeKeyser, 2000; Doughty, 2003; White, 2003; de Graff and Housen, 2009).

Instruction also affected learners' confidence in their responses, with most learners going through fewer trials after treatment. This result supports previous research which suggests that instruction does influence L2 'propensity' (motivation, anxiety and confidence; Dörnyei, 1998, 2003; Platt and Brooks 2002; Ellis, 2003). Improved linguistic performance after the treatment seems to support the claim that instruction which leads to higher levels of conscious awareness in turn increases the learner's chances of noticing forms in the input (DeKeyser, 1998; Ellis, 1995, 2001; Sharwood Smith 1991, 1993, 2008, 2009; Doughty and Long, 2003).

[24] It is important to note here the difficulty in knowing whether this was on acquisition or on learned knowledge. However, the present study did not attempt to distinguish between the two.

Based on the discussion above, we can conclude that instruction is effective in a TELL environment and leads to communicative and linguistic gains as well as affective ones. Furthermore, instruction in a TELL environment might increase the chances of the learner's noticing forms and making form-meaning connections which then lead to improvement in task completion and also to gains in linguistic knowledge. The differences between learners' performance on the pre- and post-tests were statistically highly significant. However, these results are offered with caution here as there was no uninstructed control group. (See Chapter Four for reasons why no such group was included, section 6.7 below for the shortcomings of the experiment and Chapter Seven for limitations of the study.)
One of the factors that mediates the effectiveness of instruction is type of instruction (de Graff and Housen, 2009), and we now turn to the effectiveness of type of instruction.

6.4 Effectiveness of the three types of instruction in a TELL environment

I start the discussion here by reminding the reader of the three types of instruction used in the present study and what these entailed: The Focus on Meaning/FoM group received input rich with instances of the target structure without any explicit grammar instruction expected or available. The task design was based on Ellis'(2009) 'unfocused' tasks (see Chapter 2, section 2.5.2.1). The Focus on Form/FoF group received the same rich input but were provided with a 'grammar help' button which would allow them to focus on form-meaning relationships as and when they wanted to do so. Thus focus on form in this case was contextualised in meaning, and any recourse to explicit grammar information was learner-initiated. The task design corresponded to Ellis' (2009) 'focused' tasks. Finally, the Focus on Forms/FoS group received explicit grammar explanation before the treatment which they could not avoid and were also offered the 'grammar help' button during task completion. This task design was that of a 'situational grammar' activities according to Ellis' (2009) distinction.

In terms of the type of instruction, the FoF group significantly outperformed the FoM and the FoS group on overall communicative performance score. However, there were no significant differences between the FoM and the FoS groups on these scores, and although the FoM group performed better in the post-test, the difference was not statistically significant. No significant differences were found among the groups in terms of target linguistic performance. In other words, regardless of which group they belonged to, learners attended to form at the same level. In addition, no group differences were found in relation to processing time. There were, however, significant differences between the FoS and FoM groups and the FoS and FoF groups in the trial numbers, but not between the FoF and FoM groups. There were differences among the groups in the drag and drop task but not the information-gap task. There were no differences among the three groups in relation to the modality of input. Thus, being in a particular group did not affect the learner's decision to listen to audio input. Access to the contextualised grammar help did not vary between the two form-focused groups: FoS, who received grammar help in the form of pre-task activity in addition to the contextualised during task help, and FoF, who only received the during-task explicit rule presentation. The results are summarised in Table 6.4.

I evaluate the effectiveness of the type of instruction in relation to each of the variables in light of the literature review presented in Chapters Two and Three.

Table 6.4 Type of instruction: group differences

Variable	Group Differences
Overall Communicative performance score	FoF >FoM > FoS
Target linguistic performance score	No significant differences
Processing Time	No significant differences
Number of trials	FoS > FoF >FoM >
Task Type (Drag and Drop)	FoF >FoM >FoS (communicative performance)
	FoF >FoS > FoM (linguistic performance)
	FoS > FoF > FoM (trial numbers)
Modality of input	No significant differences
Grammar access	No significant differences

6.4.1 Overall communicative performance score

As has been pointed out earlier in Chapters Two and Four, in task-based language learning, success is defined by communicative success or task completion not by formal accuracy (East, 2012). Therefore, two measures were used in the present study to examine attainment: communicative performance, which reflects task completion, and linguistic performance which indicates attending to formal aspects of the input in terms of accuracy.

There were significant differences among the three instruction groups: the FoF group outperformed the other two groups followed by the FoM and finally the FoS. This means that the type of instruction that involved learners in attending incidentally to form was the most effective, meaning-focused instruction with no attention to form was second and explicit instruction was the least effective. A possible explanation of the observed superiority of FoF group is that the contextualised, learner-initiated focus on form increased learners' opportunities of noticing the forms but at the same time encouraged them to process the input for meaning. The less improved performance by the FoS group could be interpreted using VanPatten's (2004) Input Processing model. As the learners in the FoS group received explicit instruction, their orientation would have been to focus on both meaning and form simultaneously; under VanPatten's model, learners always find it difficult to process the input for meaning and form at the same time so this would have affected their performance and depressed their scores.

The results also partially support the conclusions to which Norris and Ortega's (2000) meta-analysis led: the effects of explicit instruction are larger and more durable than those of implicit instruction. The results of the present study are more nuanced in that they showed effects of incidental focus on form/FoF to be larger followed by focus on meaning/FoM and finally explicit focus on forms/FoS. The results also support Lightbown and Spada's (1993) argument that an extreme model of getting-it-right from the beginning, i.e. teaching grammar explicitly to avoid errors, does not benefit effective communicative performance. One of the behaviours underlying the performance of the FoS group was attempts to formulate rules and apply them when dealing with the input since learners were aware that the focus of instruction was on both form and meaning. As their knowledge was not well developed (as reflected in their target linguistic performance scores), this led to worse performance. In this sense, the explicit treatment may have inhibited learning.

The results seem not to support claims made by those who have argued that comprehensible input is a sufficient type of input for acquisition to take place (Krashen, 1981, 1982). The results also go against those who found that explicit instruction is beneficial (De Graff and Housen, 2009; Spada and Tomita, 2010) and supports the research that advocates that an incidental, learner-initiated focus on form is the most effective type of instruction (Long, 1998; Doughty, 2003). The results also indicate that it is enough for the learner to know (be aware) that the focus of the task is on meaning or form or both to orient differently to task completion (see sections 6.4.7 and 6.5).

I move on now to look closely at the effects of the different types of instruction on target linguistic performance score before I fully discuss the effect of the three types of input on performance in section 6.4.3.

6.4.2 Target linguistic performance score

The two form-focused groups slightly outperformed the meaning-focused group in relation to target linguistic performance. This means that learners who received explicit (FoS) or incidental (FoF) grammar instruction scored higher than learners who received no grammar instruction at all (FoM). However, these differences were not statistically significant. There are several possible explanations for this result.

One possible explanation for these results is that a large number of learners in the form-focused groups (FoF and FoS) did not consult the grammar help (42%). This could be a direct effect of the learning environment. In the classroom, learners are faced with communication difficulties which push them to attend to form. It seems that in the self-access TELL environment where the communication is learner-machine, learners choose not to focus on form. This is also evident in the results of the time spent on the pre-treatment explicit instruction. Learners in the FoS group varied considerably in how much time they spent on the explicit instruction: Min=8 seconds, Max=283 seconds, Mean= 95 and SD= 75.2. These results could also be a consequence of the relatively small sample size (71 learners) and the high proportion of learners who chose 'no grammar' focus. A larger sample might have yielded a different outcome.

6.4.3 Type of instruction and performance

The results summarised above could be explained in light of previous studies. As has been pointed out in Chapter 2 in section 2.4.1, 'there is considerable variation in how the term *focus on form* is understood and used' (Doughty and Williams, 1998, p. 5). One of the key findings of the present study is that the type of instruction affected overall communicative rather than target linguistic performance. This could be explained in relation to different factors. One of the crucial arguments is the communicative value of the linguistic structure (Reported Speech). The target form was not communicatively salient, i.e. errors do not result in communication breakdown. FoF is claimed to be more effective if attention to form takes place incidentally as a result of communication breakdown (Long, 1996 among others). As the target form in the present study did not have this characteristic, it meant that learners attended to form regardless of whether their input was FoF, FoM or FoS. Thus in the case of less salient structures, the effect of FoF instruction could be limited. According to Long (1998) and Doughty (2003), the beneficial effect of FoF instruction is in accelerating the passage through the sequences of L2 development and extending the scope of application of grammatical rules. Thus, FoF results in improving accuracy, rate of learning and level of ultimate attainment if the target structure is communicatively salient. The same argument is echoed in Klein (1986), VanPatten (1985, 1990_ and Leow (1995): learners appear to attend to linguistic forms based on their communicative value. Long and Robinson (1998) state that when forms cause little or no communicative distress, they are less likely to be acquired without instruction. As the results of the present study indicate, the acquisition of communicatively non-salient forms is not affected by the type of instruction the learners received. As indirect speech is not of a high communicative value, it is quite possible that the choice of construction affected how the learners dealt with the input

Another explanation for the results is related to the nature of FoF instruction provided during the experiment. The inconsistency in the literature in the use of the terms means that it is important sometimes to re-establish the meanings associated with acronyms when reviewing some studies. Ellis, Basturkmen and Loewen (2002) divided FonF (their acronym for any type of focus on form) instruction into pre-emptive and reactive. They argued that reactive FonF can be conversational or didactic and implicit or explicit. For example, corrective feedback in the form of recasts is a type of reactive implicit FonF, while the teacher telling the learner what is wrong is a type of explicit reactive FonF. On the other hand, pre-emptive FonF, could be teacher initiated or student initiated. In their study of the type of focus in the classroom, Ellis et al. found that pre-emptive FonF was as common as reactive FonF. They also concluded that pre-emptive FoF is most likely student-initiated and that new language forms encountered during pre-emptive FoF were more likely to be taken up and used by the learners subsequently. The findings of the present study contradict those of Ellis et al. in the sense that most learners did not focus on form even when there was a break in communication, as indicated by the number of attempts on certain items. The results also revealed that even the FoS group, who received explicit pre-emptive focus on form - what could be described in the classroom as teacher-initiated focus on form, did not perform better than those who did not receive any focus on form. Contrary to Ellis et al.'s findings, pre-emptive focus on form does not seem to work in a TELL environment, at least for this particular construction.

Reviewing the literature on the different types of FoF, a number of empirical studies have investigated the frequency of these different types of FoF and their contribution to learner uptake. Loewen (2003) found great variability both in the frequency of FoF episodes, and also in the extent to which individual students took part in these. Similarly, Mackey, Polio and McDonough (2004) compared the use of FoF by novice and expert teachers. Their findings were that experienced teachers used significantly more pre-emptive FoF than inexperienced teachers. The inconsistencies in the findings of studies that looked at different types of FoF or used different types of instruction mean that firm conclusions are not possible (see Norris and Ortega, 2000, and Mackey and Goo, 2007). There have been calls for experimental work that uses mixed methods and longer treatments to yield more valid conclusions (see Chapter 2, section 2.3.2 for discussion and Ellis et al. 2006 for an example). These calls are strongly supported by the findings of the present study.

Moreover, the results of the study contradict the findings of meta-analyses that showed that explicit types of instruction are more effective than implicit types in that they draw the learner's attention to the target form (Schmidt, 1990, 1993, 1995, 2001; Sharwood Smith, 1991; 1993; Tomlin and Villa, 1994; VanPatten, 1996, 2002, 2004, 2007, 2009 among many). In their meta analysis, Norris and Ortega (2000) argue that the effects of explicit instruction/Focus on Forms (FoS) are larger than those of implicit instruction or Focus of Meaning (FoM). In the present study, the effects of incidental focus on form instruction were shown to be overall greater than those of either explicit (FoS) or implicit (FoM) instruction. The results are instead similar to those studies that found no facilitative effect for explicit rule presentation (e.g. Sanz and Morgan-Short, 2004), and are opposite to those that found a facilitative effect for explicit rule-presentation (e.g. Robinson 1996). The results are similar, though, to those studies that showed that any type of form-focused instruction is effective Long (1991, 1996, 1997, 1998, 2000); Ellis (1994, 2001); Spada (1997); Doughty and Williams (1998); Long and Robinson (1998); Doughty (2003) and Nassaji and Fotos (2004).

Looking at the effectiveness of the three types of input in relation to the processes that the learners exhibited, we observe that the FoF learners went through many attempts while completing the task, while the FoS learners went through fewer attempts. The FoM learners, however, went through the most attempts. This is in line with Widdowson's (2008) argument that learners do not learn by adding items of linguistic knowledge but rather by a process of continual revision and reconstruction; that is, learning is 'continual cognitive adaptation as the learner passes through different transitional stages' (Widdowson, 2008, p. 211). In a classroom, investigating Widdowson's statement is difficult as collecting data about the process of revision and reconstruction is difficult. However, in the present study, using the user-behaviour tracking technology, it was possible to link the process of reconstruction to measures of performance, and the findings strongly support Widdowson's argument. The present study's results also indicate that the FoF input resulted in the most efficient process of reconstruction where the learners went through considerable attempts that led to the highest communicative scores. This will be discussed further when looking closely at the behaviour of the three learners who received the three different types of input.

The results also do not support Chapelle's (1998) suggestion that technology offers learners better chances of noticing the forms when they have control of when and how often to focus on form. On the contrary, it seems that in a self-access TELL environment where learners have control over when and how to focus on form, they decide not to do so. The reason behind this contradiction is that Chapelle's suggestions were more directed to a blended and teacher-directed TELL environments while the environment used in the present study was a self-access one. The significant differences in overall communicative performance score and lack of such differences in target linguistic performance score are consistent with VanPatten (2004, 2007).. As noted several times above, VanPatten (2004, 2007) argues that learners always process input for meaning before they do so for form. The results in the present study are compatible with VanPatten's as most learners chose not to focus on the form even when they could have.

A possible explanation of learners' behaviour is related to the quality of input provided. The text for the task in the present study was based on one of most popular language textbooks, *Headway*. When designing the software, I adhered as much as possible to the original text as one of my interests was to identify the effectiveness of the type of input presented in typical classroom and textbooks, which is often not based on any SLA findings. In this sense, it is possible that either the input was not comprehensible or that there was not enough of it for learning to take place. However, performance by the learners on the treatment tasks was good and this would be unlikely if participants had been unable to comprehend the input. I therefore turn now to the discussion of the target items included in the input.

As was highlighted in Chapter 5, not all the target items were processed at the same level (see Table 5.9). Looking at the scores of the individual items, it was obvious that learners found the past perfect more difficult to process than the past tense. As early as the 1980s, it was pointed out that not all linguistic items in the input are attended to equally by L2 learners (e.g. McLaughlin, Rossman and McLeod, 1983). As discussed above, the observation is that learners appear to attend to linguistic forms based on their communicative value or saliency in the input or other unknown reasons (VanPatten, 1985, 1990; Klein, 1986; Leow, 1995). If, according to Long and Robinson (1998), forms which cause little or no communicative distress learners are less likely to be acquired without instruction, it could be argued that the learners in the present study indeed noticed that they needed to change the tense of the target items. However, in most cases, they decided to use the past tense as the past perfect tense did not carry a high communicative value to cause a breakdown in communication for them. As the amount of input was limited in the present study, it was not possible to draw conclusions about the effectiveness of the type instruction in relation to the individual target items and forms. This is definitely an issue worth investigating further in future experiments. Moreover, in Poole's (2005) evaluation of FoF instruction, the conclusion was that FoF is only effective in particular settings and there is a need for research that gauges the appropriateness of FoF in different settings and for different learners. Research investigating the effectiveness of FoF (or any other type of instruction) in different settings is long overdue.

Another possible explanation for the results is the lack of negative evidence in the form of feedback or information about the ungrammaticality of the target forms in the construction. A number of researchers have pointed out the potential contribution of corrective feedback of some type to learning (e.g. Lightbown and Spada, 1990; Sheen, 2004). Researchers argue that such feedback can help learners 'notice the gap' between the input and their own output, and it can increase their awareness of the target form. These potential benefits were not available to learners in the present study as negative evidence was not included as a variable. The reason behind the exclusion of negative evidence in the form of feedback or ungrammatical examples was to control for its effects. Since the FoM group were not going to have any kind of focus on formal features and the FoF group were only supposed to receive incidental focus on form, negative evidence would have only been provided to the FoS group. This would have provided the FoS group with an advantage through extra input and would have made it difficult to identify whether any differences among the groups were the consequence of the type of instruction or this extra variable.

In brief, it seems that in a TELL environment, explicit instruction did not result in raising participants' awareness and thus did not result in any of the potential benefits of this kind of input (Schmidt, 2001). In other words, the explicit form-focused treatment did not result in significantly greater linguistic intake of the forms involved in the target construction, as measured by task target linguistic performance, than the implicit meaning-focused treatment or the incidental focus-on- form treatment.

6.4.4 Processing time

Group differences in relation to processing time were not significant. In other words, the explicit form-focused treatment did not result in slower processing than the implicit meaning-focused treatment. Again, this echoes e.g. VanPatten's (2004, 2007) claim that attending to both meaning and form is difficult for learners, and that it often leads to slower processing. One possible explanation for this is that the session time, i.e. total time spent on the treatment, was used for analysis in the present study rather than time on tasks for practical reasons; that is, it was difficult to analyse and calculate the time involved in a task as learners kept switching between them.

Another possible explanation is that the treatment was too short for processing differences to emerge. However, it is worth noting here that the treatment time was not short compared to other studies that investigated the effect of explicit-implicit instruction. In their meta-analysis, Norris and Ortega (2000) reported that the treatment in the majority of studies (68%) they considered was less than two hours. The time learners took in the present study, including pre-test, treatment and post-test was 90 minutes on average. Norris and Ortega also found that shorter treatments were associated with larger effects. They attributed this finding to the intensity of instruction in shorter treatments and the 'immediacy and construct proximity of outcome' (Norris and Ortega, 2000, p.474).

6.4.5 Trials

When we look at the number of trials for learners, we do find significant group differences, with the FoS group having the fewest number of trials and the FoM group the highest. However, it is difficult to explain the results in light of the instructed second language acquisition literature as studies do not usually report on the number of trials learners go through before they decide on a response. One possible explanation is that being exposed to explicit grammar rules gave the learners in the FoS group a sense of confidence, which led to fewer trials, while receiving no grammar instruction at all made learners anxious about what exactly the goal of the task was. This then led to a less confident approach, as exhibited by a higher number of trials.

It might be that participants were unsure about the aims of the task so they kept changing their responses while the participants who received form-focused instruction were guided by their awareness of a possible focus on form. The case of the FoS group having fewer trials but being less accurate overall could be explained by arguing that in their attempts to apply their under-developed knowledge to the forms involved in the target construction along with meaning, these FoS learners failed to respond accurately. Indeed, N. Ellis'(1993) study showed that learners in the explicit rule presentation condition took more trials to arrive at an understanding of the structure. His results also showed that the meaning-focused group were the fastest but less accurate. The results in the present study do not support Ellis' since the learners who received explicit rule presentation went through the least number of trials and scored higher than the other two groups in their target linguistic performance scores but lower on overall communicative performance scores.

The results highlight the importance of investigating learners' uptake, i.e. how they perform during treatment, along with their performance, i.e. the final product. Researchers who call for such research include N. Ellis and Schmidt (1997) who state that

> SLA research aspires to understand acquisition, and acquisition results from dynamic processes occurring in real time. It is difficult to gain an understanding of learning and development from observations of the final state, when we have no record of the content of the learners' years of exposure to language nor of the developmental course of their proficiencies. If we want to understand learning we must study it directly. (N. Ellis and Schmidt, 1997, p. 146).

Many argue for the importance of investigating not only the effects of a treatment on acquisition, or the final 'product', but also on the process of language learning. Looking at both the process and the product of learning might provide insights on what really triggers and facilitates acquisition. According to R. Ellis and Sheen (2006), 'Ultimately, whether uptake facilitates acquisition must be determined empirically. It is unfortunate that, to date, there have been very few studies that have examined this issue' (p. 590).

One of the reasons for the lack of such research is the difficulty of collecting and analysing data in the classroom and particularly outside the classroom where interaction is complex. With the advances in technology, such research could be conducted much more easily using videotaping and video analysis software. The use of user-behaviour tracking technologies allows the collection of such data from large number of learners at the same time. We return to the use of such technologies further below.

6.4.6 Task Type

While there were group differences in performance on the drag and drop task, there were not on the fill-in-the gap task. The interaction between task type and the type of instruction/input affected learners' performance as well as the trial numbers (see Table 6.2). The FoF group performed better on the drag and drop, FoS was second for linguistic performance while FoM was second for communicative performance. The FoS group had the fewest number of trials followed by the FoF and finally the FoM group. Possible explanations for task effect and its relationship to the type of instruction/input are presented below.

The drag and drop task might have been easier for learners because it was cognitively less demanding. This task simply required the learners to drag an existing item to the right place while the fill-in-the-gap required the learners to write in their responses; in this sense, the drag and drop was more of a comprehension task while the fill-in-the-gap was a production one. The fill-in-the-gap was more demanding cognitively as it required learners to remember the spelling of missing items in addition to thinking about the right response. Therefore, there was likely more opportunity during the drag and drop for attention to be directed at form or meaning or both. This attention then varied according to the type of instruction the learners received. In this case, the FoF group managed to attend to both meaning and form better during the drag and drop while the FoS group was the least able to do so overall. When looking at the target linguistic performance scores, the FoF and FoS groups outperformed the FoM group. This means that the task affected the two measures of performance differently for each type of instruction.

This explanation is supported by early research on cognition. Cummins (1983) suggested that tasks that require more information to be processed at the same time are more demanding. In the present study, fill-in-the-gap required deeper processing than drag and drop. More support for this explanation emerges from the fact that task characteristics were exactly the same for the three types of instruction. The only difference was how learners dealt with the task within the experimental groups. The results can also be explained by Ellis' (2009) argument that learners orient differently to different tasks and activities. This is evident in the differences in performance and also in individual differences, as revealed by the qualitative analysis.

The results also reveal the importance of using different tasks to measure different aspects of performance. Ellis (2005) suggested that 'impossible to construct tasks that would provide pure measures' of the different types of knowledge and performance (Ellis, 2005, p.153). But the ideal (impossible) task is ruled out, and it is, therefore, important to use multiple tasks. These results also have pedagogical implications which are discussed in Chapter Seven.

6.4.7 Modality of input

The present study included input in more than one mode, aural and written. Research on the modality of input (Murphy, 1997; Lund, 1991; Johnson, 1992) shows that input presented in only one mode does not give a complete picture of the processes and knowledge that are available to learners. It was essential then to present the input in more than one mode but this also meant that the amount of input presented was limited in order to control the amount and quality of data collected through log files.

No group differences were found in relation to modality of input; in other words, the FOF, FoS and FoM groups were not different in their access to the aural and written input. However, individual differences in relation to how learners utilised the audio files were significant, and they are discussed later in section 6.5. As was pointed out in Chapter 2, section 2.2.6, most SLA models of input processing do not account for modality of input, although there is evidence that suggests that modality of input does affect processing and comprehension (Bird and Williams, 2002). Leow (1995) argues that learners' intake depends on cognitive constraints and strategies which operate and are employed differently in different modes. The same argument is presented by Wong (2001).

One possible explanation for the lack of significant results could be the restricted measure used as the pre- and post-tests only included one audio file. Although the treatment included five more audio files, the data was not analysed. This leads back to N. Ellis and Schmidt's (1997) and R. Ellis and Sheen's (2006) argument about the importance of investigating the process as well as the product.

I turn now to the evaluation of the effectiveness and patterns of access to the incidental and explicit grammar instruction.

6.4.8 Grammar access

As pointed out earlier in this chapter, two out of the three treatment groups had access to grammar instruction; the FoM group did not. The FoS group had to go through explicit grammar instruction after the two testing tasks; they also had a 'grammar help' button integrated in the treatment tasks. They had total control over how much, when and how to access this help. The FoF group were instructed to complete the tasks, but were also told that there was a 'grammar help' button over which they, too, had complete control. The results showed that 40% of the learners in the FoS group did not access grammar at all during the treatment while 60% did. Similarly, 44% of the learners in the FoF group did not while 56% did. Inferential statistics showed that learners' decision to access grammar was not related to the treatment group they belonged to. There are different possible reasons behind this result which I discuss in light of the literature.

Schmidt (1990, 1993, 1995, 2001) argues that the only sufficient condition for acquisition to take place is by raising learners' awareness by intentionally focusing attention on specific elements of the input. Other researchers have also highlighted the link between attention, intake and acquisition (Gass, 1997; VanPatten, 1990; Leow, 1995, 2001). Input needs to be noticed to be taken in and for acquisition to result. It follows from this, then, that if the type of instruction did not affect intake, it cannot be expected to affect acquisition. Under this condition, it seems at first that the results of the present study do not support Schmidt's argument as no effects were found for directing learners' attention to form. In other words, directing learners' attention to form through explicit grammar instruction did not affect intake and consequently would not have affected acquisition. One possible explanation for this is that learners did not notice the forms. This result highlights the difficulty of measuring noticing (Schmidt, 1990, 1993, 1994; Robinson,1995, 2003, 2005). In the present study, learners' decision to access the grammar help was taken as an indication of focus on form but there was no way to know if this access did actually result in learners noticing the form. This corroborates Reinders' (2009) conclusion that successful performance on a task is not necessarily a good predictor of intake and of acquisition. This is linked to the teachability of formal aspects of the language. Ellis (1989) and Pienemann (1984), among others, argue that teachability is not the same as learnability, i.e. teaching a form does not mean that learners have learnt it, and definitely it does not mean that they have learnt it at the same time it was taught.

Another possible explanation for the results regarding grammar is that the explicit treatment did not appear to have succeeded in making the target structure salient to the learners. This could be due to the small amount of exposure or to the input not being rich enough. The lack of any effect of the explicit treatment on target linguistic performance also corroborates findings by those who see no role for instruction in acquiring linguistic knowledge in an L2 (Krashen, 1984; Krashen and Terrell, 1983; Schwartz, 1993). The lack of significant outcome regarding grammar effect could also be simply, as Lightbown and Spada (1993) have put it, an extreme model of 'getting it right' from the beginning, i.e. teaching grammar explicitly from the very start so that students get things right, but this does not benefit effective communication or in this case linguistic performance.

Overarching conclusions from the above discussion on the effectiveness of different types of input in a TELL environment indicate that the explicit and implicit types of instruction resulted in lower overall communicative performance scores (FoS and FoM) than the incidental instruction (FoF). It is possible that the implicit treatment could have made participants unsure about their answers while the explicit treatment meant that the learners knew they had to pay attention to the target structure but did not have enough knowledge to successfully complete the tests and then in their endeavours to focus on form, they failed to focus on meaning. Overall, the results lend support to VanPatten's (1996) claims that learners process input for meaning before they process it for form when they are under pressure. Learners focus on being effective more than on being grammatical. It is clear that these claims are supported in TELL environments.

Slight differences among the three groups were observed in relation to some factors but they did not reach statistical significance. Possible reasons for this could be that

1. The treatments had similar effects. This is possible in the sense that they were simply not different from each other in how they affected performance.
2. All treatments had very little overall effect, and, the effect was only a practice effect and did not lead to statistical significance. In other words, the treatments may have been unsuccessful in encouraging participants to process the input for anything other than dealing with immediate task demands.
3. It was because of limited exposure; there was not enough time for the differences between the treatments to appear.
4. Differences existed but for some reason were not evident based on the calculations carried out.

This section dealt mainly with group differences which were discussed according to the three different types of instruction. I turn now to look at individual differences using quantitative (across all learners) and qualitative (three learners, one from each group) data.

6.5 Individual differences in a TELL environment

The discussion above focused on differences as moderated by the type of instruction. In this section the focus is on individual differences. First, I look at the individual differences from a quantitative perspective where all learners are included in the analysis regardless of their instruction group and their performance is analysed against contributory factors. I then look closely at the qualitative analysis of three learners, one from each group, to identify instruction-related individual differences. The results from the quantitative analysis were summarized in Table 6.3 above which is presented here again for ease of reference

Table 6.5 Factors contributing to variation in performance across all the groups

Factor	Communicative	Linguistic
Processing Time	Yes	Yes
Number of trials	No	Yes
Task Type	Drag and Drop	Drag and Drop
Modality of Input	Yes	No
Grammar Access	No	No

The results indicated that processing time was significantly associated with performance; learners who spent less time processing input scored higher on overall communicative performance and higher in the target linguistic performance. This could be explained by reference to explicit and implicit knowledge. It might be the case that learners who processed the input faster were drawing from their implicit knowledge which is believed to be the result in better/successful intake and to involve faster processing, while learners who took longer to process the input might have been drawing on their explicit metalinguistic knowledge, which typically involves slower, controlled processing. This explanation is, however, presented with caution as different measures of the types of knowledge, which are recommended by R. Ellis (2005), were not included in the present study.

There was a significant association between the number of trials during the pre-test and target linguistic performance score. Learners who had fewer trials scored higher in relation to the target linguistic performance score. As suggested above, this result indicates that learners who felt confident about their responses scored higher in relation to the target form. A possible explanation for this is that learners who felt confident perceived the treatment with a positive attitude, which helped them to attend to form as the cognitive demands were fewer. It is possible that when learners were not sure about their responses, this placed extra demands on them and this affected their ability to attend to form.

Task as a factor was a major element in the present study. The results indicate that the two task types elicited different types of performance and posed different types of cognitive and communicative demands. It was argued in Chapters Two and Three that a task can be perceived or made more complex by increasing its cognitive load and the constituent features of the task's structure. In the present study, it seems that learners perceived the drag and drop task as easier which meant the cognitive load was not large so their attention was directed to meaning or form or both.

There were significant differences in terms of overall communicative performance scores among those who listened to the audio input during the pre-test but not the post-test. Learners who listened during the pre-test scored higher in the post-test. On the other hand, listening to the audio input did not have any effect on learners' target linguistic performance scores.

The results indicate that listening to the audio input led to higher communicative performance scores. The effect of the aural input on these scores supports the finding of Nassaji's (2004) study where he found that a combination of modalities led to better gains. However, the results of the present study do not support the findings of others, namely Johnson (1992), Lund (1991) Murphy (1997) and Wong (2001), who all found greater effects for visual/written input. Neither do the results of the study mirror Leow (1995), who observed differences among learners where aural input led to better gains in relation to the linguistic form. In the present study the aural input led to communicative gains represented in better communicative task completion rather than linguistic performance scores. One justification Leow (1995) provided for his own results is that the learners in his study were able to attend to the phonological aspect of the linguistic form when exposed to aural input. This was not the case in the present study as the forms in the target construction (reported speech) are not particularly phonologically salient.

One explanation for higher overall communicative performance in response to audio input is that attentional resources are not directed in the same way during the processing of aural and written input (Wong, 2001). Leow (1995) maintains that 'readers are generally regarded to be less cognitively constrained by their exposure to L2 data than are listeners' which explains the benefits of written input. However, the results of the present study indicate that this is not always the case, which highlights the need for further research particularly when taking into account the different patterns of listening exhibited by the learners when dealing with the aural input. Control may also be an issue. Zhao (1997) found that the listening comprehension of learners who were given control over audio input improved. Since learners in the present study had complete total control over the audio input, this may be a factor. I return to this below when discussing the qualitative results.

There were no differences between learners of the FoS and FoF groups who accessed the contextualized grammar help and those who did not in terms of communicative or linguistic performance or the number of trials. Looking closely at the incidents preceding or following grammar access revealed that the participants were not using the 'grammar help' facility in a blind or random fashion; they approached the task in a strategic manner taking into account the relevance of the task goal. Nevertheless, the two groups did not differ in their access pattern. However, there were significant differences among learners in the FoS group: the more time the learners spent on grammar, the lower their scores were in communicative and linguistic performance. This is compliant with VanPatten's (1990) and Bransdorfer's (1991) findings that attending to both meaning and form in the input affects comprehension. It also supports the explanations proposed above that the provision of explicit knowledge made the learners aware that they needed to attend to form and meaning but with their under-developed knowledge, they failed to be accurate.

In N. Ellis' (1993) study of the effects of implicit and explicit learning on the acquisition of soft mutation) in Welsh, of his experimental groups (exposure only, accompanied by rule presentation or with both rule presentation and exposure/examples). The exposure only group learned faster but showed little implicit knowledge and poor performance in explicit knowledge. The 'rule presentation' group took more trials to arrive at an understanding of the structure, and the learners were able to verbalize the rules but they failed to apply them in practice. Finally, the rule-presentation-plus examples were the slowest learners; however, they were the only ones able to 'abstract a working knowledge' of the structure. They were able both to verbalize their knowledge explicit rules as well as implicitly generalize these rules to new structures. Contrary to Ellis' results that learners in the explicit rule presentation took more trials to arrive at an understanding of the structure, the learners in the FoS group in the present study had less trials than the FoF and FoM learners. The result that learners who spent more time on the grammar instruction scored less also contradicts Hulstijn and Hulstijn (1984), who found positive effects for focusing learners' attention on grammar.

The importance of the results is that they show that even when overall group differences exist, they do not reflect the actual differences among the groups. In the case of the present study, group differences were found for communicative performance scores but not for linguistic performance scores. However, when this was analysed against the type of task, it was revealed that the groups were different in both measures for the drag and drop task and in neither for the fill-in-gap task. The implications of this for pedagogy is crucial, and particularly for assessment where different types of tasks should be used to get a comprehensive picture of the learner's knowledge.

In order to have a closer look at individual learners' behaviour, data from three learners, one for FoM, FoS and FoF were compared. As was the case for the quantitative group data, the three learners' scores and trial numbers were very close. The qualitative analysis revealed that, although there were shared patterns of behaviour across the three learners, differences were evident. Even when the learners spent the same time, scored the same and had the same number of trials, the way in which they dealt with the input was different. Based on the detailed description in section 5.in Chapter Five, the following shared patterns were identified:

- Learners start dealing with the items they are most familiar with or most confident about. This is evident from the number of trials and the scores. The first items answered were most likely correct with only one trial each. These patterns lend support to VanPatten's (1996, 2007) principle that meaning is always prioritized by learners and once the meaning is comprehend, learners might direct their attention to form.
- Learners answer the questions in order when they perceive the task as easy, and they sometimes leave difficult items blank. They answer in an apparently random order when they perceive the task as difficult.

Learners exhibited varied patterns when accessing the audio input, over which they had full control. The patterns that emerged from analysing the qualitative data were that the learners

- listened to the whole audio file then dealt with the task
- listened to it bit by bit and dealt with the task while listening
- listened to it and dealt with the task simultaneously
- dealt with the task and then used the audio input to recheck answers

It is evident from the summary above and the description in Chapter Five that even when the learners scored the same in the task, they dealt with it differently. The implications of these findings will be discussed in the next chapter. Overall, the patterns exhibited by each learner could be recapitulated as below.

The qualitative analysis revealed that there were individual differences among the three learners in terms of the patterns they exhibited when dealing with the input. The FoS learner was more confident and behaved in a linear way. He hardly used the contextual grammar help. At the same time, going through the pre-task explicit grammar instruction gave him a false sense of confidence and in his attempt to apply his knowledge, he failed to be accurate. The FoF learner behaved in a linear way at the beginning and then in a non-systematic way. He resorted to using the resources when stuck, that is, not as much as he could have. Qualitative analysis also revealed that this contextual focus on form which was initiated by the learner was the most effective as it often resulted in the learner modifying his/her response in the right direction. The FoM learner behaved in a non-systematic way, and the lack of any focus on form revealed a sense of the learner being lost and not knowing what to do as he kept switching between items and tasks.

The pattern of behaviour exhibited by the FoM learner is consistent with Manning's (1996) findings. Manning's study investigated the merits of exploratory learning as opposed to the explicit and implicit approaches. The study included three groups; an exploratory group who were given control to choose explicit and implicit modes as and when they wanted, an implicit mode group who were provided with examples, exercises and revision and an explicit mode group who received explanations, examples, and exercises. The results indicated that the exploratory mode led to faster, more efficient learning, but created problems with navigation and decision making. The qualitative analysis of the data from the three learners revealed similar results in terms of navigation and decision making problems, particularly for the FoM group.

The FoS learner was more confident about his answers as he went through fewer trials and spent less time on the tasks. He listened to the audio input when there was a difficulty completing the task. The learner listened to the audio input from the start only when he perceived the whole task as difficult; this is evident in more trial numbers from the beginning. He only accessed the grammar help when the audio file did not help. When he perceived the task as easy, he did not access the audio input or the grammar instruction.

The FoF learner exhibited two distinct patterns of behaviour. In the drag and drop task, he listened to the audio file extensively and answered the questions randomly while in the information gap, he filled in all the gaps in order and even rechecked them in order. He followed a similar pattern when listening to the audio input: listen, pause and answer. He has also used the audio input as a way of checking his answers. He sometimes switched between tasks.

The FoM learner displayed a clear pattern of dealing with the audio input as an essential part of the task. He listened to the audio input wherever available and always used it in the pattern: listen, pause and answer. He seemed the less confident about his answers as he kept going back and changing them all the time even when they were correct. He checked all the answers at the end of each task. He also showed a clear pattern of switching between tasks all the times, more than once in most cases.

The qualitative analysis findings are extremely important considering that individual differences could have been easily ignored if only the quantitative data was used. This would have resulted in misleading claims about the processes learners exhibit when dealing with the input. The profiling of learners was used in the present study to get close to the process of learning in a TELL environment. The analysis showed a general pattern among learners to switch between items, or what is referred to as a ‘flip-flop’ learning behaviour according to Lai and Hamp-Lyons (2001). Lai and Hamp-Lyons claim that this flip-flop behaviour is considered a sign of active engagement in the learning process. In the present study, considering that the treatments were exactly the same apart from the grammar, it is possible that the individual differences observed were influenced by individual differences in learners’ aptitude for language learning, their motivation to seriously engage with the treatments, or various other individual differences.

There are different possible explanations for such behaviour. It could be interpreted in two ways, either the learners are lost and do not know what to do or that they are searching for the right item to do first then move to another. The right item might be the most familiar, the one they know the answer to, the one that they perceive as easy, etc. The qualitative analysis showed that the learner in the FoS group exhibited this behaviour only at the beginning of a task to answer the items the was familiar with or he knew the answer to. He then completed the task in order. The behaviour of the FoF and FoM learners was a mixture of being lost and answering the familiar items first. It was clear that the FoS learner was the most confident about his answers while the FoM was the least. This is also supported by the group results in relation to trial numbers.

Collentine (2000) used 'user-behaviour tracking technologies' to record all the events that learners generated while using a specially designed software application. He found that while learners in a TELL environment are expected to learn actively and independently, the following could be observed: learners did not exploit opportunities to engage in exploratory strategies even when they were available, most learners provided very short answers and those who did so did not benefit from instruction and learners kept moving backwards and forwards between the slides available to explore the information before answering any conscious-raising activities. Collentine also concluded that the use of tracking technologies provided insights into the processes of L2 acquisition that are difficult to observe using other techniques.

In the previous sections of the present chapter, I have looked at the group and individual differences and discussed them in light of the relevant theoretical frameworks. I turn now to the third research question, which is a methodological one.

6.6 Log files: a technology enhanced research tool

Research question three asked: ***How could the use of user-behaviour tracking technologies help us explain the variability in performance among learners?***

Log files or user behaviour tracking technologies were used as the main method of data collection. Scripts were included in the software package to record every action each learner took. Their use was evaluated in the previous chapter, and it was concluded that the use of log files was very successful in providing valid and reliable data to answer the main research questions.

It could be argued here that more than one method needed to be used to collect the data that was collected using log files only so the log files were really useful in this sense.

Thus, log files were the most reliable method to collect all the detailed information needed in a format that would allow quantitative and qualitative analysis. Moreover, by using user-behaviour tracking technologies, it was possible to link the product to the process.

These results support the advantages of log files as identified by other researchers. For example, Cubillos (1998) argues that students' logs of all kinds can provide teachers with unprecedented insights into their students' SLA processes (p.45). The same argument is echoed by Garrett (1998), who suggests that the only way to make claims about the efficacy of technology use in language teaching is by using tracking software. Garrett (1998) and Hulstijn (2000) both highlight the need for and importance of the use of log files for experimental research. Garrett (1998) maintains that what is needed is data 'on what students do with technology-based language learning materials while they are in the act of working with them' (Garrett, 1998) while Hulstijn (2000) argues that log files allow an unobtrusive observation of participants' look-up behaviour, 'with these computer-aided tools, however, researchers have the means to get closer to the processes of language acquisition and use' (Hulstijn, 2000). More recently, Chun (2013) reviewed the contributions of tracking user behaviour to SLA research. According to Chun, the key benefits of collecting and analysing data using log files is that they allow us to ascertain precisely what learners do or do not do and determine whether there is a relationship between what they do/do not do and learning. In simple terms, log files allow the documentation of learning process in a way that provides a complete picture of how knowledge is constructed.

In the present study, log files allowed the collection of a comprehensive record of all interactions and activities; they were non-invasive, so the risk of altering participants' behavior was non-existent. Moreover, they are easy to use to collect data once the coding is set and are not biased by the subjectivity of the researcher or by external variables. Yet two main problems emerged in relation to collecting and analyzing data using log files. The first is that in most cases, it requires working knowledge of programming languages. The second and the major problem, one observed in the present study, is that it is not easy to make meaning out of the vast amount of data. This problem is exacerbated by the lack of any existing models of analysis. As Bruckman (2006) has pointed out, 'log file data is more often collected than analysed' (2006, p.1449). The same argument is presented in Chun's review (2013). Chun highlights the main reasons for not incorporating user-behaviour tracking technologies in experiments as the 'sheer quantity of available data and the time required to process and analyse them' (Chun, 2013, p. 256). The other problem is data mining, in other words, identifying key actions and extracting meaningful information from the data. As Chun points out, in the future, it should become easier to collect and analyse such data but for now more research is needed to refine data mining.

6.7 Shortcomings of the experiment and their effects on the results

One shortcoming of the experiment was the lack of a control group. At the initial design stage, it was decided there was no need to include a control group. It has not been a standard practice to do so in studies of the effectiveness of explicit and implicit instruction. For example, Norris and Ortega (2000) found that only 18% of studies on the explicit and implicit instruction in their meta-analysis included control groups. However, when analysing the data and interpreting the results, it became clear that a control group would have strengthened the validity of the data in terms of measuring linguistic performance. This means that the lack of a control group in the present study or any study potentially undermines the results. For example, although the statistical tests indicate significant improvement after treatment, the absence of a no-treatment group makes it difficult to assume that this improvement was the result of the treatment. A control group would have provided a fair base to compare the results to.

This relates to the potential effect of practice. As was mentioned in Chapter 4, section 4.2.4, a practice effect is the possibility that learners have performed better in the post-test as a result of being more familiar with the test and learning from their mistakes. In order to reduce the practice effect, researchers either leave a long time between the two tests or withhold feedback and answers. In the present study, for technical reasons, it was not possible to administer the various tests with gaps between them (see Chapter 4, section 4.2.4 and 4.5). However, the practice effect was reduced by taking special steps. First, tasks were presented in different orders during the pre- and post-tests. Second, feedback was withheld and learners only received the scores after they had finished all tasks and tests. An additional measure to reduce the practice effect is counterbalancing (Brown, 1988). In such a case, more than one version of the tests would have been produced and no learners would have received the same test twice. In the present research we could have reduced the practice effect by designing two versions of the tests, A and B, and giving half the learners in each group test A as a pre-test and test B as a post-test and vice versa with the other half.

The lack of a control group was perhaps an initial design fault. Feasibility also prevented inclusion of a control group. The log file data analysis took a considerable amount of time, and when it was finished, it was unfortunately no longer possible for several reasons to carry out another experiment with a comparable, control group. First, Adobe discontinued support to the software (Macromedia Authorware) that was used to create the treatment package used in the study, which led to many compatibility problems with present operating systems (as the old software is not compatible with Windows Vista, Windows 7 or 8). Second, the English language programme that learners involved in the study were enrolled on was no longer offered by the university and all programmes were restructured. In brief, due to the technical and practical factors, it was not possible to replicate the same conditions as the original experiment.

Other aspects of the experiment that might have undermined the findings include the limited amount of input in terms of the number of target items and forms tested. As was explained in Chapter Three, and further in Chapter Four, the input was taken from a widely-used textbook. It was important to stick to the content provided in the input, as one of the goals of the experiment was to replicate a classroom and look at how learners dealt with similar content in a technology-enhanced environment. However, this meant that the amount of input was limited. This was clear in terms of the number of target items included in the pre- and post-tests and the range of forms tested. What is worth mentioning here is that the teachers might elaborate in their instruction or might use extra material in the classroom which usually provides the learners with extra input about the target construction. This, however, was not accounted for in the current study.

A final point is the length of the treatment; the experiment lasted around 90 minutes. Although in their meta-study, Spada and Tomita (2010) point out that the treatment length varied between 0.33 to 8 hours, the mean treatment length was 2.9, which is not far from the present study. Still, it is not assumed here that this time, even if doubled, would be enough for the learners for intake and acquisition of the target construction. The treatment length in the present study was limited by the fact that data collection was only possible in one session for technical reasons explained in Chapter Four. With better programming, it would be possible to deliver a longer more comprehensive treatment but this was not possible within the financial and time restrictions of the present study.

Finally, the present study is a cross sectional study. As highlighted in Chapter Four, cross sectional studies are appropriate for providing reliable information about general patterns among large samples of learners. However, since information is collected only at a single point in time, no claims can be made about time progression or across individuals.

In brief, it is important to point out that due to the shortcomings of the experiment; the findings in terms of the effectiveness of instruction are offered with caution and not assumed to be generalizable at this stage. However, the findings on the use of the user behaviour technologies to inform the learning process and to identify the features that contribute to attainment are achieved.

6.8 Study findings

What emerges from the results of present study are the following:

- In a TELL environment, explicit instruction might not be beneficial if the learners' implicit' knowledge of the linguistic forms is not developed, and until their knowledge is developed, explicit instruction leads to worse performance than when knowledge is gained from input-driven implicit learning.
- What applies in the classroom does not necessarily apply in technology-enhanced environments. Therefore, when using technology in the classroom, it is not possible to simply borrow theories, frameworks and models wholesale from ISLA, and differences in the nature of interaction, type of input, participants' roles and the environment need to be taken into account.

These findings are offered cautiously for the reasons highlighted above in section 6.7 which include the lack of a control group, limited sample size and short treatment. Nevertheless, the study achieved its aim to look closely at the learning process and the factors that contribute to the effectiveness of the different types of input in a technology-enhanced environment. It also managed to provide insights into the use of user-behaviour tracking technologies to investigate the learning processes. In the next chapter, I highlight the implications of the study in relation to research, pedagogy and methodology.

Chapter Seven
Conclusion

The present thesis has investigated the effect of three different types of grammar instruction in a TELL environment and the factors that contribute to variation in performance when using these types of instruction among learners of English as a second language. One of the main aims of the present study was to gain a better understanding of how knowledge is constructed in real time. This was done through the use of log files to track learners' behavior while dealing with the three different types of instruction. The research was a quasi-experiment where three experimental groups were compared using pre- and post-test data.

The research ultimately aimed at identifying the best ways to incorporate the teaching of form into a task-based curriculum in a self-access Technology-Enhanced Language Learning environment. Thus, in a broader sense, the study aimed at answering the question

> *When using technology in the classroom, is it possible to simply borrow theories, frameworks and models wholesale from the field of Instructed Second Language Acquisition or do differences in language forms, nature of interaction, participants' roles and environment need to be taken into account?*

The study's focus was on how L2 learners dealt with the linguistic input they received rather than on the acquisition of the target construction, reported speech. Data was collected to compare the performance and processes they exhibited of 71 L2 English learners. This concluding chapter covers the study's implications and its possible limitations and recommendations for future research. In the first section below, I look at the research, pedagogical and methodological implications of the findings emerging from the results discussed in the previous chapter. I then turn to report on the limitations of the present study. Finally, I put forward recommendations for future research.

7.1 Implications

The implications of the present study can be summarised in Garrett's words (1991) of several decades ago:

> A CALL lesson which creates an environment for some interesting language learning activity could be fitted with a program collecting data on how the learner makes use of that environment, and those data can not only feed back into improving the pedagogy but can also contribute to the development of second language acquisition theory. (Garrett, 1991, p.94)

In addition, Cook (2008) identifies three ways in which language teaching can benefit from SLA research: (1) understanding students' contributions to learning, (2) understanding how teaching methods and techniques work and (3) understanding the goals of language teaching. Based on this classification, the present study looked at (1) students' contributions to learning and (2) how teaching methods work (the third point is not relevant thesis). Therefore, the present study contributes to the body of research on linking SLA research to classroom practice.

I start with implications for research in the fields of SLA, ISLA, language pedagogy and TELL.

7.1.1 Research

The present study showed that when using technology in second language learning, we cannot simply borrow theories, frameworks and models as a whole from the field of second language acquisition. Theories of SLA, and particularly theories of Instructed SLA, need to be expanded to account for the differences in language input, nature/type of interaction, participant roles and environment. The study also provided insights on how to apply existing technology and teaching approaches in new learning contexts. As rightly pointed out by Cook (2009), SLA research is most of the time used 'to justify existing teaching methods and approaches, rather than to suggest changes to existing ways of teaching or to innovative new ways' (Cook, 2009, p. 142). The present study is one first step in the direction of suggesting changes to existing ways of teaching with respect to both Task-Based Language Learning, and to Technology-Enhanced Language Learning. The study set out to answer the call by other researchers for research that looks at the applicability of ISLA and TBLL to a technology-enhanced environment.

The findings suggest that although some aspects of theories of SLA, ISLA, TBLL and TELL might hold for different learning contexts, others aspects need to be examined.

Cook's remarks notwithstanding, there is no doubt that the findings of SLA research are crucial sources of information for ideas and theories in second language pedagogy (see e.g. Piske and Young-Scholten 2009, Whong 2011). This is and can be particularly true for research that investigates what L2 learners attend to in the input and how learners process input in real time. By using a TELL environment to identify the effects of input that learners are exposed to, it is possible to arrive at a clearer understanding of adult L2 learners' cognitive processes while interacting with certain types of L2 data. Since this study examine the factors that affect the processing of clearly delineated sets of linguistic input in the form of tasks in a TELL environment, it adds to our understanding of the mechanisms learners employ when dealing with input.

Chapelle (2009a) notes that ideas emerging from the findings of SLA theories are turned into principles. These principles are used in the design of TELL-based materials; the materials are then evaluated through theory-based perspectives. The result should be a practice-relevant theory of SLA that can inform the design and evaluation of TELL-based materials. As one of the founding principles of task-based language learning is to harmonize with SLA research under interaction-based views the ways languages are taught (Doughty and Long, 2003), and as the call for technology-based materials developed on the basis of SLA findings is growing, software that applies principles based on SLA findings to deliver technology-enhanced, task-based activities seems the best combination.

The present study provides an initial picture of how learners interact with input to begin to acquire L2 knowledge. Examining learners' acquisition of the target construction was far beyond the scope of the study. The small step taken by the present study is a crucial one given the dearth of research. I argue that things have not changed since Rast (2008) highlighted the fact that even after years of research into input and intake, we still know very little about what learners do with the input they get. Real-time processing is slowly becoming a central issue in SLA, involving cross-disciplinary work by formal-linguistics-based SLA researchers and psycholinguists. Here one of the main questions is how learners access and process in real time a type of input that was not considered in the present study, namely primary linguistic data (Carroll, 2004 and

Carroll, 2013 on the study of learners' first unconscious exposure to a new language). Clahsen and Felser point out that 'further investigation of grammatical processing in language learners is necessary before any firm conclusions can be drawn' (Clahsen and Felser, 2006, p.35). On the same grounds, Ringbom and Jarvis (2009) both argue that 'much more knowledge is needed about the mechanism through which language learning proceeds before the field is justified in pronouncing definitive statements about how languages can be taught most effectively' (Ringbom and Jarvis, 2009, p. 114). Although the present study addressed much more restricted input types than primary linguistic data, it contributed to our understanding of the process involved in the sort of input learners receive in the classroom. This study makes a contribution to fill this gap in the literature but further, larger scale research on these types of input is also needed (see future directions below).

7.1.2 Pedagogy

Technology can augment the opportunities for L2 learners to receive target-language input, and the aim of the present study was to provide insights into how learners deal with different types of exposure to English in a TELL environment to inform the design of pedagogical intervention that facilitates acquisition. The study's focus was on how L2 learners deal with types of linguistic input they receive in an instructional setting. The data collected helped in identifying what factors or actions contribute to variation in performance or processes. Such information is of great value for learning contexts. Here language professionals need to have an adequate theoretical background in order to decide when a particular tool might assist students' second language development and studies such as the present one can contribute to their understanding of the theory/research underpinning their practice.

There is no doubt that foreign language learning (in non-immersion settings) will increasingly take place through technologically-based approaches whether they are web-based, CD-ROM applications or network-based Computer Mediated Communication. This puts pressure on SLA researchers, classroom teachers and material designers to identify the most efficient and readily available ways to enhance the acquisition process by providing pedagogically principled materials that utilize advances in technology. The present study has shown that learning can on the one hand be facilitated and on the other hand be inhibited by different factors. When using or designing materials, these factors need to be taken into account for learning to be more effective.

The findings are also of relevance to task designers and teachers in providing a better understanding of the potential contribution of various task types and their accompanying instructions to both immediate task performance and learning outcomes. Furthermore, by improving our understanding of the role of modality in SLA and how learners deal with oral and written input, we can begin to propose instructional approaches and teaching materials that take into account such findings. Importantly, the findings of the study call into question the usefulness of specific tasks in reflecting the performance of individual learners. Teachers also need to be careful in interpreting immediate task performance when assessing learners. It is very likely that tasks that lead to successful performance have their advantages: they can be motivating for learners and they can facilitate interaction. However, they may not be as cognitively demanding as other activities. But learners' performance on task does not necessarily reveal learners' current linguistic and communicative competence.

7.1.3 Methodology

It is not the aim of this research to raise unrealistic expectations with respect to the role of technology in enhancing the second language acquisition process. The present study rather aimed at exploring whether and if so how the use of user behaviour tracking technologies adds any insights about the acquisition process. It is argued that the use of log files as user behaviour tracking technologies, has the potential to provide a better and clearer picture on how and when learners make form-meaning connections.

The findings also highlight another aspect of SLA research: the need for researchers to make informed choices about the types of technology to use in their research in the new and evolving technology-rich environment. It is obvious that the technology exists there to serve language acquisition research.

7.2 Limitations of the study

A number of caveats need to be noted regarding the present study. One of the limitations is the lack of a separate control group. In this case, the study might have included an uninstructed group and/or a no-technology group. However, the present study is in keeping with other such studies. Norris and Ortega (2000) point out that 82% of all studies included in their meta-study lacked a true control group. While they warn that this makes interpretation of the results difficult, in the present study it was not feasible to include a control group for the practical and methodological reasons discussed in section 6.7. I acknowledge that a control group would likely have been useful in helping to validate the findings.

In the present study, immediate task performance was adopted as a measure of intake, but, as was pointed out in Chapter Two, there is difference between performance and the intake which actually leads to –linguistic competence. That is, although the treatment effect was statistically significant, this could be an immediate effect (i.e. on performance only) rather than one that affects intake and ultimately acquisition (competence). Including a control group in the design might have been able to disentangle what could have been practice effects from learners' actual gains. It is also very possible that even when a task is performed poorly, the target structure is still detected by the learner and is available for further processing, but for unknown reasons, this is not evident when the learner's performance is measured.

Furthermore, performance could have been analysed to reflect the three measures used in classroom-based study: fluency, accuracy and complexity. If the study is replicated, further analysis will mark responses correct if the meaning is achieved but not the form. This would allow for comparison between fluency and accuracy. As most TELL environments do not involve looking at oral outcomes because of the difficulty of electronic assessment the equivalent of the fluency measure in TELL environments could be related to the number of times learners supply acceptable answers but not in the completely right form.

In addition, the study results would have been more interesting if different proficiency levels were used on the one hand, or if, on the other hand, different constructions were compared. Controlling so many variables would have been extremely difficult given the nature of the study design and the data collected.

When it comes to explanations of the patterns found, it could well be that, as already noted earlier, due to the short length of testing and treatment, processing effects did not show up. Exposure was short; however, there were practical problems in carrying out a longer treatment as the participants in the study were all enrolled in full-time language classes and the time available in their programme to complete the tasks was limited.

A final limitation is the lack of any model in the literature for analysing the data collected. For example, data was collected on learners' uptake, i.e. how they performed during the treatment, but due to time limitations and difficulty of analysis, these data were not analysed. Uptake was instead operationalised as correct suppliance of the target structure (i.e. the target linguistic performance score) during treatment. Looking at uptake as well as acquisition would allow us to determine the effects of varying types of activities. This information would then allow teachers to make better informed decisions to balance different demands

7.3 Future research

Many questions remain to be addressed by future research. The question about the use of technology in language teaching is not an either-or situation where the teacher has to choose one option. Technology use is undeniably now ubiquitous and what we need is research to guide its use and put it in the service of the acquisition/learning process. Further research is also needed on how Task Based Language Learning might realistically be operationalised in self-access or blended learning environments.

The fact that the learners did not access the 'grammar help' suggests that focusing on forms might not be effective in Technology- Enhanced Language Learning. In the classroom, where interaction is human-human and it is usually the teacher who decides on when to focus on a form, in TELL, it is the learner who decides on this. The results obtained in the present study indicate that when given a choice and even when there is a need to do so, most learners do not focus on forms. This suggests that human-human-based focus on forms needs to be rethought in terms of what individual students in a given classroom are attending to. Further research should look into what underlies the grammar access pattern exhibited by learners when dealing with the input. It should look at when learners access the grammar help and whether their decisions to access help are correlated with certain difficulties or patterns. Although information was collected about these decisions in the present study, it was not possible to execute the analysis within the time limitations of the thesis.

Decision-making tasks are neglected by TELL materials developers, but such tasks could be of great value. For example, learners could be given different information about three possible murderers and, based on the information they are provided, they need to act as a jury and decide who the killer is. Using new advances in technology, such activities are now easily executed in a TELL environment using wikis, conferencing software, etc. Such activities provide tremendous opportunities for both oral and written production.

In line with Leow (1995), the present study highlights the need for future research that investigates the role of modality on the types of cognitive processes adult L2 learners employ while exposed to the aural and written modes that may potentially affect learners' intake.

The present research looked at how learners deal with L2 input in a TELL environment in real time. The study aimed to drive forward our understanding of the complex process involved in second language learning. Although the present research is theory-led, it has been more exploratory and illuminative in nature. However, given the present state of research that tackles issues similar to the ones in the present study, attempting to construct a model may be a premature undertaking but is one which is nonetheless worth pursuing.

To put it in Rast's (2008) words, 'it is rather through collaboration and exchange of methodology, results and knowledge that we will come to discover what is actually going on in the mind of an L2 learner when exposed to TL input' (Rast, 2008, p.5). More research is needed.

Bibliography

Abbot-Smith, K. and Tomasello, M. (2006) 'Exemplar-learning and schematization in a usage-based account of syntactic acquisition', *The Linguistic Review*, 23(3), pp. 275-290.

Alderson, J., Clapham, C. and Steel, D. (1998) 'Metalinguistic knowledge, language aptitude and language proficiency', *Language Teaching Research*, 1(2), pp. 93-121.

Ammar, A., and Spada, N. (2006). ‘One size fits all? Recasts, prompts and L2 learning’. *Studies in Second Language Acquisition*, 28, 543-574.

Anderson, F. (1993) 'The enigma of the college classroom: nails that don’t stick up', in Wadden, P. (ed.) *A Handbook for Teaching English at Japanese Colleges and Universities.* New York: Oxford University Press, pp. 101-110.

Anderson, J., Davidson, N., Morton, H. and Jack, M. (2008) 'Language learning with interactive virtual agent scenarios and speech recognition: Lessons learned', *Computer Animation and Virtual Worlds,* 19(5), pp. 605-619.

Banfield, A. (1973). ‘Narrative style and the grammar of direct and indirect speech’. *Foundations of Language*, 10, 1-39.

Banfield, A. (1982) *Unspeakable sentences: Narration and representation in the language of fiction*. Boston: Routledge & Kegan.

Barbieri, F. and Eckhardt, S. (2007) ‘Applying corpus-based findings to form-focused instruction: the case of reported speech’, *Language Teaching Research*, 11 (3), pp. 319-346.

Barcroft and Wong (2013) ‘Input, input processing and focus on form’ in Herschensohn, J. and Young-Scholten, M. *The Cambridge Handbook of Second Language Acquisition*. New York: Cambridge University Press.

Bassetti, B. and Cook, V. (2011) 'Relating language and cognition: The second language user', *Language and Bilingual cognition*, pp. 143-190.

Bax, S. (2003) 'CALL—past, present and future', *System*, 31(1), pp. 13-28.

Beaman, C. and Morton, J. (2000) 'The effects of rime on auditory recency and the suffix effect', *European Journal of Cognitive Psychology*, 12(2), pp. 223-242.

Beaman, C.P. (2002) 'Inverting the modality effect in serial recall', *Quarterly Journal of Experimental Psychology: Section A*, 55(2), pp. 371-389.

Benati, A. (2004) 'The effects of structured input and explicit information on the acquisition of Italian future tense', in VanPatten, B. (ed.) *Processing instruction*. Mahwah, NJ: Erlbaum, pp. 207-25.

Bialystok, E. (1979) 'Explicit and implicit judgements of l2 grammaticality', *Language Learning*, 29(1), pp. 81-103.

Bird, S. and Williams, J. (2002) 'The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling', *Applied Psycholinguistics*, 23(04), pp. 509-533.

Blake, R. (2000) 'Computer mediated communication: A window on L2 Spanish interlanguage', *Language Learning and Technology*, 4(1), pp. 120-136.

Blake, R. (2005) 'Bimodal CMC: the glue of language learning at a distance', *CALIC*, 22(3), pp. 497-511.

Blake, R. (2008) *Brave New Digital Classroom: Technology and foreign language learning*. Washington, D.C.: Georgetown University Press.

Bland, S.K. (1996): *Intermediate Grammar: from form to meaning and use*. Oxford: Oxford University Press.

Bley-Vroman, R. (1989) 'The logical problem of foreign language learning ', *Linguistic Analysis*, 19, pp. 56-104.

Bley-Vroman, R. and Masterson, D. (1989) 'Reaction time as a supplement to grammaticality judgments in the investigation of second language learners' competence', *University of Hawai'i Working Papers in ESL*, 8(2), pp. 207-237.

Borras, I. and Lafayette, R. C. (1994) 'Effects of multimedia courseware subtitling on the speaking performance of college students of French', *The Modern Language*

Journal, 78(1), pp. 61-75.

Bowles, M. A. (2010) 'Concurrent verbal reports in second language acquisition research', *Annual Review of Applied Linguistics*, 30(1), pp. 111-127.

Bowles, M. A. (2010) *The Think-Aloud Controversy in Second Language Research*. Routledge.

Breen, M. (1987) 'Learner contributions to task design', in Candlin, C. and Murphy, D. (eds.) *Language Learning Tasks*. Englewood Cliffs NJ: Prentice-Hall.

Breen, M.P. (1984) 'Process syllabuses for the language classroom', in Brumfit, C.J. (ed.) *General English Syllabus Design, ELT Documents 118*. London: Pergamon Press/British Council, pp. 47-60.

Brinton, D. (2001) 'The use of media in language teaching', in Celce-Murcia, M. (ed.) *Teaching English as a Second or Foreign Language*. Boston, MA: Heinle and Heinle, pp. 459-476.

Brown, J. (1988) *Understanding Research in Second Language Learning*. Cambridge: Cambridge University press.

Brown, J. D. (2011). ‘Quantitative research in second language studies’. In E. Hinkel (Ed.), Handbook of research in second language teaching and learning, volume 2 (pp. 190-206). New York: Routledge

Brown, S. (2004) 'Assessment for learning', *Learning and Teaching in Higher Education*, 1(1), pp. 81-89.

Bruckman, A. (2000). Uneven Achievement in a Constructivist Learning Environment. Paper presented at the International Conference on Learning Sciences, Ann Arbor, MI.

Bruckman, A. (2006) 'Analysis of log file data to understand behavior and learning in an online community', in *The International handbook of virtual learning environments*. Springer, pp. 1449-1465.

Brumfit, C. and Johnson, K. (1979) *The Communicative Approach to Language Teaching*. Oxford University Press.

Brumfit, C., Mitchell, R. and Hooper, J. (1996) 'Grammar, language and classroom practice', in Hughes, M. (ed.) *Teaching and Learning in Changing Times*. Oxford: Blackwell, pp. 70-87.

Brunton, A. (2005). Process writing and communicative-task-based instruction: Many common features, but more common limitations? *TESL-EJ*, *9*(3), A-2.

Bruton, A. (2002) 'Recording and recognizing grammatical form by meaning: the English genitive as an example', *System* 30, pp. 237-250.

Burke Johnson, R., Onwuegbuzie, A., & Turner, L. (2007). ‘Toward a definition of mixed methods research’. *Journal of Mixed Methods Research*, 1, 112–133.

Butler, Y.G. (2005) 'Comparative perspectives towards communication activities among elementary school teachers in South Korea, Japan and Taiwan', *Language Teaching Reseach*, 9, pp. 423-446.

Bygate, M. (1987) *Speaking*. OUP Oxford.

Bygate, M. (1988) *Linguistic and Strategic Features of the Language of Learners in Oral Communication Exercises*. Institute of Education: University of London.

Bygate, M. (1988) 'Units of oral expression and language learning in small group interaction', *Applied Linguistics*, 9(1), pp. 59-82.

Bygate, M., Skehan, P. and Swain, M. (2001) *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*. Pearson Education.

Byram, M. and Mendez-Garcia, M.D.C. (2009) 'Communicative Language Teaching', in Knapp, K., Seidlhofer, B. and Widdowson, H.G. (eds.) *Handbook of Foreign Language Communication and Learning*. Berlin: Mouton de Gruyter, pp. 491-516.

Candlin, C. and Murphy, D. (1987) *Language Learning Tasks*. Prentice-Hall International.

Carless , D. (2009). Revisiting the TBLT versus P-P-P debate: Voices from Hong Kong. *Asian Journal of English Language Teaching*, 19, 49-66

Carless, D. (2004) 'Issues in teachers' reinterpretation of a task-based innovation in primary schools', *TESOL Quarterly*, 38(4), pp. 639-662.

Carless, D. (2007) 'The suitability of task-based approaches for secondary schools: Perspectives from Hong Kong', *System*, 35(4), pp. 595-608.

Carroll, J. B. (1993) *Human Cognitive Abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Carroll, S. (1996) 'Parameter-setting in second language acquisition – explanans and explanandum', *Behavioral and Brain Sciences*, 19(04), pp. 720-721.

Carroll, S. (2001) *Input and Evidence: The raw material of second language acquisition*. Amsterdam: John Benjamins Publishing Company.

Carroll, S. (2001) *Input and Evidence: The Raw Material of Second Language Acquisition*. Amesterdam: John Benjamins Publishing Company.

Carroll, S. (2006) 'Autonomous induction theory', in VanPatten, B. and Williams, J. (eds.) *Theories in Second Language Acquisition*. Mahwah, NJ: Erlbaum, pp. 155-174.

Carter, R. (1998) ‘Orders of reality: CANCODE, communication, and culture’, *ELT Journal* 52(1), pp. 43–56.

Carter, R. and McCarthy, M. (1995) ‘Grammar and the spoken language’, *Applied Linguistics*, 16(2), pp.141–58.

Chamber, A. and Davies, G. (2001) *New Technologies and Language Learning: a European perspective.* Lisse: Swets & Zeitlinger.

Chapelle, C. (1990) 'The discourse of computer-assisted language learning: toward a context for descriptive research', *TESOL Quarterly*, 24(2), pp. 199-225.

Chapelle, C. (1997) ' CALL in the year 2000: Still in search of research paradigms?', *Language Learning and Technology*, 1(1), pp. 19-43.

Chapelle, C. (1998) 'Multimedia CALL: Lessons to be learned from research on instructed SLA', *Language Learning & Technology*, 2(1), pp. 22-34.

Chapelle, C. (2001) *Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing, and Research*. Cambridge: Cambridge University Press.

Chapelle, C. (2003) *English Language Learning and Technology: Lectures on Applied Linguistics in the Age of Information and Communication Technology*. Amsterdam: Benjamins.

Chapelle, C. (2007) 'Technology and second language acquisition', *Annual Review of Applied Linguistics*, 27, pp. 98-114.

Chapelle, C. (2009) ' Computer-Assisted Teaching and Testing', *The Handbook of Language Teaching*.

Chapelle, C. (2009) 'The Relationship Between Second Language Acquisition Theory and Computer-Assisted Language Learning', *The Modern Language Journal*, 93(s1), pp. 741-753.

Chapelle, C. and Heift, T. (2009) 'Individual Learner Differences in CALL: The Field Independence/Dependence (FID) Construct', *Calico Journal*, 26(2), pp. 246-266.

Chapelle, C.A. (2009) 'The relationship between second language acquisition theory and computer-assisted language learning', *The Modern Language Journal*, 93, pp. 741-753.

Charkova, K. and Halliday, L. (2011) ‘Second and foreign-Language variation in tense backshifting in indirect reported speech’, *Studies in Second Language Acquisition*, 33, pp. 1-32.

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

Chaudron, C. (1986) 'The interaction of quantitative and qualitative approaches to research: A view of the second language classroom', *TESOL Quarterly*, 20(4), pp. 709-717.

Chaudron, C. (2000) 'Contrasting approaches to classroom research: qualitative and quantitative analysis of language use and learning', *Second Language Studies*, 19(1), pp. 1-56.

Chaudron, C. (2001) 'Progress in Language Classroom Research: Evidence from The Modern Language Journal, 1916-2000', *The Modern Language Journal*, 85(1), pp. 57-76.

Chaudron, C. (2003) 'Data collection in SLA research', *The Handbook of Second Language Acquisition*, pp. 762-828.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Massachusetts Institute of Technology MIT Press.

Chomsky, N. (1980) *Rules and Representations*. Columbia University Press.

Chomsky, N. (1985) *Knowledge of Language: Its Nature, Origins, and Use*. Greenwood Publishing Group, Incorporated.

Chomsky, N.(1957) *Syntactic Structures.* Bobbs-Merrill.

Chun, D. (2013) 'Contributions of tracking user behaviour to SLA research'. in Hubbard, P. and Schulze, M., and Smith, B. (eds) *Learner-Computer Interaction in Language Education: A festschrift in honour of Robert Fischer*. San Marcos: CALICO.pp.256-262.

Chun, D. M. and Plass, J. L. (1996) 'Effects of multimedia annotations on vocabulary acquisition', *The Modern Language Journal*, 80(2), pp. 183-198.

Clahsen, H. (1984) 'The acquisition of German word order', in Andersen, R. (ed.) *Second Languages*. Rowley, MA: Newbury House.

Clahsen, H. and Felser, C. (2006) 'Grammatical processing in language learners', *Applied Psycholinguistics*, 27(01), pp. 3-42.

Collentine, J. (1998) 'Cognitive principles and CALL grammar instruction: A mind-centered, input approach', *CALICO Journal*, 15(1-3), pp. 1-18.

Collentine, J. (2000) 'Insights into the construction of grammatical knowledge provided by user-behavior tracking technologies', *Language Learning and Technology*, 3(2), pp. 44-57.

Collentine, J. (2002) 'Review of TESOROS: a multimedia-based Spanish course on CD-ROM', *Language Learning & Technology*, 6(1), pp. 33-39.

Comrie, B. (1985) *Tense*. Cambridge: Cambridge University Press.

Cook, V. (1988) 'Designing CALL programs for communicative teaching' *ELTJ*, 4.

Accessed online at
http://homepage.ntlworld.com/vivian.c/Writings/Papers/CAllELTJ88.htm

Cook, V. (2001) *Second Language Learning and Language Teaching*. 3 edn. London: Arnold.

Cook, V. (2008) *Second Language Learning and Language Teaching*. Hodder Education.

Cook, V. and Singleton, D. (2014) *Key Topics in Second Language Learning*. Bristol: Multilingual Matters

Cook, V.(ed) (2002) 'Background to the L2 user', *Portraits of The L2 User*. Clevedon: Mulilingual Matters.

Corder, S.P. (1967) 'The significance of learners` errors', *International Review of Applied*, 5, pp. 161-169.

Corder, S.P. (1978) 'Strategies of communication', in Færch, C. and Kasper, G. (eds.) *Strategies in Interlanguage Communication*. Longman Group Limited.

Crookes, G. (1986) *Task Classification: a Cross-Disciplinary Review* Honolulu: Research, C.f.S.L.C.

Cubillos, J. (1998) 'Technology: A step forward in the teaching of foreign languages', *The Coming of Age of the Profession: Issues and emerging ideas for the teaching of foreign languages*, pp. 37-52.

Cullen, R. (2008) 'Teaching grammar as a liberating force', *ELT Journal*, 62(3), pp. 221-230.

Curtin, C., Clayton, D., Finch, C., Moor, D. and Woodruff, L. (1972) 'Teaching the translation of Russian by computer', *The Modern Language Journal*, 56(6), pp. 354-360.

Davies, G. (1989) 'Repurposing a videodisc for French language teaching'. In Kécskés I. and Agócs L. (eds.) *New Tendencies in CALL*. Hungary: Kossuth University.

Davies, G. (1993) 'CALL in the New Europe: the spirit of co-operation' In Liddell P. (ed.) *CALL: theory and application: proceedings of CCALL2/CCELAO2*, The

Second Canadian CALL Conference. Canada: University of Victoria.

Davies, G. (2002) ‘ICT and modern foreign languages: learning opportunities and training needs, *IJES,* 2(1), pp.1-18.

Davies, G. and Higgins, J. (1982) *Computers, Language and Language Learning*. London: CILT.

Davies, G. and Higgins, J. (1985) *Using Computers in Language Learning: a teacher's guide*. London: CILT

de Graaff, R. (1997) 'The Experanto experiment', *Studies in Second Language Acquisition*, 19(02), pp. 249-276.

de Graaff, R. and Housen, A. (2009) 'Investigating the effects and effectiveness of L2 instruction', in *The Handbook of Language Teaching*. Wiley-Blackwell, pp. 726-755.

de Jong, T. (2005) 'The guided discovery principle in multimedia learning', *The Cambridge Handbook of Multimedia Learning*, pp. 215-228.

de la Fuente, M. (2012) 'Learners' attention to input during focus on form listening tasks: the role of mobile technology in the second language classroom', *Computer Assisted Language Learning*, (ahead-of-print), pp. 1-16.

DeKeyser, R. (1995) 'Learning second language grammar rules: an experiment with a miniature linguistic system', *Studies in Second Language Acquisition*, 17(03), pp. 379-410.

DeKeyser, R. (1997) 'Beyond explicit rule learning', *Studies in Second Language Acquisition*, 19(02), pp. 195-221.

DeKeyser, R. (2003) 'Implicit and explicit learning', in Doughty, C.J. and Long, M.H. (eds.) *The Handbook of Second Language Acquisition*. Oxford: Blackwell, pp. 313-348.

DeKeyser, R. (2007) *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology*. Cambridge: Cambridge University Press.

DeKeyser, R. and Juffs, A. (2005) 'Cognitive considerations in L2 learning', in Hinkel, E. (ed.) *Handbook of Research in Second Language Teaching and Learning*.

Mahwah, NJ: Lawrence Erlbaum., pp. 437–454.

DeKeyser, R.(2000) 'The robustness of critical period effects in second language acquisition', *Studies in Second Language Acquisition*, 22(04), pp. 499-533.

Dewaele, J.-M. (2009) 'Individual differences in second language acquisition', in Ritchie, W.C. and Bhatia, T.K. (eds.) *The New Handbook of Second Language Acquisition*. Bingley (UK): Emerald, pp. 623-646.

Dodigovic, M. (2005) 'Vocabulary profiling with electronic corpora: A case study in computer assisted needs analysis', *Computer Assisted Language Learning*, 18(5), pp. 443-455.

Dörnyei, Z. (1998) 'What is motivation?', *Paper Presented at The AAAL 1998 Conference,*. Seattle,WA.

Dörnyei, Z. (2003) *Attitudes, Orientations, and Motivations in Language Learning: Advances in Theory, Research, and Applications*. Wiley Blackwell.

Dörnyei, Z. (2005) *The Psychology of the Language Learner*. New Jersey: Lawrence Erlbaum Associates.

Dörnyei, Z. (2006) 'Creating a motivating classroom environment', in Cummins, J. and Davison, C. (eds.) *The Handbook of English Language Teaching*. New York: Springer.

Doughty, C. (1987). Relating second-language acquisition theory to CALL research and application. In W. F. Smith (Ed.), *Modern Media in Foreign Language Education: Theory and Implementatio*n (pp. 133-167

Doughty, C. (1991) 'Second language instruction does make a difference', *Studies in Second Language Acquisition*, 13(04), pp. 431-469.

Doughty, C. (2001) 'Cognitive underpinnings of focus on form', *Cognition and Second Language Instruction*, pp. 206-257.

Doughty, C. and Long, M. (2003) 'Optimal psycholinguistic environments for distance foreign language learning'. *Forum of International Development Studies*, (23) pp. 35-73.

Doughty, C. and Williams, J. (1998) *Focus on Form in Classroom Second Language Acquisition*. Ernst Klett Sprachen.

Doughty, C.J. (2003) 'Instructed SLA: constraints, compensation, and enhancement', in Doughty, C.J. and Long, M.H. (eds.) *The Handbook of Second Language Acquisition*. Malden, MA: Blackwell, pp. 256-310.

Dulay, H. and Burt, M. (1973) 'Should we teach children syntax?', *Language Learning*, 23(2), pp. 245-258.

Dulay, H. and Burt, M. (1974) 'Natural sequences in child second language acquisition', *Language Learning*, 24(1), pp. 37-53.

East, M. (2012a) 'Taking communication to task – again: what difference does a decade make?', *The Language Learning Journal*, pp. 1-14.

East, M. (2012b) *Task-Based Language Teaching from the Teacher's Perspective: Insights from New Zealand*. John Benjamins Publishing Company.

Eastwood, J. (1999) *Oxford practice Grammar with Answers*. Oxford: Oxford University Press.

Eckerth, J. (2008) 'Investigating consciousness-raising tasks: pedagogically targeted and non-targeted learning gains', *International Journal of Applied Linguistics*, 18(2), pp. 119-145.

Eckhardt, S. (2001) *Reported speech: Empirical corpus findings compared with ESL/EFL textbook presentations.*Ames, IA: Iowa State University.

Edwards, C. and Willis, J. (2005) *Teachers Exploring Tasks in English Language Teaching*. Palgrave USA.

Egbert, J. and Hanson-Smith, E. (1999) *CALL environments: Research, practice, and Critical Issues*. Teachers of English to Speakers of Other Languages.

Egbert, J., and E. Hanson-Smith, eds. 2007. *CALL environments: Research, Practice, and Critical Issues*. 2nd ed. Alexandria, VA: TESOL.

Elbaum, S. (2001) *Grammar in Context*.3rd ed. Boston: Heinle & Heinle.

Ellis, N. (1993) 'Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge', *European Journal of Cognitive Psychology*, 5(3), pp. 289-318.

Ellis, N. (1994) *Implicit and Explicit Learning of Languages*. Academic Press.

Ellis, N. (2003) 'Constructions, chunking, and connectionism: The emergence of second language structure', in Doughty, C. and Long, M.H. (eds.) *Handbook of Second Language Acquisition*. Oxford: Blackwell, pp. 33-68.

Ellis, N. (2007) 'Dynamic systems and SLA: The wood and the trees', *Bilingualism: Language and Cognition*, 10(01), pp. 23-25.

Ellis, N. (2008) 'The dynamics of second language emergence: cycles of language use, language change, and language acquisition', *The Modern Language Journal*, 92(2), pp. 232-249.

Ellis, N. C. (1998) 'Emergentism, connectionism and language learning', *Language Learning*, 48(4), pp. 631-664.

Ellis, N.(2001) 'Memory for language', in Robinson, P. (ed.) *Cognition and Second Language Instruction.* Cambridge: Cambridge University Press, pp. 33-68.

Ellis, N.(2005) 'At the interface: dynamic interactions of explicit and implicit language knowledge', *Studies in Second Language Acquisition*, 27(02), pp. 305-352.

Ellis, R. (1989) 'Are classroom and naturalistic acquisition the same?', *Studies in Second Language Acquisition*, 11(03), pp. 305-328.

Ellis, R. (1990) *Instructed Second Language Acquisition: Learning in the classroom*. Oxford: Blackwell.

Ellis, R. (1991) *Second Language Acquisition and Language Pedagogy*. Clevedon: Multilingual Matters.

Ellis, R. (1994) *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. (1997) *Second Language Acquisition*. Oxford: Oxford University Press.

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

Ellis, R. (2000) 'Task-based research and language pedagogy', *Language Teaching Research*, 4(3), pp. 193-220.

Ellis, R. (2001) 'Investigating Form-Focused Instruction', *Language Learning*, 51, pp. 1-46.

Ellis, R. (2002) 'Does form-focused instruction affect the acquisition of implicit knowledge?', *Studies in Second Language Acquisition*, 24(2), pp. 223-236.

Ellis, R. (2003) *Task-Based Language Learning and Teaching*. Oxford University Press, USA.

Ellis, R. (2004) 'The definition and measurement of l2 explicit knowledge', *Language Learning*, 54(2), pp. 227-275.

Ellis, R. (2005) *Instructed Second Language Acquisition: A Literature Review*. New Zealand Government - Ministry of Education - Research Division.

Ellis, R. (2005) 'Principles of instructed language learning', *System*, 33(2), pp. 209-224.

Ellis, R. (2006) 'Current issues in the teaching of grammar: An SLA perspective', *TESOL Quarterly*, 40(1), pp. 83-107.

Ellis, R. (2006) 'Researching the effects of form-focussed instruction on L2 acquisition', *AILA Review*, 19(1), pp. 18-41.

Ellis, R. (2008) 'Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing', *International Journal of Applied Linguistics*, 18(1), pp. 4-22.

Ellis, R. (2009) 'Task-based language teaching: sorting out the misunderstandings', *International Journal of Applied Linguistics*, 19(3), pp. 221-246.

Ellis, R. (2010) ‘Forward’ in M. Thomas and H. Reinders (eds) Task-Based Language Learning and Teaching with technology. London: Continuum.

Ellis, R. (2010) 'Second language acquisition, teacher education and language pedagogy', *Language Teaching*, 43(02), pp. 182-201.

Ellis, R., Basturkmen, H. and Loewen, S. (2001) 'Preemptive focus on form in the ESL classroom', *TESOL Quarterly,* 35(3), pp. 407-432.

Ellis, R., Basturkmen, H. and Loewen, S. (2002) 'Doing focus-on-form', *System*, 30(4), pp. 419-432.

Ellis, R., Loewen, S. and Erlam, R. (2006) 'Implicit and explicit corrective feedback and the acquisition of L2 grammar', *Studies in Second Language Acquisition*, 28(02), pp. 339-368.

Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J. and Reinders, H. (2009) *Implicit and Explicit Knowledge in Second Language Learning, Teaching, and Testing*. UK: Multilingual Matters.

Ellis, R., Sheen, Y., Murakami, M. and Takashima, H. (2008) 'The effects of focused and unfocused written corrective feedback in an English as a foreign language context', *System*, 36(3), pp. 353-371.

Ensslin, A. and Krummes, C.. (2013). Second language development and electronic interaction. In J. Hersenschon and M. Young-Scholten (eds.) *Cambridge University Handbook of SLA*. Cambridge: CUP

Eubank, L. and Gregg, K.(1995) '‘Et in Amygdala Ego’?', *Studies in Second Language Acquisition*, 17(01), pp. 35-57.

Felser, C. and Clahsen, H. (2009) 'Grammatical processing of spoken language in child and adult language learners', *Journal of Psycholinguistic Research*, 38(3), pp. 305-319.

Felser, C., Roberts, L., Marinis, T. and Gross, R. (2003) 'The processing of ambiguous sentences by first and second language learners of English', *Applied Psycholinguistics*, 24(03), pp. 453-489.

Field, A. (2013) *Discovering Statistics Using IBM SPSS Statistics*. 4th ed. London: Sage Publications Ltd.

Finocchiaro, M. and Brumfit, C. (1983) *The Functional-Notional Approach: From*

Theory to Practice. Oxford: Oxford University Press.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students’ behavior in CALL. *Computer Assisted Language Learning*, 20(5), 409-442.

Fischer, R. (2012). Diversity in learner usage patterns. In G. Stockwell (Ed.), *Computer-Assisted Language Learning Diversity in Research and Practice.* Cambridge University Press. pp. 14-32

Fogg, B. (2003) *CHI'03 extended abstracts on human factors in computing systems*. ACM.

Foster, P. and Skehan, P. (1996) 'The influence planning time on performance in task-based learning', *Studies in Second Language Acquisition*, 18, pp. 299-334.

Foster, P. and Skehan, P. (1999) 'The influence of source of planning and focus of planning on task-based performance', *Language Teaching Research*, 3(3), pp. 215-247.

Fotos, S. (1994) 'Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks', *TESOL Quarterly*, 28(2), pp. 323-351.

Fotos, S. and Ellis, R. (1991) 'Communicating about grammar: A task-based approach', *TESOL Quarterly*, 25(4), pp. 605-628.

Fotos, S. and Nassaji, H. (2007) *Form-focused Instruction and Teacher Education: Studies in Honour of Rod Ellis*. Oxford: Oxford University Press

Fotos, S.(1993) 'Consciousness raising and noticing through focus on form: Grammar task performance versus formal instruction', *Applied Linguistics*, 14(4), pp. 385-407.

Frenck-Mestre, C. (2005) 'Eye-movement recording as a tool for studying syntactic processing in a second language: a review of methodologies and experimental findings', *Second Language Research*, 21(2), pp. 175-198.

Frenck-Mestre, C. and Pynte, J. (1997) 'Syntactic ambiguity resolution while reading in second and native languages', *Quarterly Journal of Experimental Psychology:*

Section A, 50(1), pp. 119-148.

Fries, C. (1957) *Preface to R. Lado's Linguistics Across Cultures*. Ann Arbor: University of Michigan Press.

Fuchs, M. and Bonner, M. (1995) *Focus on Grammar: A high-intermediate course for reference and practice.* London: Longman.

Fulcher, G. (1995) 'Variable competence in second language acquisition: A problem for research methodology?', *System*, 23(1), pp. 25-33.

Gall, M., Borg, W. and Gall, J. (1996) *Educational research: an introduction.* NY: Longman.

Gant Guillory, H. (1998) 'The effects of keyword captions to authentic French video on learner comprehension', *Calico Journal*, 15(1-3), pp. 89-108.

García Mayo, M. (ed.) (2007) *Investigating Tasks in Formal Language Learning*. Clevedon: Multilingual Matters.

Garrett, N. (1991). 'Technology in the service of language learning: Trends and issues'. *Modern Language Journal, **75**, 74–101.*

Garrett, N. (1998) 'Where do research and practice meet? Developing a discipline', *RECALL-HULL-*, 10, pp. 7-12.

Gascoigne, C. (2006) 'Explicit input enhancement: Effects on target and non-target aspects of second language acquisition', *Foreign Language Annals*, 39(4), pp. 551-564.

Gascoigne, C. (2006) 'Toward an understanding of incidental input enhancement in computerized L2 environments', *CALICO Journal*, 24(1), pp. 147-162.

Gass, S. (1997). *Input, Interaction and The Second Language Learner*. Mahwah, N. J.: Lawrence: Erlbaum Associates.

Gass, S. (2003) 'Input and interaction', in Doughty, C. and Long, M. (eds.) *The Handbook of Second Language Acquisition*. Malden, MA: Blackwell, pp. 224-255.

Gass, S. and Mackey, A. (2000) *Stimulated Recall Methodology in Second Language*

Research. Routledge.

Gass, S. and Mackey, A. (2005) 'Task-based interactions in classroom and laboratory settings', *Language Learning*, 55(4), pp. 575-611.

Gass, S. and Mackey, A. (2006) 'Input, interaction and output: An overview', *AILA review,* 19(1), pp. 3-17.

Gass, S. and Mackey, A. (2007) *Data Elicitation for Second and Foreign Language Research*. Routledge.

Gass, S. and Mackey, A. (2007) 'Input, interaction, and output in second language acquisition', *Theories in second language acquisition: An introduction*, pp. 175-199.

Gass, S. and Selinker, L. (2001) *Second Language Acquisition: An Introductory Course*. Second edn. Hillsdale, NJ: Lawrence Erlbaum.

Gass, S.(1988) 'Integrating research areas: a framework for second language studies', *Applied Linguistics*, 9(2), pp. 198-217.

Gass, S., Alison, M. and Pica, T. (1998) 'The role of input and interaction in second language acquisition: Introduction to the special issue', *The Modern Language Journal*, 82(3), pp. 299-307.

Gass, S., Mackey, A., Alvarez, M. and Fernandez, M. (1999) 'The effects of task repetition on linguistic output', *Language Learning*, 49(4), pp. 549–581.

Gass, S., Svetics, I. and Lemelin, S. (2003) 'Differential effects of attention', *Language Learning*, 53(3), pp. 497-546.

Gass, S.M. and Selinker, L. (2008) *Second Language Acquisition : an introductory course*. 3 edn. Taylor and Francis.

Gatbonton, E. and Segalowitz, N. (1988) 'Creative automatization: principles for promoting fluency within a communicative framework', *TESOL Quarterly*, 22(3), pp. 473-492.

Gombert, J. (1992) *Metalinguistic Development*. Harvester Wheatsheaf.

González-Lloret, M. (2003) 'Designing task-based CALL to promote interaction: En busca de esmeraldas', *Language Learning & Technology*, 7(1), pp. 86-104.

González-Lloret, M. (2007) *Revisiting Language Learning Resource*.

Grace, C. (1998) 'Retention of Word Meanings Inferred from Context and Sentence-Level Translations: Implications for the Design of Beginning-Level CALL Software', *The Modern Language Journal*, 82(4), pp. 533-544.

Granger, S. (1998) *Learner English on Computer*. London: Longman

Green, P.S. and Hecht, K. (1992) 'Implicit and explicit grammar: An empirical study', *Applied Linguistics*, 13(2), pp. 168-184.

Greenbaum, S. (1996). *The Oxford English Grammar*. Oxford: OUP.

Gregg, K. (1993) 'Taking explanation seriously; or, let a couple of flowers bloom', *Applied Linguistics*, 14(3), pp. 276-294.

Gregg, K. (1996) 'The logical and developmental problems of second language acquisition', in Ritchie, W. and Bhatia, T. (eds.) *Handbook of Second Language Acquisition*. San Diego: Academic Press.

Gregg, K. (2001) 'Learnability and second language acquisition theory', in Robinson, P. (ed.) *Cognition and second language instruction*. Cambridge: Cambridge University Press, pp. 152–182.

Gregg, K. (2003) 'SLA theory: Construction and assessment', in Doughty, C.J. and Long, M.H. (eds.) *Handbook of Second Language Acquisition*. Oxford: Blackwell.

Hampel, R. (2006) 'Rethinking task design for the digital age: A framework for language teaching and learning in a synchronous online environment', *RECALL-HULL THEN CAMBRIDGE-*, 18(1), p. 105.

Hatch, E. and Lazaraton, A. (1994) *The Research Manual: Design and Statistics for Applied Linguistics*. Boston: Heinle and Heinle Publishers.

Heift, T. and Rimrott (2012) ‘Task-related variation in computer-assisted language learning’ *The Modern Language Journal*, 96, 4.

Heift, T. and Schulze, M. (2007) *Errors and intelligence in computer-assisted language* *learning: Parsers and pedagogues*. Psychology Press.

Herschensohn, J. (2000) *The Second Time Around: Minimalism and L2 Acquisition*. Philadelphia: John Benjamins Publishing Co.

Herschensohn, J. and Young-Scholten, M. (2013) *The Cambridge Handbook of Second Language Acquisition*. New York: Cambridge University Press.

Higgins, J. (1988). *Language, Learners and Computers: Human Intelligence and Artificial Unintelligence*. Singapore: Longman.

Higgins, J. and Johns, T. (1984). *Computers in Language Learning.* London: Collins.

Hill, M. and Laufer, B. (2003) 'Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition', *International Review of Applied Linguistics in Language Teaching*, 41(2), pp. 87-106.

Hoshi, M. (2002) 'Practices, beliefs and perceptions of Japanese EFL self-access learners toward internet-based language learning', *CALL-EJ Online*, 4(1), pp. 7-1.

Housen, A. and Pierrard, M. (2005) *Investigations in Instructed Second Language Acquisition*. Berlin: Walter de Gruyter.

Housen, A., Pierrard, M. and Van Daele, S. (2005) 'Rule complexity and the effectiveness of explicit grammar instruction', in Housen, A. and Pierrard, M. (eds.) *Investigations in Instructed Second Language Acquisition*. Berlin: Mouton de Gruyter, pp. 207-241.

Howatt, A. (1984) *A History of English Language Teaching*. Oxford: Oxford University Press.

Howatt, A. and Widdowson, H. (eds.) (2004) *A History of English Language Teaching*. 2 edn. Oxford: Oxford University Press.

Hu, G. (2002) 'Psychological constraints on the utility of metalinguistic knowledge in second language production', *Studies in Second Language Acquisition*, 24(03), pp. 347-386.

Hubbard, P. and Siskin, C.(2004) 'Another look at tutorial CALL', *ReCALL*, 16(2), pp. 448-461.

Hulstijn, J. (1989) 'Implicit and incidental second language learning: Experiments in the processing of natural and partly artificial input.' in Dechert, H.W. and Raupach, M. (eds.) *Interlingual Processing*. Tubingen: Gunter Narr, pp. 49-73

Hulstijn, J. (2000) 'The use of computer technology in experimental studies of second language acquisition: A survey of some techniques and some ongoing studies', *Language Learning & Technology*, 3(2), pp. 32-43.

Hulstijn, J. (2000) 'The use of computer technology in experimental studies of second language acquisition: a survey of some techniques and some ongoing studies', *Language Learning and Technology* 3(2), pp. 32-43.

Hulstijn, J. (2002) 'What does the impact of frequency tell us about the language acquisition device?', *Studies in Second Language Acquisition*, 24(02), pp. 269-273.

Hulstijn, J. (2005) 'Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction', *Studies in Second Language Acquisition*, 27(02), pp. 129-140.

Hulstijn, J. and de Graaff, R. (1994) 'Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal', in Hulstijn, J. and Schmidt, R. (eds.) *Consciousness in Second Language Learning*, *AILA Review*. pp. 97-109.

Hulstijn, J., Hollander, M. and Greidanus, T. (1996) 'Incidental vocabulary learning by advanced foreign language students: the influence of marginal glosses, dictionary use, and reoccurrence of unknown words', *The Modern Language Journal*, 80(3), pp. 327-339.

Hulstijn, J.and Hulstijn, W. (1984) 'Grammatical errors as a function of processing constraints and explicit knowledge', *Language Learning*, 34(1), pp. 23-43.

Hulstijn, J.and Laufer, B. (2001) 'Some empirical evidence for the involvement load hypothesis in vocabulary acquisition', *Language Learning*, 51(3), pp. 539-558.

Hwu, F. (2004) 'On the applicability of the input-enhancement hypothesis and input processing theory in multimedia CALL: The case of Spanish preterite and imperfect instruction in an input application', *CALICO Journal*, 21(2), pp. 317-338.

Hymes, D. (1979) 'On communicative competence', in Brumfit, C. and Johnson, K. (eds.) *The Communicative Approach to Language Teaching*. Oxford: Oxford University Press.

Isaac, S. and Michael, W.B. (1995). *Handbook in Research and Evaluation* (3rd ed). SanDiego CA: EdITS.

Izumi, S. (2003) 'Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis', *Applied Linguistics*, 24(2), pp. 168-196.

Izumi, S. (2003) 'Processing difficulty in comprehension and production of relative clauses by learners of English as a second language', *Language learning*, 53(2), pp. 285-323.

Jagtman, M. and Bongaerts, T. (1994) 'Report-COMOLA: a computer system for the analysis of interlanguage data', *Second Language Research*, 10(1), pp. 49-83.

Jakobson, R. (1971).'Shifters, verbal categories and the Russian verb'. *Selected Writing II: Word and Languag*e, 130-147. The Hague: Mouton.

Jin, L. and Cortazzi, M. (2011) 'Re-evaluating Traditional approaches to Second language teaching and learning' in Hinkel, E. (eds) *Handbook of Research in Second Language Teaching and Learning*. Volume II. New York: Routledge. pp. 558-575

Johnson, J. (1992) 'Critical Period effects in second language acquisition: the effect of written versus auditory materials on the assessment of grammatical competence', *Language Learning*, 42(2), pp. 217-248.

Johnson, J. and Newport, E. (1991) 'Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language', *Cognition*, 39(3), pp. 215-258.

Johnson, J.and Newport, E. (1989) 'Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language', *Cognitive Psychology*, 21(1), pp. 60-99.

Jones C. and Fortescue S. (1987) *Using Computers in the Language Classroom.* Harlow: Longman

Jones, J. (2001) 'CALL and the responsibilities of teachers and administrators', *ELT Journal*, 55(4), pp. 360-367.

Jones, J. (2001) 'CALL and the teacher's role in promoting learner autonomy', *CALL-EJ Online*, 3(1), pp. 6-1.

Jones, L. and Plass, J. (2002) 'Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations', *The Modern Language Journal*, 86(4), pp. 546-561.

Jordan, G. (2004) *Theory Construction in Second Language Acquisition*. John Benjamins.

Juffs, A. (1998) 'Main verb versus reduced relative clause ambiguity resolution in L2 sentence processing', *Language Learning*, 48(1), pp. 107-147.

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), pp. 121-151.

Juffs, A. and Harrington, M. (1996) 'Garden path sentences and error data in second language sentence processing', *Language Learning*, 46(2), pp. 283-323.

Keating, G. (2008) 'Task effectiveness and word learning in a second language: The involvement load hypothesis on trial', *Language Teaching Research*, 12(3), pp. 365-386.

Kenning, M.-M. (1996) 'Creating an infrastructure for autonomous learning: the resource catalogue', *System*, 24(2), pp. 223-231.

Kessler, G. and Ware, P. D. (2013) 'Addressing the Language Classroom Competencies of the European Higher Education Area Through the Use of Technology', in *Competency-based Language Teaching in Higher Education*. Springer, pp. 93-105.

Kim, Y. (2008) 'The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition', *Language Learning*, 58(2), pp. 285-325.

Klapper, J. (2003) 'Taking communication to task? A critical review of recent trends in language teaching', *The Language Learning Journal*, 27(1), pp. 33-42.

Klein, W. (1986) *Second Language Acquisition*. Cambridge: Cambridge University

Press.

Knutsson, O., Cerratto Pargman, T. and Severinson Eklundh, K. (2003). ‘Transforming grammar checking technology into a learning environment for second language writing’ *HLT/NAACL 2003 workshop: Building Educational Applications Using NLP*, Edmonton, Canada.

Krashen, S. (1982) *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.

Krashen, S. (1977) 'Some issues relating to the Monitor Model', in Brown, H., Yorio, C. and Crymes, R. (eds.) *On TESOL '77: Teaching and Learning English as a Second Language: Trends in Research and Practice*. Washington: TESOL, pp. 144-158.

Krashen, S. (1984) *Writing: Research, Theory, and Applications*. Pearson College Division.

Krashen, S. (1985) *The Input Hypothesis: Issues and Implications*. London: Longman.

Krashen, S. (1993) *The Power of Reading.* Englewood, CO: Libraries Unlimited.

Krashen, S. and Terrell, T. (1983) *The Natural Approach: Language Acquisition in the Classroom*. London: Prentice Hall Europe.

Krashen, S.(1981) *Second Language Acquisition and Second Language Learning*. Pergamon Press.

Krashen, S.(1994) 'The input hypothesis and its rivals', in Ellis, N. (ed.) *Implicit and Explicit Learning of Languages*. London: Academic Press, pp. 45-77.

Krashen, S.and Terrell, T.(1989) *The Natural Approach: Language Acquisition in the Classroom*. Pergamon: Prentice Hall.

Kumaravadivelu, B. (2006) *Understanding Language Teaching: From Method to Postmethod*. Taylor & Francis.

Lantolf, J. (2000) 'Second language learning as a mediated process', *Language Teaching*, 33(02), pp. 79-96.

Lantolf, J. and Thorne, S.(2006) *Sociocultural Theory and the Genesis of Second*

Language Development. Oxford University Press.

Lantolf, J.(2006) 'Sociocultural theory and L2: State of the art', *Studies in Second Language Acquisition*, 28(01), pp. 67-109.

Laufer, B. (2001) 'Quantitative evaluation of vocabulary: how it can be done and what it is good for', in Elder, C., Brown, A., Grove, E., Hill, K., Iwashita, N., Lumley, T., McNamara, T. and O'Loughlin, K. (eds.) *Experimenting with Uncertainty*. Cambridge University Press, pp. 241-250.

Laufer, B. (2003) 'Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence', *Canadian Modern Language Review*, 59(4), pp. 567-587.

Laufer, B. (2005) 'Instructed second language vocabulary learning: the fault in the ‘Default Hypothesis’', in Housen, A. and Gruyter, M.P.M.d. (eds.) *Investigations in Instructed Second Language Acquisition*. pp. 286-303.

Laufer, B. (2006) 'Comparing focus on form and focus on forms in second-language vocabulary learning', *Canadian Modern Language Review,* 63(1), pp. 149-166.

Laufer, B. (2011) 'The contribution of dictionary use to the production and retention of collocations in a second language', *International Journal of Lexicography*, 24(1), pp. 29-49.

Laufer, B. and Hulstijn, J. (2001) 'Incidental vocabulary acquisition in a second language: The construct of task-induced involvement', *Applied Linguistics*, 22(1), pp. 1-26.

Lazaraton, A. (2000). Current trends in research methodology and statistics in applied linguistics. *TESOL Quarterly*, 34, 3.

Lazaraton, A. (2005) 'Quantitative research methods', *Handbook of Research in Second Language Teaching and Learning*, pp. 209-224.

Lee, J. (2000) *Tasks and Communicating in Language Classrooms*. McGraw-Hill Higher Education.

Lee, J. and Benati, A. (2007) *Second Language Processing: an analysis of theory,* *problems, and possible solutions*. Continuum.

Lee, J. and VanPatten, B. (2003) *Making Communicative Language Teaching Happen*. McGraw-Hill Education.

Lee, J.and Benati, A. (2007) *Delivering Processing Instruction in Classrooms and in Virtual Contexts: Research and Practice*. Equinox Publishing Limited.

Lenneberg, E. (1967) *Biological Foundations of Language*. Wiley.

Leow, R. (1993) 'To simplify or not to simplify: A look at intake', *Studies in Second Language Acquisition*, 15(3), pp. 333-355.

Leow, R. (1995) 'Modality and Intake in Second Language Acquisition', *Studies in Second Language Acquisition*, 17(01), pp. 79-89.

Leow, R. (1997) 'The effects of input enhancement and text length on adult L2 readers' comprehension and intake in second language acquisition', *Applied Language Learning*, 8, pp. 151-182.

Leow, R.(1998) 'Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine grained analysis of attention', *Applied Psycholinguistics,* 19(01), pp. 133-159.

Leow, R.(2000) 'A study of the role of awareness in foreign language behavior', *Studies in Second Language Acquisition*, 22(04), pp. 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), pp. 496-509.

Levy, M. (1997) *Computer-Assisted Language Learning: Context and Conceptualisation*. Oxford: Clarendon Press.

Levy, M. (1997) 'Theory-driven CALL and the development process', *Computer Assisted Language Learning*, 10(1), pp. 41-56.

Levy, M. (2007) 'Research and technological innovation in CALL', *International Journal of Innovation in Language Learning and Teaching*, 1(1), pp. 180-190.

Levy, M. (2009) 'Technologies in use for second language learning', *The Modern Language Journal*, 93(s1), pp. 769-782.

Levy, M. and Stockwell, G. (2006) *CALL dimensions: Options and issues in computer assisted language learning*. L. Erbaum Associates.

Li, C. (1986). Direct and indirect speech: A functional study. In C. Coulmas (Ed.), *Direct and Indirect Speech* (pp. 301-318). Berlin: Mouton de Gruyter.

Li, D. (1998) '‘It's always more difficult than you plan and imagine’: Teachers' perceived difficulties in introducing the communicative approach in South Korea', *TESOL Quarterly*, 32(4), pp. 677-703.

Li, S. (2010) 'The effectiveness of corrective feedback in SLA: A meta-analysis', *Language Learning*, 60(2), pp. 309-365.

Lightbown, P. (1985). Great Expectations: Second-language acquisition research and classroom teaching. *Applied Linguistics* 6 (2): 173-189.

Lightbown, P. and Spada, N. (1990) ‘Focus-on-form and corrective feedback in communicative language teaching’, *Studies in Second Language Acquisition*, 12(4), pp.429-448.

Lightbown, P. and Spada, N. (1999) *How Languages Are Learned*. 2nd edn. Oxford: OUP.

Lightbown, P. and Spada, N. (2006) *How Languages are Learned*. Oxford University Press

Lightbown, P., Spada, N. and White, L. (1993) 'The role of instruction in second language acquisition', *Studies in Second Language Acquisition*, 15(2), pp. 143-145.

Linder, D. (2004) 'The Internet in every classroom? Using outside computers', *ELT Journal*, 58(1), pp. 10-17.

Littlejohn, A. (2003) *Reusing Online Resources: a sustainable approach to e-learning*. Routledge.

Littlewood, W. (1981) *Communicative Language Teaching: an introduction*. Cambridge University Press.

Loewen, S. (2003) 'Variation in the frequency and characteristics of incidental focus on form', *Language Teaching Research*, 7(3), pp. 315-345.

Loewen, S. (2004) 'Uptake in Incidental Focus on Form in Meaning-Focused ESL Lessons', *Language Learning*, 54(1), pp. 153-188.

Loewen, S. and Reinders, H. (2011) *Key Concepts in Second Language Acquisition*. Palgrave Macmillan.

Long, M. (1980) 'Inside the ‘black box’: Methodological issues in classroom research on language learning', *Language learning*, 30(1), pp. 1-42.

Long, M. (1983) 'Does second language instruction make a difference? A review of research', *TESOL Quarterly*, 17(3), pp. 359-382.

Long, M. (1988) 'Instructed interlanguage development', in Beebe, L. (ed.) *Issues in Second Language Acquisition: Multiple Perspectives*, New York: Newbury House, pp. 115-141.

Long, M. (1990) 'Maturational constraints on language development', *Studies in Second Language Acquisition*, 12(03), pp. 251-285.

Long, M. (1997) 'Authenticity and learning potential in L2 classroom discourse', in Jacobs, G. (ed.) *Language Classrooms of Tomorrow: Issues and Responses*. Singapore: Seameo Regional Language Centre, pp. 148-169.

Long, M. (1998) 'Focus on form Theory, research, and practice ', *Focus on Form in Classroom Second Language Acquisition*, 15.

Long, M. (2000) 'Focus on form in task-based language teaching', *Language Policy and Pedagogy: Essays in Honor of A. Ronald Walton*, pp. 179-192.

Long, M. and Crookes, G. (1992) 'Three approaches to task-based syllabus design', *TESOL Quarterly*, 26(1), pp. 27-56.

Long, M. and Norris, J. (2000) 'Task-based language teaching and assessment', in Byram, M. (ed.) *Encyclopedia of Language Teaching*. London: Routledge, pp. 597-603.

Long, M. and Robinson, P. (1998) ‘Focus on form: theory, research and practice’ in

Doughty, C. and Williams, J. (eds) *Focus on Form in Classroom Second Language Acquisition*. Cambridge: CUP.pp.15-41.

Long, M.(1981) 'Input, interaction, and second language acquisition', in Winitz, H. (ed.) *Language and Foreign Language Acquisition*. New York: New York Academy of Sciences, pp. 259-278.

Long, M.(1985) ' Input and second language acquisition theory', in Gass, S. and Madden, C. (eds.) *Input in Second Language Acquisition* Rowley, MA: Newbury House, pp. 377-393.

Long, M.(1991) 'Focus on form: A design feature in language teaching methodology', in DeBot, K., Grinsberg, R. and Kramsch, C. (eds.) *Foreign Language Research in Cross-Cultural Perspective*. Amsterdam: John Benjamins, pp. 39-52.

Long, M.(1996) 'The role of the linguistic environment in second language acquisition', in Ritchie, W.C. and Bhatia, T.K. (eds.) *Handbook of Second Language Acquisition*. San Diego: Academic Press, pp. 413-468.

Long, M.and Robinson, P. (1998) 'Focus on form Theory, research, and practice', in Doughty, C. and Williams, J. (eds.) *Focus on Form in Classroom Second Language Acquisition*. Cambridge: CUP, pp. 15-41.

Lund, R. (1990) 'A Taxonomy for teaching second language listening', *Foreign Language Annals*, 23(2), pp. 105-115.

Lund, R. (1991) 'A comparison of second language listening and reading comprehension', *The Modern Language Journal*, 75(2), pp. 196-204.

Lyster, R. and Mori, H. (2006) 'Interactional feedback and instructional counterbalance', *Studies in Second Language Acquisition*, 28(2), p. 269.

Lyster, R. and Saito, K. (2010) 'Oral feedback in classroom SLA', *Studies in Second Language Acquisition*, 32(2), pp. 265-302.

Ma, Q. (2013) ‘Computer-Assisted Vocabulary Learning: Framework and tracking Users’ data’ . in Hubbard, P. and Schulze, M., and Smith, B. (eds) *Learner-Computer Interaction in Language Education: A festschrift in honour of Robert Fischer*. San Marcos: CALICO. pp.230-243.

Mackey, A. (1999) 'Input, interaction, and second language development', *Studies in Second Language Acquisition*, 21(4), pp. 557-587.

Mackey, A. (2007) *Conversational Interaction in Second Language Acquisition: A series of empirical studies*. Oxford : Oxford University Press.

Mackey, A. and Gass, S. (2011). *A Guide to Research Methods in Second Language Acquisition*. London: Basil Blackwell

Mackey, A. and Gass, S. (eds) (2012) *Research Methods in Second Language Acquisition: a practical guide*. Oxford: Wiley-Blackwell.

Mackey, A. and Goo, J. (2007) 'Interaction research in SLA: A meta-analysis and research synthesis', in Mackey, A. (ed.) *Conversational Interaction in Second Language Acquisition: A Collection of Empirical Studies.* Oxford: Oxford University Press, pp. 407-452.

Mackey, A. and Philp, J. (1998) 'Conversational interaction and second language development: Recasts, responses, and red herrings?', *The Modern Language Journal*, 82(3), pp. 338-356.

Mackey, A., Gass, S.and McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22(4), 471-497.

Mackey, A., Polio, C. and McDonough, K. (2004) 'The relationship between experience, education and teachers' use of incidental focus-on-form techniques', *Language Teaching Research*, 8(3), pp. 301-327.

MacWhinney, B. (1995) 'Language-specific prediction in foreign language learning', *Language Testing*, 12(3), pp. 292-320.

Manning, P. (1996) 'Exploratory teaching of grammar rules and CALL', *ReCALL*, 8(1), pp. 24-30.

Marinis, T. (2010) 'Using on-line processing methods in typical and atypical populations', in Blom, E. and Unsworth, S. (eds.) *Experimental Methods in Language Acquisition Research*. Amsterdam: John Benjamins, pp. 139-162.

McCarthy, M. (1998) 'So Mary was saying: Speech reporting in everyday conversation'. In McCarthy, M., *Spoken Language and Applied Linguistics*, Cambridge: Cambridge University Press. pp. 150–175.

McDonald, R. (2006) *Discriminative Learning and Spanning Tree Algorithms for Dependency Parsing*. University of Pennsylvania.

McLaughlin, B. (1987) *Theories of Second Language Learning*. London: Edward Arnold.

McLaughlin, B. and Heredia, R. (1996) 'Information-processing approaches to research on second language acquisition and use', in Ritchie, W. (ed.) *Handbook of Second Language Acquisition*. San Diego: Academic Press, pp. 213-228.

McLaughlin, B., Rossman, T. and McLeod, B. (1983) 'Second language learning: An information-processing perspective', *Language Learning*, 33(2), pp. 135-158.

Mehnert, U. (1998) 'The effects of different lengths of time for planning on second language performance', *Studies in Second Language Acquisition*, 20(01), pp. 83-108.

Metcalfe, J. (1997) ' Insight and metacognition', in Massoni, G. and Nelson, T.(eds.) *Metacognition: A neurobiological Perspective.* Erlbaum: Hillsdale, N.J, pp. 189-199.

Mitchell, R. and Myles, F. (2004) *Second Language Learning Theories*. Hodder Arnold.

Morton, H. (2005) 'Spoken Electronic Language Learning (SPELL): Results of User Trials in Scottish Schools', *Centre for Communication Interface Research*. The University of Edinburgh, 2005.

Morton, H. (2008) 'Speech Interactive CALL: A Cross-Cultural Evaluation' *Computer Assisted Language Learning,*

Murphy, R., Altman, R. and Rutheford, W. (1989): *Grammar in Use: Reference and practice for intermediate students on English*. Cambridge: Cambridge University Press.

Murphy, V. (1997) 'The effect of modality on a grammaticality judgement task', *Second Language Research*, 13(1), pp. 34-65.

Murray, G. (1999) 'Exploring learners' CALL experiences: a reflection on method', *Computer Assisted Language Learning*, 12(3), pp. 179-195.

Myles, F. (2005) 'Interlanguage corpora and second language acquisition research', *Second Language Research*, 21(4), pp. 373-391.

Myles, F. (2005) 'Interlanguage corpora and second language acquisition research', *Second Language Research*, 21(4), pp. 373-391.

Nassaji, H. (2000) 'Towards integrating form-focused instruction and communicative interaction in the second language classroom: some pedagogical possibilities', *The Modern Language Journal*, 84(2), pp. 241-250.

Nassaji, H. (2004) 'Input modality and remembering name–referent associations in vocabulary learning', *Canadian Journal of Applied Linguistics*, 7(1), pp. 39-55.

Nassaji, H. and Fotos, S. (2004) 'Current developments in research on the teaching of grammar', *Annual Review of Applied Linguistics*, 24, pp. 126-145.

Nerbonne, J. (2003). ‘Natural language processing’ in *Computer-Assisted Language Learning.* 36

Newport, E.(1990) 'Maturational constraints on language learning', *Cognitive Science*, 14(1), pp. 11-28.

Norris, J. and Ortega, L. (2000) 'Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis', *Language Learning*, 50(3), pp. 417-528.

Norris, J.and Ortega, L. (2006) 'The value and practice of research synthesis for language learning and teaching', in Norris, J.M. and L. Ortega, L. (eds.) *Synthesising Research on Language Learning and Teaching*. Amsterdam: John Benjamins, pp. 3-50.

Nunan, D. (1989) *Designing Tasks for the Communicative Classroom*. New York: Cambridge University Press.

Nunan, D. (2004) *Task-Based Language Teaching*. Cambridge: Cambridge University Press.

Nunan, D. (2005) 'Important tasks of English education: Asia-wide and beyond', *Asian EFL Journal* 7(3).

Nunan, D. (2006) *Practical English Language Teaching: Listening*. New York:

McGraw Hill.

Nunn, R. (2011) 'Improving method-in-use through classroom observation', *International Review of Applied Linguistics in Language Teaching*, 49(1), pp. 55-70.

Odlin, T. (1994) *Perspectives on Pedagogical Grammar*. Cambridge, Cambridge University Press

Ortega, L. (1999) 'Planning and focus on form in l2 oral performance', *Studies in Second Language Acquisition*, 21(01), pp. 109-148.

Ortega, L. (2005) 'Methodology, epistemology, and ethics in instructed SLA research: An introduction', *The Modern Language Journal*, 89(3), pp. 317-327.

Ortega, L. (2009) *Second Language Acquisition*. Hodder education.

Ortega, L. (2009). Interaction and attention to form in L2 text-based computer-mediated communication. In A. Mackey & C. Polio (Eds.), *Multiple Perspectives on Interaction* (pp. 226-253). New York, NY: Routledge

Papadopoulou, D. and Clahsen, H. (2003) 'Parsing strategies in l1 and l2 sentence processing', *Studies in Second Language Acquisition*, 25(04), pp. 501-528.

Paradis, M. (2004) *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.

Paradis, M. (2009) *Declarative and Procedural Determinants of Second Languages*. Amsterdam John Benjamins.

Partee, B. H. (1973). The syntax and semantics of quotation. In S. R. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 410-418). New York: Holt, Rinehart and Winston.

Pellettieri, J. (2000) 'Negotiation in cyberspace: The role of chatting in the development of grammatical competence', *Network-based language teaching: Concepts and practice*, pp. 59-86.

Penfield, W. (1959) *Speech and Brain: Mechanisms*. Princeton, NJ: Princeton University Press.

Penney, C. (1989) 'Modality effects and the structure of short-term verbal memory', *Memory and Cognition*, 17(4), pp. 398-422.

Peters, E. (2006) 'L2 vocabulary acquisition and reading comprehension: The influence of task complexity', in Garcia-Mayo, M.P. (ed.) *Investigating Tasks in Formal Language Learning* Clevedon: Multilingual Matters, pp. 178-198.

Peters, E., Hulstijn, J., Sercu, L. and Lutjeharms, M. (2009) 'Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared', *Language Learning*, 59(1), pp. 113-151.

Philp, J. (2003) 'Constraints on 'noticing the gap'', *Studies in Second Language Acquisition*, 25(01), pp. 99-126.

Pica, T. (1985) 'The selective impact of classroom instruction on second-language acquisition', *Applied Linguistics*, 6(3), pp. 214-222.

Pica, T. (1994) 'Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes?', *Language Learning*, 44(3), pp. 493-527.

Pica, T. and Doughty, C. (1986) 'Making input comprehensible: Do interactional modifications help', *ITL Review of Applied Linguistics*, 72, pp. 1-25.

Pica, T., Kanagy, R. and Falodun, J. (1993) 'Choosing and using communication tasks for second language instruction', in Crookes, G. and Gass, S. (eds.) *Tasks and Language Learning: Integrating Theory and Practice*. Clevedon: Multilingual Matters, pp. 9-34.

Pienemann, M. (1984) 'Psychological constraints on the teachability of languages', *Studies in Second Language Acquisition*, 6(02), pp. 186-214.

Pienemann, M. (1985) 'Learnability and syllabus construction', in Hyltenstam, K. and Pienemann, M. (eds.) *Modelling and Assessing Second Language Development*. Clevedon, Avon: Multilingual Matters pp. 23-76.

Pienemann, M. (1988) 'Determining the influence of instruction on L2 speech processing', *AILA Review*, pp. 40-67.

Pienemann, M. (1989) 'Is language teachable? Psycholinguistic experiments and hypotheses', *Applied Linguistics*, 10(1), pp. 52-79.

Pienemann, M. (1992) 'COALA-A computational system for interlanguage analysis', *Second Language Research*, 8(1), pp. 59-92.

Pienemann, M. (1998) *Language Processing and Second Language Development*. John Benjamins Publishing.

Pienemann, M. (2007) 'Processability theory', in VanPatten, B. and Williams, J. (eds.) *Theories in Second Language Acquisition*. Mahwah, NJ: Erlbaum, pp. 137-154.

Pienemann, M. (ed.) (2005) *Crosslinguistics Aspects of Processability Theory*. John Benjamins.

Piske, T. and Young-Scholten, M. (eds.) (2009) *Input Matters in SLA*. Clevedon: Multilingual Matters.

Plass, J. and Jones, L. (2005) 'Multimedia learning in second language acquisition', in Mayer, R. (ed.) *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press, pp. 467-488.

Plass, J., Chun, D., Mayer, R. and Leutner, D. (1998) 'Supporting visual and verbal learning preferences in a second-language multimedia learning environment', *Journal of Educational Psychology*, 90, pp. 25-36.

Platt, E. and Brooks, F. (2002) 'Task engagement: A turning point in foreign language development', *Language Learning*, 52(2), pp. 365-400.

Plonsky, L. (2011) 'The effectiveness of second language strategy instruction: A meta-analysis', *Language Learning*, 61(4), pp. 993-1038.

Poole, A. (2005) 'Focus on form instruction: Foundations, applications, and criticisms', *The Reading Matrix*, 5(1), pp. 47-56.

Prabhu, N.S. (1987) *Second Language Pedagogy*. Oxford University Press.

Prabhu, N.S. (1987) *Second Language Pedagogy*. Oxford University Press.

Raimes, A. (1998): *Grammar Troublespots: An editing guide for students*. 2nd ed.

Cambridge: Cambridge University Press.

Reber, A.S. (1967) 'Implicit learning of artificial grammars', *Journal of Verbal Learning and Verbal Behavior*, 6(6), pp. 855-863.

Reber, A.S. (1989) 'Implicit learning and tacit knowledge', *Journal of Experimental Psychology: General*, 118, pp. 219–235.

Reinders, H. (2005) 'Nonparticipation in university language support ', *Japan Association for Language Teaching Journal*, 27(2), pp. 209-226.

Reinders, H. (2009) 'Learner uptake and acquisition in three grammar-oriented production activities', *Language Teaching Research*, 13(2), pp. 201-222.

Reinders, H. (2010) *The Effects of Task Type and Instructions on Second Language Acquisition*. Newcastle: Cambridge Scholars Publishing.

Reinders, H. (ed.) (2012) *Digital Games in Language Learning and Teaching*. New York: Palgrave Macmillan.

Reinders, H. and White, C. (2010) 'The theory and practice of technology in materials development and task design', *English language teaching materials: Theory and practice*, pp. 58-80.

Richards, J.C. and Rodgers, T.S. (2001) *Approaches and Methods in Language Teaching*. New York, NY: Cambridge University Press.

Ringbom, H. and Jarvis, S. (2009) 'The importance of cross-linguistic similarity in foreign language learning', in Long, M.H. and Doughty, C.J. (eds.) *Handbook of Language Teaching*. West Sussex, UK: Wiley Blackwell, pp. 106-118.

Robinson, P. (1995) 'Attention, memory, and the’ Noticing’ Hypothesis', *Language Learning*, 45(2), pp. 283-331.

Robinson, P. (1996) 'Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions', *Studies in Second Language Acquisition*, 18(1), pp. 27-67.

Robinson, P. (1997) 'Individual differences and the fundamental similarity of implicit and explicit adult second language learning', *Language Learning*, 47(1), pp. 45-99.

Robinson, P. (1998) 'State of the art: SLA theory and second language syllabus design', *The Language Teacher*, 22(4), pp. 7-13.

Robinson, P. (2001) 'Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition', *Second Language Research*, 17(4), pp. 368-392.

Robinson, P. (2003) 'The cognition hypothesis, task design, and adult task-based language learning', *Second Language Studies*, 21(2), pp. 45-105.

Robinson, P. (2005) 'Cognitive complexity and task sequencing: studies in a triadic componential framework for second language task design', *International Review of Applied Linguistics*, 43(1), pp. 1-32.

Robinson, P. (2007) 'Criteria for classifying and sequencing pedagogic tasks', in Garcia-Mayo, M.d.P. (ed.) *Investigating Tasks in Formal Language Learning*. Clevedon, Avon: Multilingual Matters, pp. 7-27.

Robinson, P. (2009) 'Task complexity, cognitive resources and syllabus design: A triadic framework for examining task influences on SLA', in Branden, K.V.d., Bygate, M. and Norris, J. (eds.) *Task-Based Language Teaching; A Reader*. Amsterdam/Philadelphia PA: John Benjamins, pp. 193-226.

Robinson, P. (2011) *Task-based Language Learning*. Wiley-Blackwell.

Robinson, P. (ed.) (2002) *Individual Differences and Instructed Language Learning*. Amsterdam/Philadelphia: John Benjamins.

Rogers, R. (2009) *The End of The Virtual Digital Methods*. Vossiupers UvA.

Rollins, R.A. and Hendricks, R. (1980) 'Processing of words presented simultaneously to eye and ear', *Journal of Experimental Psychology: Human Perception and Performance*, 6, pp. 99-109.

Rosa, E. and Leow, R. (2004) 'Computerized task-based exposure, explicitness, type of feedback, and Spanish L2 development', *The Modern Language Journal*, 88(2), pp. 192-216.

Rosa, E. and O'Neill, M.D. (1999) 'Explicitness, intake, and the issue of awareness:

Another piece to the puzzle', *Studies in Second Language Acquisition*, 21(04), pp. 511-556.

Roussel, P. (2008) 'Method For Monitoring The Proper Operating Condition of A Computer'. EP Patent 1,924,915.

Roussel, S. (2011) 'A computer assisted method to track listening strategies in second language learning', *ReCALL*, 23(02), pp. 98-116.

Russell, J. and Spada, N. (2006) 'The effectiveness of corrective feedback for the acquisition of L2 grammar', in Norris, J. (ed.) *Synthesizing Research on Language Learning and Teaching*. Philadelphia, PA: John Benjamins, pp. 133-163.

Rutherford, W. (1988) 'Functions of grammar in a language-teaching syllabus', in Rutherford, W. and Smith, M.S. (eds.) *Grammar and Second Language Teaching: A Book of Readings.* Boston: Heinle and Heinle, pp. 231-249.

Rutherford, W. and Smith, M. (1985) 'Consciousness-raising and Universal Grammar', *Applied Linguistics*, 6(3), pp. 274-282.

Rutherford, W. and Thomas, M. (2001) 'The Child Language Data Exchange System in research on second language acquisition', *Second Language Research*, 17(2), pp. 195-212.

Salaberry, M. R. (2000) 'Pedagogical design of computer mediated communication tasks: Learning objectives and technological capabilities', *The Modern Language Journal*, 84(1), pp. 28-37.

Samuda, V. and Bygate, M. (2008) *Tasks in Second Language Learning*. Palgrave Macmillan.

Sanz, C. and Morgan-Short, K. (2004) 'Positive evidence versus explicit rule presentation and explicit negative feedback: a computer-assisted study', *Language Learning*, 54(1), pp. 35-78.

Sanz, C. and Morgan-Short, K. (2005) 'Explicitness in pedagogical interventions: Input, practice, and feedback', in Sanz, C. (ed.) *Mind and Context in Adult Second Language Acquisition: Methods, Theory, and Practice*. Georgetown University Press.

Sanz, C., Lin, H. , Lado, B., Bowden, H. and Stafford, C. (2009) 'Concurrent verbalizations, pedagogical conditions, and reactivity: Two CALL studies', *Language Learning*, 59(1), pp. 33-71.

Sasayama, S. (2011) 'Cognition hypothesis and second language performance: comparison of written and oral task performance', *Second Language Studies*, 29(2), pp. 107-129.

Savignon, S.J. (1997) *Communicative Competence: Theory and Classroom Practice: Texts and Contexts in Second Language Learning*. McGraw-Hill.

Savignon, S.J. (2002) *Interpreting Communicative Language Teaching: Contexts and Concerns in Teacher Education*. New Haven: Yale University Press.

Saville-Troike, M. (2012) *Introducing Second Language Acquisition*. 2 edn. Cambridge: Cambridge University Press.

Schachter, J. (1989) 'Testing a proposed universal ', in Gass, S. and Schachter, J. (eds.) *Linguistic Perspectives on Second Language Acquisition*. Cambridge: Cambridge University Press, pp. 73-88.

Schmidt, R. (1983) 'Interaction, acculturation and the acquisition of communicative competence', in Wolfson, N. and Judd, E. (eds.) *Sociolinguistics and Language Acquisition*. Rowley, MA: Newbury House, pp. 137-174.

Schmidt, R. (1990) 'The role of consciousness in second language learning', *Applied Linguistics*, 11(2), pp. 129-158.

Schmidt, R. (1993) 'Awareness and second language acquisition', *Annual Review of Applied Linguistics*, 13(1), pp. 129-158.

Schmidt, R. (1994) 'Implicit learning and the cognitive unconscious: Of artificial grammars and SLA', in Ellis, N. (ed.) *Implicit and Explicit Learning of Languages*. London: Academic Press, pp. 165-209.

Schmidt, R. (1995) 'Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning', in Schmidt, R. (ed.) *Attention and Awareness in Foreign Language Learning*. Honolulu, HI: University of Hawaii Press, pp. 1-64.

Schmidt, R. (2001) 'Attention', in Robinson, P. (ed.) *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press, pp. 3-32.

Schmidt, R. (2010) 'Attention, awareness, and individual differences in language learning', *Proceedings of CLaSIC 2010*. Singapore, December 2-4. National University of Singapore, Centre for Language Studies, pp. 721-737.

Schmidt, R. (2012) 'Attention, awareness, and individual differences in language learning', *Perspectives on Individual Characteristics and Foreign Language Education*, 6, p. 27.

Schrooten, W. (2006) 'Task-based language teaching and ICT: Developing and assessing interactive multimedia for task-based language teaching', *Task-Based Language*, p. 129.

Schulze, M. (1998) 'Checking Grammar-Teaching Grammar', *Computer Assisted Language Learning*, 11(2), pp. 215-227.

Schütze, C. (2005) 'Thinking about what we are asking speakers to do', *Linguistic evidence: Empirical, theoretical, and computational perspectives*, pp. 457-484.

Schütze, H. (1996) 'Ambiguity in language learning: computational and cognitive models'.

Schwartz, B. (1993) 'On explicit and negative data effecting and affecting competence and linguistic behavior', *Studies in Second Language Acquisition*, 15(2), pp. 147-165.

Scott, V. (1989) 'An empirical study of explicit and implicit teaching strategies in French', *The Modern Language Journal*, 73(1), pp. 14-22.

Seedhouse, P. (1999) 'Task-based interaction', *ELT Journal*, 53(3), pp. 149-156.

Seedhouse, P. (2009) 'How language teachers explain to student what they are supposed to Do', in Spreckels, J. (ed.) *Erklären im Kontext – Neue Perspektiven aus der Unterrichts-, Alltags- und Berufspraxis*. Schneider Hohengehren: Baltmannsweiler, pp. 66-80.

Segalowitz, N. (2003) 'Automaticity and second language acquisition', in Doughty, C.J. and Long, M.H. (eds.) *Handbook of Second Language Acquisition*. Oxford:

Blackwell, pp. 382-408.

Segalowitz, N. and Hulstijn, J. (2005) *Automaticity in Bilingualism and Second Language Learning*. Oxford, UK: Oxford University Press.

Selber, S. (2004). *Multiliteracies for a Digital Age*. Carbondale : Southern Illinois University Press.

Selinker, L. (1972) 'Interlanguage', *IRAL; International Review of Applied Linguistics in Language Teaching*, 10(3), p. 209.

Shaffer, L.H. (1975) 'Multiple attention in continuous verbal tasks', in Rabbitt, P. and Dornic, S. (eds.) *Attention and Performance V*. New York: Academic Press.

Sharwood Smith, M. and Truscott, J. (2008). MOGUL and crosslinguistic influence. In Gabrys, D. (Ed.) *Morphosyntactic Issues in Second Language Acquisition Studies* , Clevedon: Multilingual Matters: 63-85.

Sharwood Smith, M. and Truscott, J. (2014). *The multilingual mind: a modular processing perspective*. Cambridge: Cambridge University Press

Sharwood Smith, M., Snape, N. and Leung, Y-K) (eds) (2009). Representational Deficits in SLA: *Studies in Honour of Roger Hawkins* John Benjamins: Amsterdam

Shea, P. (2000) 'Leveling the playing field: A study of captioned interactive video for second language learning', *Journal of Educational Computing Research*, 22(3), pp. 243-264.

Sheen, R. (2002) ''Focus on form' and 'focus on forms'', *ELT Journal*, 56(3), pp. 303-305.

Sheen, Y. (2004) ‘Corrective feedback and learner uptake in communicative classrooms across instructional settings’, *Language Teaching Research*, 8(3), pp. 263-300.

Sheen, Y. (2004) 'Corrective feedback and learner uptake in communicative classrooms across instructional settings', *Language Teaching Research*, 8(3), pp. 263-300.

Shintani, N., Li, S. and Ellis, R. (2013) 'Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies', *Language Learning*, 63(2), pp. 296-329.

Shook, D.J. (1994) 'FL/L2 reading, grammatical information and input-to-intake phenomenon', *Applied Language Learning*, 5, pp. 57-93.

Skehan, P. (1989) *Individual Differences in Second-Language Learning*. London: Edward Arnold.

Skehan, P. (1996) 'A framework for the implementation of task-based instruction', *Applied Linguistics*, 17(1), pp. 38-62.

Skehan, P. (1998) *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Skehan, P. (2001) 'Tasks and language performance assessment', *Researching Pedagogic Tasks: Second language learning, teaching and testing*, pp. 167-185.

Skehan, P. (2002) 'Theorising and updating aptitude', in Robinson, P. (ed.) *Cognition and Second Language Instruction*. Amsterdam: John Benjamins, pp. 69-93.

Skehan, P. (2003) 'Focus on form, tasks, and technology', *Computer Assisted Language Learning*, 16(5), pp. 391-411.

Skehan, P. (2003) 'Task-based instruction', *Language Teaching*, 36(01), pp. 1-14.

Skehan, P. and Foster, P. (1997) 'Task type and task processing conditions as influences on foreign language performance', *Language Teaching Research*, 1(3), pp. 185-211.

Skehan, P. and Foster, P. (1999) 'The influence of task structure and processing conditions on narrative retellings', *Language Learning*, 49(1), pp. 93-120.

Skehan, P. and Foster, P. (2001) 'Cognition and tasks', in Robinson, P. (ed.) *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press, pp. 181-203.

Skinner, B.F. (1957) *Verbal Behavior*. Appleton-Century-Crofts.

Slabakova, R. (2009) 'Features or parameters: which one makes second language acquisition easier, and more interesting to study?', *Second Language Research*, 25(2), pp. 313-324.

Smith, B. (2003) 'Computer–mediated negotiated interaction: an expanded model', *The*

Modern Language Journal, 87(1), pp. 38-57.

Smith, B. (2004). Computer-Mediated Negotiated Interaction and Lexical Acquisition. *Studies in Second Language Acquisition*, 26, 365-398.

Smith, B. (2009) ‘Task-based learning in the computer- mediated communicative ESL/EFL classroom’ *CALL-EJ online*, vol. 11, no. 1.

Smith, M. (1981) 'Consciousness-Raising and the Second Language Learner1', *Applied Linguistics*, II (2), pp. 159-168.

Smith, M. (1991) 'Speaking to many minds: On the relevance of different types of language information for the L2 learner', *Second Language Research*, 7(2), pp. 118-132.

Smith, M. (1993) 'Input enhancement in instructed SLA', *Studies in Second Language Acquisition*, 15(2), pp. 165-179.

Sokolov, J. and Snow, C. (1994) *Handbook of Research in Language Development Using CHILDES*. Cambridge: Cambridge University Press.

Sokolov, J. and Snow, C. (1994) 'The changing role of negative evidence in theories of language development', *Input and Interaction in Language Acquisition*, pp. 38-55.

Sorace, A. (1996) 'The use of acceptability judgments in second language acquisition research', *Handbook of Second Language Acquisition*, pp. 375-409.

Spada, N. (1997) 'Form-focussed instruction and second language acquisition: A review of classroom and laboratory research', *Language Teaching*, 30(02), pp. 73-87.

Spada, N. (2005) 'Conditions and challenges in developing school-based SLA research programs', *The Modern Language Journal*, 89(3), pp. 328-338.

Spada, N. and Fröhlich, M. (1995) *COLT: Communicative Orientation of Language Teaching: Observation Scheme: Coding Conventions and Applications*. National Centre for English Language Teaching and Research, Macquarie University.

Spada, N. and Lightbown, P.M. (2002) 'L1 and L2 in the education of inuit children in Northern Quebec: Abilities and perceptions', *Language and Education*, 16(3), pp. 212-240.

Spada, N. and Tomita, Y. (2010) 'Interactions between type of instruction and type of language feature: A meta-analysis', *Language Learning*, 60(2), pp. 263-308.

Swain, J. and French, S. (2004) 'Disability and communication: Listening is not enough', in Barrett, S., Komarony, C., Robb, M. and Rogers, A. (eds.) *Communication, Relationships and Care: a reader*. Routledge.

Swain, M. (1985) 'Communicative competence: Some roles of comprehensible input and comprehensible output in its development', in Gass, S. and Madden, C. (eds.) *Input in Second Language Acquisition*. New York: Newbury House, pp. 235-256.

Swain, M. (1991) 'French immersion and its offshoots: Getting two for one', in B.Freed (ed.) *Foreign Language Acquisition: Research and the Classroom*. Lexington, MA: Heath, pp. 91-103.

Swain, M. (1995) 'Three functions of output in second language learning', in Cook, G. and Seidelhofer, B. (eds.) *Principle and Practice in Applied Linguistics: Studies in Honor of H.G. Widdowson*. Oxford: Oxford University Press, pp. 125-144.

Swain, M. (2000) 'The output hypothesis and beyond: Mediating acquisition through collaborative dialogue', in Lantolf, J.P. (ed.) *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press, pp. 97-114.

Swain, M. (2005) 'The Output Hypothesis: Theory and research', in Hinkel, E. (ed.) *Handbook of Research in Second Language Teaching and Learning*. Lawrence Erlbaum Associates, pp. 471-483.

Swain, M. and Lapkin, S. (2001) 'Focus on form through collaborative dialogue: Exploring task effects', in Bygate, M., Skehan, P. and Swain, M. (eds.) *Researching Pedagogic Tasks. Second Language Learning, Teaching and Testing*. Harlow: Longman, pp. 99-118.

Swain, M. and Lapkin, S. (2002) 'Talking it through: Two French immersion learners' response to reformulation', *International Journal of Educational Research*, 37(3), pp. 285-304.

Swain, M., Brooks, L. and Tocalli-Beller, A. (2002): Peer-peer dialogue as means of second language learning. *Annual Review of Applied Linguistics* 22: 171–85.

Swan, M. (2005) 'Legislation by hypothesis: the case of task-based instruction', *Applied Linguistics*, 26(3), pp. 376-401.

Sydorenko, T. (2010) 'Modality of input and vocabulary acquisition', *Language Learning and Technology*, 14(2), pp. 50-73.

Tarone, E. (1983) 'On the Variability of Interlanguage Systems', *Applied Linguistics*, 4(2), pp. 142-164.

Tarone, E. (1985) 'Variability in interlanguage use: A study of style-shifting in morphology and syntax', *Language Learning*, 35(3), pp. 373-403.

Tarone, E. (1987) 'Methodologies for studying variability in second language acquisition', in Ellis, R. (ed.) *Second Language Acquisition in Context*. Hemel Hempstead: Prentice-Hall, pp. 35-40.

Teeler, D. and Gray, E. (2000) *'How to Use the Intemet in ELT'*. Malaysia: Pearson Education.

Terrell, T., Baycroft, B. and Perron, C. (1987) 'The Subjunctive in Spanish interlanguage: Accuracy and comprehensibility', in VanPatten, B., Dvorak, T. and Lee, J. (eds.) *Foreign Language Learning: A Research Perspective*. Cambridge: Cambridge UP, pp. 19-32.

Thewlis, S. (2001): *Grammar Dimensions*. 3rd ed. Boston: Heinle & Heinle.

Thomas, D. (2006) 'A general inductive approach for analyzing qualitative evaluation data', *American Journal of Evaluation*, 27(2), pp. 237-246.

Thomas, M. (2013) ‘Task-based language teaching and CALL’ in Thomas, M., Reinders, H. and Warschauer , M. (eds) (2013) *Contemporary Computer Assisted Language Learning*. New York: Continuum.

Thomas, M. and Reinders, H. (2010) 'Deconstructing tasks and technology', *Task-Based Language Learning and Teaching with Technology*, pp. 1-13.

Thomas, M. And Reinders, H. (2010) Task-*Based Language Learning and Teaching with Technology*. London: Continuum.

Thomas, M., Reinders, H., Warschauer, M. (Eds.) (2013). *Contemporary Computer-*

Assisted Language Learning. New York: Continuum.

Tomasello, M. (1995) 'Language is not an instinct', *Cognitive Development*, 10(1), pp. 131-156.

Tomasello, M. (1998). The new psychology of language: Cognitive and functional approaches, Mahwah, NJ: Erlbaum.

Tomasello, M. (2003) *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Tomlin, R.and Villa, V. (1994) 'Attention in cognitive science and second language acquisition', *Studies in Second Language Acquisition,* 16(02), pp. 183-203.

Tomlinson, B. (2003) *Developing Materials for Language Teaching*. Continuum.

Tomlinson, B. (2003) 'Materials development', *The Encyclopedia of Applied Linguistics*.

Treisman, A. (1969) 'Strategies and models of selective attention', *Psychological Review*, 76(3), pp. 282-299.

Truscott, J. (1998) 'Noticing in second language acquisition: A critical review', *Second Language Research*, 14(2), pp. 103-135.

Truscott, J. and Sharwood Smith, M. (2004) 'How APT is your theory: Present status and future prospects', *Bilingualism: Language and Cognition*, 7(01), pp. 43-47.

Truscott, J. and Sharwood Smith, M. (2011). ‘Input, Intake, and Consciousness: The quest for a theoretical foundation’, *Studies in Second Language Acquisition*, 33(4), pp. 497-528.

Van den Branden, K. (2006) *Task-based Language Education: From Theory to Practice*. Cambridge University Press.

Van den Branden, K. (2009) 'Mediating between predetermined order and chaos: the role of the teacher in task-based language education', *International Journal of Applied Linguistics*, 19(3), pp. 264-285.

VanPatten, B. (1985) 'Communicative value and information processing in second language acquisition.’ in Larson, P., Judd, E. and Messerschmitt, D. (eds.) *On TESOL '84*. Washington, D.C.: TESOL, pp. 89-99.

VanPatten, B. (1990) 'Attending to form and content in the input', *Studies in Second Language Acquisition*, 12(3), pp. 287-301.

VanPatten, B. (1991) 'The foreign language classroom as a place to communicate', *Foreign Language Acquisition Research and The Classroom*, pp. 54-73.

VanPatten, B. (1993) 'Grammar Teaching for the Acquisition-Rich Classroom', *Foreign Language Annals*, 26(4), pp. 435-450.

VanPatten, B. (1996) 'Input Processing and Grammar Instruction: Theory and Research'. Norwood, NJ: Ablex.

VanPatten, B. (1997) 'How language teaching is constructed', *The Modern Language Journal*, 81(1), pp. 1-5.

VanPatten, B. (2002) 'Processing instruction: An update', *Language Learning*, 52(4), pp. 755-803.

VanPatten, B. (2003) *From Input to Output: A Teacher's Guide to Second Language Acquisition*. McGraw-Hill Higher Education.

VanPatten, B. (ed.) (2004) *Processing Instruction: Theory, Research, and Commentary*. Mahwah, NJ: Lawrence Erlbaum Associates.

VanPatten, B. and Cadierno, T. (1993) 'Explicit instruction and input processing', *Studies in Second Language Acquisition*, 15(2), pp. 225-243.

VanPatten, B. and Keating, G. (2007) 'Getting tense', *Paper presented at the annual meeting of the American Association for Applied Linguistics*. Costa Mesa, CA.

VanPatten, B. and Oikkenon, S. (1996) 'Explanation versus structured input in processing instruction', *Studies in Second Language Acquisition*, 18(04), pp. 495-510.

Vygotsky, L. (1978) *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L.(1962) *Thought and Language*. Cambridge MA: MIT Press.

Warschauer, M. (1996) 'Motivational aspects of using computers for writing and communication', *Telecollaboration in Foreign Language Learning*, pp. 29-46.

Warschauer, M. (2000) 'On-line learning in second language classrooms: An ethnographic study', *Network-Based Language Teaching: Concepts and practice*, pp. 41-58.

Webb, S. (2005) 'Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge', *Studies in Second Language Acquisition*, 27(01), pp. 33-52.

Wells, J. (2000) 'Effects of an on-line computer-mediated communication course, prior computer experience and internet knowledge, and learning styles on students' internet attitudes'.

Wesche, M. and Skehan, P. (2002) 'Communicative teaching, content-based instruction, and task-based learning', in Kaplan, R. (ed.) *Handbook of Applied Linguistics*. Oxford: Oxford University Press

White, L. (1988) 'Against comprehensible input: The input hypothesis and the development of second-language competence', *Applied Linguistics*, 8(2), pp. 95-110.

White, L. (1989) *Universal Grammar and Second Language Acquisition*. Philadelphia: John Benjamins.

White, L. (1991) 'Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom', *Second Language Research*, 7(2), pp. 133-161.

White, L. (2003) *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

Whong, M. (2011) *Language Teaching: Linguistic Theory in Practice*. Edinburgh: Edinburgh University Press.

Widdowson, H. (1978) *Teaching Language as Communication*. Oxford: Oxford University Press.

Widdowson, H. (2003) 'Expert beyond experience’: Notes on the appropriate use of theory in practice', in Newby, D. (ed.) *Mediating between Theory and Practice in the Context of Different Learning Cultures and Languages*. Strasbourg/Graz: Council of Europe Press.

Widdowson, H. (2008) 'Setting a good example', *Conference given in honour of Dr. Neil McLaren*. Granada, Spain.

Widdowson, H.G. (1972) 'The teaching of language as communication', *English Language Teaching* 27(1), pp. 15-19.

Wierzbicka, A. (1974). The semantics of direct and indirect discourse. *Papers in Linguistics,* 7, 267-307.

Wilkins, D. (1976) *Notional Syllabuses: a taxonomy and its relevance to foreign language curriculum development*. Oxford University Press.

Williams, J. (1995) 'Focus on Form in Communicative Language Teaching: Research Findings and the Classroom Teacher', *TESOL Journal* 7, pp. 6-11.

Williams, J. (2004) 'Implicit learning of form-meaning connections', in VanPatten, B., Williams, J., Rott, S. and Overstreet, M. (eds.) *Form-Meaning Connections in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 203-218.

Williams, J. (2005) 'Learning without awareness', *Studies in Second Language Acquisition*, 27(02), pp. 269-304.

Williams, J. (2013) 'Role of Instruction in Second Language Acquisition Theories', in *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd.

Williams, J. and Evans, J. (1998) 'What kind of focus and on which forms?', in Doughty, C. and Williams, J. (eds.) *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press, pp. 139–51.

Williams, J., Möbius, P. and Kim, C. (2001) 'Native and non-native processing of English wh- questions: Parsing strategies and plausibility constraints', *Applied Psycholinguistics*, 22(04), pp. 509-540.

Willis, D. and Willis, J. (2007) *Doing Task-Based Teaching: a practical guide to task-based teaching for elt training courses and practising teachers*. OUP: Oxford.

Willis, J. (1996) *A Framework for Task-Based Learning*. Longman.

Windeatt, S., Hardisty, D. and Eastment, D. (2000) *The Internet*. OUP Oxford.

Wong, W. (2001) 'Modality and attention to meaning and form in the input', *Studies in Second Language Acquisition*, 23(3), pp. 345-368.

Wright, C. (2010) 'Variation, asymmetry and working memory in the process of second language acquisition ', *Boston University Conference on Language Development* Boston. Cascadilla Press pp. 468-479.

Wright, C. (2012) 'An investigation of working memory effects on oral grammatical accuracy and fluency in producing questions in English', *TESOL Quarterly*, pp. 352–374.

Young, R. (2009) *Discursive Practice iIn Language Learning and Teaching*. Wiley-Blackwell Malden, MA.

Young-Scholten, M. 1995. The negative effects of 'positive' evidence on L2 phonology. Eubank, L., Selinker, L. and Sharwood Smith, M. (eds.) *The Current State of Interlanguage.* Amsterdam: Benjamins. pp. 107-122.

Zhao, Y. (1997) 'The effects of listeners' control of speech rate on second language comprehension', *Applied Linguistics*, 18(1), pp. 49-68.

Zuengler, J. and Miller, E.R. (2006) 'Cognitive and sociocultural perspectives: Two parallelz SLA worlds?', *TESOL quarterly*, 40(1), pp. 35-58.

Appendices

Appendix I Questionnaire

Questionnaire

Please answer the following questions.

1.Name /ID ……………………………….

2.Are you **Male** ☐ **Female** ☐

3.How old are you? …………………………………

4.What is your country of residence? …………………………………

5.What is your first language? ……………………..………..

6.What is your highest qualification? …………………………………

7.When did you start learning English? …………………………………

8.How long have you been here in the UK? …………………………………

9.How often do you use a computer?

Less than once a week ☐ Once a week ☐ 2-3 days a week ☐

Everyday ☐ More than once a day ☐

8. Where do you use a computer?

Home ☐ University ☐ Both ☐ Other ☐ please specify..................

9. Do you have a computer at home?

Yes ☐ No ☐

10. What do you use the computer for? Please rate using 1 - 5.

For example, if you use a computer mainly for games and sometimes for e-mail write 1 next to games and 2 next to e-mail.

Games ☐ Internet ☐ Learning ☐ E-mail ☐ Other ☐ please specify..

Please complete the following by ticking the answer that you feel is most appropriate

Competent = I could achieve the task without help

Good = I could do the task but would need help or instructions

Poor = I could not do the task

Task/skill	Poor	Good	Competent
11. Start the computer and log in			
12. Open an application from the start menu			
13. Log out and shut down the computer			
14. Type information using the keyboard			
15. Use the mouse to navigate in a page			
16. Use the keyboard to navigate in a page			
17. Use a mouse to select icons			
18. Use a mouse to open and close windows			
19. Single click and double click a mouse button			
20. Use a mouse to drag and drop objects			
21. Use a mouse to mark a block of text or part of an image			

22. How would you rate your current computer skills in general using the three criteria above?

Poor ☐ Good ☐ Competent ☐

Thank you

If you have any queries, please contact rola.naeb@ncl.ac.uk

Questionnaire

Please answer the following questions.

1.Name /ID ....MA........................

2.Are you **Male** ☑ **Female** ☐

3.How old are you? ....17..................

4.What is your country of residence? ....Saudi....................

5.What is your first language? ....Arabic...............

6.What is your highest qualification? ....High.....School.....

7.When did you start learning English? ..at...12..........................

8.How long have you been here in the UK? ....3.M......................

9.How often do you use a computer?

Less than once a week ☐ Once a week ☐ 2-3 days a week ☐

Everyday ☑ More than once a day ☐

8. Where do you use a computer?

Home ☑ University ☑ Both ☐ Other ☐ please specify...................

9. Do you have a computer at home?

Yes ☑ No ☐

10. What do you use the computer for? Please rate using 1 - 5.

For example, if you use a computer mainly for games and sometimes for e-mail write 1 next to games and 2 next to e-mail.

Games ☐ Internet ☑ Learning ☑ E-mail ☑ Other ☐ please specify

....

Please complete the following by ticking the answer that you feel is most appropriate

Competent = I could achieve the task without help

Good = I could do the task but would need help or instructions

Poor = I could not do the task

Task/skill	Poor	Good	Competent
11. Start the computer and log in			✓
12. Open an application from the start menu			✓
13. Log out and shut down the computer			✓
14. Type information using the keyboard			✓
15. Use the mouse to navigate in a page			✓
16. Use the keyboard to navigate in a page			✓
17. Use a mouse to select icons			✓
18. Use a mouse to open and close windows			✓
19. Single click and double click a mouse button			✓
20. Use a mouse to drag and drop objects			✓
21. Use a mouse to mark a block of text or part of an image			✓

22. How would you rate your current computer skills in general using the three criteria above?

Poor ☐ Good ☐ Competent ☑

Thank you

If you have any queries, please contact rola.naeb@ncl.ac.uk

Appendix II Consent form

Consent Form

Title of Project: **Focus on Form and Meaning**

Name of Researcher: **Rola Naeb**

1 I confirm that I understand the nature of the above study and have had the opportunity to ask questions

2 I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason

3 I understand that my participation is not going to affect my grades or any other aspect of my study

4 I agree / do not agree (delete as applicable) to take part in the above study.

________________________ __________ ___________________

Name of Participant *Date* *Signature*

________________________ __________ ___________________

Researcher *Date* *Signature*

Appendix III Overall communicative performance score analysis

Descriptives				
			Statistic	Std. Error
overall communicative performance score in the pre-test	Mean		24.2113	.54971
	95% Confidence Interval for Mean	Lower Bound	23.1149	
		Upper Bound	25.3076	
	5% Trimmed Mean		24.1792	
	Median		24.0000	
	Variance		21.455	
	Std. Deviation		4.63192	
	Minimum		14.00	
	Maximum		36.00	
	Range		22.00	
	Interquartile Range		7.00	
	Skewness		.068	.285
	Kurtosis		-.094	.563

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
overall communicative performance score in the pre-test	.106	71	.047	.985	71	.567
a. Lilliefors Significance Correction						

Overall communicative performance score –pre-test- treatment groups

Descriptives					
	treatment groups			Statistic	Std. Error
overall communicative performance score in the pre-test	FoS	Mean		24.4000	1.50683
		95% Confidence Interval for Mean	Lower Bound	21.2462	
			Upper Bound	27.5538	
		5% Trimmed Mean		24.3333	
		Median		24.0000	
		Variance		45.411	
		Std. Deviation		6.73873	
		Minimum		14.00	
		Maximum		36.00	
		Range		22.00	
		Interquartile Range		11.50	
		Skewness		.029	.512
		Kurtosis		-1.025	.992
	FoF	Mean		24.0000	.73258
		95% Confidence Interval for Mean	Lower Bound	22.4880	
			Upper Bound	25.5120	
		5% Trimmed Mean		24.0556	
		Median		24.0000	
		Variance		13.417	
		Std. Deviation		3.66288	
		Minimum		17.00	
		Maximum		30.00	
		Range		13.00	
		Interquartile Range		6.50	
		Skewness		-.133	.464
		Kurtosis		-.943	.902
	FoM	Mean		24.2692	.69627
		95% Confidence Interval for Mean	Lower Bound	22.8352	
			Upper Bound	25.7032	
		5% Trimmed Mean		24.2308	

		Median	25.0000	
		Variance	12.605	
		Std. Deviation	3.55030	
		Minimum	17.00	
		Maximum	32.00	
		Range	15.00	
		Interquartile Range	6.00	
		Skewness	.108	.456
		Kurtosis	-.442	.887

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	treatment groups	Statistic	df	Sig.	Statistic	df	Sig.
overall communicative performance score in the pre-test	FoS	.103	20	.200*	.960	20	.551
	FoF	.107	25	.200*	.965	25	.531
	FoM	.168	26	.058	.946	26	.182

a. Lilliefors Significance Correction
*. This is a lower bound of the true significance.

One-way overall communicative performance score pre--test vs. treatment groups

Descriptives

overall communicative performance score in the pre-test

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
FoS	20	24.4000	6.73873	1.50683	21.2462	27.5538	14.00	36.00
FoF	25	24.0000	3.66288	.73258	22.4880	25.5120	17.00	30.00
FoM	26	24.2692	3.55030	.69627	22.8352	25.7032	17.00	32.00
Total	71	24.2113	4.63192	.54971	23.1149	25.3076	14.00	36.00

Test of Homogeneity of Variances

overall communicative performance score in the pre-test

Levene Statistic	df1	df2	Sig.
8.414	2	68	.001

ANOVA

overall communicative performance score in the pre-test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.916	2	.958	.043	.958
Within Groups	1499.915	68	22.058		
Total	1501.831	70			

Post Hoc Tests

Multiple Comparisons

Dependent Variable:overall communicative performance score in the pre-test

	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Games-Howell	FoS	FoF	.40000	1.67547	.969	-3.7472	4.5472
		FoM	.13077	1.65992	.997	-3.9845	4.2460
	FoF	FoS	-.40000	1.67547	.969	-4.5472	3.7472
		FoM	-.26923	1.01067	.962	-2.7123	2.1739
	FoM	FoS	-.13077	1.65992	.997	-4.2460	3.9845
		FoF	.26923	1.01067	.962	-2.1739	2.7123

Homogeneous Subsets

overall communicative performance score in the pre-test			
			Subset for alpha = 0.05
	treatment groups	N	1
Tukey B[a,b]	FoF	25	24.0000
	FoM	26	24.2692
	FoS	20	24.4000
Means for groups in homogeneous subsets are displayed.			
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.			

NPar Tests overall communicative performance score pre-test vs treatment groups
Kruskal-Wallis Test

Ranks			
	treatment groups	N	Mean Rank
overall communicative performance score in the pre-test	FoS	20	37.25
	FoF	25	35.34
	FoM	26	35.67
	Total	71	

Test Statistics[a,b]	
	overall communicative performance score in the pre-test
Chi-Square	.106
df	2
Asymp. Sig.	.948
a. Kruskal Wallis Test b. Grouping Variable: treatment groups	

Descriptives				
			Statistic	Std. Error
overall communicative performance score in the posttest	Mean		28.5915	.58126
	95% Confidence Interval for Mean	Lower Bound	27.4323	
		Upper Bound	29.7508	
	5% Trimmed Mean		28.6800	
	Median		28.0000	
	Variance		23.988	
	Std. Deviation		4.89775	
	Minimum		17.00	
	Maximum		38.00	
	Range		21.00	
	Interquartile Range		8.00	
	Skewness		-.237	.285
	Kurtosis		-.410	.563

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
overall communicative performance score in the posttest	.081	71	.200*	.982	71	.392
a. Lilliefors Significance Correction *. This is a lower bound of the true significance.						

Descriptives						
	treatment groups				Statistic	Std. Error
overall communicative performance score in the posttest	FoS	Mean			26.2500	1.34140
		95% Confidence Interval for Mean		Lower Bound	23.4424	
				Upper Bound	29.0576	
		5% Trimmed Mean			26.1667	
		Median			26.0000	
		Variance			35.987	
		Std. Deviation			5.99890	
		Minimum			17.00	
		Maximum			37.00	
		Range			20.00	
		Interquartile Range			10.75	
		Skewness			.196	.512
		Kurtosis			-.971	.992
	FoF	Mean			30.8400	.89383
		95% Confidence Interval for Mean		Lower Bound	28.9952	
				Upper Bound	32.6848	
		5% Trimmed Mean			30.8778	
		Median			32.0000	
		Variance			19.973	
		Std. Deviation			4.46915	
		Minimum			23.00	
		Maximum			38.00	
		Range			15.00	
		Interquartile Range			8.00	
		Skewness			-.184	.464
		Kurtosis			-1.154	.902
	FoM	Mean			28.2308	.65235
		95% Confidence Interval for Mean		Lower Bound	26.8872	
				Upper Bound	29.5743	
		5% Trimmed Mean			28.3419	
		Median			28.0000	
		Variance			11.065	
		Std. Deviation			3.32635	
		Minimum			20.00	
		Maximum			34.00	
		Range			14.00	
		Interquartile Range			4.50	
		Skewness			-.480	.456
		Kurtosis			.224	.887

Tests of Normality							
		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	treatment groups	Statistic	df	Sig.	Statistic	df	Sig.
overall communicative performance score in the posttest	FoS	.131	20	.200*	.957	20	.492
	FoF	.126	25	.200*	.948	25	.231
	FoM	.203	26	.007	.963	26	.446

a. Lilliefors Significance Correction
*. This is a lower bound of the true significance.

Oneway overall communicative performance score post-test vs treatment groups

Descriptives								
overall communicative performance score in the posttest								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum

FoS	20	26.2500	5.99890	1.34140	23.4424	29.0576	17.00	37.00
FoF	25	30.8400	4.46915	.89383	28.9952	32.6848	23.00	38.00
FoM	26	28.2308	3.32635	.65235	26.8872	29.5743	20.00	34.00
Total	71	28.5915	4.89775	.58126	27.4323	29.7508	17.00	38.00

Test of Homogeneity of Variances

overall communicative performance score in the posttest

Levene Statistic	df1	df2	Sig.
5.047	2	68	.009

ANOVA

overall communicative performance score in the posttest

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	239.430	2	119.715	5.654	.005
Within Groups	1439.725	68	21.172		
Total	1679.155	70			

Post Hoc Tests

Multiple Comparisons

Dependent Variable:overall communicative performance score in the posttest

	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Games-Howell	FoS	FoF	-4.59000*	1.61192	.020	-8.5385	-.6415
		FoM	-1.98077	1.49161	.392	-5.6725	1.7110
	FoF	FoS	4.59000*	1.61192	.020	.6415	8.5385
		FoM	2.60923	1.10657	.058	-.0741	5.2925
	FoM	FoS	1.98077	1.49161	.392	-1.7110	5.6725
		FoF	-2.60923	1.10657	.058	-5.2925	.0741

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

overall communicative performance score in the posttest

	treatment groups	N	Subset for alpha = 0.05 1	Subset for alpha = 0.05 2
Tukey B[a,b]	FoS	20	26.2500	
	FoM	26	28.2308	28.2308
	FoF	25		30.8400

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 23.353.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

NPar Tests overall communicative performance score post-test vs treatment groups
Kruskal-Wallis Test

Ranks

	treatment groups	N	Mean Rank
overall communicative performance score in the posttest	FoS	20	27.18
	FoF	25	44.92
	FoM	26	34.21
	Total	71	

Test Statistics[a,b]

	overall communicative performance score in the posttest

Chi-Square	8.560
df	2
Asymp. Sig.	.014
a. Kruskal Wallis Test b. Grouping Variable: treatment groups	

T-Test Treatment effect overall communicative performance score

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	overall communicative performance score in the posttest	28.5915	71	4.89775	.58126
	overall communicative performance score in the pre-test	24.2113	71	4.63192	.54971

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	overall communicative performance score in the posttest and in the pre-test	71	.810	.000

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	overall communicative performance score in the posttest – and in the pre-test	4.38028	2.94893	.34997	3.68228	5.07828	12.516	70	.000

Appendix IV Target Linguistic Performance Score Analysis

Explore Target linguistic performance score pre-test all cases

Descriptives

			Statistic	Std. Error
target linguistic performance score in the pre- test	Mean		11.0423	.40887
	95% Confidence Interval for Mean	Lower Bound	10.2268	
		Upper Bound	11.8577	
	5% Trimmed Mean		11.0383	
	Median		11.0000	
	Variance		11.870	
	Std. Deviation		3.44523	
	Minimum		4.00	
	Maximum		17.00	
	Range		13.00	
	Interquartile Range		6.00	
	Skewness		.006	.285
	Kurtosis		-1.030	.563

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
target linguistic performance score in the pre- test	.104	71	.057	.963	71	.033

a. Lilliefors Significance Correction

Explore target linguistic performance score pre-test treatment groups

Descriptives

	treatment groups			Statistic	Std. Error
target linguistic performance score in the pre- test	FoS	Mean		10.6000	.91881
		95% Confidence Interval for Mean	Lower Bound	8.6769	
			Upper Bound	12.5231	
		5% Trimmed Mean		10.6111	
		Median		11.5000	
		Variance		16.884	
		Std. Deviation		4.10904	
		Minimum		4.00	
		Maximum		17.00	
		Range		13.00	
		Interquartile Range		7.00	
		Skewness		-.040	.512
		Kurtosis		-1.368	.992
	FoF	Mean		11.9200	.58572
		95% Confidence Interval for Mean	Lower Bound	10.7111	
			Upper Bound	13.1289	
		5% Trimmed Mean		11.9111	
		Median		12.0000	
		Variance		8.577	
		Std. Deviation		2.92859	
		Minimum		7.00	
		Maximum		17.00	
		Range		10.00	
		Interquartile Range		5.50	
		Skewness		.161	.464

		Kurtosis		-1.146	.902
	FoM	Mean		10.5385	.64981
		95% Confidence Interval for Mean	Lower Bound	9.2002	
			Upper Bound	11.8768	
		5% Trimmed Mean		10.4444	
		Median		10.0000	
		Variance		10.978	
		Std. Deviation		3.31338	
		Minimum		6.00	
		Maximum		17.00	
		Range		11.00	
		Interquartile Range		5.25	
		Skewness		.313	.456
		Kurtosis		-.984	.887

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	treatment groups	Statistic	df	Sig.	Statistic	df	Sig.
target linguistic performance score in the pre- test	FoS	.170	20	.130	.927	20	.136
	FoF	.144	25	.193	.944	25	.181
	FoM	.140	26	.200*	.941	26	.143

a. Lilliefors Significance Correction
*. This is a lower bound of the true significance.

Oneway target linguistic performance score pre-test vs treatment groups

Descriptives

target linguistic performance score in the pre- test

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
FoS	20	10.6000	4.10904	.91881	8.6769	12.5231	4.00	17.00
FoF	25	11.9200	2.92859	.58572	10.7111	13.1289	7.00	17.00
FoM	26	10.5385	3.31338	.64981	9.2002	11.8768	6.00	17.00
Total	71	11.0423	3.44523	.40887	10.2268	11.8577	4.00	17.00

Test of Homogeneity of Variances

target linguistic performance score in the pre- test

Levene Statistic	df1	df2	Sig.
3.146	2	68	.049

ANOVA

target linguistic performance score in the pre- test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	29.772	2	14.886	1.264	.289
Within Groups	801.102	68	11.781		
Total	830.873	70			

Post Hoc Tests

Multiple Comparisons

Dependent Variable:target linguistic performance score in the pre- test

	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Games-Howell	FoS	FoF	-1.32000	1.08962	.455	-3.9928	1.3528
		FoM	.06154	1.12537	.998	-2.6894	2.8125

	FoF	FoS	1.32000	1.08962	.455	-1.3528	3.9928
		FoM	1.38154	.87482	.264	-.7333	3.4964
	FoM	FoS	-.06154	1.12537	.998	-2.8125	2.6894
		FoF	-1.38154	.87482	.264	-3.4964	.7333

Homogeneous Subsets

target linguistic performance score in the pre- test			
			Subset for alpha = 0.05
	treatment groups	N	1
Tukey B[a,b]	FoM	26	10.5385
	FoS	20	10.6000
	FoF	25	11.9200

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 23.353.
b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

NPar Tests
Kruskal-Wallis Test

Ranks			
	treatment groups	N	Mean Rank
target linguistic performance score in the pre- test	FoS	20	33.58
	FoF	25	41.24
	FoM	26	32.83
	Total	71	

Test Statistics[a,b]	
	target linguistic performance score in the pre- test
Chi-Square	2.521
df	2
Asymp. Sig.	.284

a. Kruskal Wallis Test
b. Grouping Variable: treatment groups

Explore Target linguistic performance score post-test all cases

Descriptives				
			Statistic	Std. Error
Target linguistic performance score in the post test	Mean		14.5352	.38316
	95% Confidence Interval for Mean	Lower Bound	13.7710	
		Upper Bound	15.2994	
	5% Trimmed Mean		14.5305	
	Median		14.0000	
	Variance		10.424	
	Std. Deviation		3.22858	
	Minimum		8.00	
	Maximum		21.00	
	Range		13.00	
	Interquartile Range		5.00	
	Skewness		.077	.285
	Kurtosis		-.915	.563

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Target linguistic performance score in the post test	.101	71	.070	.968	71	.064
a. Lilliefors Significance Correction						

Target linguistic performance score in the post test

Descriptives					
	treatment groups			Statistic	Std. Error
Target linguistic performance score in the post test	FoS	Mean		14.4000	.92452
		95% Confidence Interval for Mean	Lower Bound	12.4650	
			Upper Bound	16.3350	
		5% Trimmed Mean		14.3889	
		Median		14.0000	
		Variance		17.095	
		Std. Deviation		4.13458	
		Minimum		8.00	
		Maximum		21.00	
		Range		13.00	
		Interquartile Range		8.50	
		Skewness		.070	.512
		Kurtosis		-1.469	.992
	FoF	Mean		15.2400	.61989
		95% Confidence Interval for Mean	Lower Bound	13.9606	
			Upper Bound	16.5194	
		5% Trimmed Mean		15.2556	
		Median		15.0000	
		Variance		9.607	
		Std. Deviation		3.09946	
		Minimum		10.00	
		Maximum		20.00	
		Range		10.00	
		Interquartile Range		5.00	
		Skewness		.012	.464
		Kurtosis		-1.201	.902
	FoM	Mean		13.9615	.48510
		95% Confidence Interval for Mean	Lower Bound	12.9624	
			Upper Bound	14.9606	
		5% Trimmed Mean		13.9573	
		Median		14.0000	
		Variance		6.118	
		Std. Deviation		2.47355	
		Minimum		9.00	
		Maximum		19.00	
		Range		10.00	
		Interquartile Range		3.25	
		Skewness		-.038	.456
		Kurtosis		-.340	.887

Tests of Normality							
	treatment groups	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Target linguistic performance score in the post test	FoS	.169	20	.136	.927	20	.137
	FoF	.173	25	.051	.932	25	.096
	FoM	.160	26	.085	.975	26	.759
a. Lilliefors Significance Correction							

Target linguistic performance score in the post test

Oneway target linguistic performance score post-test vs treatment groups

Descriptives								
Target linguistic performance score in the post test								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
FoS	20	14.4000	4.13458	.92452	12.4650	16.3350	8.00	21.00
FoF	25	15.2400	3.09946	.61989	13.9606	16.5194	10.00	20.00
FoM	26	13.9615	2.47355	.48510	12.9624	14.9606	9.00	19.00
Total	71	14.5352	3.22858	.38316	13.7710	15.2994	8.00	21.00

Test of Homogeneity of Variances			
Target linguistic performance score in the post test			
Levene Statistic	df1	df2	Sig.
6.803	2	68	.002

ANOVA					
Target linguistic performance score in the post test					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21.340	2	10.670	1.024	.364
Within Groups	708.322	68	10.416		
Total	729.662	70			

Post Hoc Tests

Multiple Comparisons							
Dependent Variable:Target linguistic performance score in the post test							
	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Games-Howell	FoS	FoF	-.84000	1.11311	.733	-3.5661	1.8861
		FoM	.43846	1.04406	.908	-2.1390	3.0159
	FoF	FoS	.84000	1.11311	.733	-1.8861	3.5661
		FoM	1.27846	.78714	.246	-.6280	3.1850
	FoM	FoS	-.43846	1.04406	.908	-3.0159	2.1390
		FoF	-1.27846	.78714	.246	-3.1850	.6280

Homogeneous Subsets

Target linguistic performance score in the post test			
			Subset for alpha = 0.05
	treatment groups	N	1
Tukey B[a,b]	FoM	26	13.9615
	FoS	20	14.4000
	FoF	25	15.2400
Means for groups in homogeneous subsets are displayed.			
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.			

NPar Tests target linguistic performance score post-test vs treatment groups
Kruskal-Wallis Test

Ranks			
	treatment groups	N	Mean Rank
Target linguistic performance score in the post test	FoS	20	34.92
	FoF	25	40.18
	FoM	26	32.81
	Total	71	

Test Statistics[a,b]	
	Target linguistic performance score in the post test
Chi-Square	1.719
df	2
Asymp. Sig.	.423
a. Kruskal Wallis Test b. Grouping Variable: treatment groups	

T-Test target linguistic performance score treatment effect pre vs post test scores

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Target linguistic performance score in the post test	14.5352	71	3.22858	.38316
	target linguistic performance score in the pre- test	11.0423	71	3.44523	.40887

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Target linguistic performance score in the post test and target linguistic performance score in the pre- test	71	.765	.000

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Target linguistic performance score in the post test - target linguistic performance score in the pre- test	3.49296	2.29828	.27276	2.94896	4.03695	12.806	70	.000

Appendix V Mixed ANOVA Analysis

General Linear Model- Overall communicative performance score

Within-Subjects Factors

Measure: MEASURE_1

Test	Dependent Variable
1	spretest
2	sposttest

Between-Subjects Factors

		Value Label	N
treatment groups	1.00	FoS	20
	2.00	FoF	25
	3.00	FoM	26

Descriptive Statistics

	treatment groups	Mean	Std. Deviation	N
overall communicative performance score in the pretest	FoS	24.4000	6.73873	20
	FoF	24.0000	3.66288	25
	FoM	24.2692	3.55030	26
	Total	24.2113	4.63192	71
overall communicative performance score in the posttest	FoS	26.2500	5.99890	20
	FoF	30.8400	4.46915	25
	FoM	28.2308	3.32635	26
	Total	28.5915	4.89775	71

Box's Test of Equality of Covariance Matrices[a]

Box's M	16.736
F	2.669
df1	6
df2	81409.619
Sig.	.014

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + inputtype

Within Subjects Design: Test

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Test	Pillai's Trace	.793	260.802[b]	1.000	68.000	.000
	Wilks' Lambda	.207	260.802[b]	1.000	68.000	.000
	Hotelling's Trace	3.835	260.802[b]	1.000	68.000	.000
	Roy's Largest Root	3.835	260.802[b]	1.000	68.000	.000
Test * inputtype	Pillai's Trace	.466	29.708[b]	2.000	68.000	.000
	Wilks' Lambda	.534	29.708[b]	2.000	68.000	.000
	Hotelling's Trace	.874	29.708[b]	2.000	68.000	.000
	Roy's Largest Root	.874	29.708[b]	2.000	68.000	.000

a. Design: Intercept + inputtype
Within Subjects Design: Test

b. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Test	1.000	.000	0	.	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + inputtype
Within Subjects Design: Test

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Test	Sphericity Assumed	622.994	1	622.994	260.802	.000
	Greenhouse-Geisser	622.994	1.000	622.994	260.802	.000
	Huynh-Feldt	622.994	1.000	622.994	260.802	.000
	Lower-bound	622.994	1.000	622.994	260.802	.000
Test * inputtype	Sphericity Assumed	141.930	2	70.965	29.708	.000
	Greenhouse-Geisser	141.930	2.000	70.965	29.708	.000
	Huynh-Feldt	141.930	2.000	70.965	29.708	.000
	Lower-bound	141.930	2.000	70.965	29.708	.000

Error(Test)	Sphericity Assumed	162.436	68	2.389		
	Greenhouse-Geisser	162.436	68.000	2.389		
	Huynh-Feldt	162.436	68.000	2.389		
	Lower-bound	162.436	68.000	2.389		

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Test	Type III Sum of Squares	df	Mean Square	F	Sig.
Test	Linear	622.994	1	622.994	260.802	.000
Test * inputtype	Linear	141.930	2	70.965	29.708	.000
Error(Test)	Linear	162.436	68	2.389		

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
overall communicative performance score in the pretest	8.414	2	68	.001
overall communicative performance score in the posttest	5.047	2	68	.009

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + inputtype

Within Subjects Design: Test

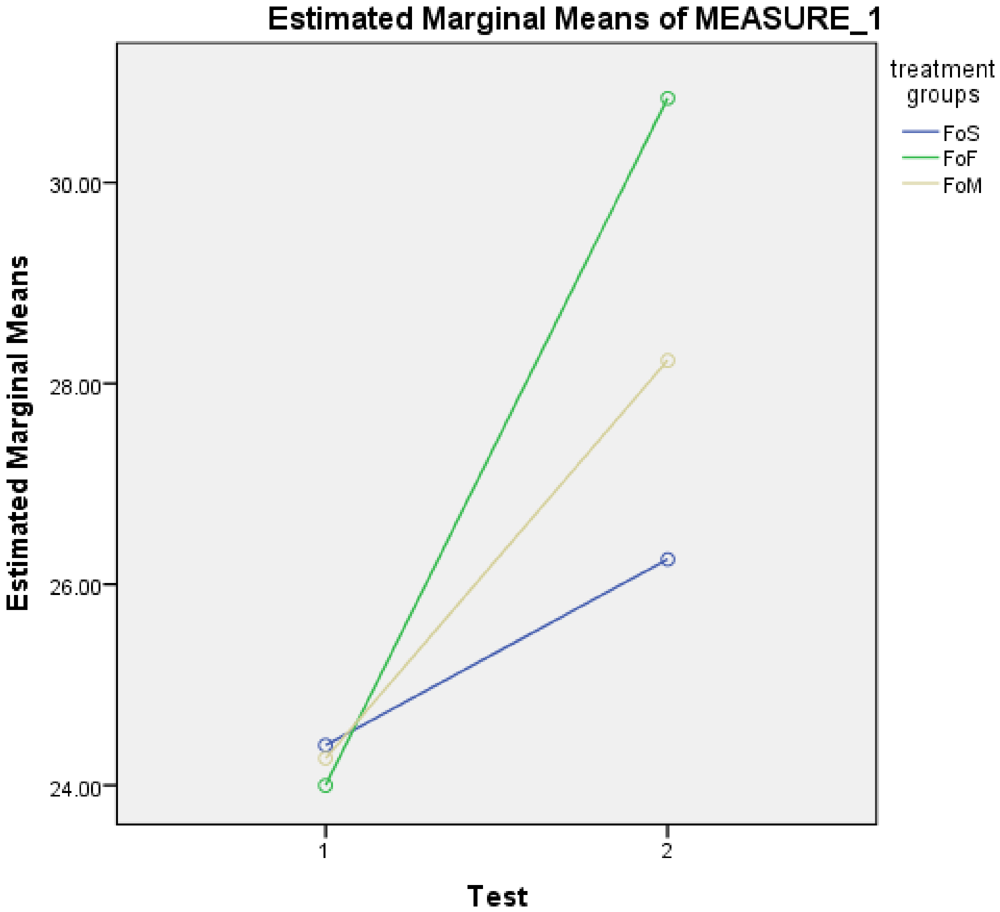
Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	97152.970	1	97152.970	2378.795	.000
inputtype	99.415	2	49.707	1.217	.302
Error	2777.205	68	40.841		

Profile Plots


Estimated Marginal Means of MEASURE_1
Estimated Marginal Means
30.00
28.00
26.00
24.00
1
2
Test
treatment groups
FoS
FoF
FoM

General Linear Model- Target linguistic performance score

Within-Subjects Factors

Measure: MEASURE_1

Test	Dependent Variable
1	TSpretest
2	TSpostest

Between-Subjects Factors

		Value Label	N
treatment groups	1.00	FoS	20
	2.00	FoF	25
	3.00	FoM	26

Descriptive Statistics

	treatment groups	Mean	Std. Deviation	N
target linguistic performance score in the pre test	FoS	10.6000	4.10904	20
	FoF	11.9200	2.92859	25
	FoM	10.5385	3.31338	26
	Total	11.0423	3.44523	71
Target linguistic performance score in the post test	FoS	14.4000	4.13458	20
	FoF	15.2400	3.09946	25
	FoM	13.9615	2.47355	26
	Total	14.5352	3.22858	71

Box's Test of Equality of Covariance Matrices[a]

Box's M	13.025
F	2.077
df1	6
df2	81409.619
Sig.	.052

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + inputtype

Within Subjects Design: Test

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Test	Pillai's Trace	.702	160.332[b]	1.000	68.000	.000
	Wilks' Lambda	.298	160.332[b]	1.000	68.000	.000
	Hotelling's Trace	2.358	160.332[b]	1.000	68.000	.000
	Roy's Largest Root	2.358	160.332[b]	1.000	68.000	.000
Test * inputtype	Pillai's Trace	.007	.256[b]	2.000	68.000	.775
	Wilks' Lambda	.993	.256[b]	2.000	68.000	.775
	Hotelling's Trace	.008	.256[b]	2.000	68.000	.775
	Roy's Largest Root	.008	.256[b]	2.000	68.000	.775

a. Design: Intercept + inputtype
 Within Subjects Design: Test

b. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Test	1.000	.000	0	.	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + inputtype
 Within Subjects Design: Test

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Test	Sphericity Assumed	432.645	1	432.645	160.332	.000
	Greenhouse-Geisser	432.645	1.000	432.645	160.332	.000
	Huynh-Feldt	432.645	1.000	432.645	160.332	.000
	Lower-bound	432.645	1.000	432.645	160.332	.000
Test * inputtype	Sphericity Assumed	1.380	2	.690	.256	.775
	Greenhouse-Geisser	1.380	2.000	.690	.256	.775
	Huynh-Feldt	1.380	2.000	.690	.256	.775
	Lower-bound	1.380	2.000	.690	.256	.775
Error(Test)	Sphericity Assumed	183.493	68	2.698		
	Greenhouse-Geisser	183.493	68.000	2.698		
	Huynh-Feldt	183.493	68.000	2.698		
	Lower-bound	183.493	68.000	2.698		

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Test	Type III Sum of Squares	df	Mean Square	F	Sig.
Test	Linear	432.645	1	432.645	160.332	.000
Test * inputtype	Linear	1.380	2	.690	.256	.775
Error(Test)	Linear	183.493	68	2.698		

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
target linguistic performance score in the pre test	3.146	2	68	.049
Target linguistic performance score in the post test	6.803	2	68	.002

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + inputtype

Within Subjects Design: Test

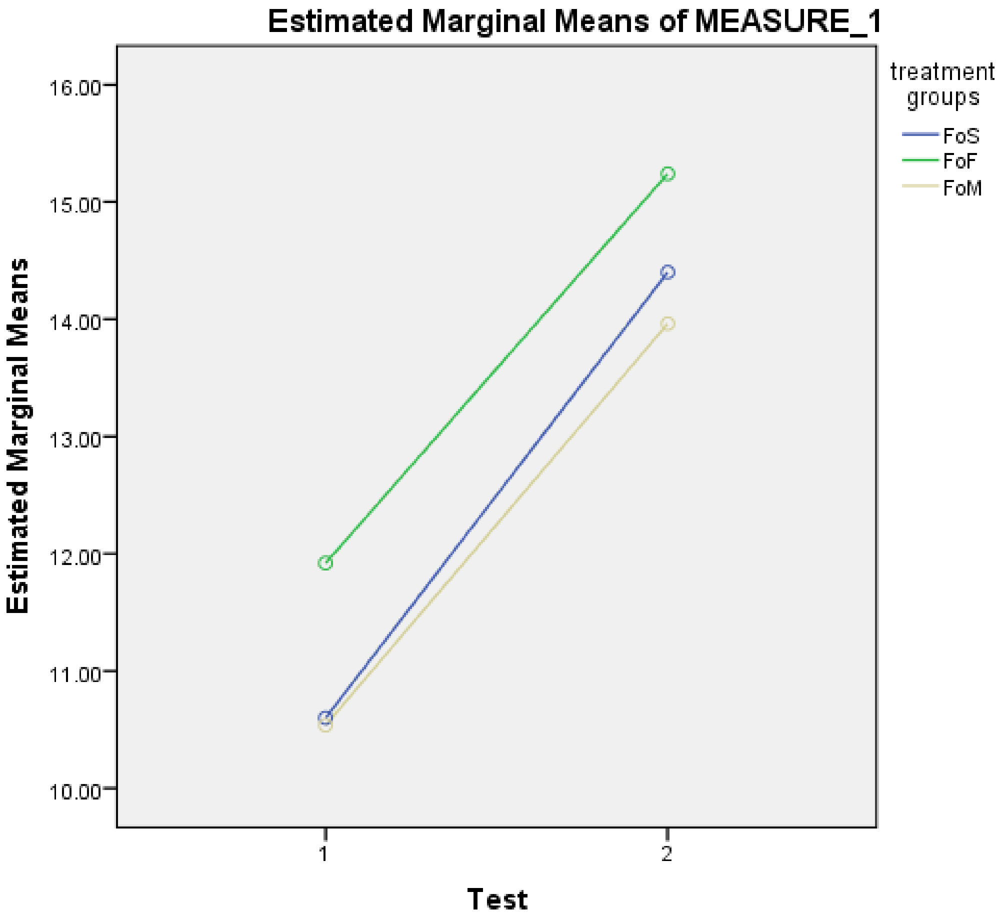
Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	22873.600	1	22873.600	1173.067	.000
inputtype	49.732	2	24.866	1.275	.286
Error	1325.930	68	19.499		

Profile Plots



Appendix VI Processing time analysis

Explore normality distribution treatment groups

Descriptives

	treatment groups			Statistic	Std. Error
processing time for whole session	FoS	Mean		3418.80	161.710
		95% Confidence Interval for Mean	Lower Bound	3080.34	
			Upper Bound	3757.26	
		5% Trimmed Mean		3406.39	
		Median		3134.00	
		Variance		523005.116	
		Std. Deviation		723.191	
		Minimum		2287	
		Maximum		4774	
		Range		2487	
		Interquartile Range		1142	
		Skewness		.579	.512
		Kurtosis		-.819	.992
	FoF	Mean		3475.68	173.364
		95% Confidence Interval for Mean	Lower Bound	3117.87	
			Upper Bound	3833.49	
		5% Trimmed Mean		3477.84	
		Median		3277.00	
		Variance		751374.143	
		Std. Deviation		866.818	
		Minimum		1851	
		Maximum		5040	
		Range		3189	
		Interquartile Range		1477	
		Skewness		.150	.464
		Kurtosis		-.696	.902
	FoM	Mean		3901.23	184.234
		95% Confidence Interval for Mean	Lower Bound	3521.79	
			Upper Bound	4280.67	
		5% Trimmed Mean		3908.42	
		Median		3908.50	
		Variance		882497.705	
		Std. Deviation		939.413	
		Minimum		1952	
		Maximum		5917	
		Range		3965	
		Interquartile Range		1499	
		Skewness		-.204	.456
		Kurtosis		.036	.887

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	treatment groups	Statistic	df	Sig.	Statistic	df	Sig.
processing time for whole session	FoS	.179	20	.093	.919	20	.096
	FoF	.128	25	.200*	.967	25	.572
	FoM	.097	26	.200*	.973	26	.689

a. Lilliefors Significance Correction
*. This is a lower bound of the true significance.

Processing time for whole session
Oneway processing time vs treatment groups input effect

Descriptives								
processing time for whole session								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
FoS	20	3418.80	723.191	161.710	3080.34	3757.26	2287	4774
FoF	25	3475.68	866.818	173.364	3117.87	3833.49	1851	5040
FoM	26	3901.23	939.413	184.234	3521.79	4280.67	1952	5917
Total	71	3615.49	873.562	103.673	3408.72	3822.26	1851	5917

Test of Homogeneity of Variances			
processing time for whole session			
Levene Statistic	df1	df2	Sig.
.343	2	68	.711

ANOVA					
processing time for whole session					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3385252.491	2	1692626.246	2.300	.108
Within Groups	50032519.255	68	735772.342		
Total	53417771.746	70			

Post Hoc Tests

Multiple Comparisons							
Dependent Variable: processing time for whole session							
						95% Confidence Interval	
	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Games-Howell	FoS	FoF	-56.880	237.076	.969	-632.41	518.65
		FoM	-482.431	245.138	.132	-1077.01	112.15
	FoF	FoS	56.880	237.076	.969	-518.65	632.41
		FoM	-425.551	252.977	.222	-1037.01	185.90
	FoM	FoS	482.431	245.138	.132	-112.15	1077.01
		FoF	425.551	252.977	.222	-185.90	1037.01

Homogeneous Subsets

processing time for whole session			
			Subset for alpha = 0.05
	treatment groups	N	1
Tukey B[a,b]	FoS	20	3418.80
	FoF	25	3475.68
	FoM	26	3901.23
Means for groups in homogeneous subsets are displayed.			
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.			

Correlations processing time vs communicative performance

Descriptive Statistics			
	Mean	Std. Deviation	N
overall communicative performance score	28.5915	4.89775	71

in the posttest			
processing time for whole session	3615.49	873.562	71

Correlations			
		overall communicative performance score in the posttest	processing time for whole session
overall communicative performance score in the posttest	Pearson Correlation	1	-.316**
	Sig. (2-tailed)		.007
	N	71	71
processing time for whole session	Pearson Correlation	-.316**	1
	Sig. (2-tailed)	.007	
	N	71	71
**. Correlation is significant at the 0.01 level (2-tailed).			

Correlations processing time vs linguistic performance

Descriptive Statistics			
	Mean	Std. Deviation	N
processing time for whole session	3615.49	873.562	71
Target linguistic performance score in the post test	14.5352	3.22858	71

Correlations			
		processing time for whole session	Target linguistic performance score in the post test
processing time for whole session	Pearson Correlation	1	-.314**
	Sig. (2-tailed)		.008
	N	71	71
Target linguistic performance score in the post test	Pearson Correlation	-.314**	1
	Sig. (2-tailed)	.008	
	N	71	71
**. Correlation is significant at the 0.01 level (2-tailed).			

Appendix VII Trial number analysis

Oneway Trial number pre-test vs treatment groups

Descriptives

Trial number for pre- test

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
FoS	20	60.2500	8.55862	1.91377	56.2444	64.2556	46.00	79.00
FoF	25	61.7200	15.05136	3.01027	55.5071	67.9329	26.00	91.00
FoM	26	62.1154	9.32235	1.82826	58.3500	65.8808	44.00	75.00
Total	71	61.4507	11.36635	1.34894	58.7603	64.1411	26.00	91.00

Test of Homogeneity of Variances

Trial number for pre- test

Levene Statistic	df1	df2	Sig.
3.663	2	68	.031

ANOVA

Trial number for pre- test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	42.134	2	21.067	.159	.853
Within Groups	9001.444	68	132.374		
Total	9043.577	70			

Post Hoc Tests

Multiple Comparisons

Dependent Variable:Trial number for pre- test

	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Games-Howell	FoS	FoF	-1.47000	3.56710	.911	-10.1586	7.2186
		FoM	-1.86538	2.64670	.762	-8.2924	4.5617
	FoF	FoS	1.47000	3.56710	.911	-7.2186	10.1586
		FoM	-.39538	3.52197	.993	-8.9694	8.1787
	FoM	FoS	1.86538	2.64670	.762	-4.5617	8.2924
		FoF	.39538	3.52197	.993	-8.1787	8.9694

Homogeneous Subsets

Trial number for pre- test

	treatment groups	N	Subset for alpha = 0.05
			1
Tukey B[a,b]	FoS	20	60.2500
	FoF	25	61.7200
	FoM	26	62.1154

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 23.353.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

NPar Tests Trial number pre-test vs treatment groups

Kruskal-Wallis Test

Ranks			
	treatment groups	N	Mean Rank
Trial number for pre- test	FoS	20	33.53
	FoF	25	35.94
	FoM	26	37.96
	Total	71	

Test Statistics[a,b]	
	Trial number for pre- test
Chi-Square	.523
df	2
Asymp. Sig.	.770

a. Kruskal Wallis Test
b. Grouping Variable: treatment groups

Oneway Trial number post-test vs treatment groups

Descriptives

trial number for post test

					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
FoS	20	43.6500	6.17529	1.38084	40.7599	46.5401	30.00	54.00
FoF	25	48.7600	7.46257	1.49251	45.6796	51.8404	38.00	64.00
FoM	26	52.0000	9.66644	1.89574	48.0956	55.9044	31.00	65.00
Total	71	48.5070	8.60876	1.02167	46.4694	50.5447	30.00	65.00

Test of Homogeneity of Variances

trial number for post test

Levene Statistic	df1	df2	Sig.
2.204	2	68	.118

ANOVA

trial number for post test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	790.636	2	395.318	6.113	.004
Within Groups	4397.110	68	64.663		
Total	5187.746	70			

Post Hoc Tests

Multiple Comparisons

trial number for post test
Tukey HSD

					95% Confidence Interval	
(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
FoS	FoF	-5.11000	2.41241	.094	-10.8903	.6703
	FoM	-8.35000*	2.39170	.002	-14.0807	-2.6193
FoF	FoS	5.11000	2.41241	.094	-.6703	10.8903
	FoM	-3.24000	2.25246	.327	-8.6371	2.1571
FoM	FoS	8.35000*	2.39170	.002	2.6193	14.0807
	FoF	3.24000	2.25246	.327	-2.1571	8.6371

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

trial number for post test			
Tukey HSD[a,b]			
		Subset for alpha = 0.05	
treatment groups	N	1	2
FoS	20	43.6500	
FoF	25	48.7600	48.7600
FoM	26		52.0000
Sig.		.083	.359
Means for groups in homogeneous subsets are displayed.			
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.			

T-Test Trial number treatment effect comparing trials in pre- and post-tests

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	trial number for post test	48.5070	71	8.60876	1.02167
	Trial number for pre- test	61.4507	71	11.36635	1.34894

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	trial number for post test and Trial number for pre test	71	.209	.081

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	trial number for post test - Trial number for pre-test	-12.94366	12.74574	1.51264	-15.96053	-9.92679	-8.557	70	.000

Correlations between pre-test trial number and communicative performance

Descriptive Statistics			
	Mean	Std. Deviation	N
Trial number for pre- test	61.4507	11.36635	71
overall communicative performance score in the posttest	28.5915	4.89775	71

Correlations			
		Trial number for pre- test	overall communicative performance score in the posttest
Trial number for pre- test	Pearson Correlation	1	-.140
	Sig. (2-tailed)		.246
	N	71	71
overall communicative performance score in the posttest	Pearson Correlation	-.140	1
	Sig. (2-tailed)	.246	
	N	71	71

Correlations between post-test trial number and communicative performance

Descriptive Statistics			
	Mean	Std. Deviation	N
trial number for post test	48.5070	8.60876	71
overall communicative performance score in the posttest	28.5915	4.89775	71

Correlations			
		trial number for post test	overall communicative performance score in the posttest
trial number for post test	Pearson Correlation	1	.051
	Sig. (2-tailed)		.672
	N	71	71
overall communicative performance score in the posttest	Pearson Correlation	.051	1
	Sig. (2-tailed)	.672	
	N	71	71

Correlations between pre-test trial number and linguistic performance

Correlations			
		Trial number for pre- test	Target linguistic performance score in the post test
Trial number for pre- test	Pearson Correlation	1	-.218
	Sig. (2-tailed)		.068
	N	71	71
Target linguistic performance score in the post test	Pearson Correlation	-.218	1
	Sig. (2-tailed)	.068	
	N	71	71

Correlations

Correlations			
		Target linguistic performance score in the post test	trial number for post test
Target linguistic performance score in the post test	Pearson Correlation	1	.040
	Sig. (2-tailed)		.741
	N	71	71
trial number for post test	Pearson Correlation	.040	1
	Sig. (2-tailed)	.741	
	N	71	71

Appendix VIII Task type Analysis

Descriptives
treatment groups = FoS

Descriptive Statistics[a]					
	N	Minimum	Maximum	Mean	Std. Deviation
Zscore(Spre-1)	20	-2.65425	1.91982	.0749462	1.48888377
Zscore(Spost2)	20	-2.97195	1.00934	-.4978618	1.27211187
Zscore(Spre2)	20	-2.32198	2.31604	-.0240512	1.16458982
Zscore(Spost1)	20	-2.60634	2.06235	-.2507702	1.16417327
Zscore(TSpre1)	20	-1.58626	1.24635	.0000000	1.00000000
Zscore(TSpost2)	20	-1.58769	1.13406	.0000000	1.00000000
Zscore(TSpre2)	20	-.97468	1.94936	.0000000	1.00000000
Zscore(TSpost1)	20	-1.50940	2.00083	.0000000	1.00000000
Zscore: Trial number in the pre- test task 1	20	-1.46834	3.12023	.0000000	1.00000000
Zscore: Trial number in the post test task 2	20	-2.28336	1.67864	.0000000	1.00000000
Zscore: Trial number in the pre- test task 2	20	-1.32938	2.21563	.0000000	1.00000000
Zscore: Trial number in the post test task 2	20	-.86715	2.66569	.0000000	1.00000000
Valid N (listwise)	20				
a. treatment groups = FoS					

treatment groups = FoF

Descriptive Statistics[a]					
	N	Minimum	Maximum	Mean	Std. Deviation
Zscore(Spre1)	25	-1.73944	1.00501	.0536005	.72760516
Zscore(Spost2)	25	-1.55006	1.57810	.4519608	.90481089
Zscore(Spre2)	25	-1.47871	2.31604	-.1631921	.94532408
Zscore(Spost1)	25	-1.33306	2.06235	.2797627	.92501580
Zscore(TSpre1)	25	-1.66701	2.04663	.0000000	1.00000000
Zscore(TSpost2)	25	-2.12843	1.13889	.0000000	1.00000000
Zscore(TSpre2)	25	-.80557	1.94068	.0000000	1.00000000
Zscore(TSpost1)	25	-1.01740	2.01058	.0000000	1.00000000
Zscore: Trial number in the pre test task 1	25	-1.34039	2.55398	.0000000	1.00000000
Zscore: Trial number in the post test task 2	25	-1.92704	1.65815	.0000000	1.00000000
Zscore: Trial number in the pre test task 2	25	-2.09469	1.81032	.0000000	1.00000000
Zscore: Trial number in the post test task 2	25	-1.15273	1.84137	.0000000	1.00000000
Valid N (listwise)	25				
a. treatment groups = FoF					

treatment groups = FoM

Descriptive Statistics[a]					
	N	Minimum	Maximum	Mean	Std. Deviation
Zscore(Spre1)	26	-1.43450	1.00501	-.1091898	.76632389
Zscore(Spost2)	26	-1.26568	1.00934	-.0516071	.61550703
Zscore(Spre2)	26	-2.32198	1.89441	.1754164	.92317221
Zscore(Spost1)	26	-2.18191	1.63793	-.0761025	.90218923
Zscore(TSpre1)	26	-1.45541	1.95365	.0000000	1.00000000
Zscore(TSpost2)	26	-1.53743	1.91704	.0000000	1.00000000
Zscore(TSpre2)	26	-1.04168	2.24119	.0000000	1.00000000
Zscore(TSpost1)	26	-1.19785	1.39749	.0000000	1.00000000
Zscore: Trial number in the pre test task 1	26	-1.71888	2.37999	.0000000	1.00000000
Zscore: Trial number in the post test task 2	26	-1.89503	1.97884	.0000000	1.00000000

Zscore: Trial number in the pre test task 2	26	-1.71156	1.39641	.0000000	1.00000000
Zscore: Trial number in the post test task 2	26	-1.67883	1.72436	.0000000	1.00000000
Valid N (listwise)	26				
a. treatment groups = FoM					

Descriptives
treatment groups = FoS

Descriptive Statistics[a]					
	N	Minimum	Maximum	Mean	Std. Deviation
Spre1	20	8.00	23.00	16.9500	4.88257
TSpre1	20	4.00	14.00	9.6000	3.53032
Trial number in the pre-test task 1	20	30.00	55.00	38.0000	5.44832
Spost2	20	9.00	23.00	17.7000	4.47331
TSpost2	20	7.00	16.00	12.2500	3.30669
Trial number in the post test task 2	20	17.00	36.00	27.9500	4.79556
Spre2	20	2.00	13.00	7.4500	2.76205
TSpre2	20	.00	3.00	1.0000	1.02598
Trial number in the pre test task 2	20	14.00	36.00	22.2500	6.20590
Spost1	20	3.00	14.00	8.5500	2.74293
TSpost1	20	.00	5.00	2.1500	1.42441
Trial number in the post test task 2	20	13.00	24.00	15.7000	3.11364
Valid N (listwise)	20				
a. treatment groups = FoS					

treatment groups = FoF

Descriptive Statistics[a]					
	N	Minimum	Maximum	Mean	Std. Deviation
Spre1	25	11.00	20.00	16.8800	2.38607
TSpre1	25	7.00	16.00	11.0400	2.42350
Trial number in the pre-test task 1	25	26.00	69.00	40.8000	11.04159
Spost2	25	14.00	25.00	21.0400	3.18172
TSpost2	25	9.00	16.00	13.5600	2.14243
Trial number in the post test task 2	25	21.00	37.00	29.6000	4.46281
Spre2	25	4.00	13.00	7.1200	2.24202
TSpre2	25	.00	3.00	.8800	1.09240
Trial number in the pre-test task 2	25	.00	39.00	20.9200	9.98716
Spost1	25	6.00	14.00	9.8000	2.17945
TSpost1	25	.00	5.00	1.6800	1.65126
Trial number in the post test task 2	25	13.00	29.00	19.1600	5.34384
Valid N (listwise)	25				
a. treatment groups = FoF					

treatment groups = FoM

Descriptive Statistics[a]					
	N	Minimum	Maximum	Mean	Std. Deviation
Spre1	26	12.00	20.00	16.3462	2.51304
TSpre1	26	5.00	15.00	9.2692	2.93336
Trial number in the pre-test task 1	26	27.00	58.00	40.0000	7.56307

Spost2	26	15.00	23.00	19.2692	2.16440
TSpost2	26	9.00	16.00	12.1154	2.02636
Trial number in the post test task 2	26	18.00	50.00	33.6538	8.26047
Spre2	26	2.00	12.00	7.9231	2.18949
TSpre2	26	.00	4.00	1.2692	1.21845
Trial number in the pre-test task 2	26	10.00	32.00	22.1154	7.07857
Spost1	26	4.00	13.00	8.9615	2.12567
TSpost1	26	.00	4.00	1.8462	1.54123
Trial number in the post test task 2	26	7.00	30.00	18.3462	6.75836
Valid N (listwise)	26				
a. treatment groups = FoM					

Oneway Drag and Drop communicative performance treatment groups

Descriptives								
Spost2								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
FoS	20	17.7000	4.47331	1.00026	15.6064	19.7936	9.00	23.00
FoF	25	21.0400	3.18172	.63634	19.7267	22.3533	14.00	25.00
FoM	26	19.2692	2.16440	.42447	18.3950	20.1434	15.00	23.00
Total	71	19.4507	3.51645	.41733	18.6184	20.2830	9.00	25.00

Test of Homogeneity of Variances			
Spost2			
Levene Statistic	df1	df2	Sig.
6.665	2	68	.002

ANOVA					
Spost2					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	125.302	2	62.651	5.755	.005
Within Groups	740.275	68	10.886		
Total	865.577	70			

Post Hoc Test

Multiple Comparisons							
Dependent Variable:Spost2							
						95% Confidence Interval	
	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Tukey HSD	FoS	FoF	-3.34000*	.98984	.003	-5.7117	-.9683
		FoM	-1.56923	.98134	.253	-3.9206	.7821
	FoF	FoS	3.34000*	.98984	.003	.9683	5.7117
		FoM	1.77077	.92421	.142	-.4437	3.9853
	FoM	FoS	1.56923	.98134	.253	-.7821	3.9206
		FoF	-1.77077	.92421	.142	-3.9853	.4437
Games-Howell	FoS	FoF	-3.34000*	1.18552	.022	-6.2483	-.4317
		FoM	-1.56923	1.08660	.334	-4.2704	1.1320
	FoF	FoS	3.34000*	1.18552	.022	.4317	6.2483
		FoM	1.77077	.76493	.065	-.0874	3.6290
	FoM	FoS	1.56923	1.08660	.334	-1.1320	4.2704
		FoF	-1.77077	.76493	.065	-3.6290	.0874
*. The mean difference is significant at the 0.05 level.							

Homogeneous Subsets

Spost2				
	treatment groups	N	Subset for alpha = 0.05	
			1	2
Tukey HSD[a,b]	FoS	20	17.7000	
	FoM	26	19.2692	19.2692
	FoF	25		21.0400
	Sig.		.242	.166
Means for groups in homogeneous subsets are displayed.				
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.				

NPar Tests Drag and drop communicative performance treatment groups
Kruskal-Wallis Test

Ranks			
	treatment groups	N	Mean Rank
Spost2	FoS	20	28.60
	FoF	25	45.28
	FoM	26	32.77
	Total	71	

Test Statistics[a,b]	
	Spost2
Chi-Square	8.339
df	2
Asymp. Sig.	.015
a. Kruskal Wallis Test b. Grouping Variable: treatment groups	

Oneway Drag and drop linguistic performance treatment groups

Descriptives								
TSpost2								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
FoS	20	12.2500	3.30669	.73940	10.7024	13.7976	7.00	16.00
FoF	25	13.5600	2.14243	.42849	12.6756	14.4444	9.00	16.00
FoM	26	12.1154	2.02636	.39740	11.2969	12.9339	9.00	16.00
Total	71	12.6620	2.54077	.30153	12.0606	13.2634	7.00	16.00

Test of Homogeneity of Variances			
TSpost2			
Levene Statistic	df1	df2	Sig.
7.728	2	68	.001

ANOVA					
TSpost2					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	31.323	2	15.662	2.532	.087
Within Groups	420.564	68	6.185		
Total	451.887	70			

Post Hoc Tests

Multiple Comparisons							
Dependent Variable:TSpost2							
	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	FoS	FoF	-1.31000	.74608	.192	-3.0977	.4777
		FoM	.13462	.73967	.982	-1.6377	1.9069
	FoF	FoS	1.31000	.74608	.192	-.4777	3.0977
		FoM	1.44462	.69661	.103	-.2245	3.1138
	FoM	FoS	-.13462	.73967	.982	-1.9069	1.6377
		FoF	-1.44462	.69661	.103	-3.1138	.2245
Games-Howell	FoS	FoF	-1.31000	.85458	.290	-3.4129	.7929
		FoM	.13462	.83943	.986	-1.9359	2.2052
	FoF	FoS	1.31000	.85458	.290	-.7929	3.4129
		FoM	1.44462*	.58440	.044	.0318	2.8575
	FoM	FoS	-.13462	.83943	.986	-2.2052	1.9359
		FoF	-1.44462*	.58440	.044	-2.8575	-.0318
*. The mean difference is significant at the 0.05 level.							

Homogeneous Subsets

TSpost2			
	treatment groups	N	Subset for alpha = 0.05 1
Tukey HSD[a,b]	FoM	26	12.1154
	FoS	20	12.2500
	FoF	25	13.5600
	Sig.		.124
Means for groups in homogeneous subsets are displayed.			
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.			

NPar Tests Drag and drop linguistic performance treatment groups
Kruskal-Wallis Test

Ranks			
	treatment groups	N	Mean Rank
TSpost2	FoS	20	34.08
	FoF	25	42.96
	FoM	26	30.79
	Total	71	

Test Statistics[a,b]	
	TSpost2
Chi-Square	4.760
df	2
Asymp. Sig.	.093
a. Kruskal Wallis Test b. Grouping Variable: treatment groups	

Oneway information gap communicative performance treatment groups

Descriptives
Spost1

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
FoS	20	8.5500	2.74293	.61334	7.2663	9.8337	3.00	14.00
FoF	25	9.8000	2.17945	.43589	8.9004	10.6996	6.00	14.00
FoM	26	8.9615	2.12567	.41688	8.1030	9.8201	4.00	13.00
Total	71	9.1408	2.35612	.27962	8.5832	9.6985	3.00	14.00

Test of Homogeneity of Variances			
Spost1			
Levene Statistic	df1	df2	Sig.
1.130	2	68	.329

ANOVA					
Spost1					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	18.680	2	9.340	1.717	.187
Within Groups	369.912	68	5.440		
Total	388.592	70			

Post Hoc Tests

Multiple Comparisons							
Dependent Variable:Spost1							
	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	FoS	FoF	-1.25000	.69971	.182	-2.9266	.4266
		FoM	-.41154	.69370	.824	-2.0737	1.2506
	FoF	FoS	1.25000	.69971	.182	-.4266	2.9266
		FoM	.83846	.65332	.409	-.7269	2.4039
	FoM	FoS	.41154	.69370	.824	-1.2506	2.0737
		FoF	-.83846	.65332	.409	-2.4039	.7269
Games-Howell	FoS	FoF	-1.25000	.75245	.234	-3.0896	.5896
		FoM	-.41154	.74160	.845	-2.2266	1.4035
	FoF	FoS	1.25000	.75245	.234	-.5896	3.0896
		FoM	.83846	.60315	.354	-.6195	2.2964
	FoM	FoS	.41154	.74160	.845	-1.4035	2.2266
		FoF	-.83846	.60315	.354	-2.2964	.6195

Homogeneous Subsets

Spost1			
			Subset for alpha = 0.05
	treatment groups	N	1
Tukey HSD[a,b]	FoS	20	8.5500
	FoM	26	8.9615
	FoF	25	9.8000
	Sig.		.167
Means for groups in homogeneous subsets are displayed.			
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.			

Oneway

Descriptives							
TSpost1							
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for	Minimum	Maximum

					Mean			
					Lower Bound	Upper Bound		
FoS	20	2.1500	1.42441	.31851	1.4834	2.8166	.00	5.00
FoF	25	1.6800	1.65126	.33025	.9984	2.3616	.00	5.00
FoM	26	1.8462	1.54123	.30226	1.2236	2.4687	.00	4.00
Total	71	1.8732	1.53929	.18268	1.5089	2.2376	.00	5.00

Test of Homogeneity of Variances			
TSpost1			
Levene Statistic	df1	df2	Sig.
.735	2	68	.483

ANOVA					
TSpost1					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.485	2	1.242	.517	.599
Within Groups	163.375	68	2.403		
Total	165.859	70			

Post Hoc Tests

Multiple Comparisons							
Dependent Variable:TSpost1							
	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Tukey HSD	FoS	FoF	.47000	.46501	.573	-.6442	1.5842
		FoM	.30385	.46102	.788	-.8008	1.4085
	FoF	FoS	-.47000	.46501	.573	-1.5842	.6442
		FoM	-.16615	.43418	.923	-1.2065	.8742
	FoM	FoS	-.30385	.46102	.788	-1.4085	.8008
		FoF	.16615	.43418	.923	-.8742	1.2065
Games-Howell	FoS	FoF	.47000	.45882	.566	-.6440	1.5840
		FoM	.30385	.43910	.769	-.7625	1.3702
	FoF	FoS	-.47000	.45882	.566	-1.5840	.6440
		FoM	-.16615	.44769	.927	-1.2486	.9163
	FoM	FoS	-.30385	.43910	.769	-1.3702	.7625
		FoF	.16615	.44769	.927	-.9163	1.2486

Homogeneous Subsets

TSpost1			
	treatment groups	N	Subset for alpha = 0.05
			1
Tukey HSD[a,b]	FoF	25	1.6800
	FoM	26	1.8462
	FoS	20	2.1500
	Sig.		.557
Means for groups in homogeneous subsets are displayed.			
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.			

Descriptives across all task variables

Descriptive Statistics			
	N	Mean	Std. Deviation
Spre1	71	16.7042	3.27935
Spost2	71	19.4507	3.51645

TSpre1	71	9.9859	3.01659
TSpost2	71	12.6620	2.54077
Spre2	71	7.5070	2.37170
Spost1	71	9.1408	2.35612
TSpre2	71	1.0563	1.11979
TSpost1	71	1.8732	1.53929
Trial number in the pre- test task 1	71	39.7183	8.46030
Trial number in the post test task 2	71	30.6197	6.57781
Trial number in the pre- test task 2	71	21.7324	7.93178
Trial number in the post test task 2	71	17.8873	5.54604
Valid N (listwise)	71		

T-Test across all treatment groups exploring task type and treatment effect

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Spost2	19.4507	71	3.51645	.41733
	Spre1	16.7042	71	3.27935	.38919
Pair 2	Spost1	9.1408	71	2.35612	.27962
	Spre2	7.5070	71	2.37170	.28147
Pair 3	TSpost2	12.6620	71	2.54077	.30153
	TSpre1	9.9859	71	3.01659	.35800
Pair 4	TSpost1	1.8732	71	1.53929	.18268
	TSpre2	1.0563	71	1.11979	.13289

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Spost2 and Spre1	71	.808	.000
Pair 2	Spost1 and Spre2	71	.762	.000
Pair 3	TSpost2 and TSpre1	71	.758	.000
Pair 4	TSpost1 and TSpre2	71	.841	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Spost2 - Spre1	2.74648	2.11605	.25113	2.24562	3.24734	10.937	70	.000
Pair 2	Spost1 - Spre2	1.63380	1.63217	.19370	1.24747	2.02013	8.435	70	.000
Pair 3	TSpost2 - TSpre1	2.67606	1.98404	.23546	2.20644	3.14567	11.365	70	.000
Pair 4	TSpost1 - TSpre2	.81690	.85038	.10092	.61562	1.01818	8.094	70	.000

Descriptives

Descriptive Statistics

	N	Mean	Std. Deviation
Spre1	71	16.7042	3.27935
TSpre1	71	9.9859	3.01659
Trial number in the pre- test task 1	71	39.7183	8.46030
Spost2	71	19.4507	3.51645
TSpost2	71	12.6620	2.54077
Trial number in the post test task 2	71	30.6197	6.57781
Spre2	71	7.5070	2.37170

TSpre2	71	1.0563	1.11979
Trial number in the pre- test task 2	71	21.7324	7.93178
Spost1	71	9.1408	2.35612
TSpost1	71	1.8732	1.53929
Trial number in the post test task 2	71	17.8873	5.54604
Valid N (listwise)	71		

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Trial number in the post test task 2	30.6197	71	6.57781	.78064
	Trial number in the pre-test task 1	39.7183	71	8.46030	1.00405
Pair 2	Trial number in the post test task 1	17.8873	71	5.54604	.65819
	Trial number in the pre- test task 2	21.7324	71	7.93178	.94133

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Trial number in the post test task 2 and Trial number in the pre- test task 1	71	.314	.008
Pair 2	Trial number in the post test task 1 and Trial number in the pre- test task 2	71	.179	.135

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Trial number in the post test task 2 - Trial number in the pre- test task 1	-9.09859	8.93813	1.06076	-11.21421	-6.98297	-8.577	70	.000
Pair 2	Trial number in the post test task 1 - Trial number in the pre- test task 2	-3.84507	8.82634	1.04749	-5.93423	-1.75591	-3.671	70	.000

Oneway

Descriptives

Trial number in the post test task 2

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Minimum	Maximum
FoS	20	27.9500	4.79556	1.07232	25.7056	30.1944	17.00	36.00
FoF	25	29.6000	4.46281	.89256	27.7578	31.4422	21.00	37.00
FoM	26	33.6538	8.26047	1.62001	30.3174	36.9903	18.00	50.00
Total	71	30.6197	6.57781	.78064	29.0628	32.1767	17.00	50.00

Test of Homogeneity of Variances

Trial number in the post test task 2

Levene Statistic	df1	df2	Sig.
5.639	2	68	.005

ANOVA

Trial number in the post test task 2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	407.898	2	203.949	5.292	.007
Within Groups	2620.835	68	38.542		
Total	3028.732	70			

Post Hoc Tests

Multiple Comparisons							
Dependent Variable:Trial number in the post test task 2							
	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Tukey HSD	FoS	FoF	-1.65000	1.86246	.651	-6.1126	2.8126
		FoM	-5.70385*	1.84647	.008	-10.1281	-1.2795
	FoF	FoS	1.65000	1.86246	.651	-2.8126	6.1126
		FoM	-4.05385	1.73898	.058	-8.2206	.1129
	FoM	FoS	5.70385*	1.84647	.008	1.2795	10.1281
		FoF	4.05385	1.73898	.058	-.1129	8.2206
Games-Howell	FoS	FoF	-1.65000	1.39518	.470	-5.0476	1.7476
		FoM	-5.70385*	1.94276	.015	-10.4268	-.9809
	FoF	FoS	1.65000	1.39518	.470	-1.7476	5.0476
		FoM	-4.05385	1.84962	.085	-8.5612	.4535
	FoM	FoS	5.70385*	1.94276	.015	.9809	10.4268
		FoF	4.05385	1.84962	.085	-.4535	8.5612
*. The mean difference is significant at the 0.05 level.							

Homogeneous Subsets

Trial number in the post test task 2				
	treatment groups	N	Subset for alpha = 0.05 1	2
Tukey HSD[a,b]	FoS	20	27.9500	
	FoF	25	29.6000	29.6000
	FoM	26		33.6538
	Sig.		.637	.073
Means for groups in homogeneous subsets are displayed.				
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.				

NPar Tests
Kruskal-Wallis Test

Ranks			
	treatment groups	N	Mean Rank
Trial number in the post test task 2	FoS	20	28.10
	FoF	25	34.48
	FoM	26	43.54
	Total	71	

Test Statistics[a,b]	
	Trial number in the post test task 2
Chi-Square	6.570
df	2
Asymp. Sig.	.037
a. Kruskal Wallis Test b. Grouping Variable: treatment groups	

Oneway

Descriptives								
Trial number in the post test task 1								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
FoS	20	15.7000	3.11364	.69623	14.2428	17.1572	13.00	24.00
FoF	25	19.1600	5.34384	1.06877	16.9542	21.3658	13.00	29.00
FoM	26	18.3462	6.75836	1.32542	15.6164	21.0759	7.00	30.00
Total	71	17.8873	5.54604	.65819	16.5746	19.2001	7.00	30.00

Test of Homogeneity of Variances			
Trial number in the post test task 1			
Levene Statistic	df1	df2	Sig.
11.119	2	68	.000

ANOVA					
Trial number in the post test task 1					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	141.654	2	70.827	2.394	.099
Within Groups	2011.445	68	29.580		
Total	2153.099	70			

Post Hoc Tests

Multiple Comparisons							
Dependent Variable:Trial number in the post test task 1							
						95% Confidence Interval	
	(I) treatment groups	(J) treatment groups	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Tukey HSD	FoS	FoF	-3.46000	1.63163	.093	-7.3695	.4495
		FoM	-2.64615	1.61762	.238	-6.5221	1.2298
	FoF	FoS	3.46000	1.63163	.093	-.4495	7.3695
		FoM	.81385	1.52345	.855	-2.8365	4.4642
	FoM	FoS	2.64615	1.61762	.238	-1.2298	6.5221
		FoF	-.81385	1.52345	.855	-4.4642	2.8365
Games-Howell	FoS	FoF	-3.46000*	1.27554	.026	-6.5656	-.3544
		FoM	-2.64615	1.49716	.195	-6.3015	1.0092
	FoF	FoS	3.46000*	1.27554	.026	.3544	6.5656
		FoM	.81385	1.70265	.882	-3.3060	4.9337
	FoM	FoS	2.64615	1.49716	.195	-1.0092	6.3015
		FoF	-.81385	1.70265	.882	-4.9337	3.3060
*. The mean difference is significant at the 0.05 level.							

Homogeneous Subsets

Trial number in the post test task 1			
			Subset for alpha = 0.05
	treatment groups	N	1
Tukey HSD[a,b]	FoS	20	15.7000
	FoM	26	18.3462
	FoF	25	19.1600
	Sig.		.083
Means for groups in homogeneous subsets are displayed.			
a. Uses Harmonic Mean Sample Size = 23.353. b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.			

NPar Tests
Kruskal-Wallis Test

Ranks			
	treatment groups	N	Mean Rank
Trial number in the post test task 1	FoS	20	28.88
	FoF	25	41.64
	FoM	26	36.06
	Total	71	

Test Statistics[a,b]	
	Trial number in the post test task 1
Chi-Square	4.290
df	2
Asymp. Sig.	.117
a. Kruskal Wallis Test b. Grouping Variable: treatment groups	

Appendix IX Modality of input analysis

Crosstabs chi square comparing input groups with modality of input group

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
inputtype * Audiopre	71	100.0%	0	.0%	71	100.0%

inputtype * Audiopre Crosstabulation					
			Audiopre		
			Didn't listen	Listened	Total
inputtype	FoS	Count	8	12	20
		Expected Count	7.3	12.7	20.0
		% within inputtype	40.0%	60.0%	100.0%
	FoF	Count	11	14	25
		Expected Count	9.2	15.8	25.0
		% within inputtype	44.0%	56.0%	100.0%
	FoM	Count	7	19	26
		Expected Count	9.5	16.5	26.0
		% within inputtype	26.9%	73.1%	100.0%
Total		Count	26	45	71
		Expected Count	26.0	45.0	71.0
		% within inputtype	36.6%	63.4%	100.0%

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.738[a]	2	.419
Likelihood Ratio	1.773	2	.412
Linear-by-Linear Association	.954	1	.329
N of Valid Cases	71		
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.32.			

Crosstabs chi square comparing input groups to access to audio post-test

inputtype * Audiopost Crosstabulation					
			Audiopost		
			Didn't listen	Listened	Total
inputtype	FoS	Count	13	7	20
		Expected Count	13.2	6.8	20.0
		% within inputtype	65.0%	35.0%	100.0%
	FoF	Count	18	7	25
		Expected Count	16.5	8.5	25.0
		% within inputtype	72.0%	28.0%	100.0%
	FoM	Count	16	10	26
		Expected Count	17.2	8.8	26.0
		% within inputtype	61.5%	38.5%	100.0%
Total		Count	47	24	71
		Expected Count	47.0	24.0	71.0
		% within inputtype	66.2%	33.8%	100.0%

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	.641[a]	2	.726
Likelihood Ratio	.648	2	.723
Linear-by-Linear Association	.091	1	.762
N of Valid Cases	71		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.76.

T-Test comparing the modality of input in pre-test groups to performance

Group Statistics					
	Audiopre	N	Mean	Std. Deviation	Std. Error Mean
sposttest	Didn't listen	26	26.9615	5.42203	1.06335
	Listened	45	30.0889	4.18776	.62428

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
sposttest	Equal variances assumed	3.021	.087	-2.717	69	.008	-3.12735	1.15110	-5.42372	-.83098
	Equal variances not assumed			-2.536	42.345	.015	-3.12735	1.23306	-5.61516	-.63954

T-Test comparing the modality of input in post-test groups to performance

Group Statistics					
	Audiopost	N	Mean	Std. Deviation	Std. Error Mean
sposttest	Didn't listen	47	29.0851	4.80412	.70075
	Listened	24	28.6667	5.12171	1.04546

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
sposttest	Equal variances assumed	.043	.835	.340	69	.735	.41844	1.23241	-2.04016	2.87704
	Equal variances not assumed			.332	43.881	.741	.41844	1.25859	-2.11828	2.95516

T-Test comparing the modality of input in pre-test to linguistic performance

Group Statistics					
	Audiopre	N	Mean	Std. Deviation	Std. Error Mean
TSpostest	Didn't listen	26	13.7308	3.56155	.69848
	Listened	45	15.0000	2.96188	.44153

Independent Samples Test		
	Levene's Test for Equality of Variances	t-test for Equality of Means

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
TSpostest	Equal variances assumed	2.463	.121	-1.614	69	.111	-1.26923	.78637	-2.83799	.29953
	Equal variances not assumed			-1.536	44.898	.132	-1.26923	.82633	-2.93365	.39519

T-Test

Group Statistics					
	Audiopost	N	Mean	Std. Deviation	Std. Error Mean
TSpostest	Didn't listen	47	14.8511	3.26363	.47605
	Listened	24	13.9167	3.13350	.63962

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
TSpostest	Equal variances assumed	.524	.472	1.156	69	.252	.93440	.80806	-.67764	2.54643
	Equal variances not assumed			1.172	48.151	.247	.93440	.79733	-.66862	2.53741

T-Test Treatment effect Modality of input

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Access to audio file during post test	.3380	71	.47641	.05654
	Access to audio file during pre-test	.6338	71	.48519	.05758

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Access to audio file during post test and Access to audio file during pre-test	71	.296	.012

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Access to audio file during post test - Access to audio file during pre- test	-.29577	.57057	.06771	-.43083	-.16072	-4.368	70	.000

Appendix X Grammar access analysis

Crosstabs grammar access descriptives

treatment groups * GramAcc2 Crosstabulation					
			GramAcc2		Total
			.00	1.00	
treatment groups	FoS	Count	8	12	20
		% within treatment groups	40.0%	60.0%	100.0%
	FoF	Count	11	14	25
		% within treatment groups	44.0%	56.0%	100.0%
Total		Count	19	26	45
		% within treatment groups	42.2%	57.8%	100.0%

Chi-Square Tests					
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.073[a]	1	.787		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.073	1	.787		
Fisher's Exact Test				1.000	.514
Linear-by-Linear Association	.071	1	.790		
N of Valid Cases	45				
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.44. b. Computed only for a 2x2 table					

T-Test grammar access groups and communicative performance

Group Statistics					
	GramAcc2	N	Mean	Std. Deviation	Std. Error Mean
overall communicative performance score in the posttest	.00	19	28.7895	5.87442	1.34768
	1.00	26	28.8077	5.57149	1.09266

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
overall communicative performance score in the posttest	Equal variances assumed	.003	.958	-.011	43	.992	-.01822	1.72043	-3.48780	3.45136
	Equal variances not assumed			-.011	37.710	.992	-.01822	1.73498	-3.53139	3.49495

T-Test grammar access groups and communicative performance

Group Statistics					
	GramAcc2	N	Mean	Std. Deviation	Std. Error Mean
Target linguistic performance score in the post test	.00	19	14.3684	3.70001	.84884
	1.00	26	15.2308	3.51349	.68905

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Target linguistic performance score in the post test	Equal variances assumed	.015	.903	-.795	43	.431	-.86235	1.08435	-3.04915	1.32445
	Equal variances not assumed			-.789	37.739	.435	-.86235	1.09331	-3.07614	1.35144

T-Test grammar access groups and trial number

Group Statistics					
	GramAcc2	N	Mean	Std. Deviation	Std. Error Mean
trial number for post test	.00	19	47.5263	7.65483	1.75614
	1.00	26	45.7308	7.10244	1.39290

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
trial number for post test	Equal variances assumed	.332	.568	.811	43	.422	1.79555	2.21495	-2.67133	6.26242
	Equal variances not assumed			.801	37.178	.428	1.79555	2.24147	-2.74538	6.33647

Descriptives time spent on grammar instruction

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
time spent on main grammar by FoS group	20	8.00	283.00	95.6500	75.20937
Valid N (listwise)	20				

Explore time spent on main grammar instruction.

Descriptives				
			Statistic	Std. Error
time spent on main grammar by FoS group	Mean		95.6500	16.81733
	95% Confidence Interval for Mean	Lower Bound	60.4509	
		Upper Bound	130.8491	
	5% Trimmed Mean		90.1111	

	Median	93.0000	
	Variance	5656.450	
	Std. Deviation	75.20937	
	Minimum	8.00	
	Maximum	283.00	
	Range	275.00	
	Interquartile Range	104.25	
	Skewness	.840	.512
	Kurtosis	.391	.992

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
time spent on main grammar by FoS group	.160	20	.192	.922	20	.107
a. Lilliefors Significance Correction						

Time spent on main grammar by FoS group
Correlations time spent on grammar instruction vs. communicative performance

Descriptive Statistics			
	Mean	Std. Deviation	N
time spent on main grammar by FoS group	95.6500	75.20937	20
overall communicative performance score in the posttest	26.2500	5.99890	20

Correlations			
		time spent on main grammar by FoS group	overall communicative performance score in the posttest
time spent on main grammar by FoS group	Pearson Correlation	1	-.442
	Sig. (2-tailed)		.051
	N	20	20
overall communicative performance score in the posttest	Pearson Correlation	-.442	1
	Sig. (2-tailed)	.051	
	N	20	20

Correlations time spent on grammar instruction with linguisitc performance

Descriptive Statistics			
	Mean	Std. Deviation	N
time spent on main grammar by FoS group	95.6500	75.20937	20
Target linguistic performance score in the post test	14.4000	4.13458	20

Correlations			
		time spent on main grammar by FoS group	Target linguistic performance score in the post test
time spent on main grammar by FoS group	Pearson Correlation	1	-.467*
	Sig. (2-tailed)		.038
	N	20	20
Target linguistic performance score in the post test	Pearson Correlation	-.467*	1
	Sig. (2-tailed)	.038	
	N	20	20
*. Correlation is significant at the 0.05 level (2-tailed).			

Correlations

Descriptive Statistics			
	Mean	Std. Deviation	N
time spent on main grammar by FoS group	95.6500	75.20937	20
trial number for post test	43.6500	6.17529	20

Correlations			
		time spent on main grammar by FoS group	trial number for post test
time spent on main grammar by FoS group	Pearson Correlation	1	-.083
	Sig. (2-tailed)		.726
	N	20	20
trial number for post test	Pearson Correlation	-.083	1
	Sig. (2-tailed)	.726	
	N	20	20